Advances in Integrated Design and Manufacturing in Mechanical Engineering

Edited by
Alan Bramley
Daniel Brissaud
Daniel Coutellier
Chris McMahon

Springer

ADVANCES IN INTEGRATED DESIGN AND MANUFACTURING IN MECHANICAL ENGINEERING

Advances in Integrated Design and Manufacturing in Mechanical Engineering

Edited by

ALAN BRAMLEY
University of Bath, Bath, UK

DANIEL BRISSAUD
University of Grenoble, Grenoble, France

DANIEL COUTELLIER
University of Valenciennes, Valenciennes, France

and

CHRIS McMAHON
University of Bath, Bath, UK

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-3481-4 (HB)
ISBN-10 1-4020-3482-2 (e-book)
ISBN-13 978-1-4020-3481-7 (HB)
ISBN-13 978-1-4020-3482-4 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.
www.springeronline.com

Printed on acid-free paper

Printed in the Netherlands.

Contents

KEYNOTES

PART 1: DESIGN STRATEGIES AND METHODOLOGIES

PART 2: INTEGRATED DESIGN OF MANUFACTURING PROCESSES

PART 3: DESIGN TOOLS FOR PARTICULAR APPLICATIONS

Introduction

This book presents a selection of papers related to the fifth edition of book further to the International Conference on Integrated Design and Manufacturing in Mechanical Engineering. This Conference has been organized within the framework of the activities of the AIP-PRIMECA network whose main scientific field is Integrated Design applied to both Mechanical Engineering and Productics. This network is organized along the lines of a joint project: the evolution, in the field of training of Integrated Design in Mechanics and Productics, in quite close connection with the ever changing industrial needs over the past 20 years. It is in charge of promoting both exchanges of experience and know-how capitalisation. It has a paramount mission to fulfil, be it in the field of initial and continuous education, technological transfer and knowledge dissemination through strong links with research labs.

For the second time, in fact, the IDMME Conference has been held abroad and, after Canada in 2000, the United Kingdom, more particularly Bath University, has been retained under the responsibility of Professor Alan Bramley, the Chairman of the Scientific Committee of the conference. The Scientific Committee members have selected all the lectures from complete papers, which is the guarantee for the Conference of quite an outstanding scientific level. After that, a new selection has been carried out to retain the best publications, which establish in a book, a state-of-the-art analysis as regards Integrated Design and Manufacturing in the discipline of Mechanical Engineering.

The 33 papers presented in the book, were selected from the 138 papers presented during the Conference.

The unique focus of this series of conferences is the close symbiotic relationship between design and manufacture and it is within this overall focus that the following themes were highlighted during the conference:

Integrated design of manufacturing processes

The manufacturing process or system has to be designed and optimised to afford benefit. The aim is to ensure a high quality product meeting all requirements, rapidly and at optimum cost. In this stream we consider some of the implications of cutting and forming processes for this optimisation.

The Design/Manufacturing interface

In this stream we will discuss Design for Manufacture (DFM) and the importance of geometric modelling and virtual reality (VR) in this area. It has also become apparent that handling tolerances and inspection is still of critical importance. The issue of process planning is also dealt with.

Design models and tools

The designer has to cope with a considerable number of inputs to ensure that the product, machine or system being developed is optimal. This topic area deals with a number of the support strategies that are available at a process level, a design support tool level, and at a more detailed representational level.

Design Structures and Approaches

Increasingly all aspects of the product development and delivery process are undertaken in a widely dispersed manner. This applies to design and development teams, suppliers who both design and manufacture, and manufacturing operations, which can be anywhere. It is critical that underlying structures are in place. Thus this topic area will deal with the critical element of knowledge in design and manufacturing processes and the methods with which designing can be undertaken collaboratively in an integrated manner.

Innovation/Change and Flexibility

The key demand for the 21st century is that customers, however defined, require what they want when they want it. In this topic area, in addition to the main focus of flexibility and agility, the issues of change and change management has emerged as an important focus, along with how design understanding emerges and interacts.

The book is divided into sections reflecting the above themes and will be of interest to academics, students and practitioners specialising in design and manufacturing issues in mechanical engineering. We hope that you will find it of the greatest interest to compare your various points of view within the fields broached throughout the Conference. I hope you all enjoy reading this book, which aims to be a reference textbook for all researchers in this particular field and for the teaching staff confronted with training methodologies in integrated design and productics. It will allow you to assess the scope of the development prospects in an extremely wide ranging field.

In a practical way the texts selected have been divided in 3 main sections to cover the all topics developed in the conference. 2 keynote presentations were written in texts to present the general synergies and relationships between the design and manufacturing processes through the concepts of production knowledge and integration. Part 1deals with **Design strategies and methodologies**, Part 2 with **Integrated design of manufacturing processes** and Part 3 with **Design tools for particular applications**. It covers a large number of actual issues in the field by authors from the main research groups involved in developing methods, models and tools for the improvement of design and manufacturing processes in companies.

We would like to highlight the very significant input of the members of the organising committee for the success of the conference: Mrs Libby Brodhurst from the Institution of Engineering Designers, Steve Culley, Dr Elies Dekoninck, Chris McMahon, Professor Tony Medland, Professor Tony Mileham, Dr Linda Newnes and Dr Geraint Owen from the University of Bath, UK.

Alan Bramley,

Daniel Brissaud,

Daniel Coutellier,

Chris McMahon.

AIP-PRIMECA network

is a nation-wide structure

In 2001, under the impulse of the Ministry of National Education, the AIP-PRIMECA competence network was created in the fields of Integrated Design in Mechanics and Productics. It results from the merger between the AIPs (Inter-Institute Productics Workshops) - regional resource centres used as experimental support of in-depth training courses in the field of Productics, set up in 1984 and initiated by The Ministry of National Education, in cooperation with the Ministry of Industry and with the regions concerned - and PRIMECA (Computer Science Resource Pole for Mechanics), created in 1991 on the instigation of the Science Academy Application Committee in order to promote the use of computer tools in the design of mechanical products and thus to create a training channel in Computer Aided Mechanical Design.

Organized around resource poles, the local missions give impetus to training in the fields of Integrated Design in Mechanics and Productics. The association of pedagogical and scientific competences thus generates a training area around the themes of integrated engineering, favouring the setting up not only of in-depth teaching modules in disciplines but also of inter-disciplinary teaching packages aimed at training the industrial world future executives. These are transfer centres of research to training bodies and companies, by using industrial scale equipments.

The will to mutualize knowledge, strengthen the competence of all the poles and create strong synergies between the different actors of the network was concretized by the creation of an open or virtual institute which gives impetus to innovative pedagogical developments often associated with the new information and communication technologies, thus facilitating the nation-wide dissemination of the Poles' scientific specificities as well as the participation in Numerical Campus projects.

The principal missions of the AIP-PRIMECA network are the training. In fact, within each pole, means of an industrial type permits to train specialists

* Cognisant with an overall approach of the integrated design and production system
* Mastering the methods and techniques of modelling, analysis, production, computerization, automation
* Able to get involved into multidisciplinary activities

* Perfectly aware of the necessary coherence within the company's different functions

The AIP-PRIMECA network is open to the industrial world, more particularly :

* By organizing national or international events (IDMME) accessible to a large public which allow to maintain technological lookout.
* By making available feasibility studies facilities.
* By organizing thematic day-meetings with professionals
* By offering industrial scale platforms aimed at research cooperation schemes.
* By proposing specific training courses in high technology fields.

The AIP-PRIMECA centres are competences and resources in the fields of Integrated Design in Mechanics and Productics to be used for education, research and industry. The 9 centers are: Auvergne, Dauphiné-Savoie, Franche-Comté, Ile de France, Lorraine, Pays de la Loire, Nord-Pas de Calais, Rhône-Alpes Ouest, Toulouse.

Professor Daniel Coutellier,
Head of AIP-PRIMECA network

Keynotes

A VIRTUAL RESEARCH LAB FOR A KNOWLEDGE COMMUNITY IN PRODUCTION (VRL-KCIP)

Serge Tichkiewitch
3S Laboratory - Domaine Universitaire, BP°53, 38041 Grenoble Cedex 9
phone: +33-476825141, fax: +33-476827043, Serge.Tichkiewitch@inpg.fr

Abstract: The globalization of the production systems will result in Europe losing its know-how in production technologies. However, high quality & low cost products, accommodating sustainable development and the need for shorter lead times can only be obtained by integrating the production phase into the design stage. It has also been requested to integrate new production technologies in order to design innovative products. This can be achieved by implementing a virtual production system.

Through the VRL-KciP network of excellence, Europe will be able to federate many research teams in the field of design, production and innovation and thus preserve the common inheritance of manufacturing processes while carrying out a collaborative system of design integrating a module of virtual production.

20 research units, gathering 200 researchers, have decided to work together in order to provide a strong integration of their teams and thus break the fragmentation of European research in this area and work more efficiently. A new organization of the research system must lead to a clearer view of the research orientations, research quality improvement by making the most of complementarities, avoiding duplication, and finally to a better impact on our industry for the welfare of our society

Key words: Virtual Research Lab; Knowledge Community; Network of Excellence; Innovation; Production.

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 3-16.

1. INTRODUCTION

Within the 6th Framework Program, the European Commission has launched new instruments called "Integrated Project" (IP) and "Network of Excellence" (NoE). With the common support of the "Information Society Technologies" (IT) [1] and the "Nano-technologies and nano-sciences, Knowledge based multifunctional materials, and new production processes and devices" (NMP) [2] priorities, a NoE called "Virtual Research Lab for a Knowledge Community in Production" (VRL-KCiP) aims at linking up 20 research units thus gathering some 200 researchers [see Table 1].

2. NOE'S OBJECTIVES

The purpose of establishing the VRL-KCiP NoE is to reduce the research fragmentation in the field of production technologies and to bring about a network-based approach to avoid effort duplication in modeling and manufacturing processes simulation by joint partners. Our objective is to support dynamic organizations, inter-enterprise operability, and necessary standardization.

Incorporation of advances in virtual production, supply chain and life-cycle management, interactive decision-aid systems, development and rapid manufacturing will be the VRL-KCiP network's driving force. We also aim at benefiting from the different approaches of the multi-cultural teams towards common manufacturing problems and to promote successful technology transfer. This should be achieved through the incorporation of emerging technologies driving new production paradigms in all phases of the complete/extended value-chain (design, production, distribution, use, maintenance, recycling) to allow development of new knowledge-based, added value and quality products and services in traditional sectors.

Europe has developed a strong industrial base in particular in the field **of high added value products manufacturing and services** that have provided a successful economy up to now. Europe has demonstrated that innovative design of products, advanced production processes and the right marketing startegiescan give a **competitive edge to the economy**. However, in the context of the new global marketplace and the rapidly changing environment, in order to remain competitive, Europe must keep up-to-date with the latest advances in design methods and production techniques, translate its research results into economic value, build up a critical mass of expertise in order to face the competition from other industrialized countries, and still concentrate on developing the high added-value products that offer greater profit margins.

European production systems have been under constant change over the past 20 years because they have had to adapt to two major factors:

- Globalization of the economy
- Need for industrial innovation

Concerning the first factor, more than 85% of the parts constituting a computer are from Asian origin, which covers the whole keyboard. Actually, the candidate countries on the waiting list to join the EU have tended to specialize in low-cost production – a move reflected in limited production transfers from the current Member States to the candidate countries, which made it possible to retain activities in Europe that might otherwise have been re-located outside Europe. In any case, the **traditional manufacturing processes are now well-established outside European countries** and associated technologies knowledge will also be lost in the near future if nothing is done.

In relation to the second factor, the innovation of products mainly lies with **technology transfer**, by using unusual materials or non-traditional manufacturing processes to obtain the parts making up the product. In innovation using new processes, large companies must seek co-partners, by using the competences of the latter in specific manufacturing processes and by integrating these competences during the design process. For example, in France when a car manufacturer wanted to introduce the People-Carrier concept, the first batch quantities did not allow the use of a traditional body made in steel and it was necessary to find specialists in composite materials manufacture. Here, innovation is the result of integration of new partners in the co-design chain.

As a result of these two factors, policy makers have not paid enough attention to manufacturing, assuming, often wrongly, that in the knowledge economy and the information and service companies manufacturing industries are no longer playing a key role. At the same time, thanks to productivity gains, the relative price of manufactured products has been constantly declining. In order for Europe to maintain a leading position in an ever-competitive, constantly innovating world marketplace, the VRL-KCiP wants to develop collaborative technologies and methodologies for extended service and product development approaches, including associated services and distributed global manufacturing organization. EC funding should hopefully help us integrate, in a global context, our fragmented European laboratories and give us access to the international R&D efforts in product and process design, in order to focus on new holistic product/service concepts

In order to keep the key production competences in Europe, it is paramount that the specialists be associated in a network allowing them to

share the competences and the knowledge which they still hold, giving them the means of integrating these competences in new design systems. The creation of the VRL-KCiP NoE is an answer to this problem, bringing together the research **specialists in production processes** in order to share and utilise their knowledge within a common structure.

The manufacturing phase decentralization must be accompanied by an **organizational change**, making it possible to remain in control of the parameters: cost, delivery , quality, service, customer and recycling, either by integrating the manufacturing process modeling and simulation during design, or by integrating the collaborating companies in the context of the extended enterprise. This dual integration entails obvious new **communication and information technology challenges** on the one hand and on the other, **sharing cultures, competences, responsibilities and risks** of a new order. Only focused, collaborative research is today able to bring to the European enterprises the new tools and organizations necessary for the successful resolution of these problems.

How can Europe preserve and still enhance its key competences?

- Establish a European **virtual laboratory of excellence** in design integration for sustainable production with a durable structure.
- Create an Integrated **European task force** in production technology research.
- Co-ordinate and **orientate European production research** according to society's needs.
- Create lasting **synergies** between researchers from different disciplines in order to overcome the technical challenges of production engineering.
- Promote a **common language** among scientific groups in different countries by developing common procedures and tests.
- Create a European dimension of **research careers** in design integration for production engineering .
- Establish permanent **interactions** between researchers, industries and end-users and accelerate new discoveries and industrial applications for the extended enterprise.
- Initiate a **new training and educational program** in order to support the development of the future generation of European researchers in production engineering.
- Provide a **permanent vehicle** for the interplay **between scientists, engineers and students** and the expansion of the overall scientific understanding.
- Foster **international** information **exchange** on the development and application of production engineering methods and tools for the benefit of all.

The virtual laboratory of excellence will be a legal structure allowing the partners to construct a European laboratory with decentralized teams, sharing research strategies, knowledge and resources, responsibilities, rights and duties, and able to contract with industry. This ideal European production research organization will permit a better research output than today, leading to a clearer view of the research orientations and to a better overall quality of research.

2.1 Addressing the scientific, technical, socio-economic and policy objectives

One of the main objectives of the NoE VRL-KCiP proposal is to create a collaborative integrated design platform allowing the different members of the network to participate either in a synchronous or asynchronous mode in collective design projects. Each member will bring in the knowledge related to his own expertise as part of a larger entity. Therefore, each member has to be connected to a common data base and has to be able to understand in detail the part of the content, which he/she needs to use and also the scope of the knowledge which can be delivered by other partners involved in the network. Sharing information in the right context needs a transformation of information into knowledge, in order to disseminate the same meaning to the different actors.

Knowledge may be universal, vehicular or vernacular. Universal knowledge is normally shared by all. This is for example the case with geometrical knowledge. Vernacular knowledge is only used by a specific actor who is only concerned with his own job. It does not need to be shared. Vehicular knowledge is the type of knowledge which can be exchanged between two or more actors, allowing them for instance to perform collaborative design based on a common understanding. Therefore, the latter type of knowledge is very important to initiate a dialogue between two partners. A threaded plug in a product requires a manufacturing process, an assembly process (screwing phase) and finally a spatial structure to enable adequate access when screwing. At least, three different people are concerned with the decision to put a thread plug in a product during design; they all need to use the same information, and to apply it in their own context.

Much of the knowledge concerning different fabrication processes is the intellectual property of different laboratories. NoE VRL-KciP's goal is to integrate knowledge for the purpose of process integration covering at least the various specific production domains in which the different members are active. The knowledge involved ranges from the determination of product specifications to the recycling of the material at the end of the product-life. It also includes the knowledge of the production process design, the resources and finally that of the enterprise as a whole. The integration will be the basis for the development of methods for management support of dynamic enterprises, inter-enterprise operability and the necessary standardization for communication and control. This represents exactly the goal of the European call for ***Creation of "knowledge communities" in production technologies*** in its chapter ***3.4.3.1 Development of new processes and flexible, intelligent manufacturing systems.***

The application of effective methods from knowledge integration to design integration actually addresses topics like interactive decision-support, process navigational support, workflow control, supply chain and life-cycle management, including virtual production and simulation.

Apart from knowledge generated in their own specific research areas, all the network members will need to participate in a joint research activity dealing with the development of knowledge management methodologies and tools for the sharing of knowledge and the demonstration of integrated applications inside the network, such as

- An *ontology* library
- A *distributed*, agent- and constraint-based *collaboration* framework,
- *Artificial intelligence (AI) tools* for knowledge management in production, including data-mining and machine learning techniques.

The library, framework, methods and tools will be used for the setting up of a design platform and toolboxes for various applications allowing new methods for co-design, co-production and co-innovation in support of sustainable production.

Today, the mastering of knowledge, communication and decision support in global supply networks composed of various kinds of suppliers and customers is a critical issue. The solution to this problem largely determines the margins for innovation and economic performance regarding the ever-challenging but critical world market.

The unavoidable development of the European industry towards clusters of extended enterprises requires the development of a coherent competence in the field of knowledge management and collaboration. This must be strongly supported by a research community capable of producing the necessary knowledge, methods and tools allowing the enterprises' e-transformation.

The VRL-KCiP Network of Excellence will cover the wide-ranging research field and develop new concepts, methodologies and tools in ordre to master innovative technologies. Particular attention will be paid to human and socio-economic aspects in terms of psychological, philosophical, cybernetics, educational aspects and their relation to human capabilities as regards work, communication, knowledge structuring, user interfaces, self-organizing capabilities and learning. Attention will also be paid to the impact of new technologies and to e-(R)evolution on local economic conditions, education and training. All these subjects will have a significant and direct influence on long term competitiveness of European industry and will lead to an economically sustainable society.

The performance of cooperation within NoE will be promoted by the fact that the different laboratories will participate as team members in a common legal structure and, apart from sharing knowledge and research goals, they will commit themselves to sharing responsibilities based on contractual rules.

2.2 Generate knowledge, reduce the fragmentation and create a progressive and durable integration

In 1990, globalization was only starting to emerge as a policy issue. It is now recognized as one of the key change factors in contemporary economic systems and societies. Most of the world, including China and Russia, is now taking part in the market-based international economy. This opens up **new markets for EU products and services**: as highlighted above, EU companies are selling a growing share of their production in foreign markets. At the same time, it increases competition from imports and can lead to the re-location of productive activities to countries with lower cost bases. As Europe cannot compete on costs alone, **knowledge has a central role to play** in helping industries adjust to the pressures of globalization, in all sectors – whether high technology or not. In this context, the VRL-KCiP has to foster innovation and retain, develop and/or attract highly skilled people. These are essential factors, especially if the enlarged EU wants to keep high-value added knowledge-based activities within its borders. The training of new scientists able to think about sustainable products will be a challenge in Europe and is a key for a progressive and sustainable integration for our academic system.

As previously described, the problem of the integration of actors into teams performing co-design, co-innovation and co-production is to enable them to discuss and negotiate about design and production processes in order to find and agree on the different constraints and to facilitate the emergence of the result. The creation of vehicular knowledge is the key to starting such integrated processes. The other very important aspect is the facilitation of the communication between the different actors. A possible solution, which will be investigated into, is the development of an ontology library as part and parcel of a generic model of a so-called “Global Production Interface”. This model is simultaneously meant to serve as the basis for NoE's joint knowledge base.

The creation of vehicular knowledge demands that researchers document their knowledge, and expertise, discuss the context in which the knowledge may be used and accept a common view for continuing research, development and/or industrial knowledge application. It is clear that the creation of vehicular knowledge may reduce the willingness for cooperation between team members, because each actor may feel that he/she may loose some power or status by surrendering some personal competencies. The creation of a joint knowledge base is possible only if the actors can foresee a personal benefit from the common effort and the success of synergies resulting from it. However, we can expect that the creation of vehicular knowledge by members of the NoE will be successful because such a network will be backed-up by a strong management structure. The different researchers, members of the network, will be able to attach their status to collaboration and for instance recognition by industry, instead of personal knowledge in competition with colleagues.

We claim that the implementation of the VRL-KCiP network gives us the chance to initiate a Virtual Research Laboratory in the meaning of a decentralized laboratory without 'walls', and not in the meaning of an unreal laboratory. The use of the web and tools such as video-conferencing will reduce the need for face-to-face contact by NoE's members .

The capacity for solving design and production problems and supporting dynamic organizations, inter-enterprise operability and new standardization efforts will enable the VRL-KCiP to position itself as the best partner for industrialists. The progressive creation of a common knowledge database and of a good legal structure for the network will allow more and more involvement by in industrial partners, which in turn will secure the network and make it self-sufficient.

2.3 NoE's Overall structure

A conventional laboratory team participating in the network should consist of one or two senior researchers, some permanent researchers and PhD students. The senior researchers have the authority to direct team work into new research directions. The permanent researchers and PhD students will mainly be involved in carrying out R & D activities. They may also be involved in knowledge dissemination by the different training programs they are involved in, or by specific programs set up to disseminate the results to industry. The integrating strategies will mainly be designed by senior researchers, who can define collaborative research projects by linking up complementary research groups together. The senior researchers will also be greatly involved in the ‘dissemination of excellence’ activities.

Normally, a network of research laboratories have collaborative research activities. These activities are normally supported by regional, national or industrial programs in order to develop synergies between partners and to increase the task force on a specific subject. However, the consortium is only valid during the program’s lifetime and once the project has been completed, each partner returns to his/her own particular interest and activities.

The 6th framework program introduced the new instrument: “Network of Excellence”, in order to transform the way collaborative research is conducted, with the goal of establishing a sustainable network that evolves into a new legal European structure. In creating such a system, the laboratories have to devote resources to some activities in order to integrate their resources and to join the common structure, as shown in Fig. 1. These activities have to facilitate the integration of the teams, interaction with global structure, reinforcing the communication system, sharing the tools and platforms for a common use, managing the integration of knowledge, and suggesting exchanges of staff or students to foster cross fertilization of ideas.

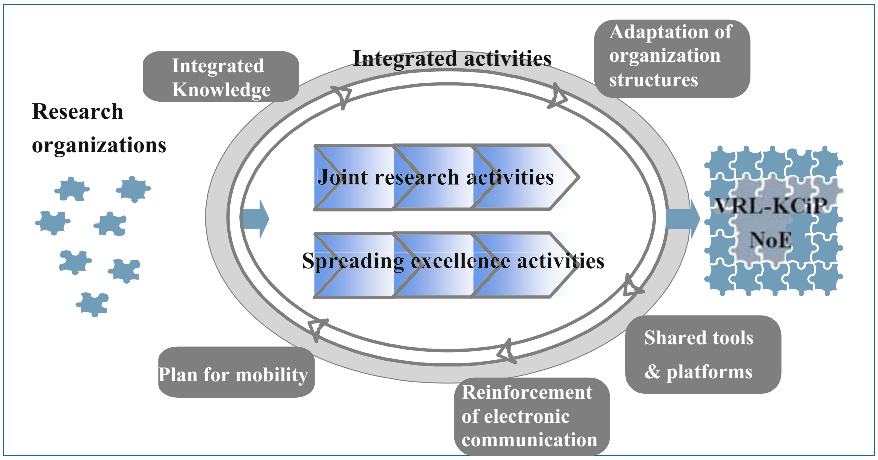


Figure 1. Overall structure of the NoE

Such integration activities are to be conducted with the conviction of a benefit: the group is stronger than the individual. Communication and dissemination activities may involve in-house partners, academic and industrial trainings, industries and non profit making organisms, for the benefit of science and society.

3. APPROPRIATENESS OF USING A NOE

The competitiveness of companies has passed successively from the customer-supplier system to the "supply-chain" to the contemporary extended enterprise. In an extended enterprise, companies cluster temporarily in an ad-hoc industrial structure to meet a common business objective. This mode of management is drawn by the obligation to quickly make products of high quality, at lower costs at the same time as being mindful of other environmental issues. European companies must adapt themselves to this new operating mode to remain competitive on the world market.

In order to maximize the success of companies operating in extended enterprises, it is absolutely necessary that tools and methodologies for the integration of the companies are developed through academic research. These research programs can only be truly successful if the laboratories are aware of all the issues concerning integration, and if they put forward a holistic approach to the complex problem. It is not only a question of asking the research team to work on common themes, but to work in a multidisciplinary approach in a structure close to the extended enterprise. The network of excellence is a very powerful tool that allows the full integration of the research teams and breaks up the actual current research fragmentation in Europe.

The network must consider the following issues:

- A discourse in the fragmentation of the European research,
- A lateral thinking and a holistic approach to address common problems in different fields.
- A multidisciplinary approach in order to overcome the variety of technical challenges.
- A high level of European education in sciences.
- A creation of synergies with industry and end-users in order to supply the market with highly innovative and competent products.

To achieve this goal, the network integrates 20 teams of researchers whose expertise covers the entire field of production engineering: design, technology, simulation, rapid prototyping, digital mock-up, various production processes, maintenance, recycling and knowledge management. It also integrates some educational networks in order to train engineers and members of some organizations in charge of technology transfer. The network also includes teams of researchers outside Europe, making it possible to make use of links with similar projects outside Europe and to carry out a worldwide strategic survey.

4. POTENTIAL IMPACT

4.1 Integration of research capacities and long-term structuring impact

One of the goals of the network VRL-KCiP is the creation of a persistent structure acting as a European laboratory. An integrated activity will be the development and evolution of a new legal entity in line with future industrial and research needs. The NoE VRL-KCiP will be able to build attractive common research programs for the industrial partners due to the synergy among these excellent laboratories. The VRL-KCiP has to develop authority and recognition in its field of research in order to disseminate its results and its newly created tools to the European countries and to be a driving force of innovation. Its authority can influence the strategic attention of other laboratories.

With a view towards developing a long term vision and having an ongoing impact on the European Research Area, the network will put focus on understanding the real industrial needs in knowledge based production/engineering and merging the underlying future research and industrial strategies To achieve this, it will “integrate” a number of key industrial partners throughout Europe, through appropriate links to the academic nodes. SMEs will also be addressed in this procedure. Alternative options for this “integration” involve:

- creation of an industrial board that will define/confirm the research directions and validate results, and/or
- provide industrial test beds to carry out real life pilots of NoE’s joint research results.

The VRL-KCiP will facilitate the effective merging of European research priorities with the programs and initiatives funded by the Research Councils of individual member states. This way, we foresee that a bi-directional relationship will be established, harmonizing and co-ordinating initiatives between state agencies and creating a dynamic presence in the European Research Area.

Through such industrial links, VRL-KCiP aims at becoming a EU advisor for its research policy in terms of vision and strategic reports. The new research structure will greatly influence the national, regional or industrial agenda, due to the reduction in time to report the results of research. The decrease in throughput times will increase the enterprises’ dynamic behaviour. One impact on industry may be the development of spin- off companies using the knowledge created by the network.

4.2 Addressing the problems of knowledge of tomorrow

Particular attention has been paid to achieve an extended trans-nationality in structuring the NoE's consortium. The network can therefore track a multi-cultural input on the specific knowledge-based concerns and needs of the European society. This input might become a catalyst for merging these needs with the network's research strategies and effectively addressing tomorrow's problems of knowledge from a representative societal point of view.

At an enterprise level, knowledge problems will be addressed by the close links of the network with numerous European enterprises. A multi-sectorial understanding of industrial reality will be a sound background for effectively structuring and integrating the R&D knowledge framework for the future European knowledge-based enterprises. Given this background, the network will build up its excellence addressing knowledge problems from all the different phases of a product's life-cycle, such as design, planning, manufacturing, recycling phases etc., as perceived from all the different points of view, such as product, process, resource, system, and enterprise. The distribution of complimentary knowledge is expected to highly contribute to the effectiveness of the network's approach.

Knowledge management in Design and Production is one of the central aspects of competitiveness for factories or enterprises. It concerns the enterprise/ technological/ product life-cycles as used in extended enterprises. The network will also understand how knowledge life cycle will be affected: Indeed, the reduction in knowledge life-cycle imposes a change to the traditional static view of knowledge, to work on knowledge flow that is changing in a dynamic way. This allows uncompetitive knowledge to be changed very quickly, adapted or even rejected whenever necessary.

In the field of design knowledge framework, we not only have to exchange and store a lot of information but also to build a common perception of this information with the construction of the common knowledge in order to use and share it. The creation of ontologies is an answer to the creation of knowledge-based enterprises.

4.3 Spreading excellence

Each member of the Network is a university professor in charge of research and well experienced in the management of major international conferences related to the NoE. We consider this participation in the management of conferences as an excellent method for disseminating the results of the common work. This will also be the case in the worldwide CIRP seminars or in IMS, ASME or JSME conferences. In IMS, the NoE will propose to create a CCI (Community of Common Interest) with other similar networks in the IMS regions (USA, Canada, Japan, Switzerland, Australia, and Korea). The CCI is a new instrument of the IMS program with intentions similar to those of the NoE of this 6^{th} FP. Our colleague and associated members in charge of the IMS program in Japan, such as Prof. Kimura and Prof. Ueda, or in the USA such as Prof. Lu will be particularly concerned with the creation of a CCI.

At the end of each year, a general assembly will be organized and the executive committee will meet to present the results of the year and to propose the activities for the following year. A second proposed way to spread excellence of the results is to open this general assembly to our industrial partners. Publishing a scientific review of the Network, under the name of a prestigious editor, will serve these objectives. We will also use the associated partners of the Network to open the door to a larger diffusion of results inside their international organizations.

At a national level, the participation in the Network of the groups such as Berliner Kreis, Aip-Primeca or the Netherlands Research School of Integrated Manufacturing is designed to amplify the dissemination of the results in local training. Likewise, the participation of technology transfer centres such as Pôle Productique Rhône-Alpes in France or Fraunhofer in Germany will permit the research to reach SMEs more easily.

5. CONCLUSION

The VRL-KCiP wants to develop authority and recognition in its field of research, in order to be the leader and to be able to propose normalization and directives including the field of intellectual property rights. A better understanding of common and distant work will be used for contributing to standardization on communication and on human interface definition. Several Laboratories have strong links with International and US standardization bodies and these will facilitate the effective representation of European Research in the creation of new standards.

The use of a real network of international teams working on the cooperative integrated design area will be a very good field of experimentation in order to test diverse communication protocols, data exchange processes and sharing heterogeneous models. We wager that these experiments will open up towards new standards and new regulations giving an advance to the EC !

REFERENCES

1. Information Society Technologies, 2003-2004 Workprogramme, http://www.cordis.lu/ist
2. Nano-technologies and nano-sciences, knowledge-based multifunctional materials, and new production processes and devices, 2002-2003 Workprogramme, Edition December 2002, Commission Decision C(2002)4789, 9.12.02

Table 1. Institution engaged in the VRL-KCiP

F	Institut National Polytechnique de Grenoble
NL	The University of Twente
G	Fraunhofer – IPK Berlin
I	Istituto di Tecnologie Industrialie Automazione – CNR – Milano
UK	University of Bath
E	TEKNIKER Eibar
GR	University of Patras
S	Kungliga Tekniska Högskolan-Stockholm
HU	Hungarian Academy of Sciences, SZTAKI, Budapest
SL	University of Ljubljana
G	University of Stuttgart
IS	Institute Israel of Technology - Technion – Haïfa
F	Ecole Centrale Nantes
F	Université. Technologique de Troyes
RO	PoliItehnica University of Timisoara
CH	Ecole Polytechnique Fédérale de Lausanne .
UK	University of Durham
NL	Delft University of Technology
NL	Eindhoven University of Technology
PL	Poznam University of Technology
G	Berliner Kreis
F	AIP-PRIMECA
NL	The Netherlands Research School of Integrated Manufacturing, IPV
F	Pôle Productique Rhône-Alpes
SA	University of Stellenbosch

INTEGRATED DESIGN FOR MANUFACTURE, SERVICE AND ENVIRONMENT

Winston. A. Knight
Department of Industrial and Manufacturing Engineering, University of Rhode Island, Kingston, RI, 02881, USA, and Boothroyd Dewhurst Inc., 138 Main St., Wakefield, RI., 02879, USA. knight@egr.uri.edu

Abstract: This paper outlines the development of product analysis procedures for manufacture, service and environmental performance. The paper further describes how these procedures can be integrated to give a predictive assessment of the life cycle performance of products. This integration enables design teams to evaluate design proposals and effect design changes to meet the specific life cycle goals that are judged to be the most important for the particular product under consideration.

Key words: Product design, Design for manufacture, Life cycle design, Design to cost.

1. INTRODUCTION

The author and his colleagues have been involved in research and developments in the area of product design for manufacture and assembly (DFMA) for the past 20 years or more [1]. Much of the basic research in support of these developments has been carried out through a series of industry-funded projects. The underlying philosophy of this work is that early design decisions on product configurations, together with material and process selection, lock in a high proportion of the subsequent manufacturing costs. Thus suitable predictive analysis tools are required that can enable design teams to evaluate the assembly and manufacturing aspects of product design proposals at the early stages of design. This quantification of design decisions facilitates the development of simpler and more easily manufactured products, with reduced costs. This work has resulted in

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 17-30.

software that has become widely used in industry for the development of more competitive products [2]. In addition these tools can be used for product benchmarking and design to cost programs [2-5].

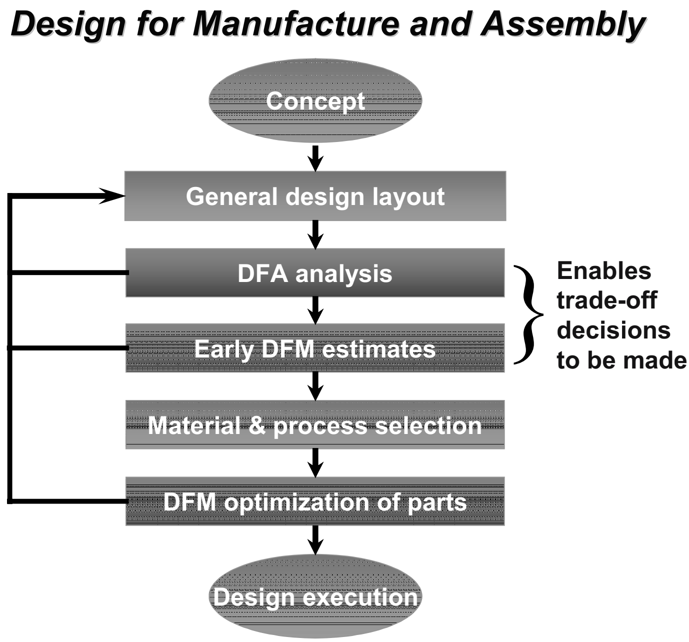


Figure 1. General Sequence for Design for Manufacture and Assembly

Over the years the projects undertaken have been widened to include consideration of other life-cycle aspects of products, including design for service and design for environment [6-8]. Again early design decisions on product configurations determine the serviceability and environmental performance of products. The initial emphasis on design for assembly leads logically to considerations of design for disassembly. Servicing of a product involves careful disassembly and then reassembly, after carrying out some service task. Design for service (DFS) evaluation tools effectively allow the disassembly and reassembly tasks to be simulated, so that realistic estimates of service times and costs can be obtained. This facilitates the redesign of products for ease of service [6]. Interest amongst the industrial collaborators also evolved towards evaluating products for disassembly, recycling and overall environmental performance. Several projects were carried out with the aim of developing analysis tools that incorporate these aspects of product design [8, 9]. Some of this work was carried out in collaboration with the TNO Industry Centre in the Netherlands and resulted in the development of a software tool for product design for environment [10]. This tool is somewhat unique in that both a financial and environmental evaluation of the end-of-life disposition of a product is obtained from the basic disassembly analysis, together with an environmental assessment of the initial manufacture of the product. This program has also become used by a number of industrial companies [11].

As these analysis procedures have been developed they have been converted into separate software tools, but recent developments have enabled these separate tools to be integrated such that the total life cycle costs of a product can be evaluated at the early stages of design and this is the focus of this paper.

2. DESIGN FOR ASSEMBLY AND MANUFACTURE

Serious developments in product analysis procedures for DFMA began in mid 1970s. The most widely used systems for DFMA analysis are those developed by Boothroyd and co-workers [1], but other assembly analysis systems have also been applied [12-15]. Development of design for assembly (DFA) analysis procedures built on earlier work on analysis of feeding and orienting of parts for automatic assembly [16]. The basic aim is to predict assembly costs in the early stages of design, through time standard databases accessed by means of a classification of part features. These procedures have become widely used in industry and numerous case studies of product improvements exist [1].

Fig. 1 summarizes the steps in DFMA, consisting of the two main stages, DFA and DFM. DFA analysis is carried out first, aimed at product simplification and part count reduction, together with detailed analysis of the product for ease of assembly. The basic input to a DFA analysis is the product structure; a listing of items in assembly order, including the main assembly and all subassemblies. Each item is then considered in order of assembly. An essential feature of DFA is the analysis of the proposed product by multidisciplinary design teams as part of the design review process. Product simplification and part count reduction are facilitated by applying simple criteria for the existence of separate parts, including:

- During operation, does the part move relative to other parts already assembled?
- Must the part be of a different material, or be isolated from other parts already assembled? Only fundamental reasons, e.g., material properties are acceptable.
- Must the part be separate from other parts assembled because necessary assembly or disassembly of other parts would otherwise be impossible?

- Assembly fabrication
- Die casting
  - Hot chamber
  - Cold chamber
- Sand casting
  - Automatic
  - Manual
  - Semi-automatic
- Investment casting
- Metal Injection Molding
- Powder metallurgy
- Plastics Processing
  - Blow molding
  - Extrusion
  - Injection Molding
  - Structural foam molding
  - Thermoforming
- Machining
- Hot Forging
- Sheet metal cutting
  - Laser
  - plasma
- Sheet metal deep drawing
  - Progressive die
  - Separate operations
  - Transfer press
- Sheet metal stamping
  - Compound die
  - Multi-stage
  - Progressive die
  - Separate operations
  - Transfer press
  - Turret press

Figure 2. Processes Covered by Predictive Cost Models

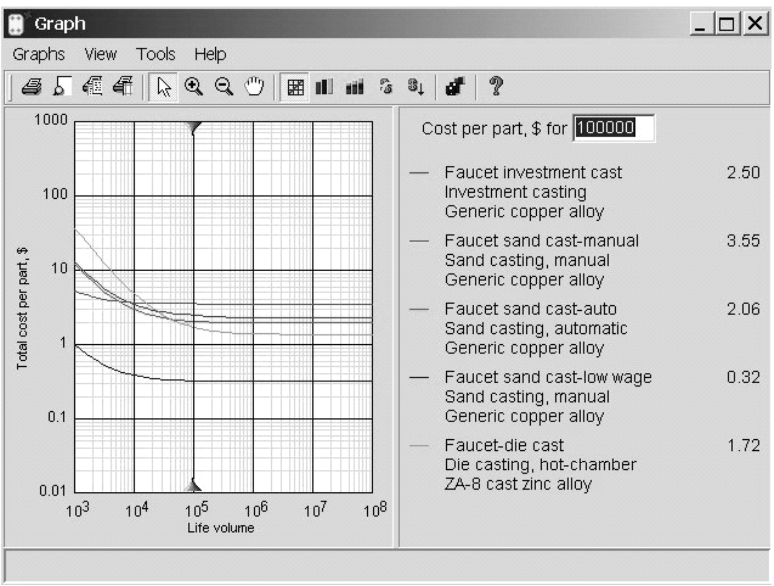


Figure 3. Assessment of Different Material Process Combinations for a Part

If an item fulfills none of these criteria, then it is theoretically unnecessary. This analysis leads to the establishment of a theoretical minimum part count for the product. It does not mean that the product can practically or economically be built of this number of items, but it represents a level of simplification to be aimed for. DFA is not a design system: the innovation must come from the design team. However team members are challenged and stimulated to consider simplified product structures with fewer parts. Detailed DFA analysis considers the geometry and other features which influence assembly difficulties, and hence assembly times. To estimate assembly time, each part is examined for two considerations: how the part is to be acquired, fetched, oriented, and made ready for insertion, and how it is inserted and/or fastened. The difficulty of these operations is rated, and, from this rating, standard times are determined for all operations necessary to assemble each part, from a time standard database accessed through a classification of part features that affect assembly difficulty. Besides enabling total assembly time to be estimated, items that contribute the most to this time are also identified, which again leads design teams to consider ways of improving ease of assembly. From the assembly time and minimum part count, a design efficiency or index can be determined, which can be used as a metric to gage product improvements [1].

2.1 Design for Manufacture

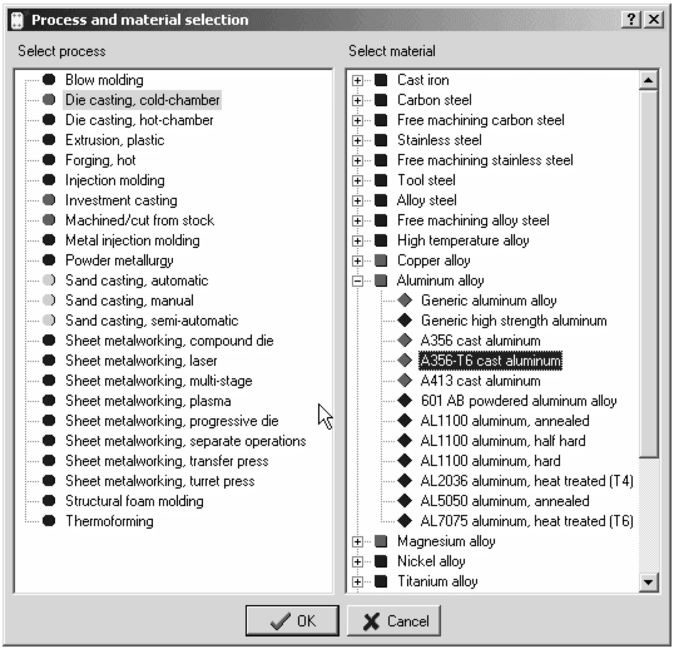


Figure 4. Materials Database Indicating Process/Material Limitations for a Part

An outcome of DFA analysis will be alternative proposals for product simplification though elimination or integration of parts, including alternative combinations of materials and processes. These alternatives need to be evaluated economically as soon as possible. Thus Design for Manufacture (DFM) analysis is focused on the early assessment of component manufacturing costs and selection of best process/material based on realistic cost estimates. The objective is to influence the early decisions on choice of material and process for a part, before detailed design is undertaken. This is because, once this decision is made, the items must be designed in detail to meet the processing requirements of the chosen process. The overall aim of DFM is to enable design teams to make cost trade-off considerations while these decisions are being made.

This has led to considerable effort on the development of predictive cost models for a wide range of processes (see Fig. 2). The development of early cost estimating procedures is somewhat similar to the development of computer-aided process planning (CAPP) systems. The process must be studied in detail to determine factors that influence processing sequences and to determine the main cost drivers, thus establishing how part features influence these. These models use general part information that is available during early product design to predict both piece part costs and tooling investment. Initially these models were converted into separate software packages [1], but recently these have been integrated into a single package (Concurrent Costing) [15], that links directly to DFA. This enables a comparison of costs for a given part configuration for different material/process combinations to be readily obtained. For example Fig. 3 shows a typical comparison of costs versus quantity produced (life volume) for several different processes. The package has been designed to work in conjunction with DFA, such that total product manufacturing costs can be evaluated, but the program has also been used separately for effective cost management initiatives [e.g. 5].

The DFM cost program uses an extensive materials database that includes the materials related parameters for each process. This database also includes processing limits data that aid material and process selection by indicating combination suitability for a given part. This is illustrated in Fig. 4, where a green symbol indicates an acceptable combination and red unacceptable, with amber indicating possible combinations but at increased cost or difficulty.

2.2 Design for Service

The serviceability of products has a major influence on customer satisfaction, warranty repair costs for manufacturers, and often, the useful life of products, since products are often discarded when service costs become higher than purchasing a replacement product. In addition for some products, such as transfer line equipment, the lifetime service costs may be more significant than the initial manufacturing costs. Product analysis tools for DFS enable design teams to evaluate the efficiency of service tasks during design and compare the serviceability of alternative design configurations [6]. This is done through the effective simulation of each service task. The output of a DFS analysis is estimates of service times and costs, determined from time standard databases of disassembly times, accessed by a classification of the disassembly and assembly attributes of the items removed during a service task. The reassembly times are obtained from the DFA time standard database. In the current version of the DFS analysis tool, the user selects items from the DFA structure chart to be removed to carry out the service task. After this additional operations may be added and the reassembly sequence generated automatically using the DFA database.

Two other aspects considered are, a service efficiency index and an importance ranking of service tasks/repairs. A time-based service task efficiency index can be formulated, similar to the DFA efficiency index [6]. Three criteria, to establish the theoretical minimum number of parts that must be removed or unsecured to access a service item, or carry out a service procedure, have been proposed. A part, subassembly or operation is justified for disassembly if:

- The part/subassembly removed is the service item or contains the item and will be replaced as a component; or the operation performed is the service operation.
- The part/subassembly must be removed, or operation performed to isolate the service item, or subassembly containing the service item, from other parts and subassemblies with which it has functional connections, which typically transfer fluids, electricity, electrical signals, motion or force. An example may be a wire harness connection, drive belt or a fan blade connected to a motor.
- The part/subassembly removed is a cover part, or the operation performed reorients a cover part, which must completely obstruct access to the service item and related parts to seal or protect it from the environment. A cover part may also protect the user from potential injury.

Multiple identical operations that have the same purpose must be accounted for separately. For example, a set of connectors released to isolate a service item must be considered independently. If several of the connectors could be combined, into one larger connector, then only one of the group may be counted as necessary in the disassembly sequence. Again consideration of these criteria challenges the design team to simplify the disassembly processes required for service tasks. From the minimum disassembly steps and the total service time estimate an efficiency rating for service tasks can be determined [6].

Important factors that should be considered in the design of any product are the frequency (or probability) of different failures, and the consequences of failures. If an item is expected to fail frequently, then the required service procedure should clearly be designed for easy execution. Furthermore, if a failure may result in grave consequences then easy preventive maintenance procedures are necessary even if the chance of the failure is small. To this end, a Frequency Ranking and Consequence Ranking for service tasks can be developed [6], based on similar principles to Failure Modes and Effect Analysis (FMEA). From these a Service Task Importance Ranking can be derived. Service tasks with a high importance ranking should be designed with high service efficiency.

The DFS methodology has been successfully applied to a wide range of products including domestic appliances and computer peripherals [18].

2.3 Design for Disassembly and Environment (DFE)

Materials have significant recycled value when they are divided into clean, separate types, and viable recycling depends greatly on the efficiency with which materials can be separated. The separation and recovery of items and materials from products can be approached in two main ways, which can be used together. These are careful disassembly, and bulk recycling, which involves shredding of mixes of items and using separation techniques on the stream of particles produced. In the long term, however, recycling can be made more effective by design of products for greater ease of disassembly and recycling. This requires development of product analysis tools to enable design teams to evaluate ease of disassembly and recycling, in the early stages of design, similar to DFMA and DFS.

Development of analysis procedures for disassembly has received some attention [e.g. 19]. These procedures concentrate mainly on the economics of disassembly, and not on environmental impact. Recently, tools that enable consideration of a balance between the financial and environmental aspects have been developed [8-10, 20]. Product analysis procedures for disassembly are aimed at simulation of disassembly at end-of-life disposal and quantifying the resulting costs, benefits and environmental impact. Disassembly costs are determined from time-standard databases developed for end-of-life disassembly processes, which are accessed by classifying disassembly attributes of items. Each item in the assembly is allocated to an appropriate end-of-life destination (recycle, reuse, regular or special disposal and so on) based on its material content, and this information, together with the item weight, enables the possible costs or profits to be determined from a materials database. The user also provides information on disassembly precedence of each item (items that must be removed immediately prior to the component). This information enables the disassembly sequence to remove valuable items as early as possible to be determined.

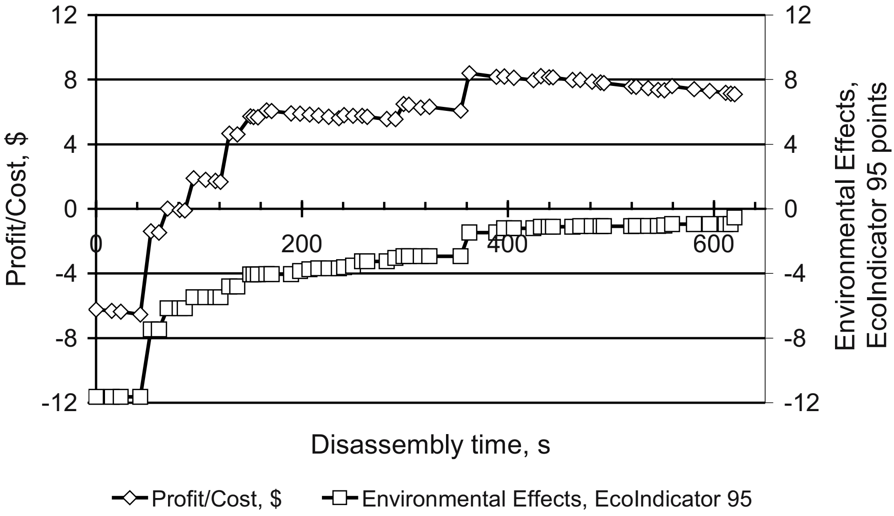


Figure 5. End-of-Life Financial and Environmental Assessment

An initial disassembly sequence can be entered directly by the user or is generated automatically from a DFA analysis, by reversing the initial assembly list. Based on the resulting disassembly sequence, two main analyses are performed: 1) Financial return assessment of disassembly and disposal, including remanufacturing and recycling, and, 2) Environmental impact assessment, resulting from initial product manufacture and disposal, including remanufacturing and recycling.

The financial assessment, for the disassembly of a personal computer (PC), is shown by the upper line in Fig. 5 as the return or cost as disassembly progresses. This is determined as the difference between costs incurred to disassemble each item and recovered value, plotted against disassembly time. Disassembly times and costs are obtained from time standard databases accessed by a classification of disassembly attributes of each item removed. A point on this curve represents the profit or net cost if disassembly is stopped at this stage.

To develop the environmental line, materials and manufacturing processes are allocated environmental indicators based upon life cycle assessment (LCA). At present two alternative databases can be selected by the user; the first uses EcoIndicator 95 [21] and the second MET points [22].

The environmental impact line (lower line in Fig. 4) shows the environmental effect at any stage of disassembly. A specific point on the curve represents the net environmental impact of the product if disassembly is stopped at this stage. Note that all indicator points are negative, since they measure effects on the environment through emissions, use of scarce materials, etc. Re-manufacturing or recycling of items reduces the negative effects of product initial manufacture and end-of-life disposal. The first point on the curve represents the environmental impact from initial manufacture of the whole product plus environmental impact of disposal of the whole product, by regular or special waste treatment. As disassembly proceeds, the curve moves up if items are recycled or remanufactured, and some of the indicator points are effectively 'recovered.'

2.3.1 Allocation of Disassembly Precedence

A procedure for the user to allocate the precedence for disassembly has been developed [7, 20]. In this process, starting at the top of the list, the items or operations that must be removed or carried out immediately prior to the item under consideration are selected. In this way, the disassembly precedence tree for the product is established. The assignment of precedence is useful in highlighting to designers complex disassembly sequences for specific items.

2.3.2 Optimisation of Disassembly Sequences

In complex products, there may be numerous possible disassembly sequences, limited by disassembly precedences assigned to each item. For example many items in subassemblies can be removed without removing the subassemblies themselves from the product first or the subassembly can be removed before disassembling further to reach the required item. Determination of the most appropriate disassembly sequence for a product is therefore of interest.

Procedures have been developed [7, 20] that will reorder the critical items (those which produce significant steps in the financial curves) as early as possible in the disassembly sequences, but limited by the precedence constraints input with the initial disassembly sequences. The order in which items are moved forward in the disassembly sequence is determined from the greatest yield (rate of improvement).

3. INTEGRATION FOR THE LIFE CYCLE COST ASSESSMENT OF PRODUCTS

The product assessment tools for DFA, DFM, DFS and DFE have been developed as separate packages, although facilities exist for the linking of DFA and DFM, so that the total manufacturing and assembly costs of a product can be accumulated into the DFA cost summary. A scheme has been developed for integration of these separate analysis tools into one integrated package that enables the total life time costs for a product to be estimated during the early stages of product design.

The complete analysis is driven from the DFA structure chart, which is a tree structure listing of all items in the product in assembly order (Fig. 6). For the DFA analysis a response screen for identifying assembly attributes and minimum part criteria is used for each item. Item manufacturing costs can be entered by the user or estimated by from the Concurrent Costing models that link directly to each item as required. This combination determines the predicted manufacturing costs of the product, as the sum of the assembly costs and the manufacturing costs of all items in the product.

All materials related data is added to the Concurrent Costing materials library, including recycling values and environmental indicators. Sections of the library have been added for use-phase items, such as consumables, energy inputs and replacement items. From these the use-phase costs and environmental impact for the product can be determined.

The environmental impact assessment of the product is determined from additional responses in the DFA analysis for each item in the assembly. The initial environmental impact of the materials and processes is then determined, together with the end-of-life impact resulting from the end-of-life scenario entered for the items in the product. The end-of-life analysis includes the possibility of bulk recycling the product or selected subassemblies. The user can select whether bulk recycling is available for processing any portions of the assembly with mixed materials that are not disassembled. A disassembly sequence is generated automatically by reversing the initial DFA assembly lists.

From this the user can assign the disassembly precedence and then the best disassembly sequence for maximum financial return at end-of-life is determined, as is shown in Fig. 5. The end-of-life costs are added to the DFA product cost summary, together with the use phase costs. A separate summary of the environmental impact for all life cycle phases is determined.

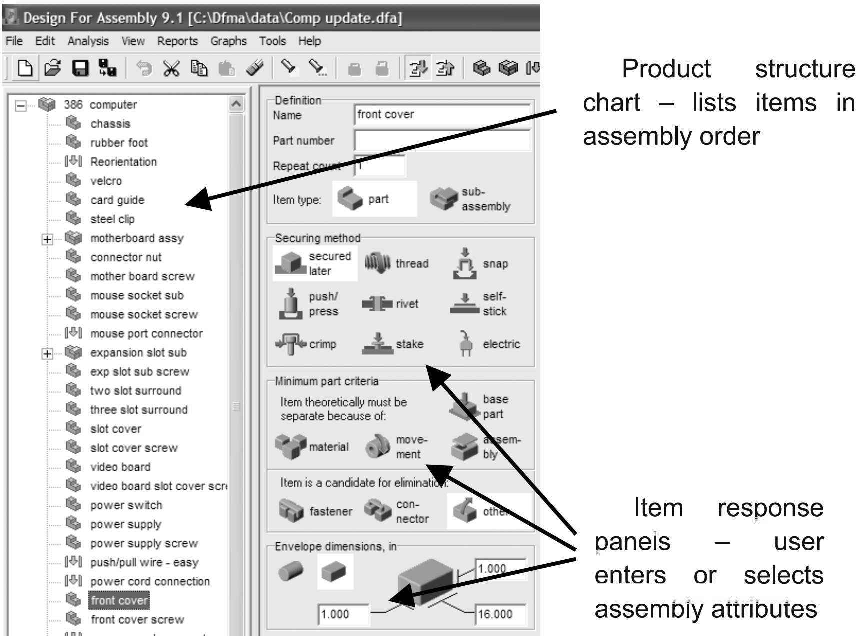
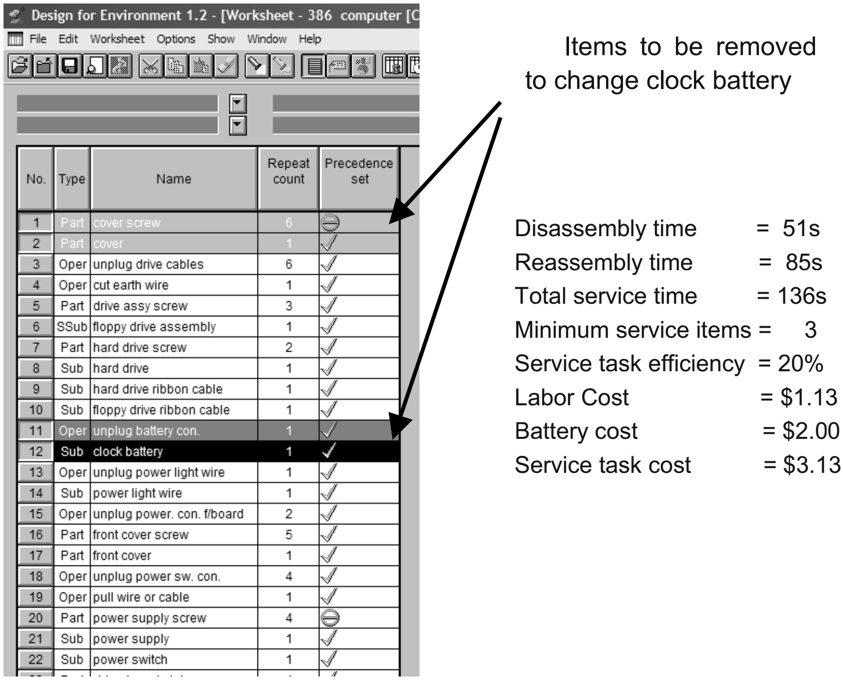


Figure 6. DFA Structure Chart and Some Item Response Panels

Once the disassembly precedence relationships have been established, the times and costs for any service tasks can be automatically determined. The item to be removed for service or the last item to be removed for a service task can be selected from the DFA structure chart. The shortest disassembly sequence is then generated automatically, with the associated times determined from the disassembly time standard database. The reassembly sequence and times are then determined from the DFA analysis data. Fig. 7 shows the items that must be removed in order to change the battery in the PC, together with an assessment of the service task costs and service task efficiency. The estimated service task efficiency is 20 percent, largely because 6 screws are used to hold the cover in place. Elimination of these screws would result in a service task efficiency of about 44 percent as well as improving the initial assembly of the product. The user can then select or enter the service task costs, to give the total service costs. All service costs evaluated can then be added to the DFA product cost summary to give the total life cycle costs of the product.



No.	Type	Name	Repeat count	Precedence set
1	Part	cover screw	6	⊖
2	Part	cover	1	✓
3	Oper	unplug drive cables	6	✓
4	Oper	cut earth wire	1	✓
5	Part	drive assy screw	3	✓
6	SSub	floppy drive assembly	1	✓
7	Part	hard drive screw	2	✓
8	Sub	hard drive	1	✓
9	Sub	hard drive ribbon cable	1	✓
10	Sub	floppy drive ribbon cable	1	✓
11	Oper	unplug battery con.	1	✓
12	Sub	clock battery	1	✓
13	Oper	unplug power light wire	1	✓
14	Sub	power light wire	1	✓
15	Oper	unplug power. con. f/board	2	✓
16	Part	front cover screw	5	✓
17	Part	front cover	1	✓
18	Oper	unplug power sw. con.	4	✓
19	Oper	pull wire or cable	1	✓
20	Part	power supply screw	4	⊖
21	Sub	power supply	1	✓
22	Sub	power switch	1	✓

Figure 7. Service Analysis of Clock Battery Replacement

4. CONCLUDING REMARKS

A scheme has been developed for integration of previously separate analyses for DFA, DFE and DFS into one package. This development will enable the life cycle costs of a product to be determined during the early stages of product design, together with a life cycle environmental impact assessment. Design teams do not need to always carry out the complete analysis, but can choose those sections of the analysis that are deemed the most important for the specific product under consideration.

REFERENCES

1. G. Boothroyd P. Dewhurst and W.A. Knight, Product Design for Manufacture and Assembly, 2nd, Edition (Marcel Dekker, NY, 2000).
2. G. Boothroyd and P. Dewhurst, Computer-Aided Design for Manufacture, Manufacturing Engineering, April (1988).

3. G. Boothroyd, Design for Producibility: The Road to Higher Productivity, Assembly Engineering, March (1992).
4. W.A. Knight, Complex Design Problems Vanish with DFMA, Medical Design News, 1(1), January/February (2001).
5. Anon, Software Tool Helps Harley Drive Out Costs, American Machinist Harley Davidson 100th Anniversary Commemorative Issue, 24-26 (2003).
6. P. Dewhurst and N. Abatiello, Design for Service, in Design for X: Concurrent Engineering Imperatives, edited by G.Q. Huang (Chapman Hall, , 1996) pp 289-717.
7. B. Rapoza, T. Harjula, W.A. Knight, G. Boothroyd, Design for Disassembly and Environment, Annals of CIRP, 45(2), 109-113 (1996).
8. W.A. Knight, Disassembly Analysis of Consumer Products, in Product Innovation and Eco-efficiency, edited by J.E.M. Klostermann and A.Tukker (Kluwer Academic Publishers, 1997)
9. W.A. Knight, Product Benchmarking using DFE Analysis Tools, Proc. of IEEE Conference on Electronics and the Environment, Darien, Ma. (1999).
10. W.A. Knight and M. Curtis, Measuring Your Ecodesign, Manufacturing Engineer, 81(2), 64-68 (2002).
11. J. Jancsura, Cool, Stylish and Green, Appliance Manufacture, 47(11), 50-54 (2000).
12. S. Miyakawa, T. Ohashi and M. Iwata, The Hitachi New Assemblability Evaluation Method, Trans. North American Manufacturing Research Institution (NAMRI), (SME, Dearborn, Michigan, 1990).
13. S. Arimoto, T. Ohashi, M. Ikeda and S. Miyaka, Development of Machine Productibility Evaluation Method, Annals of the CIRP, 41(1), 119-112 (1993).
14. B.L. Miles, Design for Assembly- a key element within design for manufacture, Proc. Inst. Mech. Engineers., Part D: J. Automobile Engineering, 203, 29-38 (1989).
15. G.F. Dagleish, G.E.M. Jared and K.G. Swift, Design for Assembly: Influencing the Design Process, Proc. Inst. of Mech. Engineers, Part B: Journal for Engineering Manufacture, 214(B1), 81-88 (2002).
16. G. Boothroyd, C. Poli and L.E. Murch, Handbook of Feeding and Orienting Techniques (University of Massachusetts, Amherst, 1978)
17. T. Stevens, DFM Concurrent Costing Version 2.0, Industry Week, 251(12), 64-66 (2002).
18. A. Lee, Design for Service and Damage Repair, (unpublished Thesis, University of Rhode Island, 1995).
19. M.R. Johnson and M.H. Wang, Planning product disassembly for material recover, Int. J. Production Research, 43, 3119 (1995).
20. A. Girard and G. Boothroyd, Design for Disassembly, Proc. 10th Int. Forum on Product Design for Manufacture and Assembly, Newport, RI June 12-13 (1995)
21. Netherlands NOH, Eco-Indicator 95 – Manual for Designers, Report No. 9524 (Netherlands NOH, Amsterdam, 1995).
22. S. Kalisvaart and J. Remmerswaal, The MET-Points Method: a new single figure environmental performance indicator, Proc. Integrating Impact Assessment into LCA (SETAC, Brussels, Oct., 1994).

Part 1

DESIGN STRATEGIES AND METHODOLOGIES

THE CHALLENGE OF ENVIRONMENTAL IMPROVEMENT IN DIFFERENT TYPES OF INNOVATION PROJECTS

Johan Tingström

University of Kalmar, Department of Technology, SE-391 82 Kalmar, phone: +46 480 446322, fax: +46 480 446330,johan.tingstrom@hik.se and Royal Institute of Technology, Machine Design, tingan@kth.se

Abstract: Product development tends to strive towards radical innovations of products to gain an advantage in a competitive market. Which tool should be used to evaluate environmental performance in order to reach the market faster with radically improved products? This paper discusses radical and incremental innovation and combines them with tools for evaluating environmental performance of products. The environmental tools examined or used in this paper are design handbook, life cycle assessment (LCA), and environmental effect analysis (EEA). This paper is based on a literature review in the field of innovation and environmental research. Primarily journal articles and conference papers have been used but also books and dissertations. The result of the discussion is that a radical innovation should use EEA as a supporting tool, and incremental innovation should use LCA. This is based on known facts about EEA and LCA. A design handbook can easily be used as a support for both types of innovation.

Key words: Incremental and radical innovation, Product development, Design for Environment (DfE), Life Cycle Assessment (LCA), Environmental Effect Analysis (EEA)

1. INTRODUCTION

One of the major difficulties in product development is the design-paradox as shown in Fig. 1. In the beginning when a development project starts the knowledge about the final product is limited. As the project

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 33-44.

progresses and decisions are taken the freedom decreases. This is a paradox that crucial decisions need to be taken in the beginning with little knowledge of the final product. This is often called the *design paradox* [1] and is frequently used in design research. If wrong decisions have been taken and modifications are later needed, the cost of this increases rapidly with the project time. As long as the product is at the drawing stage changes are not so costly, but if production has started the cost of changing manufacturing tools and re-educating staff is high.

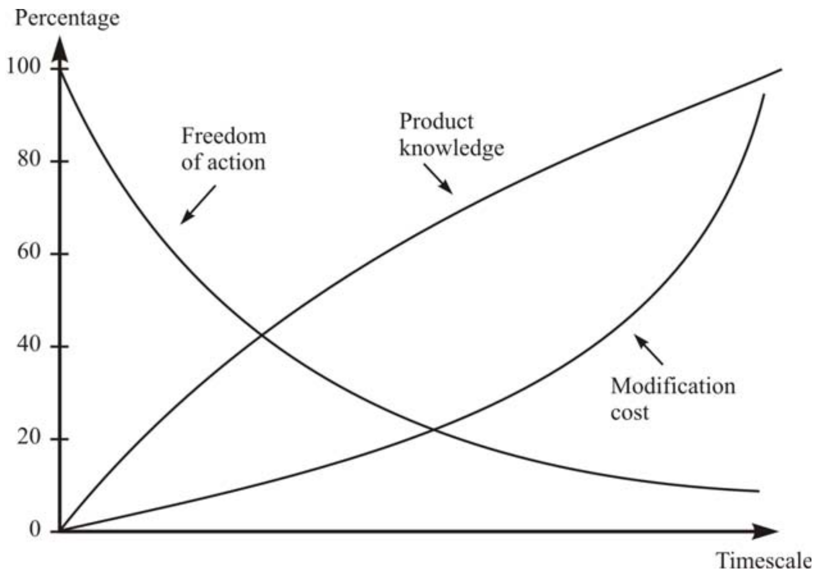


Figure 1. The design paradox. Extended with the knowledge of the product [1-5]

Many ideas need to be evaluated in order to find one that can enter the market as a success. According to Steven and Burley [6] as many as 3000 raw ideas are examined in order to achieve one commercial success for a new product, where the raw ideas are generated during brainstorming. A commercial success is not only when someone is willing to buy the product, the product should also generate profit for the company developing the product.

2. AIM AND SCOPE

The aim of this paper is to examine which environmental tools are best suited to different types of innovations. In this paper different eco-design tools will be related to different types of innovation. The different innovation approaches are radical and incremental. A short literature study has been made to find the core approaches in both radical and incremental innovation respectively. The different tools used in this study are Design Handbook (DH), Environmental Effect Analysis (EEA), and Life Cycle Assessment (LCA). They have been chosen because they have different levels of complexity and they use different information (qualitative and quantitative). This paper gives a short description of these different tools, together with some key advantages and disadvantages from the literature. The product development process in this paper shall be seen as an ongoing activity and not as a development of one product as an isolated action.

3. METHOD AND MATERIAL

The research method used in this paper has been deductive. Deduction is about deriving statements from other statements. Going from theories and statements, explanations are sought that later can be confirmed in reality [7]. In summary, deduction has its base in theory. This paper is based on information found in the literature. The literature study focused on the innovation and tools for product development processes and different support tools that can be useful when developing products.

4. INNOVATION

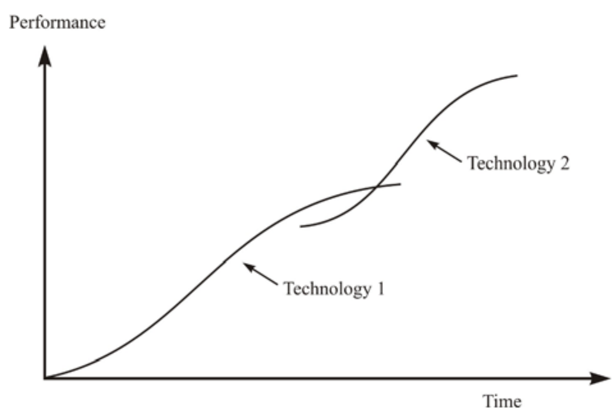


Figure 2. Differences between incremental and radical innovation as S-curves [9]. Incremental is Technology 1 and Technology 2 as individuals. Radical is the jump from Technology 1 to Technology 2.

There are many different definitions of innovation and according to Tidd [8] "*[...] innovation is a process of turning opportunities into new ideas and putting these into widely used practice*". There are two main kinds of innovation depending on the extent of what is new for the product. This section will further describe radical and incremental innovation. The major difference can be described with S-curves as shown in figure 2.

Radical innovation may offer opportunities for new companies, while incremental innovation is more likely to benefit established firms [10].

4.1 Radical

Radical innovations take the innovation to a higher level than incremental in order to create new industries, products, and markets [11]. According to figure 2, radical innovation is the change between technology 1 and technology 2. The technological improvement of the product is significantly higher as compared to a change in efficiency, scale, and design could not show the same result [12]. Tuchman and Andersson [12] and Koberg et al. [13] have defined radical innovation as strategic changes in products/services, markets served, and technological breakthroughs in order to produce a product or render a service based on significant innovation.

A radical innovation often needs more resources i.e. money and time than other types of innovation. To develop a breakthrough innovation can require a ten-year period along with investment of millions of dollars [14]. Radical innovations are often based on new technology, knowledge etc., which can be difficult to find in a company that has focused on developing one single product.

A radical innovation of a product does not necessarily need to be better than other products on the market but it has large improvement potential and could in the future take a large market share.

4.2 Incremental

Tushman and Romanelli [15] describe incremental innovations as those that encourage the status quo. Incremental innovation can be seen as an update of software or a minor change of a physical product. The product has still potential to be developed and reach better performance along the S-curve. Incremental innovation means small improvements between generations of the product. According to Hall and Kerr this is technology one or technology two studied as a unique technology. If the innovation improves the knowledge of both component and architecture it is incremental, if it destroys both it is radical [16].

According to Abernathy and Utterback [17] a technological evolution of a product in a branch tends to emerge as a "dominant design", which Suarez and Utterback [18] have defined as "[…] a specific path, along an industry's design hierarchy, which establishes dominance among competing design paths". Following the development trajectory of this dominant design product shows an incremental innovation.

5. ECODESIGN TOOLS

There are many different tools that can be used by companies that want to handle environmental considerations in their product development. Different tools have different complexities and this makes it difficult for companies to choose which one to use. To make the choice easier some classifications have been made. Kortman et al. [19] classifies tools into analysis or improvement tools. According to Kortman [19], analytical tools are used to identify environmental impacts and are used for comparing different design alternatives. One well-known tool for this is LCA. The aim of improvement tools is to assist the designer in taking decisions. This can be done with checklists, guidelines and design principles. Wrisberg et al. [20] has made a distinction between analytical and procedural tools. With the terminology from Wrisberg et al. [20] all three tools considered in this study are analytical tools based on physical parameters. According to the terminology from Kortman [19] the handbook containing guidelines is an improvement tool, and LCA and EEA are considered as analytical.

It can be difficult for companies to know which tool to use and how to use it. Results from a survey show that too many tools can limit creativity [21]. Also companies are not aware of the usefulness of a tool they have not tested. To get the best result out of a tool it should be one that fits together with the product and the development process of the company. Nowadays many companies use whichever tool or method is in fashion at the moment and which may not fit their company or their product.

5.1 Design handbook

It is important that the designer can find information easily during the development. The handbook should contain information such as representative design solutions, steering documents, checklists, and guidelines. Examples of steering documents can be found at Volvo where they have three "lists" for choosing materials. The "blacklist" [22], contains substances that are not allowed in Volvos products. The "greylist" [23] describes materials that are limited in Volvo's products. As a complement to these two lists, the "whitelist" [24] gives recommendations for materials to substitute from the "black-" and "grey-list".

As an aid for designers, suggestions of design solutions can be included in the handbook. This can be information of on how to design for ease of assembly and disassembly, some specific parts to use, which surface properties to have, etc. Guidelines are tools, which are easy to use and apply in the product development process. The aim of these tools is to give some general directions when it comes to material selection, guidelines for suppliers, design options, etc.

5.2 Environmental Effect Analysis (EEA)

EEA has been developed from the quality tool FMEA (Failure Mode and Effect Analysis), and uses the same structured way of working. EEA is a qualitative method that can be used early in the development process since it does not need quantitative data on a detailed level. This is an advantage early in the development process when the designer does not have sufficient product data or knowledge [4, 5].

EEA is carried out in five steps. The first step is preparation, which contains the goal and scope of the study, creating a team that will conduct the study, and identification of environmental demands and regulations related to the product. This team represents different sectors or parts in the lifecycle of the product. Positions that can be useful in the team are, marketing, project leader, designer, manufacturing, after sales service etc. EEA uses the skill, competence and experience in the group that carries out the analysis. EEA focuses on the environmental requirements that authorities, internal and external actors put on the product. Requirements can be both absolute (legislation, decrees, etc.) and future legislation. Future demands do not have to be met immediately but it is usually an advantage to fulfil them nevertheless. Policy documents taken by the company have the same value as an absolute requirement from an authority.

Step two is the inventory of the product's lifecycle. Depending on the size of the studied system and the need for detailed information, this phase can be very time consuming. The inventory starts with the knowledge that the group possesses. In addition to this extra information is usually needed. When extra information is searched for it is important to choose sources with high reliability. The third step analyses the data from the inventory. This phase comprises an assessment of the data and proposals for actions. The assessment can be conducted using several different evaluation methods [25]. From these a number of hot spots are extracted which will be taken care of later. During this phase proposals on how to improve the product are chosen.

Next step is to implement the chosen proposals followed by evaluation of the implemented proposals. An analysis is made according to the same pattern as before to see whether the proposed changes gave the expected result.

EEA is a qualitative tool, meaning that the tool focuses on improvements or a trend in the right direction for the product instead of assigning absolute values. The focus is on how to do something better instead of finding figures and values as measures of the current environmental impact of the product. The level of detail, as well as the time frame for EEA, can be varied depending on chosen definitions of goal and system boundaries, i.e. it is a flexible method. Assumptions and sources of data are accounted for in a transparent and understandable way.

Since EEA is a qualitative method the result is intended for internal use and especially for product development. One aim of EEA was to allow it to be used as early as possible in the development process. It is important to transform information to a result that can be used to make decisions early in the design process.

When working with EEA, the focus is on the available data, as compared to LCA that has a focus on the data missing to be able to make a complete analysis. An analysis is carried out by a group of people forming a multifunctional team. As with all methods, EEA has some advantages and disadvantages, as shown in Table 1.

Table 1. Advantages and disadvantages with EEA [3]

Advantages	**Disadvantages**
Can be integrated in an EMS	Two technical systems cannot be compared
It is a qualitative method, can be used early in the development process	Does not give a quantitative result
It is based on environmental requirements	The result cannot be communicated outside the company
Easy to find data	
Less time consuming than LCA	

5.3 Life Cycle Assessment (LCA)

LCA is the most well-known and well-documented academic method for evaluation and assessment of environmental performance of a product, or service. This method is quantitative; to be able to use LCA numerical data are needed [26]. LCA is described in the ISO 14 040-43, and in many textbooks and research publications.

An LCA study is conducted in a number of phases. In the first phase the goal and scope of the study are defined. Issues to be defined here are what data quality is needed, what is the system that shall be studied, and the aim of the study etc. The second phase is to collect all data needed for the study. This is usually the most time consuming part of an LCA study.

The impact assessment comprises classification, characterisation and evaluation. This last step can be looked upon as a funnel. The collection phase provides a variety of data, which in the third part is transformed into structured and aggregated information.

A search in the literature gives some of the key advantages and disadvantages for LCA, table 2.

Table 2. Advantages and disadvantages with LCA [20, 27-33]

Advantages	**Disadvantages**
Unveils material- and energy-flows upstream and downstream that could have been hidden with an other method	Performing an LCA on a new product/ process is expensive and difficult in relation to the result
Gives support for decisions for new ideas to meet the demands of the function with less environmental impact	Data is often missing and needs to be based on short series of measures and theoretical calculations
A base for making checklists/guidelines used in Design for Environment	To be able to carry out an LCA specialist knowledge is needed
The result can provide solutions to environmental controversy	The data quality is not always sufficient for the purpose and this can make a big difference in the final result
The result is given in values, based on activities and flows	The time aspect makes it difficult to use in a product development process because it takes time and resources
Possible to communicate to a third person	There is a lack of comparable and reliable LCA/LCI data
The possibility to compare environmental performance	No weighting-method is generally accepted

6. DISCUSSION

To be able to extract one idea that can be a success (from 3000 ideas), tools are needed that can yield a result quickly and efficiently. Research shows that EEA can evaluate the environmental performance of a new product concept faster than LCA [4, 5] and this can be a reason for using EEA to a higher extent in a radical innovation process. A design handbook can be used as support during the development process. This method favours experience from the designer. The most important success factor pointed out in early stages in the new product development is intuition [34, 35]. Since EEA is based on qualitative information, intuition and experience from the designer is an important factor, this may be found subjectively. The team also creates networks within and outside the organisation that are important for moving a radical innovation project forward.

For making incremental changes, models and data from earlier products can be used. Less work needs to be carried out. This favours LCA since the tool works with a model containing quantitative data built in a model that can be modelled in software. Changing this model and evaluating different ideas can be done with a minimal expenditure of time and resources.

One extra problem can be that problems are never "black or white" but continuous scales. It can be illustrated how incremental and radical changes can be further divided into major and minor innovation changes. These changes depend on whether new knowledge in the product is high or low, and whether the new technology has been added or substituted into the product. If the different types of innovation are further divided the difference between them becomes less and less. When the difference becomes small it is difficult to provide clear guidelines on which tool to use in the different types.

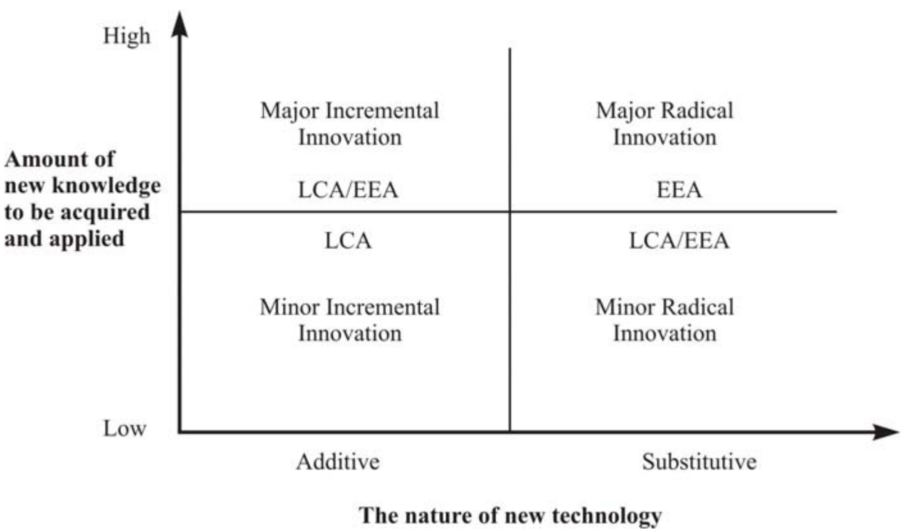


Figure 3. The "innovation plot" describing radical vs incremental [36]

For minor radical changes and major incremental changes both EEA and LCA can be used. For minor radical changes LCA can be useful because the existing numerical model from the initial design process can be used. For major incremental changes it is the other way around, major changes favour a screening method to achieve a rapid evaluation of concepts.

Since a design handbook can contain information such as design solutions, steering documents, checklists, and guidelines, it can be useful regardless of which kind of innovation companies have. It is important that a design handbook should be a dynamic tool where experiences from earlier projects are collected. The lone designer needs support, which is then easy to find. In the end the quality of the work is often dependent on decisions taken by a single designer or by the design team in the design process. The mere existence of a structure or a model to follow does not automatically result in a sustainable product [37]. But the possibility is higher when a company is working according to well-defined procedures and with tools that are easy to implement and easy to learn [38].

It has been stated before that it is not easy to integrate environmental concerns into the product development process. According to Handfield [39] and Magnusson [40] management must set up environmental targets and goals and these shall be formulated in the same way as other targets in the technical specification. Another way of reducing the gap between theory and practice when it comes to implementing environmental concerns is to put a cost on creating a bad environment. This technique may be applied regardless of whether the development is radical or incremental.

REFERENCES

1. D.G. Ullman, 2002, The Mechanical Design Process. New York, USA, McGraw-Hill Higher Education, 0-07-112281-8.
2. B. Bergman, and B. Klefsjö, 1994, Quality from Consumer Needs to Consumer Satisfaction, Studentlitteratur, Lund, ISBN 91-44-46331-6.
3. M. Lindahl, 2000, Environmental Effect Analysis (EEA) – An approach to design for environment, Licentiate thesis, Dept. of Chemical Engineering and Technology, KTH, Stockholm, ISBN 91-973906-2-3.
4. M. Lindahl, C. Jensen, and J. Tingström, 2000, A Comparison between the Environmental Effect Analysis and the Life Cycle Assessment methods – Based on four case studies, Proceedings of the 7th International Seminar on Life Cycle Engineering, Life Cycle Planning, Design and Management for Eco-Products and Systems, November 27-29, Tokyo.
5. M. Lindahl, J. Tingström, and C. Jensen, 2000, A small textbook about Environmental Effect Analysis (in Swedish), Department of Technology, University of Kalmar, ISBN 91-973906-0-7.
6. G. Stevens, and J. Burley, 1997, 3000 raw ideas = 1 commercial success! Research Technology Management, 40:16-27.
7. A.F. Chalmers, 1999, What is this thing called science?, Open University Press, ISBN 0-335-20109-1.
8. J. Tidd, J. Bessant, and Pavitt, 1997, Managing innovation: integrating technological, market and organisational change, Wiley, New York, ISBN 0-471-49615-4, 1997.
9. C.M. Christensen, 1992, Exploring the limits of the technology S-curve, Part 1 Component technology, Production and Operational Management, 1:334-357.

10. A. Afuah, 1998, Innovation management, strategies, implementation and profits, Oxford university press, New York.
11. A. Meyer, G. Brooks, and J. Goes, 1990, Environmental jolts and industry revolution: organisational responses to discontinuous changes, Strategic Management Journal, 11:93-110.
12. M.L. Tuchman, and P. Andersson, 1986, Technological discontinuities and organisational environment, Administrative Science Quarterly, 31:439-465.
13. C. Koberg, D. Detienne, and K. Heppard, 2003, An empirical test of environmental, organisational, and process factors affecting incremental and radical innovation, Journal of high Technology, 14:21-45.
14. J. Monroe, 1993, Winning in high-tech markets, Harvard School Press, Boston, 1993.
15. M.L. Tuchman, and E. Romanelli, 1985, Organisation evolution: a metamorphosis model of convergence and reorientation, in Cummins, L.L. and Staw, B.M. (Ed,) Research in organisation behaviour, 7:171-222, JAI press, Greenwich, CT.
16. J. Hall, and R. Kerr, 2003, Innovation dynamics and environmental technologies: the emergence of fuel cell technology, Journal of Cleaner Production, 11:459-471.
17. J.A. Abernathy, and J.M. Utterback, 1978, Patterns of Industrial Innovation, Technology Review, 80:40-47.
18. F. Suarez, and J. Utterback, 1995, Dominant design and survival of firms, Strategic Management Journal, 16:415-431.
19. J. Kortman, R. van Berkel, and M. Lafleur, 1995, Towards an environmental design toolbox for complex products; Preliminary results and experiences from selected projects, CONCEPT, Clean Electronics Products and Technology, 9-11 October, Edinburgh, Scotland.
20. N. Wrisberg, H.A. Udo de Haes, U. Triebswetter, P. Eder, and R. Clift, 2000, Analytical tools for environmental design and management in a systems perspective, Eds. Report to the European Union. CHAINET - European Network on Chain Analysis for Environmental Decision Support, Leiden, NL, 326.
21. L. Skoglund, and J. Svensson, 2002, Environmental driven design in SMEs – a need survey (in Swedish), Master thesis, Dept. of Technology, University of Kalmar.
22. Volvo, 1999a, Corporate Standard STD 1009, 1, Chemical substances which must not be used within the Volvo Group (black list), Gothenburg, Sweden.
23. Volvo, 1999b, Corporate Standard STD 1009, 11, Chemical substances whose use within the Volvo Group shall be limited (grey list), Gothenburg, Sweden.
24. Volvo, 1999c Corporate Standard STD 1009, 2, Substitutes for hazardous chemical substances (white list), Gothenburg, Sweden.
25. J. Tingström, and M. Lindahl, 2001, Evaluation methods for the Environmental Effect Analysis, CIRP, seminar on life cycle engineering, life cycle planning, design and management for eco-products and systems, June 18-20, Varna, Bulgaria.
26. C.J. Rydh, M. Lindahl, and J. Tingström, 2003,. Life cycle assessment – a method for evaluating products and services (in Swedish), Studentlitteratur, Lund, ISBN 91-44-02447-1-9.
27. P.R. Portney, 1993-94, The price is right: Making use of Life Cycle Analyses, Issue in Science and Technology, 10:69-75.
28. L. Lave, E. Cobas-Flores, C. Hendrickson, and F. McMichael, 1995, Using Input – Output Analysis to estimate Economy – wide Discharges, Environmental Science & Technology, 29:420A-426A.
29. J. Fiksel, 1996, Design for Environment, McGraw-Hill, New York, ISBN 0-07-020972-3.

30. C.T. Hendrickson, A. Horvath, S. Joshi, M. Klausner, L.B. Lave, and F.C. McMichael, 1997, Comparing Two Life Cycle Assessment Approaches: A Process Model- vs. Economic Input-Output-Based Approach. Proceedings of the 1997 IEEE International Symposium on Electronics and the Environment, San Francisco, CA
31. E.G. Hertwich, P.S. William, and C.P. Koshland, 1997, Evaluating the environmental impact of products and processes: a comparison of six methods, The Science of the Total Environment, 196:13-29.
32. C. Luttropp, 1997, Design for Disassembly – Environmentally Adapted Product Development Based on Prepared Disassembly and Sorting, Doctoral thesis, Royal Institute of Technology, Stockholm, ISRN KTH/MMK/R – 97/5 – SE.
33. S. Ritzén, J. Bäckmar, and M. Norell, 1997, Product development - Integration of environmental aspects, Proceedings 4th International Seminar on Life Cycle Engineering, Berlin, Germany.
34. G. Stevens, J. Burley, and R. Divine, 1999, Creativity + business discipline = Higher profits faster from new product development, Journal of Product Innovation Management, 16:455-468.
35. G. Stevens, and J. Burley, 2003, Piloting the rocket of radical innovation, Research Technology Management, 46:16-25.
36. R. Hall, and P. Andriani, 2003, Managing knowledge associated with innovation, Journal of Business Research, 56:145-152.
37. S. Poole, M. Simons, A. Sweatman, T.S. Bhamar, S. Evans, and T.C. McAloone, 1999, Integrating Environmental Decisions into the Product Development Process: Part 2 The Later Stages, EcoDesign 1999, pp 334-337, Tokyo.
38. J. Tingström, 2003, Environmental Adapted Design – with focus on Environmental Effect Analysis, Licentiate thesis, Royal Institute of Technology, Stockholm, ISBN 91-973906-5-8, 2003.
39. R. Handfield, S. Melnyk R. Calantone, and S. Curkovic, S., 2001, Integrating Environmental Concerns into the Design Process: The gap between Theory and Practice, IEEE transactions on Engineering Management, 48:189-208.
40. T. Magnusson, 2003, Architectural or modular innovation? Managing discontinuous product development in response to challenging environmental performance targets, International Journal of Innovation Management, 7:1-26.

EVOLUTION OF COLLECTIVE PRACTICES IN DESIGN BY INTEGRATION OF NEW METHODS

Flore Perrin, Dominique Millet, andAméziane Aoussat
Laboratoire CPI, Ecole Nationale Supérieure d'Arts et Métiers, Paris, France.Ameziane.Aoussat@paris.ensam.fr

Abstract: This article explores the problem of the evolution of collective practices in terms of design and offers the solution of doing this operation by integrating new methods. In order to guarantee the success of the integration of a new method, the process of change is viewed as evolving from being one that is imposed to one that is participatory. An experiment, carried out at Bourjois, a cosmetics company, illustrates this method through the integration of functional analysis.

Key words: Collective practices, Integration, Method, Learning, Co-construction.

1. INTRODUCTION

At the beginning of the twentieth century, the performance and productivity of a company were closely related. Then the idea of design came to occupy a central place in the performance of the company's productivity. This evolution generated important organizational changes. Interdisciplinary work groups appeared, including product development and production engineers, quality control, marketing etc. Today, innovation has become a very important part of a company's performance. The number of disciplines, and therefore of participants in the design of a product, has grown considerably (ergonomics, sensorial evaluation etc.). These actors of design and innovation, specialists in their respective disciplines, have to learn to work together. Thus, the performance of a company is increasingly linked to the ability to work collectively. In order to do so, the individuals have to evolve new collective practices.

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 45-56.

We know of two kinds of evolution of collective practices: the improvement of collective practices that does not imply a radical change in the designer's mode of operation, and the transformation of collective practices for which the design team has to abandon its old practices to acquire new ones. This type of evolution implies the evolution of each individual's behaviour. It is more difficult to put in place.

There are several ways of operating an evolution of collective practices. In this document, it is proposed to work specifically on the transformation of collective practices in matters of design by the integration of new methods.

The first stages of change, namely the diagnostic and the choice of method, have been dealt with by numerous researchers [1-3]. They will therefore not be discussed here.

The first part of this article will explore the impact of the integration of a new method on the learning of individuals. The second part will envisage the kind of learning process that is favourable for the evolution of collective practices. Thirdly, the integration of a new method through a process of evolutionary change will be considered. Lastly, a fourth part will be devoted to the testing of the hypotheses laid out in the two preceding parts based on a trial carried out at Bourjois, a cosmetics company, on the integration of functional analysis.

2. THE IMPACT OF THE INTEGRATON OF A NEW METHOD ON THE LEARNING PROCESS OF INDIVIDUALS

The introduction of a new method in a design process provokes a number of disturbances. Individuals resist change; they are puzzled and try to project into the future. After a phase of imbalance, they very quickly adopt a "mode of regulation". Guillevic has identified three modes in the context of the integration of a new tool: *restructuring*, *articulation* and *rejection* [4].

Restructuring corresponds to a reorganization of individuals' schemes of action. They have put aside their old mode of action and have adopted a new one, in agreement with the new method.

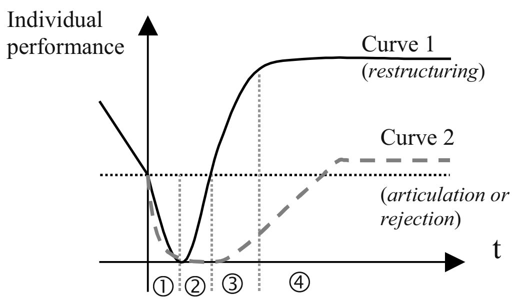
Articulation corresponds to a partial evolution of the individuals. Here: "the project is technically completed, but culturally incomplete"[1] [5]. Under the pressure of the evolving environment during the method's integration, individuals try to find a balance between the old procedures and those specified by the new method.

[1] Quote from a manager of SOFREL [5]

Finally, the last mode of regulation identified is ***rejection***, in which the team tries to find its way and repeats familiar patterns before giving up.

Articulation and *rejection* are the two modes of regulation that need to be avoided, the team's performance running the risk of never being improved.

At the beginning of integration, the individuals experience many disturbances due, according to Guillevic, to the "search for an explanation of mistakes in the contradictions that emerge between anterior operational knowledge and the operational needs of the new tool" [6]. During this period, the individual is unsettled. He oscillates between his old and his new landmarks. This imbalance acts directly on his performance. The evolution of each individual's performance during the integration of a new method is represented in Fig. 1, in a very qualitative way.



t = 0 beginning of the integration of a new method

Figure 1: Evolution of individual performance during the integration of a new method.

The phase of imbalance corresponds to phase ① (see curve 1). Performance drops. In phase ②, the team evolves, learns and individual performance improves. However, the benefits of the method cannot yet be observed (in relation to the initial situation t = 0). In phase ③ the benefits of a well-integrated method can be monitored and effectiveness increases. Lastly, in phase ④, the individual performance is stabilized, the latter having mastered the method.

The first curve illustrates a successful integration, corresponding to a *restructuring* mode of regulation. The second curve represents the evolution of the performance of an individual who adopted *articulation* or *rejection*. The mode of regulation can be detected at the very beginning of training. This is why phases ① and ② are fundamental. They remain, however, the most dangerous, each individual's performance being inferior to their initial performance.

The impact of integration of new methods on the learning of individuals having been explored, a second part will now address the type of training that is most favourable to *restructuring*.

3. THE EVOLUTION OF COLLECTIVE PRACTICES REQUIRES DOUBLE LOOP LEARNING

3.1 The Governing Variables: a lever for the evolution of collective practices

Many authors include, in the representation of a company's Organization, a fundamental notion. Mintzberg labels it "*ideology"* [7], Tushman and Nadler "*core values*" [8] and Argyris and Schön "*governing variables*" [9]. For the remainder of the article, the concept of Argyris and Schön "*governing variables"* will be used.

The *governing variables*, strongly determined by a brand's strategy, constitute an important lever for initiating a process of change. For Tushman and Nadler, the *core values*, when they are in agreement with the processes at play in the company, enable the "motivation of the different behaviours" [8]. Mintzberg explains that thanks to *ideology*, "individuals and the organization generate an "*esprit de corps*", a "sense of mission" and the integration of individuals in the organization to produce a synergy" [7].

But in many companies, the *governing variables* are not explicit. Kotter has observed that many failures to implement change were due to a "*lack of vision*". He observed indeed the wide diffusion of plans, directives and programs, but never that of *vision* [10].

Argyris and Schön have observed the same phenomenon and link it to an incomplete mode of learning [9].

3.2 Double loop learning.

According to Argyris and Schön, all ***Consequences*** to be achieved are obtained by the unfolding of an ***Action Strategy***, itself informed by the company's ***Governing Variables*** (Fig. 2) [9].

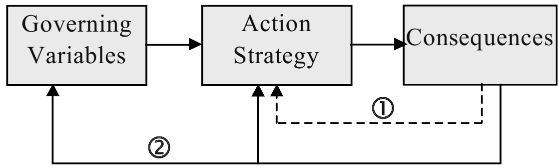


Figure 2. Argyris and Schön double loop learning [9].

The ***Governing Variables*** correspond to the teams' representation of the product, of the organization, but also of the market. It is those representations that determine their approach of a product's design. If the company does not address the transmission of a common vision of the product and the organization, individuals will retain their own representations of the product, linked to their past experience, and their own representation of the organization, linked to their job description and their role in the company.

Action Strategy, for its part, is made up of methods, processes and procedures in operation in the company. It includes, moreover, all the individuals' habits, directly linked to their *Governing Variables*. Indeed, if a member of the design team perceives the product to be conceived as an addition of elements, he will in the first place design its component parts, and will then assemble them in order to create a product. On the other hand, if he perceives the product as a sensory experience, he will conceive the product around the sensory outputs aimed for and will work first and foremost on the search for materials. He is not always conscious of the approach he has taken, he is working with habits.

Lastly, the ***Consequences*** correspond to products and to the organizational structure of the company.

Each type of evolution of collective practices is linked to a type of learning process:

The **improvement of collective practices** does not require the modification of the *governing variables*. The objective of integration of a new method is, in this case, to provide a strategy and tools for the design team so that its current practices are improved. This does not necessitate a change of habits. Only a*ction strategy* needs to be changed. For Argyris and

Schön, this amounts to generating a "***simple loop learning***" (see loop ① Fig. 2).

The **transformation of collective practices** requires a change of habits. It is necessary to act on the *action strategy* while integrating a new method, but mostly to evolve the *governing variables* so that they are in agreement with the new aims to be achieved (*consequences*). This corresponds to "***double loop learning***" (see loop ② Fig. 2).

Since this article concentrates on the transformation of collective practices, it will only consider *double loop learning*. A third part will now address the type of processes of change favourable to *double loop learning*, the appropriation of change and of method.

4. THE INTEGRATION OF A METHOD: PROGRESSION FROM A PROCESS OF CHANGE IMPOSED FROM ABOVE TO A PARTICIPATIVE ONE

To guarantee the success of the integration of a method and therefore the transformation of collective practices, it is necessary to ascertain the assimilation of change on one hand and the appropriation of the method[2] and the tool[3] on the other. It is therefore necessary to put in place a process of change favourable to these two appropriations.

In the literature, three main types of change have been identified [11], [12]. Claret's classification is used throughout this article, namely: the imposed, participative or negotiated processes of change [12].

The imposed process of change corresponds here to a coerced integration of the new method. The ***participative process of change***, for its part, unfolds wholly from the operational team, from the diagnostic to the implementation of the chosen solution. The operational team being actors of change, this type of change generates, according to Claret, a "true collective dynamic of change" [12]. Lastly, the third process is the ***negotiated process of change*** for which "management provides the impetus, by indicating the objectives and by choosing a frame of thought for the procedure to be followed" [12]. A workgroup made up of operational actors is then set up to implement the objectives and the proposals chosen by management.

[2] Method is a rational undertaking that needs to be followed in order to achieve desired effects. For example, FMECA, value analysis, functional analysis, QFD are all methods.

[3] A tool is a work tool. It is the technical consequence of method. It can be software, a form, an Excel sheet, etc. It is made up of a scenario of use, an architecture and of components.

While certain researchers propose to choose the kind of process to be followed before change takes place, this article proposes to begin with the integration of a new method along an imposed process of change, in order to transform it into a participative one, via different stages of negotiated change.

In the first instance, each phase of individual learning (Fig. 1) was associated with a specific process of change:

Phase ① according to an *Imposed Process of Change*:

At the beginning of integration, the design team goes through a difficult phase of change during which individuals are unsettled and see their performance drop. It corresponds, for everyone, to a period of de-motivation and discouragement that can provoke the *articulation* or the *rejection* of the method. In order to prevent these two modes of regulation, it seems reasonable to lead the design team strongly in order to help it achieve a new balance. The imposed process of change does seem to be appropriate to begin the integration of a new method.

In order to achieve a new balance and initiate *restructuring*, it is necessary for the team to appropriate change. Sonntag and Oget have shown that the "appropriation of change necessarily goes through a phase of speech, discussion, of necessary articulation into words" and it is therefore necessary for "the subjects to be able to deconstruct and reconstruct their representation through symbolic exchanges" [13]. This stage, in a process of change, corresponds to what Lewin has called *de-crystallization* [14, 15]. It seems necessary for the appropriation of change and has to be integrated in the first phase of the process of change that is then in the "imposed" form.

Phase ② according to a *Negotiated Process of Change*:

During this phase, the double loop learning has begun, the team appropriates change and sees its performance improve. In order to begin appropriating the method, and then the tool, it seems necessary to make the team participate more actively in the process of change. However, its performance still remains inferior to its initial performance. It is therefore necessary to accompany and guide the team through this change. This is the reason why the negotiated process of change seems to be appropriate to this phase of learning.

Phase ③ according to a Participative Process of Change.

In this phase, the design team has appropriated the change, the double loop learning is completed, individual performance has risen above the initial performance, the benefits of the method are manifest and the design team is convinced of the usefulness of the method. In this favourable context, the design team can freely finalize change. A participative process of change is therefore proposed to complete the integration of the new method.

For phases ② and ③, it is necessary to ascertain the appropriation of the method and the tool. As Parise recommends [16], it is proposed in this article to quickly associate the future users of the method (here the design team) to the design of the tool. It is necessary to put in place **the co-construction** of the tool by the latter.

At the source of the process of change, the person in charge of the integration of the new method (the integrator), associating with the top design management in a *Guiding Coalition* [17] will have to elaborate his integration strategy.

From the diagnostic, the study of the individual and collective consequences of the introduction of the new method, and the proposed evolutionary process of change, the *Guiding Coalition* will be able to conceive the process of integration of the method. It will then choose the different stages of the process that will allow the *de-crystallization* and the co-construction, and then will identify the stages favourable to the transfer from an imposed, to a negotiated and finally to a participative process. Lastly, it will meticulously prepare the first phases of integration of which it will be the main actor, the process of change being first imposed and then negotiated.

In this fourth part, the results of a trial of this hypothesis carried out on the integration of the functional analysis at Bourjois, a cosmetics company, will be exposed.

5. HYPOTHESIS TRIAL : INTEGRATION OF THE FUNCTIONAL ANALYSIS AT BOURJOIS

5.1 The project and the strategy undertaken

Following a diagnostic that revealed usability witness of certain Bourjois products, it was decided to integrate functional analysis (FA). This diagnostic showed that the design teams seldom integrated the consumer at the beginning of the design and therefore did not anticipate the functional problems of the products they designed. They were in a mode of operation where they designed the product, controlled it and modified it in an iterative way.

It was therefore necessary to change the collective practices of the design teams so that they would take into account the consumer restraints from the outset of the project.

Functional analysis was chosen as the best method to engender the transformation of collective practices.

In the first instance, the *Guiding Coalition* in charge of the project undertook the construction of the integration process along the guidelines provided in 2 and 3. In the context of the co-construction of the tool, it had planned the formation of a workgroup composed of a representative of each component of the project design group namely: a product manager, a member of the R&D Formulation team, a packaging engineer and a production engineer. For reasons of availability, the latter was unable to attend the workgroup (the effects of this absence will be observed in 4.2). Then the *Guiding Coalition* deconstructed the integration process in three steps: de-crystallization, the global design of the tool, and lastly the detailed design.

1. De-crystallization

In agreement with the recommendation made in 3., the integration had to begin with an imposed process of change. The *Guiding Coalition* was the principal actor of this step. It was therefore necessary for it to prepare it.

It began by decomposing the FA method in order to adapt it to the design of cosmetic products. From the 6 tools and the 13 steps contained in a theoretical FA, it chose 2 tools and 4 steps. It then constructed four scenarios for the use of the method, that it analysed in order to select that most appropriate.

The chosen scenario is based on the use of generic requirement file and unfolds in two steps:

1. *Exchanges in the work groups around the Generic Requirement File.*
2. *Choice of packaging.*

Once the method and the scenario of its use were chosen, the *guiding coalition* built its training plan. It consisted of a brief theoretical training session, followed by the application of the method by the design teams on their own projects.

Once the preparation for this step was done, the integration of the FA began with the complete presentation of the project to the design teams. This presentation described the origins of the project, its objectives, the benefits of the chosen method, the chosen scenario and the steps to be followed. The exchanges that took place during this meeting initiated the de-crystallization, which was pursued in an “analysis of the project needs” carried out by the work group, with the help of the *Guiding Coalition*. The work group then undertook the training.

2. Co-construction : global design.

During this stage, it was envisaged to follow a negotiated process of change. The *Guiding Coalition* was still present at this stage, and the future users of the method were to become actors of change. Still in a phase of imbalance, they did not have full degree of freedom. In order to associate them with the process of change, the *Guiding Coalition* built between four and eight possible architectures of tool, one for each stage of the scenario. For example for the first stage "*1- Exchanges in the work groups around the generic requirement file"*, the *Guiding Coalition* proposed to provoke this exchange by undertaking a study of competition using the generic requirement file as a filter, or by filling in generic requirement file to turn it into the specific requirement file for the product to be designed, or else by introducing a role-play situation etc. Thus the future users constructed a more precise scenario of the use of the tool. The latter became:

1. *Specifying the Generic requirement file for the product to be designed.*
2. *Evaluation of the three packaging solutions / choice of a solution.*

3. Co-construction : detailed design.

The *Guiding Coalition*, after having checked whether the double loop learning was clearly engaged, moved on to a participative process of change. It left the work group free to elaborate the detailed design of the tool. The group then chose freely the type of tool, its software application etc.

5.2 Results and conclusions of this trial

This trial is still being evaluated, and here are the first results:

Evolution of the collective practices.

The indicator to be observed for this criterion was the evolution of the use of FA in the process of design for successive projects. At the beginning of integration, we observe that FA was only applied at the end of the design of a product, before the transition to production. It was then use as a mode of control. The design team tried to find the requirements of the product that it had failed to address. This observation was made on only one project. For the three following projects, FA was used during the choice of packaging. The method was then used as a helping tool for the decision making process. Finally, for the last four projects, FA was used as early as the second meeting of the work group. The design teams had therefore switched to the anticipation of the product's requirements. The evolution of the collective practices therefore seems to have taken place.

The appropriation of the tool.

The first indicator of the tool's appropriation by its users is how often it is used. Since the beginning of the integration, FA was applied to eight of the nine new projects. Moreover, out of twelve FA meetings (stage 1 and 2 of the scenario of the use of method), product managers were in attendance at every meeting, packaging engineers at eleven meetings, members of the R&D Formulation team at ten of them, while the production engineers at only five meetings. We saw that they were no production engineers in the work group. This seems to confirm that participation of the future users in the design of the tool encourages the appropriation of the tool. Finally, the last indicator is recourse to the method for another use. During the integration of FA, it was observed that some product managers took inspiration from the generic requirement file to write the briefs of new projects, that some packaging engineers used it to prepare and structure the risk analyses, while some members of the R&D Formulation team took inspiration from them to prepare consumer tests. Once more, there was no sign of appropriation of the tool by production engineers as far as this indicator is concerned.

The first results of the integration of functional analysis are positive for now, and seem to confirm the hypotheses outlined in 2 and 3.

CONCLUSION

In this article, the evolution of collective practices in design through the integration of new methods was proposed. After having discussed the impact of the integration of a new method on the learning process of individuals as well as the kind of learning (double loop) necessary for the transformation of collective practices, the evolution of the process of change from an imposed mode to a participative mode was proposed, via various levels of negotiated modes.

A first trial seems to show that this kind of process is favourable to the transformation of collective practices on one hand and the appropriation of the tool on the other.

However, the choice of an integration strategy by the *Guiding Coalition* (design and choice of a scenario of use of the method, construction of a integration process etc.) remains tied to its intuition. In order to accompany it in the initial stage of change, the next proposed step is to work on a piloting tool of integration of new methods. This tool would include, in the first instance, a study of the impact of integration of the method, which would allow the integrator to make the right choices during the preparation of the process of integration of the new method in product design process.

REFERENCES

1. V. Chanal, (1995), Le management de l'innovation de produit industriel : Mise en œuvre d'une démarche de diagnostic pour améliorer notre compréhension du processus. Thèse de doctorat de l'Ecole Supérieure des Affaires. Université Pierre Mendès-France de Grenoble.
2. P. Vadcard, (1996), Aide à la programmation de l'utilisation des outils en conception de produit. Thèse de doctorat de l'Ecole Nationale Supérieure d'Arts et Métiers. N° 96-33.
3. A. Chauvet, (1995), Méthode de management : le guide, Les Editions d'Organisation, Paris.
4. C. Guillevic, (1991), Distinction between two Types of Competences : Expertise Centred on the Means or the Goal, in Y. Queinnec & al., Designing of Everyone, Taylor & Francis, London.
5. T. Bertrand, (2003), Actes de modéliser et appropriation du changement : leçons tirées d'une recherche intervention, Actes du 5e congrès International de Génie Industriel, Québec, octobre 2003.
6. C. Guillevic, (1992), La psychologie du travail, Paris, Nathan.
7. H. Mintzberg, (1989), Le management. Voyage au centre des organisations, Ch.6-14, Les Editions d'Organisation, Paris.
8. M. Tushman, D. Nadler, (1986), Organizing for Innovation, California Management Review, Volume XXVIII.
9. C. Argyris, D. Schon, (2002), Apprentissage organisationnel. Théorie, méthode, pratique, Ch.1, DeBoeck University, Bruxelles.
10. J. Kotter, (1995), Leading Change: Why Transformation Efforts Fail, Harvard Business Review, March- April 1995, pp. 59-67.
11. M. Liu, (1997), Fondement et pratiques de la recherche action. L'Harmattan, Paris.
12. N. Claret, (1995), Proposition d'une méthode d'évaluation des processus de changement pilotés par des consultants, Thèse de doctorat de l'Ecole Centrale de Paris.
13. M. Sonntag, D. Oget, (2003), « la compétence collective. Condition théoriques requises et résistance du « facteur humain » », Actes du 5e congrès International de Génie Industriel, Québec, octobre 2003.
14. K. Lewin, (1947), Frontiers in group Dynamics II, Human Relations, Vol.1.
15. M. Combes, C. Verzat, (1999), La question de l'organisation : le défi de l'autonomie, Actes du 3e congrès International de Génie Industriel, Montréal.
16. R. Parise, (1997), Les processus d'appropriation de l'outil: Hypothèses sur le rôle du collectif de travail, Thèse de doctorat de l'Université de Toulouse Le-Mirail. n° TR 1015-1997-87.
17. J. Kotter, L. Schlesinger, (1979), Choosing strategy for change, Harvard Business Review, March- April 1979, pp.106-114.

EFFECTIVE STRATEGIES IN THE REDESIGN PROCESS

Hui-Hsia Chen[1&2] and Manlai You[1]
1Graduate School of Design, National Yunlin University of Science and Technology, Taiwan
2Department of Information and Design, Taichung Healthcare and Management Unversity

Abstract: This paper illustrates how to conduct preference-based design research properly before redesigning a pleasurable product for success in the market. The objective of this paper is to ascertain certain elements of the redesigning process of a mouse. We would propose an effective approach from two well-known manufacturers; Logitech and KYE, to understand the critical properties manipulated to develop a pleasurable product. First, in this paper we discuss the properties and factors which contribute to a newly redesigned product, providing a customer-oriented specification to meet market demands. Additionally, this study proposes to minimize non-value variations to meet a broader market need. Second, results from a questionnaire, asking college students to answer based on their reactions to a particular mouse, were also adopted as guidelines in the formulation of a newly redesigned mouse intended for sale on the market. Using a computer mouse case study, we present a simple yet powerful method for (VERB) designing a pleasurable product. Third, we develop the "four pleasures" metrics to construct a product specification and provide insight into and trace the customer's mode of thinking when selecting a given product in the market. This paper therefore has generalized a design strategy for a product through an appropriate case study of a computer mouse product. Finally, companies can provide designers with a quick overview of the impact of the products and assist them in the design of a pleasurable product effectively and economically. It is hoped that the intention of this study will help provide effective metrics and design strategies for redesign progress before a product reaches the shelf.

Key words: Redesign; Pleasurable product; Preference-based design; Design strategy; Product development.

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 57-68.

1. INTRODUCTION

The importance of preference and attractiveness in a pleasurable product design has recently been increasing. To develop attractive products and systems, the research proposes creating rational approaches for “Preference-based Design”.

1.1 Identifying customer needs

The process of identifying customer needs is an integral part of the larger product development process and is most closely related to generating concepts, selecting concepts, competitive benchmarking, and the establishment of product specification. The *customer-needs activity* shown in Fig. 1, in relation to other front-end product development activities, can be thought of collectively as the concept development phase [1]. User needs must be at the forefront of thinking for the designer during the development of a successful product. In order to give the product relevance to operation, and help the designer to prioritize ways in which the product meets expectations for its performance, analysis of user needs must be critical and reflect precisely what the user does based on direct observation and feedback [2]. Pugh [3] emphasizes the importance of addressing user needs based on consultation in the form of user interviews and questionnaires; the examination of trends which may include buying habits, market data and investigation of user behaviours. Features that differentiate each product from the rest of the market, while still meeting identified needs, may be seen as an added value by the customer and give it a further competitive edge.

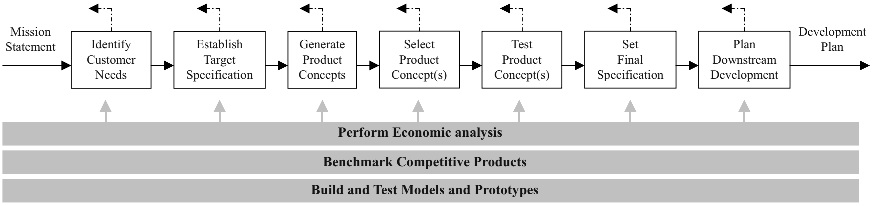


Figure 1. Customer-needs activity in relation to other concept development activities [1]

1.2 Organize the needs hierarchically

Functional analysis is a key component in determining the requirements of a product that will be included in the product design specification. Conducting a *needs analysis* for a product will reveal a number of preferred performance requirements based largely on market data and expert opinion. A functional analysis is essentially a method of breaking down the product into its component parts. Identifying and elaborating technical processes that serve to achieve sub-functions and the main function can develop that function analysis further [4].

1.3 Evaluating the competitive products

Cross-referencing needs analysis and functional analysis for a product may indicate design conflicts and leave questions unanswered. Solutions must be found and evaluated before a suitable concept design can be generated. The search for solutions begins with an analysis of what has already taken place and an evaluation of the effectiveness of existing products relative to their current situation. This can be achieved by mapping design features of current products using matrix analysis and parametric analysis [5]. Parametric analysis can be carried out as a 'desk top study' comparing specifications from different products. This requires an availability of comparative material for a range of products from different manufacturers. Matrix analysis enables a study to set its own parameters which can be measured and recorded for a range of existing products. Technical processes taken from function analysis for the product are in turn considered and solutions from existing products are identified and evaluated to determine their effectiveness.

1.4 Exploring the solutions for concept generation

Controlled convergence of design solutions can be achieved as the concept is 'narrowed down' to the most suitable solution through the addition or removal of ideas. All compatible solutions should be evaluated by cross-referencing needs and functional requirements [6]. Through this process those features that are considered to be impractical, inefficient or incompatible with other features can be discarded, leaving a host of preferable solutions for inclusion in a concept design.

1.5 Organize the needs into a hierarchy-the four pleasurable

Tiger [7] discovered and identified a useful way of classifying different types of pleasures. His framework models four conceptually distinct types of pleasures: physical, social, psychological and ideological. A Physio-pleasure includes pleasures related with touch, taste and smell as well as feelings of sensual pleasure. In a Socio-pleasure, this enjoyment derives from the relationships with others. A Psycho-pleasure pertains to people's cognitive and emotional reactions, whereas Ideo-pleasure pertains to people's values. Tiger posits that a structured approach facilitated using a framework can help ensure that possible benefits are not overlooked.

2. METHOD

In this paper, a computer mouse has been chosen as our model product. This study investigates how to conduct pre-design research properly before redesigning a product form to increase the success rate of a redesigned product in the market. It is important to note how to increase the interaction between a customer and a product. The chosen scenario entertains a product having already been designed and produced where, in the next phase; non-value variations from the previous one will be reduced.

The following methods can be divided into quantitative and qualitative characteristics. The quantitative method includes three stages. First, the study collected and classified existing information about mice produced by two well-known manufacturers, Logitech and KYE. Next, analysis of style formulation techniques was used to compare the differences in order to ascertain how certain elements were manipulated to create a newly redesigned mouse. The analyzed elements included shape, color, material/texture, and detailed treatment/design features, all of which are essential elements to style formulation.

Third, according to the above data and literature, we compiled a questionnaire listing all statements. It was answered by the 140 college students regarded as the target consumers in this particular study. The questionnaire contained three parts; the first and second parts included two kinds of response scales. Scales were also classified as categorical (rating) scale and comparative (ranking) scale, respectively. During the first part, the respondent was asked to agree or disagree with each statement, each response given a numerical score reflecting its degree of agreement, and the scores were totalled to measure the respondent's preference. We would measure the degree of agreement in size, weight, texture, material, form factor, color and size of button using the Likert 5 scale. The above properties are critical elements for a computer mouse, embodying the essence of the very product for most designers. The numbers indicate values assigned to each possible answer with 1 indicating a strong disagreement, 5 indicating a strong agreement, and 3 indicating a neutral position. This section also includes four sets of contrasting pairs of adjectives: large vs. small, heavy vs. light, geometry vs. curve, and symmetry vs. asymmetry. The second part of the questionnaire analyzed the importance of function for a mouse. We asked whether the respondents had a stronger agreement when deciding which function was important to a mouse, including the scrolling ball, rolling ball, wireless or optical enabled and high-tech images. We also skewed the degree of agreement using the Likert 5 scale. The third part provided the respondents with an open-ended question. Disregarding the questionnaire statements, what features or benefits are important for you when choosing to a mouse? We constructed an effective metric of designing a pleasurable mouse to meet the needs of the user for new product development.

Under the qualitative method, there were three designers, two senior college student designers, and one expert designer with six year's experience. During the redesign session, all designers were asked to model a computer mouse consistent with the four pleasures from the above analysis results.

3. RESULTS AND DISCUSSION

3.1 Sample descriptive statistical analysis

As part of the study, we can determine the sex and the design experience from the 140 college students (Table 1). All together, 35 males and 105 females were represented in the case study. Overall, there was a good ratio of design experience between males (54.3%) and females (45.7%) (Table 1).

Table 1. The cross-analysis between sex and design experience

			Design experience		
			No	Yes	Total
SEX	Male	Number	17	18	35
		Total percentage (%)	12.1%	12.9%	25.0%
	Female	Number	59	46	105
		Total percentage (%)	42.1%	32.9%	75.0%
Total		Number	76	64	140
		Total percentage (%)	54.3%	45.7%	100.0%

3.2 Preferred properties of a mouse

We can now locate and define the preferred properties of a mouse. First, *Volume* refers to its size, big or small. Second, *Weight* refers to its mass, heavy or light. Third, the *Property* of *Texture* refers to the smoothness of its surface or the roughness of its surface when holding it. Fourth, its *Material* refers to its outer casing, whether its plastic or metal. Its *color* refers to a mouse with the same color or devoid of color completely. Finally, its *Button* refers to the symmetric or asymmetric buttons.

As a part of the study, it was possible to extrapolate a figure for the property of *Form*, most being preferential to a higher mean score (i.e. 4.01) (Tables 2 and 3). For *Color*, the lowest mean score (i.e. 2.75) indicates less effectiveness to a mouse. Furthermore, the properties of *Weight*, *Texture*, and *Volume* have the closest margin. It is evident that the *Weight* is higher than the others in mean score. The property *Button* is 3.84 and M*aterial* is 3.72 respectively.

Table 2. Statistics of effective properties

		Volume	**Weight**	**Texture**	**Material**	**Form**	**Colour**	**Button**
Number	Valid	139	139	140	138	133	140	131
	False	1	1	0	2	7	0	9
	Mean	3.92	3.97	3.93	3.72	4.01	2.75	3.84
	Std.	.75	.68	.82	.85	.75	.93	.77
	Total	545	552	550	514	533	385	503

Table 3. Statistics of preference-based properties

	Volume		**Weight**		**Material**		**Texture**		**Form**		**Colour**	
Likert Scale	Freq.	Per. (%)	Freq.	Per. (%)	Freq.	Per. (%)	Freq.	Per. (%)	Freq.	Per. (%)	Freq.	Per. (%)
1			1	7	2	1.4	1	7	1	8	2	1.5
2	8	5.8	2	1.4	7	5.1	7	5.0	4	3.0	5	3.8
3	21	15.1	22	15.8	41	29.7	25	17.9	19	14.3	24	18.3
4	**84***	60.4	**89***	64.0	**65***	47.1	**75***	53.6	**78***	58.6	**81***	61.8
5	26	18.7	25	18.0	23	16.7	32	22.9	31	23.3	19	14.5
Total	139	100.0	139	100.0	138	100.0	140	100.0	133	100.0	131	100.0

3.3 The preferred functions of a mouse

Tables 4 and 5 illustrate the mean scores of those functions as being from 3.72 to 3.95. Forever, they indicate a close degree and the function “Optical” is preferred to a mouse.

Table 4. Statistics of function

		Scrolling ball	**Rolling ball**	**Wireless**	**Optical**
Number	Valid	137	140	140	140
	False	3	0	0	0
	Mean	3.88	3.89	3.72	3.95
	Std.	.99	.98	1.01	.88
	Total	531	544	521	553

Table 5. Statistics of preference-based functions

Likert Scale	Scrolling ball		Rolling ball		Wireless		Optical	
	Freq.	Per. (%)	Freq.	Per. (%)	Freq.	Per. (%)	Freq.	Per. (%)
1	1	.7	1	.7	4	2.9	1	.7
2	13	9.5	15	10.7	10	7.1	5	3.6
3	31	22.6	24	17.1	42	30.0	37	26.4
4	***49**	35.8	***59**	42.1	***49**	35.0	***54**	38.6
5	43	31.4	41	29.3	35	25.0	43	30.7
Total	137	100.0	140	100.0	140	100.0	140	100.0

3.4 Effective factors

After producing a correlating matrix and rotated matrix, the extraction of components is illustrated in Table 6. The 5 factors, according to factor eigenvalue >1 and factor loading > 0.5, have been extracted from 12 factors and these five factors account for 70.19 % of total variance explained. Therefore, we will name these five factors by their commonality. First, component 1 identifies the factor of product image that includes form, color and high-tech image. Second, component 2 identifies the tactile factor including material and texture. Third, component 3 identifies the physical factor and the technical factor including volume and weight.

Table 6. Principle factor analysis

	Component				
	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Form	***0.828**	0.131			
Color	***0.809**				0.149
High-tech	***0.587**		0.402	0.278	
image		***0.868**	0.146	0.101	
Material		***0.748**	0.302		0.191
Texture	0.212	0.184	***0.803**	-0.108	0.156
Volume		0.315	***0.800**		
Weight				***0.867**	
Wireless	0.491	0.120		***0.719**	
Optical	0.123	0.161	-0.125	-0.106	***0.816**
Rolling ball		-0.382	0.245	0.327	***0.580**
Scrolling Ball		0.443	0.276	0.130	0.485
Button					
Cumulative%	70.19%				
Factor name	**Product image**	**Tactus**	**Physics**	**Technology**	**Operation**

Fourth, component 4 identifies the technical factor, including wireless and optical. Finally, component 5 identifies operational factor including rolling ball and scrolling ball.

3.5 Correlation of different properties pertaining to sex and design experience

3.5.1 Correlation between sex and preference-based properties

Tables 7 and 8 show the different preferences between males and females in their respective properties: volume, weight, form style, button size, button number, button form and scrolling ball location. Sex vs. volume and sex vs. scrolling ball location have considerable significance. Males prefer large size mice (i.e. 14%: 10%) while females prefer smaller sized mce (i.e. 27%: 47%). Meanwhile, it is significant to note that a χ- square value of 0.032 <α (0.05) shows the selection of volume are different in regard to sex. Furthermore, males and females both prefer scrolling balls located on the topside of their mice. Also, it is quite significant with an χ- square value of 0.010 <α (0.05). Females however, (i.e. 81%: 17%) prefer a scrolling ball on the top more often than males (i.e. 14%: 12%). Finally, other properties make little obvious difference. Males and females prefer equal characteristics when choosing a mouse, i.e. light, curved form, large buttons, symmetric buttons, a scrolling ball on the topside and two buttons.

Table 7. Correlation between sex and preference-based properties

	Properties							
	Volume size		**Weight**		**Mouse's outlined form**		**Button size**	
SEX	Large	Small	Heavy	Light	Geometry	Curve	Large	Small
Male	20(14%)	15(10%)	30(21.8%)	5(3.5%)	6(4.5%)	28(20.8%)	31(22.2%)	4(2.9%)
Female	38(27%)	66(47%)	94(68.2%)	9(6.5%)	31(22.9)	70(51.8%)	88(62.8%)	17(12.1%)
Total	58(41%)	81(57%)	124(90%)	14(10%)	37(27.4%)	98(72.6%)	119(85%)	21(15%)
χ^2	0.032		0.348		0.14		0.494	

Table 8. Correlation between sex and preference-based properties

	Properties						
	Button number			**Button form**		**Scrolling ball location**	
SEX	2	3	4	Symmetry	Asymmetry	Up	Side
Male	18(12.8%)	15(10.7%)	2(1.4%)	24(17.3%)	11(7.9%)	18(14%)	12(9.3%)
Female	71(50.7%)	33(23.5%)	1(0.7%)	86(62.4%)	17(12.4%)	81(63%)	17(13.7%)
Total	89(63.5%)	48(34.2%)	3(2.1%)	110(79.7%)	28(20.3%)	99(77%)	29(23%)
χ^2	0.88			0.058		0.010	

3.5.2 Correlation between design experience and other properties

Tables 9 and 10 illustrate the different preferences between designers and non-designers in regard to properties: volume, weight, form style, button size, number of buttons, button form and scrolling ball location. Design experience vs. button form has much significance and a χ- square value is of 0.008 <α (0.05). Males and females generally prefer symmetric buttons (i.e. 47.8% and 6.9%), as well as non-designers (i.e. 31.8%: 17.1%). Moreover, the additional properties make little difference. Finally, these findings can determine males and females preferences with regard to differences in size, weight, form factor curvature, button size, button symmetry, presence of a scrolling ball on the topside and the presence of two buttons.

Table 9. Correlation between design experience and preference-based properties

	Properties							
Design Experience	**Volume size**		**Weight**		**Mouse's outlined form**		**Button size**	
	Large	Small	Heavy	Light	Geometry	Curve	Large	Small
No	28(20.1%)	48(34.6%)	7(5%)	68(49.5%)	24(17.8%)	49(36.25)	61(43.5%)	15(10.7%)
Yes	30(21.6%)	33(23.7%)	7(5%)	56(40.5%)	13(9.6%)	49(36.25%)	58(41.5%)	6(4.3%)
Total	58(41.7%)	81(58.3%)	14(10%)	124(90%)	37(27.4%)	98(72.5%)	119(85%)	21(15%)
χ^2	0.2		0.73		0.122		0.087	

Table 10. Correlation between design experience and preference-based properties

	Properties						
Design Experience	**Button number**			**Button form**		**Scrolling ball location**	
	2	3	4	Asymmetry	Symmetry	Up	Side
No	52(37.1%)	23(16.4%)	1(0.7%)	9(6.5%)	66(47.8%)	51(39.8%)	19(14.9%)
Yes	37(26.4%)	25(17.8%)	2(1.4%)	19(13.7%)	44(31.8%)	48(37.5%)	10(7.8%)
Total	89(63.5%)	48(34.2%)	3(2.1%)	28(20.2%)	110(79.7%)	99(77.3%)	29(22.7%)
χ^2	0.381			0.008		0.183	

3.6 The effective metric of the preference-based design

As noted, pleasure enhancement can be divided into four sub-categories: Physio-pleasure, Socio-pleasure, Psycho-pleasure and Ideo-pleasure. The four-pleasure framework may be used as a means of structuring thought with respect to pleasure with a mouse (Table 11).

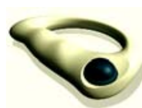
Table 11. The four criterions for experiencing pleasure using a mouse

Physio-pleasure	Physics	Useful functions (scrolling ball, wireless, optical, rolling ball) Ergonomics Easy to use (large button, 2 buttons, scrolling ball location- top (finger), operation)
	Tactus	Good in hand (light-weight, texture or sensorial pleasing feeling, volume size: small, form) Easy to hold (weight, volume, texture) Feel comfortable against the hand (material, hand sweat-resistance, for use for over several hours)
Psycho-pleasure	Technology	Wireless, Optical
	Operation	Easy to use at the first attempt (guessable, compatibility) Fun to use Scrolling ball location- up (finger)
Socio-pleasure	Product image	Color, Form, High-tech image
Ideo-pleasure	Color	Aesthetic pleasure (curve & streamline mouse form, scrolling ball location – up (finger))

3.7 Concept Generation and Concept Selection

Table 12 considers these three groupings: generating a concept, final concept selection and computer modelling. Table 12 illustrates three designers' different models for a mouse based on the effective metric of the preference-based design. We find that these sketches of redesigned mice match the four aforementioned pleasure criterions. Designer A focuses on texture and a product with emphasis on pleasing the senses. Its form closely resembles that of a cell phone. Designer B begins with a lateral thinking model which emphasizes the shape of the hand. More importantly, it focuses on operation and ergonomics. Moreover, the expert designer shows a wholly different representation. He focuses on the high-tech image and the size of its base. Finally, we can find these redesigned sketches of the two novices and the one expert to follow a process of redesigning that are consistent with the four criterions for experiencing pleasures.

Table 12. The redesigned sketches

	Redesigned sketches		
	Senior College Student Designer: A	**Senior College Student Designer: B**	**The Expert Designer: six years of experience**
Concept Generation		 	
Final Concept Selection		 	
Computer Modelling		 	

4. CONCLUSION AND FUTURE WORK

This study has first described how using a design strategy could provide structure to the problem of redesign, allowing a degree of logic to be used in the search for solutions. Second, the holistic pre-design approach could help to make it easier for those involved in the design process to consider the full spectrum of pleasures and functions that a product can offer. Moreover, it is not a theory of the design process, but simply a tool that can aid when taking a structured approach to a newly redesigned product. Third, each of the four distinctions of pleasure can provide further insight into and trace the customer's mode of thinking in the design strategy. In addition, it is hoped that this study will provide effective metrics and better design strategies when redesigning a product which will ultimately find the market. Furthermore, a product designer and/or planner should always keep them in mind.

Future research is that the effective strategy will be required to be refined and validated. This work may be extended to cover conceptual testing and comparison with the existing products in the market. Also, the scope may be broadened to encompass a setting where greater amounts of noise and greater amounts of information are incorporated.

5. REFERENCES

1. K.T. Ulrich and S.D. Eppinger, 2000, Product Design and Development, Boston, Irwin & McGraw-Hill, pp.60-61.
2. S. Pheasant, 1991, Ergonomics Work and Health, Macmillan, New York, pp.13-15.
3. S. Pugh, 1990, Total Design, Addison-Wesley, Reading, MA, pp.38-41.
4. V. Papanek, 1976, Twelve methodologies for design, Proceedings of the ISCIO "Design for need" symposium, Royal College of Art, London, pp. 33-37.
5. M. Cottam and I. Sherrington, 2002, Development of a design strategy for an established semi-technical product. A case study of a safety harness for tree workers. Design Studies, Vol.23, No.1, pp.41-65.
6. N. Cross, 1994, Engineering Design Methods, Wiley, New York, pp.106-119.
7. L. Tiger, 1992, The Pursuit of Pleasures, Boston: Little, Brown & Company, pp.52-60.

COSTS MODELS IN DESIGN AND MANUFACTURING OF SAND CASTING PRODUCTS

Nicolas Perry, Magali Mauchand and Alain Bernard
IRCCyN, 1 rue de la Noé, BP.92101, 44321 Nantes, France

Abstract: In the early phases of the product life cycle, cost controls became a major decision tool in the competitiveness of companies due to world competition. After defining the problems related to these control difficulties, we will present an approach using a concept of cost entity related to the design and realization activities of the product. We will try to apply this approach to the fields of the sand casting foundry. This work will highlight the enterprise modelling difficulties (limits of a global cost modelling approach) and some specifics limitations of the tool used for this development. Finally we will discuss the limits of a generic approach.

Key words: Cost management, Enterprise – Product – Cost modelling, Cost entity

1. INTRODUCTION

In the early Seventies, studies in the United Kingdom and in the United States highlighted the strategic role of design activities. The conclusions led both companies and authorities towards new approaches in order to improve the economic performances of companies. At the end of the Eighties, the paramount role of quality in design was reinforced in the United States by the Made-in-America report from the MIT "Commission on the Productivity" [1]. The Commission on Engineering and Technical Systems study (from the United States' Committee on Engineering Design Theory and Methodology) confirmed these conclusions in the "Designing for Competitive Advantage" report [2] in 1991.

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 69-80.

As described by Perrin [3], the design phase is the key factor of the product development process. The ability to produce new products with a high quality, a low cost and which fit with the customer requests is fundamental to improve national competitiveness. Consequently, costs (and cost management from the early design to the end delivery) become as important as the other technical requests.

Due to the global market and the worldwide competition, reactivity and agility are the only way to maintain company competitiveness. This can be characterized by the ability of a company to change its products and/or processes in a very short time and at minimal cost. Cost control, at the early stages of design, becomes a key factor of success, since this phase fixes an average up to 70 to 80% of the end product costs (depending on the kind of production).

Moreover, the distribution of costs (respectively direct and non direct) is changing: more time and services are dedicated to the studies for smaller product batches and shorter product life. The former cost-sharing methods and the analytic or analogical cost-accounting methods no longer give efficient results. Thanks to studies from CAM-I (Computer Aided Manufacturing-International [4]) and authors like Johnson and Kaplan [5], the increasing gap between "traditional methods" of cost estimation and the new management requirements have been highlighted.

All these works lead to new approaches integrating the complete cost and promote accounting methods based on the enterprise activities (Activity Based Costing for instance). The French economist, Perrin [6], since the Sixties, has also developed a method based on one single cost-inductor through all the steps of the product development process (Added Value Unit method). We implemented such a costing management in a French sand casting foundry in order to allow a multi-level management approach, based on indicators linked to the exact costs of the product to be delivered [7, 8]. In this work we validated the concepts and also the methodology required for a complete numerical traceability.

The work that will be presented in this paper concerns this former study and uses a concept called "cost entity" [9]. It includes several concepts: cost inductors from activity-based accounting methods, features from computer-aided design (CAD) and homogeneity from analytical cost accounting. Consequently, in order to define a cost entity, it is necessary to fill in several attributes linking technical and economical variables. The product model uses the concept of manufacturing feature. Costs are evaluated on the base of specific knowledge and reasoning models with the tool "Cost Advantage", giving information on costs to the CAD model. This is adding a cost semantic level to the CAD model. This model (called costgramme) makes the expertise of the manufacturing cost available to the designer.

Some models, dedicated to the sand casting production of primary parts, were created with the wish to evaluate a meta-model that could be deployed in all sand casting industries. Thus, the goals of this study were, on the one hand, to create the most generic model related to sand casting jobs, and on the other hand, to determine how it is transposable from one company to another (or from a production line to another…). So we will define and discuss which are the limits of the concepts from the triptych product/process/cost, and what level of detail is necessary to implement the methodology in most industrial environments.

This study is realized in collaboration with Cognition Europe, software developer of the tool Cost Advantage, and based on an industrial experimentation in the SMC Colombier Fontaine casting enterprise.

2. VALUE AND COST MANAGEMENT

The concept of cost or value is strongly connected with the context. Today, the selling price depends on the price the customer is ready to pay according to the value that it appreciates. Consequently, the margin is not used to calculate the selling price, it results from it. The selling margin and price must be defined beforehand; they determine an "objective maximum cost", known as "objective cost".

According to Perrin the indicators relating to costs are mainly to be classified among the results' indicators and not driving indicators [3]. The two criteria considered are on the one hand the cost of the product and on the other side the respect of the budget:

- The respect of the product cost is proposed within the framework of the development of the target costing methods (management by cost-target) and Activity Based Costing (ABC). The idea is to fix a target cost as an objective to which one must make the cost of the product correspond. The indicator then used is the relationship between the effective and the objective product cost,
- The respect of the budgets (in particular for the studies' and industrialization services), is the second element to be taken into account at the financial level. The indicator is then naturally the ratio between the actual exceeding and the initially fixed budget.

Evolving industrial practice is changing the nature of the costs in the company. Formerly, the final costs included a large majority of direct costs (often about 70%), i.e. directly assigned to the products (labour or raw materials). The other costs could be the subject of global distributions; the choice of the scale hardly influenced the result. The method of cost accounting with indirect expenses shared through allocation of a fixed percentage overhead could be admitted as long as their proportion was not too big. But the proportions are now reversed, and direct costs often do not constitute more than 30 % of the total cost. Not only do indirect expenses create an arbitrary factor in the calculation of the costs, but they make control difficult. These two concerns gave place to much research that ended up in deeply modifying the systems of cost evaluation.

Let us analyse the product life cycle and the impact of decisions in terms of generated costs. The early phase of design implies between 70 and 85% of the product cost and, at the end of the detailed design, the margin on the final cost is thus very limited (Fig. 1.). Contrary, the evolution of the costs really engaged by the company is very limited in these upstream phases. However, it is there that the control of the costs should be really exerted [10].

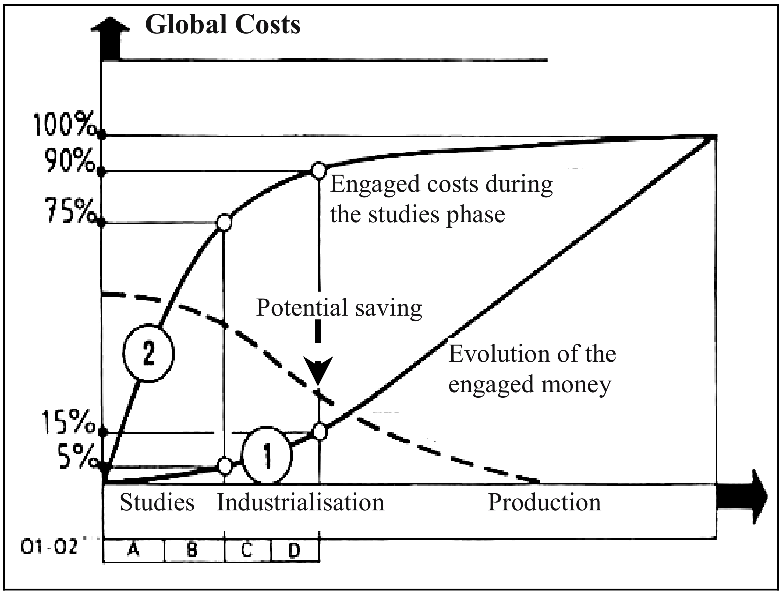


Figure 1. Cost evolution during the PLM

Consequently, the more the product is defined in detail, the less straightforward is cost reduction. Modification becomes increasingly expensive. There is thus an obvious economic interest to have product optimization at the early step of the process. The profit potential is significant and the committed costs are weak. Design choices, inexpensive in terms of resource consumption, can be very expensive in production, when it becomes extremely difficult to carry out modifications [11, 12].

3. DESIGN TO COST

Design to cost can be defined as a principle of action aiming at establishing rigorous objectives. It allows compromises between performance and cost [13, 14]. The "cost" constraint becomes capital and this data should be managed on the same level as technical performances. Then the cost objectives become constraints and the technical performances, variables.

The interest in integrating cost management at the design phase is shared by the customer and the seller since it makes it possible to control development according to exact needs for the future users (it stimulates the search for new ideas required by the economic constraints), in addition, it makes it possible to prepare and organize its production very early and to better control its margin [15].

Once again, this method of design takes the steps of activities or tasks with a focus on the cost limits. This results in an extension of the design time due to the successive iterations needed to agree on the solutions after negotiation (technical or economic). This leads enhance the relation within the company of economical, engineering and production services and can be used jointly with methods of costs analysis like ABC. This also enables a very thorough analysis of the product functions linked to its costs [12].

One can however point out that the design phase takes longer. But the debugging step (pre-production) is no longer disturbed by modifications as it occurs in a traditional step. We then accept, in fact, to ”waste“ time in the design phase, therefore with the quasi-certainty to regain it downstream before the delivery date.

3.1 The Cost Entity concept and the modelling logic

The aim of our study is to carefully manage the costs (direct and indirect) during the production of sand casting parts. As illustrated previously, it is imperative to give a tool to the engineers of the engineering and design department in order to help them to control the costs of part design. In collaboration with the company Cognition Europe, and on the basis of the tool *Cost Advantage*, we worked on cost models to apply in the case of the sand cast steel parts. We based the approach on previous work, proposing an integrated approach for the sand foundry, realized within the framework of a thesis in partnership with SMC Colombier Fontaine (France) of group AFE Metal [7]. This work allowed the formalisation of the knowledge base necessary to control the product life cycle in a foundry company. In addition, we validated an approach, a methodology and a deployment leading to an exact knowledge of the part costs and their impact on the output of the company [8].

Let us start with the concept of cost entity and context, which are our modelling bases.

3.1.1 Costs Entities

A Cost Entity is a grouping of costs associated with the resources consumed by an activity. The general condition is due to the homogeneity of the resources, which makes it possible to associate a single inductor: the entity cost [5]. The model allows expertise formalisation, knowledge capitalisation and to have, at the early design phase, some information about the production steps. Moreover, it helps communication between all collaborators during the product life cycle.

3.1.2 Contexts

The contexts specify defined entities at three levels in our model. The first is defined at a process level, the second at a material level and the last is directly related to the feature. This context is a cross between a process, a material and a feature, connected to an environment (cf. Fig. 2.).

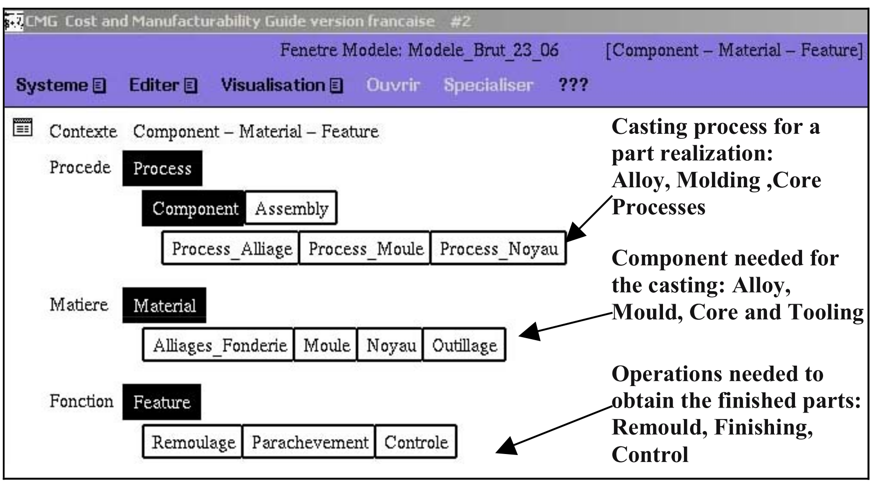


Figure 2. Sand casting modelling

3.2 Sand casting modelling

Based on this analysis, a generic model using Cost Advantage Software was created. The first step closely defines the production process dedicated to this industry. The master parameters acting upon the product cost must be identified in order to enrich the cost semantics of the model.

The generic approach quickly highlights the problems of the context characteristics, which are how to define a significant cost for the part and at which level of detail in order to be generic? We will not answer this question, but present the paths or solutions used.

First, let us describe the sand casting organization through the part life cycle in the enterprise. The cost model will be reduced to the production phase, without taking into account the tooling design and part industrialisation phases.

3.3 Step induced by the use of Cost Advantage

Fig. 3 presents a functional view of the process with the compound (raw material, tooling) and the elements needed to manufacture a part linked with the major indicators dealing with the final cost (loss, scrap ratio, production rate).

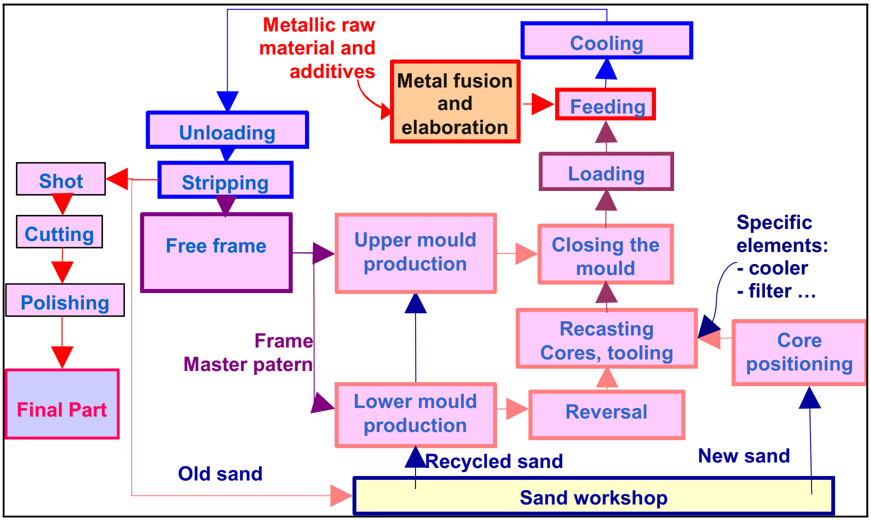


Figure 3. Process view of a sand casting production cycle

Fig. 4 represents a transposition under the concepts of Cost Advantage of this model grouping together the three levels of entities defined in the software. To illustrate this, for example, the mould, the tooling and the cores are components required to carry out the assembly named moulding by the operation (feature) of remould. It is thus necessary to define the final part, to carry out the two assemblies, which are the moulding then the casting.

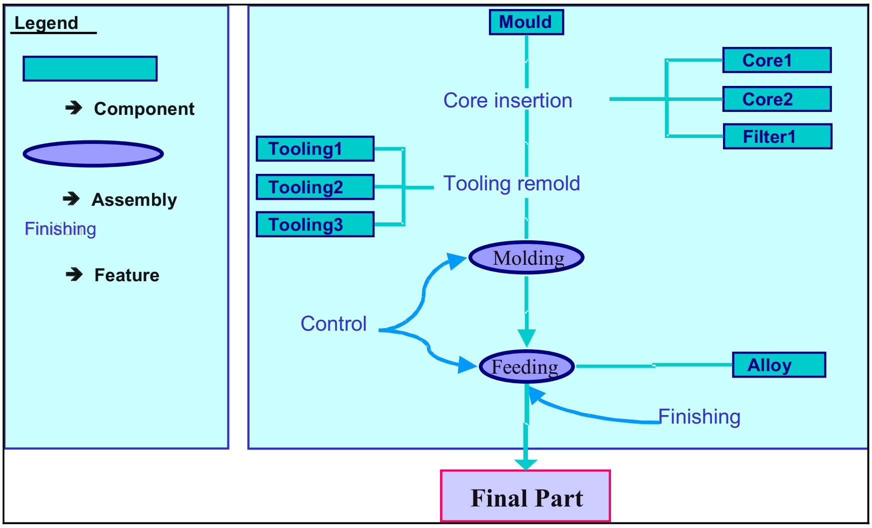


Figure 4. Cost Advantage modelling example, at the assembly level

In terms of model design, the functional view identifies the assemblies needed, it is then necessary to define the components and choose and define the related operations. An ascending step must be practiced, starting with the components up to the definition of the assemblies. The cost calculation is presented in Fig. 5 and the structure of the implemented data in Fig. 6.

Calculations simply take into account volumes of material, rates of production, losses and the machine and labour costs. To say nothing about the difficulty in knowing the exact parameters, the problems we met were model and data organization, more than process modelling.

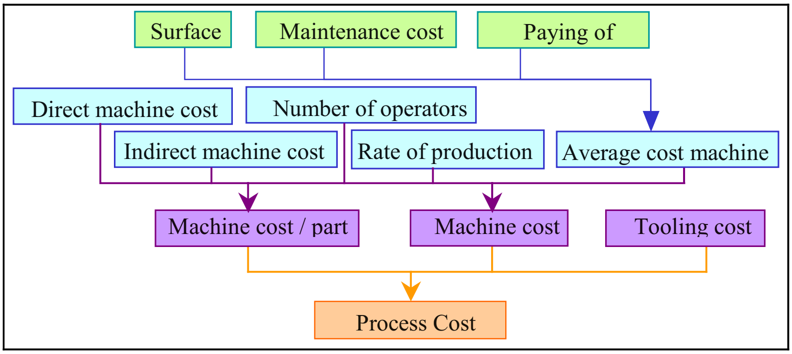


Figure 5. Process cost structure example

The rules of calculation then implemented will make possible for the future user to inform only the relevant data about his or her study. Indeed, only the operational process, rates, dimensions, numbers of cores (etc.) will be required (or deduced directly in a CAD software) to allow an automatic calculation of the cost of the part according to its particular characteristics.

4. DISCUSSION

The principal remark on this model, implemented with the Cost Advantage platform, is that it does not integrate the global aspect of the costs management. For instance, the indirect share due to the development and the design of the tools (master pattern, core boxes…) is not taken in account.

But let us keep in mind the framework of the use of such a tool: to help the designer to achieve a cost objective related to expected technical functions. He can then propose design modifications (joint section, cores, quality…) or process (number of parts in the mould…) to achieve its goal. Moreover, when the whole partners validate this design, the parameter costs are fixed (in agreement or not with the objective laid down initially), the cost is well known and will no longer be a consequence of later decisions.

During this work a principal difficulty has been identified, which for this modelling impacts on the multiplicity of the elements on the availability for their characterization and for their organization. Even if the manufacturing sand casting process seems simple, it uses many components (alloy, cores, mould…), and we limited the definition in terms of model refinement since each one of these components could be the subject of a finer-detail modelling. The difficulty was to choose this limitation in order to represent the general process without going closer into enterprise specificities.

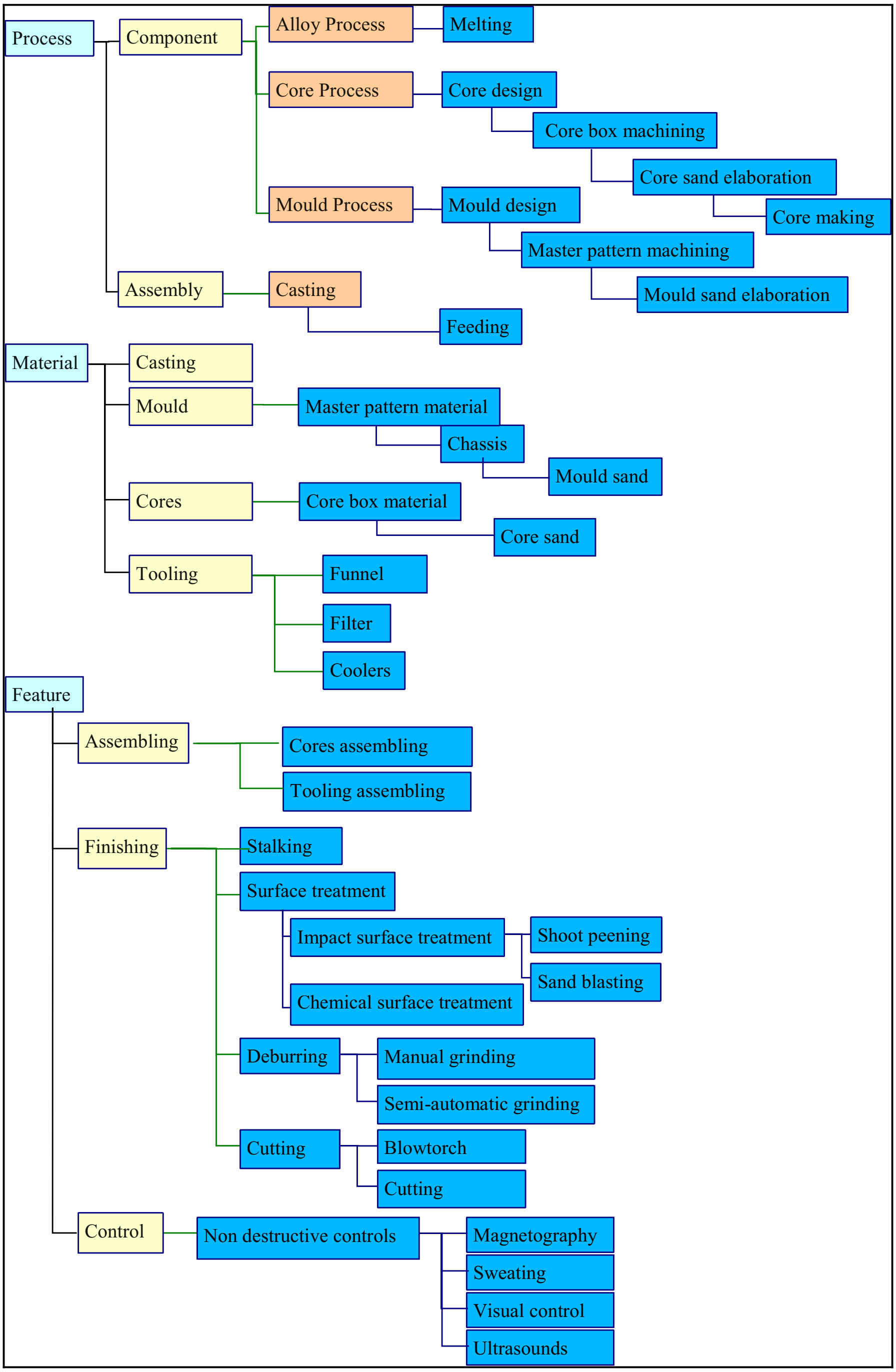


Figure 6. Structure of data with Cost Advantage

For example, the cores or material melting process are not completely defined since they depend on the machines, uses and other specificities of the workshop. In fact this remark is common for all the components entering into the realisation of the finished part. Finally, we think that this model, according to our understanding, define a basic minimal skeleton, transposable from one company to another using the sand casting process.

5. CONCLUSION

To conclude this work, our approach was:

- First, understand the logic of cost oriented modelling through a tool using the concept of cost entity.
- Then, in order to ensure a generic aspect of our work, limit the details of the operations, components and assemblies. Indeed, the development of these elements takes into account many parameters that it seemed to us initially overflowing to define.
- Finally define a structure that we think minimal, as well as the indicators necessary to evaluate the whole costs without the indirect part.

The next two steps of this study are, on the one hand, among other sand casting companies, to apply this modelling, configure the model with the existing processes and inform the exact values of the indicators. But also on the other hand, to calibrate the model and the results on real studies that have already been done. These two last steps enable us to compare the effectiveness of the various companies and could be used as a benchmark. A foreseeable difficulty is the possibility of obtaining this information. Moreover, these factors are often managed in a total cost accounting approach. As a result, the efficiency indicators may be sunk in a not very transparent accounting system or may be aggregated with other ones, that are not relevant to the particular issue of concern.

The other significant continuation to give to this work is the taking into account of the global costs mainly related to the indirect shares (structural). Our introduction highlights the lack of management of these aspects and our first approach did not result in a better control of these factors. However the work is done and the workers must be paid (designer, maintenance, buyers, logistics…) even if their work is not as well managed as through a cost management system. A better specification (by means of indicators, of some metrics) of the tool design phase, tool lifespan, etc. could integrate a real cost of the complete series. The question of the relevance of the tool used for this type of approach then arises. Some solutions come from the use of single or very limited number of cost inductors such as time and they define a minimal global enterprise cost per hour to balance its financial objectives. Such an approach enables a multi-level management of the impact of parts and gives real time information to assess the enterprise objectives and manage strategies and operational decisions.

REFERENCES

1. M.L. Dertouzos, R.K. Lester, R.M. Solow, and The MIT Commission on Industrial Productivity. Made in America: Regaining the Productive Edge. Cambridge, Massachusetts: The MIT Press, 1989.
2. Committee on Engineering Design Theory and Methodology, Improving Engineering Design: Designing for Competitive Advantage, Manufacturing Studies Board - National Academy Press, Washington D.C., ISBN 0-309-04478-2, 1991
3. J. Perrin, Cohérence, pertinence et évaluation économique des activités de conception, in Cohérence, Pertinence et Evaluation, ECOSIP, Economica, 1996.
4. C. Berliner, and J.A. Brimson, Cost Managment for Today's Advanced Manufacturing, The CAM-I Conceptual Design, Harvard Business School Press, 1988
5. H.T. Johnson, R.S. Kaplan. 1987. Relevance Lost: The Rise and Fall of Management Accounting, Boston, Harvard Business School Press. 1987
6. G. Perrin, Prix de revient et contrôle de gestion par la méthode GP, Dunod, 1963.
7. A.Bernard, N. Perry, J.C. Delplace, S. Gabriel, Optimization of complete design process for sand casting foundry, Proceedings of IDMME'2002 conference, Clermont-Ferrand, France, 2002.
8. N. Perry, A. Bernard, Cost objective PLM and CE", CE: the Vision for the future generation in research and applications, ISBN 90 5809 622 X, 2003, p.817-822.
9. F. H'Mida, L'approche entité coût pour l'estimation des coûts en production mécanique, PhD Thesis, LGIPM-ENSAM Metz-Univ. Metz, Mars 2002.
10. C. Alcouffe, M-P. Bes, Standardisation et évaluation des activités de conception dans le secteur aérospatial, Deuxième Congrès International Franco-Québécois de Génie Industriel, ALBI, France, 1997.
11. S. Bellut, Estimer le coût d'un projet, Saint Denis de La Plaine : AFNOR Gestion, Edition 2002.
12. P. Mevellec, Le coût global, nouvelle frontière du calcul de coûts, FINECO (en attente de publication), 2003.
13. G. Delafollie, Analyse de la Valeur, Paris: Hachette Technique, ISBN 2 01017151 9, 1991.
14. C. Petitdemange, La maîtrise de la valeur, La gestion de projet et l'ingénierie simultanée – Paris-La Défense: Afnor Gestion, ISBN 2 12475021 6, 1991.
15. N. Giannopoulos, R. Roy, V. Taratoukhine, T. Griggs, Embedded systems software cost estimating within the concurrent engineering environment, CE : the Vision for the future generation in research and applications, ISBN 90 5809 622 X, 2003, p.353-357.
16. J-P. Kieffer, P. Pujo, Méthodes du pilotage des systèmes de production, Hermès Sciences Lavoisier, ISBN 2 74620514 9, 2002.

UNDERSTANDING MACHINE-MATERIAL INTERACTION FOR IMPROVED DESIGN METHODS

Colin McPherson, Ben Hicks, Chris Berry, Tony Medland and Glen Mullineux
Innovative Manufacturing Research Centre, Department of Mechanical Engineering, University of Bath, BA2 7AY, UK, Phone +44(0)1225 385937, fax +44(0)1225 386928, e-mail c.j.mcpherson@bath.ac.uk

Abstract: Market price demands and legislation are forcing fast-moving consumer goods manufacturers to use thinner and lighter weight packaging materials to convey their products to the consumer. In general these thinner, lighter materials, and materials containing high levels of recycled substrate perform less well on the present designs of packaging machinery. This problem is further complicated by the fact that customers are demanding improved operating efficiencies and expect machines to be able to handle an increasing number of product types and pack sizes. The packaging machinery manufacturers are therefore looking to develop machinery that is more flexible and responsive to changes in material properties as well as better matching new materials to current processes and specific machine designs. In order to achieve this, a fundamental understanding of the interaction between machine systems and the packaging materials used on them is required. This paper presents a design methodology which supports the identification and consideration of the machine-material interactions, and the assessment of the impact of these interactions on the process performance and the machine capabilities. From this assessment, a set of design rules for the particular packaging operation and class of packaging material can be generated. These design rules can be used to support the design of responsive packaging machines, or to reverse engineer package design (shape and size) and material specification for improved processibility.

Key words: Packaging machinery, Material properties, Design rules

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 81-94.

1. INTRODUCTION

In the UK alone the packaging industry represents an annual spend of £11 billion of which £9 billion is spent on materials and £2 billion is spent on machinery [1]. The manufacturers of packaging machinery do not only have to design machinery that is capable of handling these materials but also more intricate and often novel pack designs. Market price demands and legislation through the European Packaging Waste Directive are forcing fast-moving consumer goods manufacturers to use thinner and lighter weight packaging materials to convey their products to the consumer. In general these thinner, lighter materials, and materials containing high levels of recycled substrate, perform less well on the present designs of packaging machinery. Furthermore, these rapidly evolving requirements need to be handled by machinery with improved performance capabilities, reliability and reduced waste [2]. The packaging machinery manufacturers are therefore looking to develop machinery that is more flexible and responsive to changes in material properties as well as better matching new materials to current processes and specific machine designs.

The ability of many packaging machinery manufacturers to deliver these requirements is frustrated by the fact that the fundamental design principles of many packaging machines are the result of incremental improvements made over the last few decades. In many sectors of the industry, organisations often lack the fundamental understanding of the process and of increasing importance is an understanding of the machine-material interface, which more often than not dictates the performance capabilities of the machine and a given packaging process [3]. As a consequence of these factors, machine development in the packaging industry is typically undertaken with a trial-and-error approach. This time-consuming activity is necessary to attempt to match existing machines and tooling to new materials, although this process can often be unsuccessful, requiring a number of iterations, or leads to a very sensitive process setup. In addition to this, rather than supporting the design of the next generation of machinery capable of handling new material variants and pack designs, the process can add considerable complexity. All of this seriously frustrates the manufacturers' ability to be responsive and agile in terms of their ability to deliver in the required lead-time a packaging solution that meets the desired performance requirements and materials handling capabilities.

In order to address these issues a design methodology has been created and validated using industrial studies. The methodology supports the identification and consideration of the machine-material interactions, and the assessment of the impact of these interactions on the process performance and the machine capabilities. This assessment involves a detailed experimental program and where appropriate theoretical modelling. From this assessment, a set of design rules for the particular packaging operation and class of packaging material can be generated. These design rules can be used to support the design of responsive packaging machines, or reverse engineer (match) pack design and material for improved processibility.

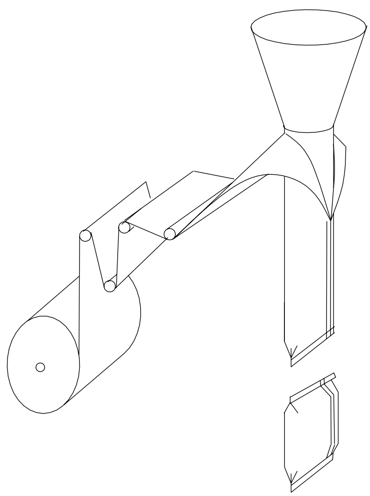


Figure 1. Vertical form, fill and seal process

This paper presents an overview of the methodology and its individual activities. A detailed discussion of the application of the design method to an industrial problem is provided. The high-speed packaging process investigated in this paper involves the production of a bag or pouch from a reel of packaging material, which is then filled with product and sealed. This process is shown in Fig 1 and is more widely known as a form, fill and seal operation. Such a process is commonly used to package crisps, snacks and pasta. The method is applied to this problem in order to understand the interaction between the materials and tooling and the influence of key machine components during setup and run-up.

2. UNDERSTANDING MACHINE-MATERIAL INTERACTION

The design methodology presented in this paper has been created to improve the design of packaging machinery, and in particular, provide for the development of an understanding of, and the evaluation of, the interactions between packaging material and packaging machines. The specific aim of the methodology is to expedite the characterisation of the machine-material interactions and the identification of critical material and machine parameters. The relationships between these parameters are explored through theoretical modelling and experimental investigation. This process assists in the identification of limiting factors and the design bounds for critical parameters in order that the process efficiency is acceptable. These factors can be used to generate design rules and provide the basis for creating and evaluating possible design solutions for the next generation of more flexible and responsive machinery. For example, solutions may be sought from various combinations of materials and mechanisms that maintain the bounded conditions over the full performance range of the machine. An overview of the full methodology is shown in Fig. 2, and the key phases are discussed in the following sections. The application of the methodology to a vertical form, fill and seal machine is described in detail in section 3.

2.1 Evaluation of machine system

This first phase of the methodology (Fig. 2) involves the evaluation of the machine system, the product and the process. In particular, the aim of this phase is to identify the key interactions that occur between the machine, materials and product during the process. An interaction occurs where all or only part of the product/material is in contact with one or more machine elements in order to achieve, or support the undertaking of a particular function. For example, a number of interactions may occur during material or product transport, manipulation, filling, closing and sealing. Where an interaction occurs it is necessary to identify the properties or attributes that impact or influence the success of the operation. These properties or attributes may be related to the pack, the material, environmental conditions, machine operation or individual machine elements. This process can be undertaken for the complete system, part of a process or a particular operation. This phase culminates with the identification of relationships between the properties and attributes of two or more interacting elements.

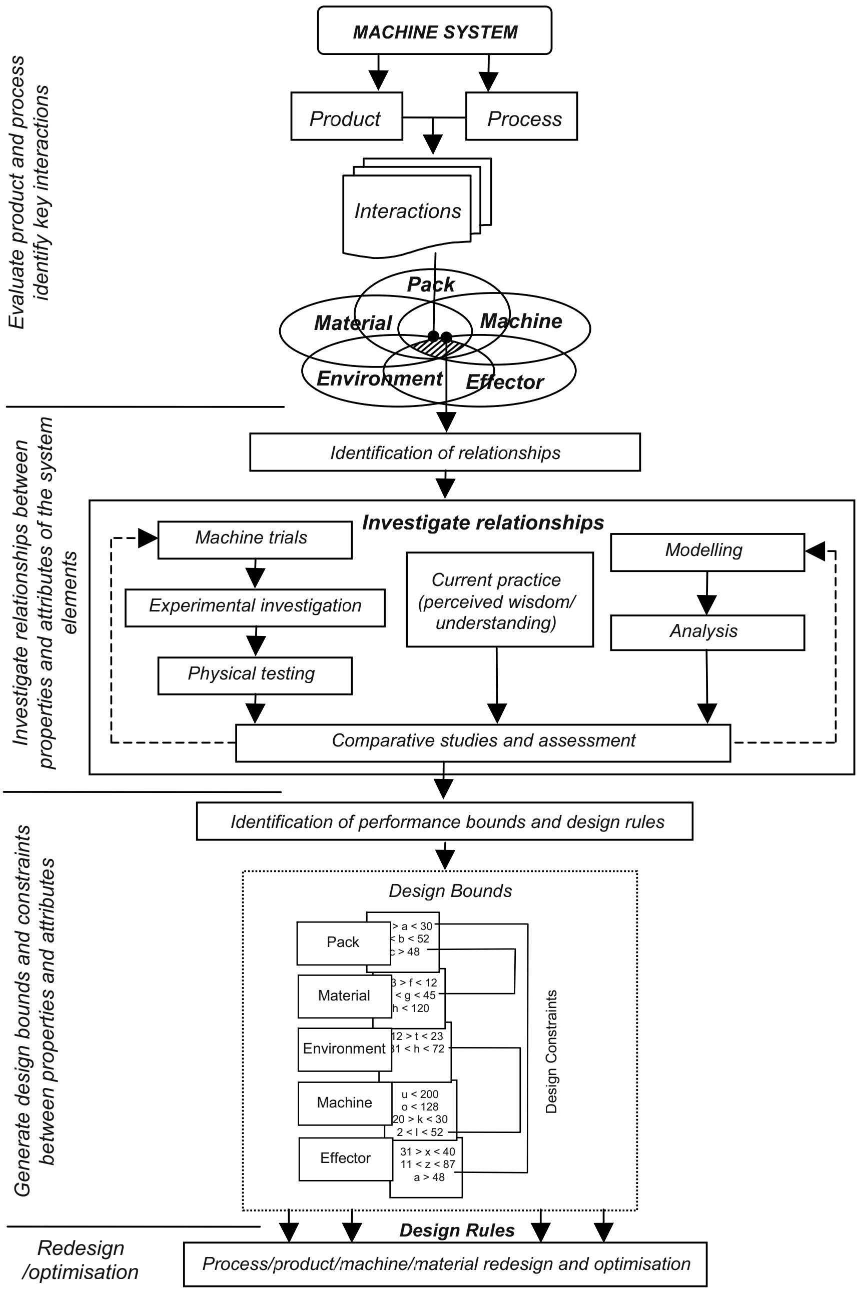


Figure 2. Design methodology for understanding machine-material interaction

2.2 Investigation of relationships

During the second phase of the methodology the relationships between the various properties of the pack, material, environmental conditions, machine system and effector during interaction are investigated. In this work, the effectors are those elements which physically interact with the pack or material. This phase of the methodology may involve machine trails, experimental assessment (such as Taguchi) and physical testing of machine elements and materials. The testing may involve known parameters and use accepted metrics or may demand the development of new more representative test procedures and metrics [3]. In addition to the practical studies, theoretical modeling and analysis may be undertaken to understand the nature of the interaction [4]. The results of the practical studies and theoretical modeling are compared and contrasted with current practice and where available, perceived wisdom. In this paper perceived wisdom encompasses the current understanding held by the manufacturer or industry. The results of comparative assessment may either substantiate the perceived wisdom, and assist in the expression and quantification of the relationships, or highlight shortcomings or even anomalies. For the latter case, further modeling or practical studies might be necessary and are denoted by the iterative paths in the methodology.

The objective of this phase is to be able to quantify the permissible range of values within which it is possible to achieve successful undertaking of the process or of a particular operation. This is obtained when an acceptable performance is achieved; that is to say that products fall within acceptable quality limits and there are no adverse effects for other operations. It is also important to understand the reasons why it is successful, and it is often this fundamental understanding that is lacking. This deficiency or incompleteness complicates the transfer of knowledge to different machines, processes, materials and products. It also frustrates an organization or even an industry's ability to be flexible and responsive. In Fig. 2, this region of successful operation is the intersection of the five sets shown, denoted by the hatched area. Such relationships may be dependent upon factors such as speed of operation, load distribution, point of application, material compliance or surface friction. During this phase it is also important that these dependencies between properties/attributes are identified across the five elements of the pack, material, environment, machine and end-effector. The outcome of this phase of the methodology is the implementation and quantification of relationships that govern the machine-material interaction.

2.3 Generation of design rules

The penultimate phase of the methodology supports the generation of what can be thought of as design rules. These design rules are a collection of bounded parameters and constraints that describe the relationship necessary for successful processing. The design rules define the influence of the pack attributes, the material, the environment, the machine system and the end-effector upon the success of a particular operation. These rules are necessary to provide an understanding of the interaction and are critical for the development of responsive and agile packaging machinery.

2.4 Redesign/optimization

The design rules, generated through the application of the methodology, define bounds and relations between product, materials, machine elements and tooling that must be satisfied to achieve acceptable process efficiency. These rules can be used to support the redesign of machinery, the refinement of the process, the matching of materials to tooling or even govern the alteration of the pack design. The critical parameters of all of these elements may be altered to improve processiblity and the impact of their alterations considered. For example, during the task of redesign these rules can be used to test a redesign concept. Or in the case of optimizing an existing machine system, the design rules can be used to formulate the objective function(s) to be satisfied by the optimiser.

3. INDUSTRIAL PROBLEM (VERTICAL FORM-FILL-SEAL PROCESS)

The industrial problem investigated in this paper is the class of packaging operation known as form-fill and seal. The process involves the formation of a bag or pouch from a flat web of packaging material, which is then filled in with product and sealed. This process is shown in Fig. 1 and is commonly used to package crisps, snacks and pasta. In this process the critical material handling elements are the material transport mechanism and the forming shoulder. A schematic overview of the machine system is shown in Fig. 3. The primary function of the forming shoulder is to guide the material from the flat roll to a tubular shape around the product feed tube. The edges of the film are overlapped and are either heat sealed or glued to form a tube into which the product to be packaged can be inserted in measured quantities. The tube of packaging material is then cross-sealed to form a closed bag.

The final seal also forms the base seal of the next bag. This approach has been adapted to produce a variety of different packages such as packs designed with bottom gussets or flat bottoms [2].

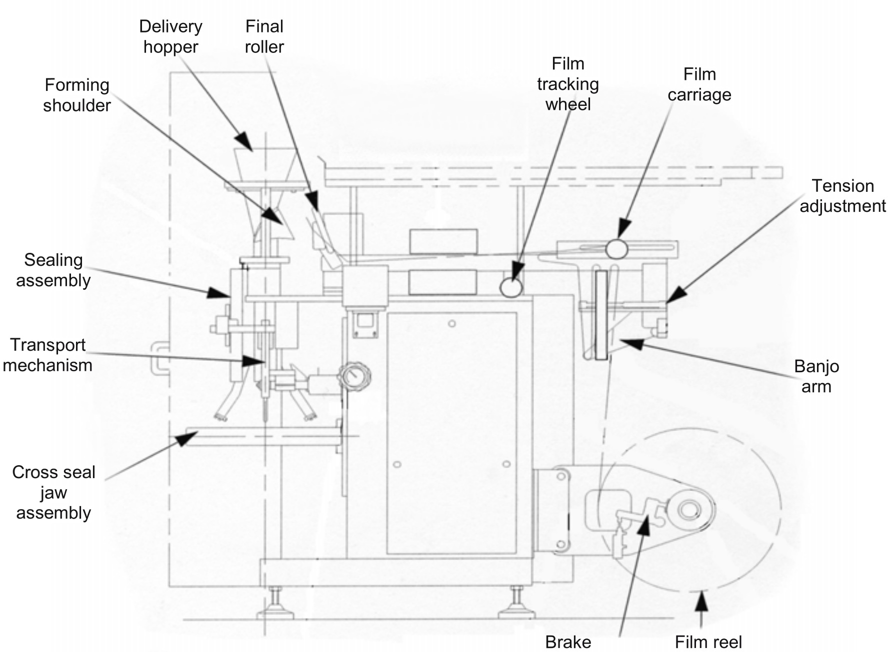


Figure 3. A schematic overview of a vertical form, fill and seal machine

The successful creation of the pack is largely dependent upon the performance of the forming shoulder, which in turn is highly dependent upon the surface geometry of the shoulder [5]. A different forming shoulder is required for not only different bag widths but also different materials and sometimes print finishes. The size of the former is directly related to the bag width whilst the surface geometry depends upon the material properties and the properties that govern the interaction between the former and the material. For example, the upper portion of Fig. 4 shows various forming shoulder geometries for the same size of bag. Here the key parameter is what can be thought of as the height to radius ratio. Effectively, altering the height of the former has the effect of reducing the included angle between the vertical portion of the former and the fin. This is necessary to satisfy the conditions for a continuous developable surface [6]. The primary reason for altering the height to radius ratio is to reduce the forces necessary to draw the material over the shoulder and hence the magnitude of the tension field in the material. Such an approach is common practice within the industry, however, there is little or no quantitative data. Rather manufacturers will produce a number of standardised shoulders that induce various levels of stress the material. These are necessary to accommodate materials with different levels of compliance. This matching of former geometry to material is very important. However, it is widely accepted that the greater the height:radius ratio the worse the tracking. As a consequence, a trade-off has to be made. Achieving the optimum trade-off is further frustrated by current production techniques.

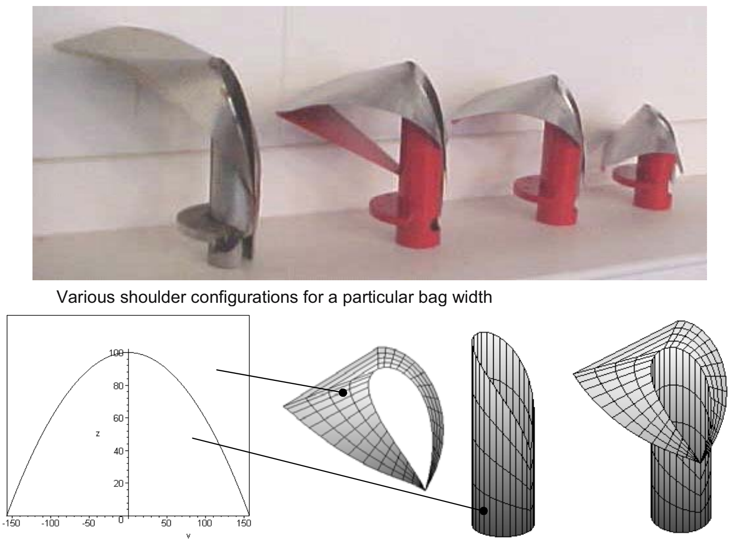


Figure 4. The geometry of the forming shoulder

3.1 Current design practice

Although some attempts have been made at casting or moulding forming shoulders the majority of formers are manufactured from a flat metal sheet cut into in two parts. For the purpose of this paper, these sections are referred to as the tube and the collar, shown in the lower portion of Fig. 4. The tube portion of the forming shoulder has a circumference equal to twice the full bag width. The process for constructing these parts is best described in terms of a single curve, referred to as the bending curve. This bending curve is inscribed on a flat sheet. The portion below the curve is rolled into a cylinder to form the tube, whilst the section above the bending curve is formed around the top of the tube. This is shown in the lower portion of Fig. 4. The shape of the resulting forming shoulder depends on the bending curve and the tube radius. The current production method results in a degree of imprecision in the final geometry of the forming shoulder. Such errors or deviations are corrected during trial-and-error testing by the engineers. This typically involves slight alterations to the surface of the former.

As previously mentioned in addition to ensuring that the material passes over the shoulder without excessive shearing or distortion, the engineer must also ensure that the material tracks uniformly over the former. In order to fulfil these requirements and produce a suitable former, a number of shoulders may be made from different bending curves and different height to radius ratios. Each of these is fine tuned through trial-and-error testing to obtain acceptable tracking without distorting the material. This can be a costly and time-consuming process. But, if the material and shoulder geometry are not suitably matched, process efficiency can be seriously affected. In addition to this, machine setup can be altered to attenuate errors and improve tracking. In particular, the 'banjo arm' and film carriage assemblies, shown in Fig. 3, can be used to increase web tension and alter the inclined angle and orientation of the web respectively. However, there is little understanding of the complex relationship between materials, the web tension, orientation, the forming shoulder and tracking.

4. INVESTIGATION OF RELATIONSHIPS

In order to develop a more in-depth and quantitative understanding of the relationship between, materials, web tension and the forming shoulder the current understanding or perceived wisdom held by the manufacturer was assimilated. Two practical studies were devised: the physical testing of the former and the evaluation of the sensitivity of tracking to the setup of the 'floating arm' assembly.

4.1 Current understanding

The current understanding is discussed in the earlier sections and is summarised in Fig. 5. Here, as the material stiffness reduces the height to radius ratio is increased so as to decrease the stresses in the material and hence minimise any stretching/distortion. However, with large height to radius ratios the tracking of the material is less reliable and tends to be more sensitive to changes in setup. In order to overcome this, the web tension is increased and this appears to reduce the sensitivity.

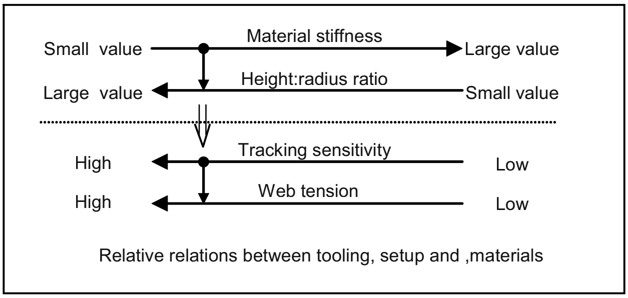


Figure 5. A qualitative understanding of the relationship between machine and materials

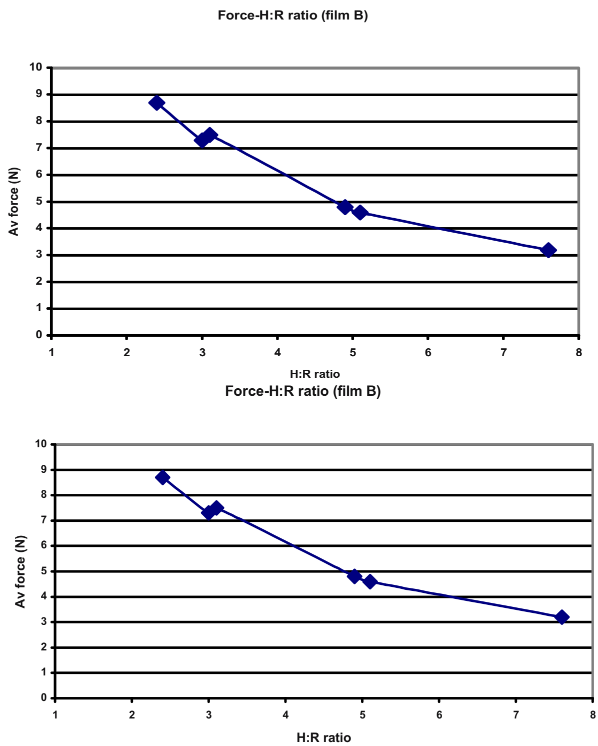


Figure 6. Force against height to radius ratio

4.2 Material testing

This study investigates the force required to draw (pull) material over a range of forming shoulders. Two different films (foil and polyethylene laminates) were considered and in total six forming shoulders. The forming shoulders were selected (three shoulders were manufactured specifically for these trials) to represent the continuum of feasible height/radius ratios for a particular bag width. The study was undertaken to provide a practical understanding of the relationship between the force necessary to draw material over a forming shoulder and the geometry of the former. In particular, the relationship between the height/radius ratio and the force is considered. The trials were repeated a large number of times. The characteristics of the height/radius ratios against average force for each film are shown in Fig. 6. It is seen that the force decreases as the ratio increases.

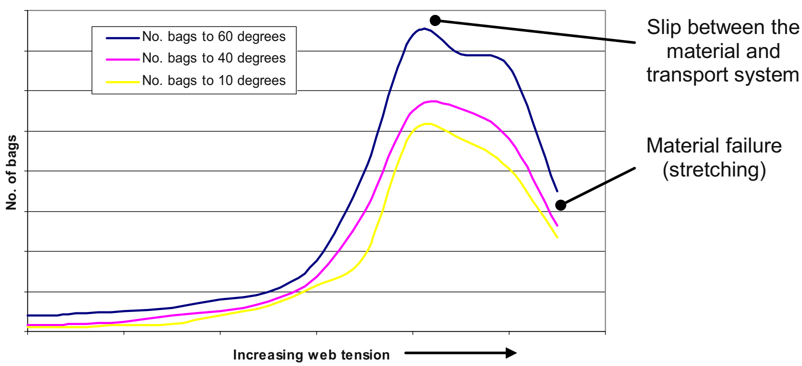


Figure 7. Web tension against time taken to unacceptable tracking offset

4.3 Machine investigation

The machine investigation involves assessing the sensitivity of the tracking to changes in web tension, the inclined angle of the web and the orientation of the web. The approach adopted is to set the machine up for a given material and obtain acceptable process performance/efficiency. These machine settings are then used as the datum points and the sensitivity of changes to the settings are evaluated by introducing small changes to each. The results of the investigation into web tension are summarized in Fig. 7. In particular, it was found that web tension was a key factor in influencing tracking and is discussed in more detail in section 4.4. The machine investigation also showed that altering the orientation of the web influences tracking and could be used to correct small deviations.

4.4 Comparative study and assessment

The practical studies undertaken agree with the perceived wisdom and provide a more detailed quantitative understanding of the relationships that govern the interaction. In particular, the materials testing program revealed that between the extremes of feasible height/radius ratios almost a two hundred percent increase in applied force is necessary. The machine investigation revealed that web tension can be used to influence tracking. Increasing web tension will reduce the rate at which tracking offset occurs, which in turn increases the meantime between machine reset, and improves process efficiency. At very high tensions the tracking problems appear to be eliminated, however, slip between drive belts and rollers can cause other problems. Furthermore, at these high tensions the material can exhibit stretching and distortion. As a consequence of these findings, further work is necessary to evaluate quantitatively the impact of increased tension on the forces necessary to draw material over the former. This can be used with the understanding already developed to determine quantitatively the optimum trade-off. Arguably with the current level of understanding it would be possible to cancel out the benefits of using a large height/radius ratio by excessively increasing web tension.

5. CONCLUSIONS

In order for packaging machinery manufacturers to be agile and responsive, in terms of their ability to deliver a solution in the required lead time that meets the desired performance requirements and materials handling capabilities, a fundamental understanding of machine-material interaction is required. This fundamental understanding must involve a qualitative and, where possible, quantitative relationships. This in-depth understanding provides the basis for knowledge to be transferred across different machines and processes as well as the capability to evaluate the ability of these machines or processes to handle different packs and materials. To address these important research issues a design methodology is presented. The methodology supports the characterisation of the machine-material interactions and the identification of critical material and machine parameters. The relationships between these parameters are explored through theoretical modelling and experimental investigation. This process assists in the identification of limiting factors and the design bounds for critical parameters in order that the process efficiency is acceptable. These factors can be used to generate design rules and provide the basis for creating and evaluating possible design solutions for the next generation of more flexible and responsive machinery.

The overall methodology is applied to an industrial problem to demonstrate the key activities. For the case considered, it is shown how the methodology can be used to assess current design practice (perceived wisdom) and provide a more quantitative understanding of the relationships that govern machine-material interaction. This understanding enables the manufacturer to better match tooling to particular materials and understand the impact of what is a complex setup procedure on tracking and ultimately process efficiency.

ACKNOWLEDGEMENTS

This work was undertaken as part of the EPSRC Innovative Manufacturing Research Centre at the University of Bath (GR/R67507/01). Support was also provided by PIRA, PPMA and industrial collaborators.

REFERENCES

1. J.F. Hanlon, R.J. Kelsey, H.E. Forcinio, 1998, Handbook of package engineering, Chapter 1, Technomic Publishing Company, Inc, Pennsylvania, U.S.A
2. D. Twede, R. Goddard, 1998, Packaging Materials, Pira International, Leatherhead, UK, 1998.
3. B.J. Hicks, C. Berry, G. Mullineux, C.J. McPherson, A.J. Medland, 2004, An energy based approach for modelling the behaviour of packaging material during processing, IMechE Journal of Mechanical Engineering Science, Part C, ISSN 0954-4062, Vol. 218, pp. 1-14, ISSN 0954-4062.
4. B.J. Hicks, C. Berry, G. Mullineux, C.J. McPherson, A.J. Medland, 2003, Energy method for modelling delamination buckling in geometrically constrained systems, IMechE Journal of Mechanical Engineering Science, Part C, Vol. 217, pp. 1015-1026, ISSN 0954-4062.
5. C.J. McPherson, G. Mullineux, C. Berry, B.J. Hicks, A.J. Medland, 2004, The performance envelope of forming shoulders and its implications for design and manufacture", IMechE Journal of Engineering Manufacture, Part B, to appear.
6. M. Lipschultz, 1969, Theory and Problems of Differential Geometry, Schaum's Outline Series, McGraw-Hill, New York.

AN INTEGRATED METHODOLOGY FOR PROPAGATING THE VOICE OF THE CUSTOMER INTO THE CONCEPTUAL DESIGN PROCESS

Bernard Yannou[1] and Jean-François Petiot[2]

[1]Ecole Centrale Paris, Laboratoire de Génie Industriel - Grande Voie des Vignes 92295, Châtenay-Malabry, France, yannou@lgi.ecp.fr; [2]Ecole Centrale Nantes, IRCCyN, 1 rue de la Noë, BP 92101, 44321 Nantes Cédex 3, France

Abstract: The elementary conceptual design loop is made of understanding and measuring the customer need, proposing product concepts and propagating the customer need along the product deployment, and finally assessing concepts in a multi-criteria way in regards to the need. Many methods from various disciplines (industrial/design engineering, psychophysics, multi-criteria decision analysis, marketing, artificial intelligence, statistical analysis) presently contribute partially to support this design loop. But few methods propose an integrated and coherent framework to deal with this elementary conceptual design loop. In this article, we propose such a methodology which is a coherent combination of classical methods in psychophysics, marketing and decision theory, namely multidimensional scaling, semantic differential method, factor analysis, pairwise comparison and Analytical Hierarchy Process. Our approach provides designers with a tool which helps the definition of the semantic part of the need, it rates and ranks the new product prototypes according to their proximity to the "ideal product", and it underlines the particular semantic dimensions that could be improved. To illustrate our approach, we have performed users' tests and applied our methodology to the design of table glasses. For clarity, each stage of the methodology is presented in detail on this particular example.

Key words: Product semantics; Conceptual design; Multidimensional scaling; Pairwise comparison; AHP.

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 95-108.

1. INTRODUCTION TO THE PRODUCT PARAMETER TRAIL ISSUE

That is an evidence to say that capturing the customer needs, propagating these needs all along the product parameter cascading[1] and comparing or assessing product alternatives in a multi-criteria way in regards to the need, represent a major issue in design engineering. Indeed, the challenge is to better decode the intricate customer needs and to link them to the product engineering specifications so as to monitor that the resulting product emerges from the voice of the customer. Many methods are assumed to fulfil a part or even the totality of these requirements. One can regroup them according to the disciplines they come from:

- **Methods from industrial/design engineering**. They are used to perform a *target cascading* (or a *parameter trail*); the two most famous are the *Quality Function Deployment* (QFD) [1] and *Axiomatic Design* (AD) [2]. QFD assists the designer to explicitly list the *customer requirements*, to establish qualitative mappings between successive levels of parameters, i.e.: the *customer attributes*, the *functional requirements*, the *design parameters* and finally the *process variables*, to identify tradeoffs and to compare product alternatives. Nevertheless, after Aungst et al [3], QFD suffers from a number of severe limitations. First, there is no reason to a priori guess the customer requirements since it requires to be discovered with a non trivial measurement procedure. Second, QFD uses at any level raw subjective rankings. Third, the 3 mappings between the 4 levels of parameters could be known quantitatively as precise engineering relations instead of being expressed qualitatively. In a similar manner, Aungst et al [3] criticize the *Axiomatic Design* approach because it only considers linear mappings between levels of parameters, allowing only sensitivity analyses around narrow ranges.
- **Methods from psychophysics** that are based on user's tests. *MultiDimensional Scaling* (MDS) [4] aims at identify the perceived dissimilarity between products to a distance in a metric space. *Pairwise Comparison* (PC) methods [5] are based on the subjective comparison of the ratio between two levels of a given property for a pair of products. These methods may be used to build an understanding of the product perception with no or few pre-existing assumptions on the perceptual dimensions. But they do not cover the entire spectrum of the problem like the previous industrial engineering methods could pretend to.
- **Methods from Multi-Criteria Decision Analysis (MCDA).** The *Expected Utility-Value Theory* (EUVT) has been developed in the

[1] One also finds the *parameter deployment* or *trail*.

beginnings of the fifties by Von Neumann and Morgenstern [6] to express in an aggregated formula the satisfaction of a customer regarding a good. It is a mathematically grounded theory for quantifying the contributions of the utility drivers. It is a compensatory theory, in the sense that any non-satisfaction may be compensated by more satisfaction elsewhere. Despite this compensation cannot be generalized between performances in engineering, the EUVT is become the basic theory to express an objective function in engineering and the basis of *Decision-Based Design* (DBD). DBD (see the DETC/DTM/Open Workshop on Decision-Based Design http://dbd.eng.buffalo.edu/) is a general topic of engineering design that puts the emphasis on evaluation and decision in the design process. In this context, more general aggregation formulas, possibly non-compensatory ones, are considered under the name of *preference aggregation models* [7] for the theory and the SPEC approach [8] for a performance and risk monitoring system in design. Also, the traditional multi-criteria rating and ranking methods can be used in design. Let us mention the most famous: the *Analytic Hierarchy Process* (AHP) [9] and the outranking approach developed in the *ELECTRE* methods [10]. A sub-topic linked to DBD is *design selection* [11].

– **Methods from marketing.** The *Semantic Differential Method* (SDM)c[12] and *Conjoint Analysis* (CA) [13]. The *Semantic Differential Method* (SDM) consists of listing the semantic attributes of the product to analyse, and carrying out user-tests in which the user must assess the product according to these attributes. *Conjoint Analysis* [13] is a more technical and structured marketing approach also based on consumer tests. Recently, researchers have begun to merge market analysis and the design stages of requirement definition and concept generation [14]. Complete marketing and conceptual design platforms have been proposed in [15, 16]. Despite the authors claim they provide platforms to *design for profit*, it turns out that such approaches appear to us to be paradoxical and not adapted to a conceptual stage. Indeed, a set of product configurations with their corresponding production costs should be enumerated in advance before really starting the design stage with the best list of functional requirements with options and variants.
– **Methods from artificial intelligence.** The *subjective evaluation* of a product utility by fuzzy logic is a well developed discipline [17]. For instance, the *Method of Imprecision* (MoI) has been developed to fuzzily capture requirements on the product performances and to perform tradeoffs between design parameters and performance variables [6, 18]. The MoI provides primitives to express non-compensatory preference aggregation models, well adapted to engineering assessments. This

fuzzy-based method is also well adapted to the *management of uncertainty in design* research topic. Japanese researchers investigate the customer's ergonomic feeling under the name *Kansei Engineering Methods* [19]. These methods aim at performing a forward mapping of the customer's feeling for the product to the design elements and a backward mapping of design elements to a feeling assessment. It is made by the way of databases and artificial intelligence tools.

– **Methods from statistical analysis**. *Factor Analysis* (FA), *Principal Component Analysis* (PCA), *Design of Experiments* (DOE) and *Metamodels* (MM) [20] are statistical analyses methods for identifying approximated mathematical models between for instance a set of *design parameters* and a set of *functional requirements*. After Aungst et al [3], these methods may be advantageously used to establish non-trivial (i.e. quantitative and non-linear) mappings between parameters of different levels in the target cascading instead of qualitative mappings of QFD and linear mappings in AD.

2. A NEED FOR AN INTEGRATED METHOD IN CONCEPTUAL DESIGN

The previous review of the literature reveals a multitude of available methods for dealing with the representation of the customer need, the successive mappings with the product and process parameters and the elementary and comparative solution assessment issues. These methods are efficient in a certain context to tackle a part of the problem. But, there is no or few methods that conveniently integrate all three aspects in a conceptual design stage.

Aungst et al [3] have proposed to solve this problem in extending QFD with a fourth mapping, named *House of Quality #0*, between a perceptual space and the customer attributes space. Indeed, the user's or consumer's perception of a product is by essence subjective and hazardous to a priori express whereas other mappings may be more easily captured as technical relations. Product semantics, related to how a person perceives the appearance, the use and the context of a product, is an important problem in industrial design [21]. This is particularly the case when people is more sensible to subjective/esteem functions of a product (including aesthetics or styling attributes) than to objective/usage functions, which is the case of luxury goods and even apparently technical products such as cars, house equipments, clothes, etc. Some studies surprisingly showed that a number of trendy goods were much more designed to satisfy designers' trends than the customers' ones [22]. This is due to the fact that subjective functions and criteria are often neither named nor objectively assessed. Aungst et al [3] use a *semantic analysis* of the product and a *Principal Component Analysis* (PCA) to build this first perceptual model of the consumer.

In comparison to Aungst et al's [3], we present in this article a complementary integrated approach more dedicated to the conceptual design stage. We adopt a bit more sophisticated combination of methods to build this consumer perceptual model and we provide multi-criteria elementary and comparative solution assessment facilities that are more theoretically grounded than the ones imbedded in QFD. Our method is decomposed into 8 stages, each of them including users' tests performed by a panel of subjects.

3. A CASE STUDY OF TABLE GLASSES

The starting point is a set of representative existing products which all answer the same usage functions, but differ from a perception/esteem point of view. A study on wine glasses has already been proposed [23], where the authors presented a method for form generation. For our study, we imagine a company, which build a range of glasses (shapes given Fig. 1), and which wants to design a new glass in order to diversify its products portfolio. We propose to show in this paragraph how our method can be used to assess in a solid way product semantics of several design solutions.

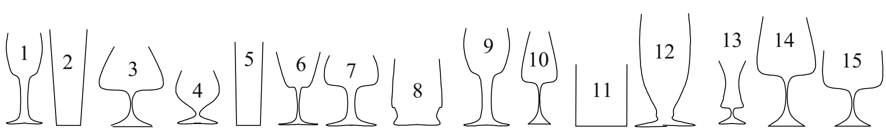


Figure 1. *Shapes of the 15 glasses proposed for the study.*

3.1 Stage 1: definition of the semantic attributes

The 15 glasses have been physically proposed to 11 subjects (10 males, 1 female) for a detailed evaluation. Subjects are asked to verbally describe their perceptions about the glasses freely. A list of relevant semantic criteria is extracted from this description. An analysis of their descriptions has led to the setting up of 17 antonymous adjective pairs (**v1** to **v17**) (Table 1).

3.2 Stage 2: determination of the perceptual space

Multidimensional Scaling method uses dissimilarity assessments to create a geometrical representation of the perceptual space related to a family of objects [24]. This method, developed initially for psychometric analysis [4], is a process whereby a distance matrix among a set of stimuli is translated into a representation of these stimuli inside of a perceptual space. The lowest dimension of this space is computed so as to interpret at best the dissimilarities as Euclidean distances [25]. For each pair of glasses, subjects were asked to sort the products into mutually exclusive groups based on their similarities. No constraint is given on the number of classes to make. The assumption underlying this method is that products occurring in the same group are more similar than products occurring in different groups. The sorting data for any subject consists of a matrix of 0 and 1, indicating whether the subject grouped two glasses together or not. Individual *dissimilarity matrices* are then summed for all subjects, leading to the group's dissimilarity matrix (Fig. 2). Here, one assumes, for the moment, that the group members behave in a somewhat similar manner, i.e. we do not deal with clustering considerations of the group. Given this matrix as input, non metric MDS has been used to calculate the perceptual coordinates of the glasses. A 2-dimensional space turned out to be sufficient (Fig. 3).

Table 1. The 17 adjective pairs of the SDM test and the corresponding semantic attributes

Adjective pair	Semantic attribute	Adjective pair	Semantic attribute
v1: Traditional-modern →	Modernity	v12: Classy-vulgar →	Smartness
v2: Easy for drinking/not… →	Ease of drinking with	v7: Common-particular ↘	
v3: Decorative-practical →	Decorativeness	v13: Unoriginal-creative →	Originality
v4: Unstable/stable →	Stability	v14: Existing-new ↗	
v6: Complicated-simple →	Simplicity	v15: Good perceived quality-bad… →	Quality
v10: Multiusage-occasional →	Ordinariness	v16: Strong-fragile ↘	
v8: Easy to fill-not… →	Ease of filling	v5: Masculine-feminine →	Fragility
v9: Flashy-discreet →	Showiness	v17: Coarse-delicate ↗	
v11: Easy to handle-not… →	Ease of handle		

$$\Delta_{ij} = \begin{bmatrix} 0 & 10 & 6 & 6 & 9 & 6 & 5 & 8 & 1 & 3 & 10 & 7 & 9 & 4 & 5 \\ & 0 & 9 & 9 & 1 & 10 & 9 & 6 & 7 & 7 & 4 & 4 & 10 & 6 & 9 \\ & & 0 & 3 & 10 & 5 & 6 & 8 & 5 & 7 & 9 & 8 & 10 & 6 & 4 \\ & & & 0 & 9 & 7 & 7 & 8 & 6 & 7 & 9 & 7 & 8 & 7 & 5 \\ & & & & 0 & 10 & 10 & 5 & 7 & 8 & 5 & 4 & 10 & 5 & 7 \\ & & & & & 0 & 3 & 9 & 7 & 5 & 9 & 9 & 6 & 8 & 4 \\ & & & & & & 0 & 9 & 6 & 6 & 9 & 8 & 10 & 8 & 3 \\ & & & & & & & 0 & 5 & 8 & 3 & 9 & 9 & 7 & 7 \\ & & & & & & & & 0 & 4 & 8 & 5 & 9 & 1 & 5 \\ & & & & & & & & & 0 & 9 & 6 & 5 & 6 & 6 \\ & & & & & & & & & & 0 & 6 & 10 & 6 & 8 \\ & & & & & & & & & & & 0 & 9 & 4 & 6 \\ & & & & & & & & & & & & 0 & 8 & 8 \\ & & & & & & & & & & & & & 0 & 7 \\ & & & & & & & & & & & & & & 0 \end{bmatrix}$$

Figure 2. Group's dissimilarity matrix.

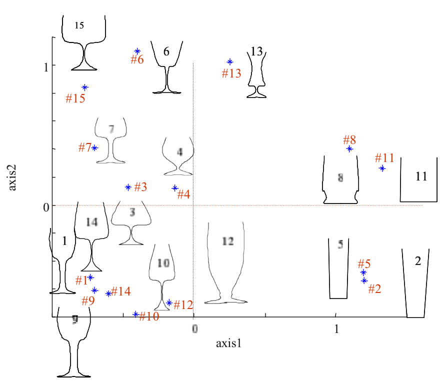


Figure 3. 2D perceptual space.

3.3 Stage 3: raw determination of the semantic space

Next, *Semantic differential method* (SDM) is applied. It consists in carrying out user-tests in which the subjects were asked to assess each glass on a 7 levels Likert scale according to the list of adjective pairs (Table 2). A cluster analysis has been performed on these data in order to find the panel as much homogeneous as possible. One subject, whose the assessment was very different[2] to the rest of the group's assessment, has been removed. We have then calculated the average of the assessment for 10 subjects only.

The *Principal Component Analysis* on the average data allowed the research of underlying dimensions of the semantic space (Fig. 4). Factor 1 and factor 2 account for 64% and 17 % of the variance respectively. So, 91 % of the variance is accounted by a two-dimensional factorial space. Each adjective pair is represented in the factorial space by a "vector", the scalar product between 2 vectors being the correlation coefficient between 2 adjective pairs. After an analysis of the correlations between adjective pairs (colinearity of the vectors), we have extracted a minimal list of semantic attributes (Table 1). For example, adjective pairs **v16**, **v5** and **v17** have been merged because they are highly correlated (Fig. 4), and they are furthermore synonyms.

Table 2. A typical SDM evaluation sheet filled by a subject

	-3	-2	-1	0	1	2	3	
Traditional					×			Modern
Easy for drinking			×					Not easy for drinking
Decorative							×	Practical
Unstable				×				Stable
Masculine						×		Feminine
Complicated						×		Simple
Particular		×						Common
Easy to fill				×				Not easy to fill
Flashy				×				Discreet
Multi-usage						×		Occasional
Easy to handle					×			Not easy to handle
Classy			×					Vulgar
Unoriginal		×						Creative
New				×				Existing
Bad perceived quality					×			Good perceived quality
Fragile					×			Strong
Coarse					×			Delicate

[2] The subject's understanding of the meaning of several adjective pairs was "opposite" to the group's one.

3.4 Stage 4: fine determination of the semantic space

Instead of assessing a particular *score* for the performance of a product on a scale in an absolute manner, subjects were asked to estimate for some pairs of products the relative importance of the *scores* for a given semantic attribute. A qualitative 7-levels scale (<<, <, <~, =, >~, >, >>) is defined. With the pairwise comparison (PC) matrix as input, a score vector summing 100% of importance over the 15 glasses is next computed (Table 3).

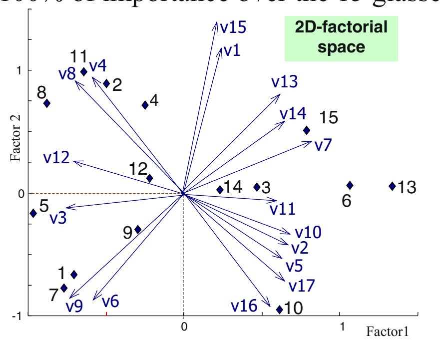


Figure 4. *Position of glasses and adjective pairs in the factorial space.*

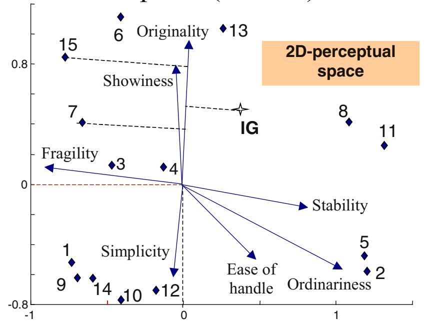


Figure 5. *Position of glasses and semantic attributes in the perceptual space.*

Table 3. A typical PC matrix to pairwise qualitatively compare the "originality" of row product i with column product j's. Next, an "originality" score vector is computed

PC matrix:

« originality »	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1			<					=		<				
2			<<	=						=				
3							>			<<				
4					>>		=	=		=	=	>>	<~	
5							=			=				
6										<<	=			=
7										<~				>>
8										=				
9									=	<				
10										<				
11											>~	>>	<	<<
12														
13														
14														

→

Glass #	*Originality score*
1	4.6
2	7.7
3	7
4	5.7
5	4.7
6	9.2
7	7
8	4.5
9	4.4
10	5.1
11	5.1
12	5.2
13	14.2
14	5.1
15	10.5

Pairwise comparisons are known to be easily administrated because decision makers (DMs), or customers assessing the products, in our case, only focus on a pair of products and on a criterion instead of brutally facing the whole multi-attribute issue. So as not to compel DMs to fill the overall PC matrix like in the well known eigenvector method [26], we used the *Least Squares Logarithmic Regression* (LSLR) PC method proposed by de Graan and Lootsma [27]. Sparse PC matrix are then tolerated, which is preferable for the relative assessment of numerous products (more than eight). Each product is then assessed more precisely and robustly (information is given in excess) with pairwise comparison tests. By this process, for each attribute, a percentage of 100% of importance is shared among the set of 15 glasses. In practice, each of the 10 people is asked to fill at least 30 pairwise comparisons in each of the 13 comparison matrices (corresponding to semantic attributes) of 15×15 size (for the 15 products). Next the LSLR pairwise comparison method [27] is used to calculate the scores relative to a given semantic attribute from the superimposition of the 10 corresponding PC matrices.

In a second step, in order to infer the meaning of the perceptual axis, and to find which semantic attributes are determinant for the perceptions, the semantic space is mapped onto the perceptual space. This is done by a *multiple regression*, using the perceptual axes as independent variables and the semantic attribute as the dependent variable. The outputs of this method, called PROFIT (for PROperty FITting), are the correlation coefficients and the direction cosines (rescalings of the regression coefficients). The attributes for which the multiple regression is significant (according to Fisher-Snedecor table with P-value = 0.05) are called the *determining semantic attributes* (in grey in Table 1); it is assumed that they play an important role for user's perceptions. The vector model of these attributes is plotted in the perceptual space. The origin of the vector is located arbitrarily in the origin of the frame, the values of the direction cosines give the orientation of the arrow, the arrowhead points in the direction of increasing attribute values and the norm of the vector is proportional to the regression coefficient (Fig. 5).

3.5 Stage 5: perceptual products positioning and synthesis aid for a new specification

When the company decides to develop a new glass, it should be made possible to position it in the perceptual space as a perceptual specification. Indeed, this perceptual positioning is determined to avoid the cannibalisation of the other successful products and to fight against competitors. It must be possible to determine from which existing product the new specified glass is perceptually close and to synthesize the vector of semantic attributes of the *ideal glass* (called IG) as a specification for the design department. Here, we propose to specify ideal glass IG in a very flexible and intuitive way by qualitative comparisons under each semantic attribute. Then, a new row and a new column are added (for IG) to the 13 pairwise comparison matrices that are somewhat completed and recomputed into new semantic scores. This is one of the strong point of the methodology: It's easy and intuitive to give specifications by comparisons, particularly when we have to deal with semantic attributes. For example, an absolute value of "originality" = 8/10 doesn't make much sense. On the other hand, a specification of "originality" formulated as "less original than glass #8 but more than #2" is interesting and more easily understandable. A group session is particularly suitable for this specification stage, where each participant can bring a particular light on what could be the new product: The perceptual space and the vector models give for that a convenient support for discussions. Let us recall that our chosen PC method allows to omit some comparisons when no particular specification has to be made. Again this last facility strengthens the flexibility of our methodology.

It has been proposed the following orientations for the ideal glass to develop: A creative and original glass, for occasional usage, which suggests a feeling of solidity, but neither massive nor rough. The corresponding positioning of IG is represented in Fig. 5. It has been specified by instance that the originality of "IG" is "less than #15 but more than #7". The resulting score vectors looks like the ones presented in Table 4.

Finally, in order to adapt the product to the market's segment, pairwise comparison is used again to weight the importance of each semantic attribute. This step is classical in value analysis where the importance of functions is weighted so as to take various aspects of the customer need into account. The processing of the pairwise comparison matrix has led to the semantic attributes' *weights* w_i presented in table 4 (penultimate column).

3.6 Stage 6: design stage

Two glass prototypes (a real and a virtual) are candidates in Fig. 6 for the new design.

Figure 6. Glasses candidates for the new design: **N1** (left) and **N2** (right).

3.7 Stage 7: assessment of the candidate products

Again, new glasses **N1** and **N2** are added to the PC matrices and some relative assessments are provided by the subjects. The scores of all glasses are recalculated (Table 4) entailing a slight modification of the previous scores for the #1 to #15 and IG glasses.

Table 4. Semantic scores for the existing glasses (#1 to #15), ideal glass IG, and new proposed concepts N1 and N2. The two last columns complete the specification

scores (%)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	IG	N1	N2	w_i %	type
Stability	6.5	6.1	6.7	6.8	7.3	4.0	4.4	7.0	5.5	3.9	8.0	5.3	3.4	4.7	5.3	5.5	4.6	5.0	12	**1**
Fragility	6.8	1.6	7.0	2.7	4.5	8.8	6.7	1.3	7.8	8.7	2.6	5.3	7.5	6.0	6.8	5.4	5.4	5.4	7	**3**
Quality	4.8	5.6	6.6	6.7	4.5	5.8	4.5	5.7	5.6	4.5	7.0	5.9	6.3	5.2	7.1	5.1	5.1	4.1	5	**1**
Originality	2.6	5.6	6.2	5.0	3.2	8.8	3.2	4.5	3.1	4.2	4.7	5.6	8.7	5.9	8.5	7.9	6.8	5.6	12	**3**
Smartness	4.4	5.2	10	2.2	3.4	10.3	1.8	2.5	7.2	10.4	6.0	2.5	10.4	6.3	9.1	2.7	2.7	2.7	5	**3**
Ease of handle	6.3	6.7	4.3	4.6	7.6	5.1	6.4	7.6	6.8	5.1	5.9	6.6	4.0	4.8	6.2	5.2	3.4	3.4	10	**1**
Ordinariness	6.8	9.0	3.0	4.2	10.2	3.0	8.1	10.4	6.3	3.2	10.8	5.3	1.2	6.2	2.9	3.0	3.0	3.3	6	**2**
Ease of filling	6.6	8.0	5.0	6.6	7.7	5.3	5.4	7.7	7.6	2.2	7.1	6.1	2.0	5.7	6.5	3.6	2.7	4.1	6	**1**
Showiness	3.0	5.9	5.8	6.7	4.1	7.7	3.6	4.4	4.2	4.9	5.3	5.4	7.1	5.2	7.7	6.7	6.7	5.8	6	**3**
Simplicity	9.0	4.0	5.1	7.9	7.6	1.9	8.5	6.7	8.7	7.3	5.2	6.9	4.2	5.7	3.4	2.2	2.2	3.4	9	**3**
Decorativeness	3.5	5.3	6.2	3.5	2.9	9.4	3.3	3.7	4.1	6.9	4.1	5.3	9.0	6.0	9.2	6.8	5.7	5.1	10	**3**
Ease of drink.	6.1	7.3	4.5	5.4	7.2	4.5	6.2	6.9	7.2	3.1	6.9	6.9	2.3	6.7	5.2	4.4	4.4	4.4	7	**1**
Modernity	3.3	6.0	4.7	5.0	5.3	5.1	3.1	7.4	4.3	5.1	6.5	4.7	7.8	7.4	6.1	5.6	5.6	7.0	5	**3**

3.8 Stage 8: rating and ranking of the products

The first step of this process is to calculate *satisfaction vectors* for each product: existing products (i.e. #1 to #15 glasses) as well as new products (i.e. N1 and N2 glasses). This satisfaction is an estimation of the "closeness" between a given product and the ideal product IG. A satisfaction must be maximum (i.e. equal to 1) when the score of a given product perfectly matches the score of IG, and it must be minimum (i.e. null) when the scores are so different that they are considered as not acceptable.

So as to express that closer the scores, higher the satisfaction, we have defined 3 types of *satisfaction curves* (Fig. 7). Then, a *specification type* must be attributed to each semantic attribute (see the last column of table 4). In Fig. 7, *Ci* is the "target value", which represents the performance (the score) of the ideal product, given in column "IG" of table 4. x% and y% allow the definition of limits of validity and flexibility, classical in functional analysis. Different values of the percentages x% and y% for each semantic attribute allow the definition of appropriate satisfaction curves. The affectation of the type of satisfaction curves to an attribute depends on the meaning of the attribute and affects how the need is formulated.

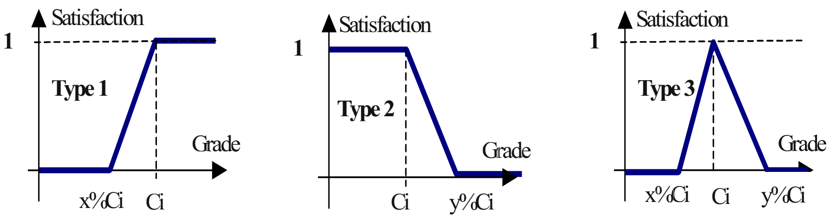


Figure 7. Definition of the 3 satisfaction curves / specification types.

Starting from the scores, the weights and the specification types given in Table 4, a *satisfaction table* is computed (Table 5) through a basic spreadsheet. Note that obviously IG obtains satisfaction grades equal to1 everywhere. For the final grade of each product, a classical multicriteria evaluation procedure, adapted from the AHP, is used. Let *Sij* be the *satisfaction* for product *j* and semantic attribute *i*, let *sij* be the *normalised satisfaction*, N the number of products:

$$s_{ij} = S_{ij} / \sum_{j=1}^{N} S_{ij} \tag{1}$$

Let w_i be the relative weight of semantic attribute *i*, M the number of attributes: The final evaluation of product *j*, *grade(j)*, is given by:

$$grade(j) = \sum_{i=1}^{M} w_i . s_{ij} \tag{2}$$

A rank of the products can then be established, ideal product IG being of course ranked number one. In our case, the new design N1 is the best ranked by the group (2nd, after IG) and glass N2, ranked 6th, has to be improved. The semantic criteria to work on for glass N2 are highlighted (Ease to handle, Simplicity, originality, decorativeness and modernity). Various propositions can be made with CAD systems and virtual prototyping during group sessions.

4. DISCUSSION AND CONCLUSIONS

We proposed in this paper a new integrated and consistent methodology that help designers in the four following conceptual design stages:

- Understanding the need related to product semantics
- Finding relevant criteria to characterise and express the need
- Specifying the requirements of a new product
- Assessing the performances of new solutions

Table 5. Satisfaction's grades and final rank of the glasses

Satisfaction S_{ij}	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**IG**	**N1**	**N2**
Stability	1	1	1	1	1	.11	.33	1	1	.06	1	.89	0	.5	.89	1	.44	.72
Fragility	.11	0	0	0	.47	0	.2	0	0	0	0	.93	0	.6	.13	1	1	1
Quality	.78	1	1	1	.6	1	.6	1	1	.6	1	1	1	1	1	1	1	.33
Originality	0	.05	.29	0	0	.62	0	0	0	0	0	.05	.67	.14	.76	1	.52	.05
Smartness	0	0	0	.33	.17	0	0	.67	0	0	0	.67	0	0	0	1	1	1
Ease of handle	1	1	.44	.67	1	.94	1	1	1	.94	1	1	.22	.78	1	1	0	0
Ordinariness	0	0	1	0	0	1	0	0	0	.83	0	0	1	0	1	1	1	.67
Ease of filling	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	.17	1
Showiness	0	.57	.52	1	0	.52	0	0	0	.1	.29	.33	.81	.24	.52	1	1	.52
Simplicity	0	0	0	0	0	.5	0	0	0	0	0	0	0	0	0	1	1	0
Decorativeness	0	.28	.72	0	0	0	0	0	0	.94	0	.28	0	.61	0	1	.44	.17
Ease of drink.	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1
Modernity	0	.75	.5	.67	.83	.75	0	0	.25	.75	.42	.5	0	0	.67	1	1	.17
grade	2.2	4.4	6.6	4.5	3.5	6.7	1.9	3.4	2.4	4.8	2.9	6.3	4.5	5.0	6.6	15.9	12	6.3
rank	17°	12°	4°	11°	13°	3°	18°	14°	16°	9°	15°	7°	10°	8°	5°	1°	2°	6°

The methodology is efficient to grasp subjective assessments of people. Indeed, the building of the perceptual space and its relation with the subset of relevant semantic attributes is made along several progressive stages. This allows a better comprehension of the market logic. Consequently, a choice in the perceptual positioning of a new product to develop entails an automatic determination of a semantic specification vector for an ideal product. Finally, our methodology performs a multi-criteria subjective evaluation of new design solutions in a precise and flexible way. Subjective evaluations of existing products (pairwise comparison matrices) can be capitalized into a database that can be enriched throughout the new projects. It takes time to complete the tests. Fortunately, the pairwise comparison method that is used runs without all the comparisons, letting believe that our method is tractable in practice. This evaluation procedure is particularly suitable to use in a group session, during which a unique answer of the group is recorded for each comparison, after discussions and negotiations. The next step is to incorporate synthesis tools of design parameters, like in *Kansei* Engineering approaches [19], so as to support the design stage itself.

REFERENCES

1. M. Vigier, La pratique du QFD, La petite encyclopédie de la qualité (Les éditions d'organisation, 1992).
2. N. Suh, The Principles of Design (Oxford University Press, Oxford, 1993).
3. S. Aungst, R. Barton, D. Wilson, The Virtual Integrated Design Method, Quality Engineering, 565 - 579 (2003).
4. R.N. Shepard, K. Romney, S.B. Nerlove, Multidimensional scaling: Theory and applications in the behavioral sciences (Volume I: Theory) (New York, 1972).
5. F. Limayem, B. Yannou, Generalization of the RCGM and LSLR Pairwise Comparison Methods, to appear in Computers and Mathematics with Applications (2004).
6. K. Otto, E. Antonsson, The Method of Imprecision Compared to Utility Theory for Design

Selection Problems, in ASME/DTM '93: Design Theory and Methodology (1993).
7. J.F. Allen, On the Aggregation of Preferences in Engineering Design, in ASME Design Engineering Technical Conferences (Pittsburgh, USA/PA, sept 9-12, 2001).
8. B. Yannou, F. Limayem, S. Hajsalem, Comparaison de la méthode SPEC et de l'Analyse de la Valeur pour l'aide à la conception préliminaire de produits, Mécanique et Industries 3(2), 189-200 (2002).
9. T.L. Saaty, The Analytic Hierarchy Process (McGraw-Hill, New-York, 1980).
10. B. Roy, Méthodologie multicritère d'aide à la décision, SERIE : Production et techniques quantitatives appliquées à la gestion, ed. Gestion C., (Economica, 1985).
11. T. L'Eglise, P. De Lit, J. Danloy, B. Rekiek, A. Delchambre, A multi-criteria decision-aid method for material handling equipment selection, in IDMME2000 : Third International Conference on Integrated Design and Manufacturing in Mechanical Engineering (Montréal, 16-19 may, 2000).
12. C.E. Osgood, G.J. Suci, P. Tannebaum, The measurement of meaning (Illinois Press, 1957).
13. P.E. Green, V. Srinivasan, Conjoint Analysis in Consumer Research: Issues and Outlook, Journal of Consumer Research 5, 103-123 (1978).
14. G.L. Urban, J.R. Hauser, Design and Marketing of New Products (Prentice Hall, New Jersey, 1993).
15. S.K. Gupta, A.K. Samuel, Integrating Market Research with the Product Development Process: a Step towards Design for Profit, in ASME Design Engineering Technical Conferences (Pittsburgh, USA/PA, sept 9-12,, 2001).
16. H. Li, S. Azarm, Product Design Selection under Uncertainty and with Competitive Advantage, Journal of Mechanical Design 122, 411-418 (2000).
17. ECRIN, Evaluation subjective: Méthodes, Applications et Enjeux, Les cahiers du club CRIN (Paris, 1997).
18. E.K. Antonsson, K.N. Otto, Imprecision in Engineering Design, Journal of Mechanical Design 117(B), 25-32 (1995).
19. M. Nagamachi, Kansei Engineering: A new ergonomic consumer-oriented technology for product development, International Journal of Industrial Ergonomics 15, 3-11 (1995).
20. T.W. Simpson, Peplinski, J.D., Koch P.N. and Allen, J.K., Metamodels for Computer-based Engineering Design: Survey and recommendations, Engineering with Computers 17, 129-150 (2001).
21. K. Krippendorff, R. Butter, Product semantics: Exploring the symbolic qualities of form, The Journal of the Industrial Designers Society of America, 4-9 (1984).
22. S.H. Hsu, M.C. Chuang, C.C. Chang, A semantic differential study of designers' and users' product form perception, International Journal of Industrial Ergonomics 25, 375-391 (2000).
23. Y. Matsuoka, Method for constructing an inverse reasoning system for form generation from product image objectives, in ICED 99 (Munich, 1999).
24. R. Lin, C.Y. Lin, W. J., An application of multidimensional scaling in product semantics, International Journal of Industrial Ergonomics 18, 193-204 (1996).
25. J-F. Petiot, B. Yannou, How to Comprehend and Assess Product Semantics - A Proposal for an Integrated Methodology, in International Conference on Engineering Design: ICED'03 (Stockholm, Sweden, 19-21 August, 2003).
26. T.L. Saaty, G. Hu, Ranking by the eigenvector versus other methods in the analytical hierarchy process, Applied Mathematical Letter 11(4), 121-125 (1998).
27. F.A. Lootsma, Performance evaluation of nonlinear optimization methods via multi-criteria decision analysis and via linear model analysis, in Nonlinear Optimization, Powell M.J.D. Editor, vol 1 (Academic Press, London, 1982).

A COMPARISON OF METHODS IN PROBABILISTIC DESIGN BASED ON COMPUTATIONAL AND MODELLING ISSUES

Yee Mey Goh[1], Julian Booker[1], and Chris McMahon[2]
[1]*Department of Mechanical Engineering, University of Bristol;* [2]*Department of Mechanical Engineering, University of Bath*

Abstract: The result of a study on the incorporation of uncertainty into engineering design analysis processes is reported. The work presented comprises a review of deterministic and probabilistic design methods in approximating a design performance function and a comparison of their strengths and weaknesses as tools to aid designers in producing optimum designs. Probabilistic design has existed for more than 40 years, but the culture of deterministic design still dominates in engineering manufacture. An important question for the engineer concerns how probabilistic ideas may be used in design, which is the most appropriate probabilistic approach to use and what are the advantages and disadvantages of each approach? The authors seek to contribute towards an answer to these questions through a comparative assessment of probabilistic design techniques and two methods of deterministic design.

Key words: Probabilistic design; Uncertainty; Design performance; Modelling.

1. INTRODUCTION

A deterministic approach to engineering analyses may result in inconsistent products (over- or under-design) owing to limitations in the information about design, service, material and manufacturing related parameters inherent in the approach [1]. This situation may in the future be unacceptable due to increased customer concern about reliability, the need to improve energy utilisation and increased global market competition in terms of cost and time to market. Probabilistic design offers the required realism needed in engineering analysis and provides the opportunity for improvement in product and business performance in terms of:

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 109-122.

- Economics of design (avoiding material wastage, excessive space)
- Improved reliability of products (failure probability predicted, greater robustness, safety)
- Reduced failure costs (redesign, rework, warranty, liability claims)
- Increased confidence in analysis (more design information, sensitivity relationships)
- Reduction in need for prototypes (which are both expensive and time-consuming).

Many probabilistic design techniques have been developed to facilitate the use of probability theory and to allow knowledge of uncertainty to be applied more effectively in the design process. However, their full use in manufacturing industry is still limited owing to many factors including: the computational intensity of several of the techniques, the requirement for strong statistical knowledge by the user, lack of confidence in the application of the techniques, limitations in the availability and difficulty in data collection [2]. Nevertheless, high volume and customer sensitive manufacturing sectors, such as the automotive industry, would potentially benefit greatly from the use of probabilistic design. However, a cultural and educational step-change, and significant improvement in available data, are necessary for it to be fully integrated into design activities. A first stage in this process is understanding by the engineer of the capabilities and limitations of the available techniques in different problem domains. Current improvements in computer performance, identification of potential failure modes and in computer databases for materials, dimensional and loading conditions indicate that the feasibility of probabilistic design keeps increasing. Increased confidence in the use of probabilistic design techniques will facilitate a move towards fully simulation-based design leading to eventual cost savings on prototype validation prior to the production stage. Key concerns designers have in using probabilistic design techniques are how reliable are the techniques when they are incorporated into the design process, and which of the many available techniques is most appropriate to apply at different stages of the process. This author seeks to contribute to answers to these concerns, firstly through a review of deterministic and probabilistic design techniques, and then through presentation of a method for comparing the strengths and weaknesses of various techniques in application at different stages in the design process.

2. DETERMINISTIC DESIGN

Deterministic design accounts for uncertainties by introducing Factors of Safety (FS) and using a lower bound strength value and an upper bound load value. The selection of FS value greatly depends upon the designer's engineering knowledge and experience, which may not be available for new designs. Excessive conservatism in the selection of FS will lead to over-design, which is uneconomic, whereas lack of conservatism causes under-design, which results in low reliability and high failure costs. Often, both will be present in a single product. Deterministic design, by definition, results in the loss of probability information.

2.1 Interval Analysis

Interval analysis is the simplest way of representing uncertainty when no information concerning the likelihood or probability of any particular value within the range is available. A common application of interval analysis is in tolerance stack analyses. In an Absolute Worse Case (AWC) design, for instance, the variables are either set to the lowest or largest expected value depending on the respective partial derivatives of the variables and on maximising or minimising conditions. AWC is thus an approach to account for the statistical nature of a design environment in a deterministic way [3]. The prerequisite of this approach is a monotonic performance function, i.e. its first derivative is of a constant sign (+ or -).

2.2 Partial Safety Factor (PSF)

PSF is a widely used approach in civil engineering. It employs the limit state concept (see section 3 below) with separates safety factors on loading (γ_f) and strength (γ_m) in order to account for the different amounts of variability associated with each of these. The factors are available in standards such as BS5950 and Eurocodes [4], the latter providing more conservative values. Although variability in load and strength are accounted for by the use of partial factors, this conventional design method tends to over-design owing to conservatism in the factors applied. In addition, there is no estimate of probability of design failure, only an indication of whether the design is in a feasible region or otherwise.

3. PROBABILISTIC DESIGN

The limit state theory is the foundation for the development of many probabilistic design techniques. These generally attempt to solve the limit state functions, sometimes called the performance functions. Limit state theory provides a framework within which the performance of components can be assessed against various limiting conditions, such that the component is no longer capable to fulfill its intended function [5]. Owing to its importance in structural safety, the limit state is often defined in terms of load and resistance or stress and strength, but in principle any performance parameter can be used in its formulation. Mathematically, the performance function for a strength limit state is defined as the difference between the resistance, R and load, P as in Eq. (1).

$$g(\mathbf{x}) = R - P = 0 \tag{1}$$

where $\boldsymbol{x}$ is the vector of basic random variables, $g(\mathbf{x}) < 0$ represents the failure region and $g(\mathbf{x}) > 0$ represents the safe region. The Probability of Failure (POF) can be evaluated by the integral of the joint probability function, $f_x(\mathbf{x})$ of the random variables in the system over the entire failure region, given by Eq. (2).

$$POF = \int_{g(x)<0} \cdots \int f_x(x)dx \tag{2}$$

3.1 Coupling Formula

When the limit state function is a combination of more than two statistically independent variables, x_i the variance of response $g(x_i)$ can be approximated by Taylor expansion. Ignoring second order and higher terms, the variance for a function $g(x_i)$, σ_g can be analytically approximated by Eq. (3), where σ is the variable standard deviation and μ is its mean.

$$\sigma_g^2 \approx \sum \left(\frac{\partial g}{\partial x_i} \right)^2 \sigma_{x_i}^2 \tag{3}$$

The standard normal variate can be calculated from the coupling formula given by Eq. (4) [6]. This method is sometimes known as the First Order Second Moment (FOSM) method. The POF is then determined from the standard normal distribution function, Φ as in Eq. (5).

$$z = -\frac{\mu_R - \mu_P}{\sigma_g} \tag{4}$$

$$\mathrm{POF} = \Phi(z) \tag{5}$$

3.2 Monte Carlo Simulation (MCS)

In many practical problems it is not possible to directly evaluate the POF, and in these cases Monte Carlo Simulation (MCS) allows its evaluation by simulation involving repeated computation of the performance function for randomly generated combinations of input variables and generation of the POF from the distribution of these calculated values. MCS is effective in representing the variability in design parameters, however the number of iterations must be relatively large (~10^5-10^7) to produce reasonably accurate results. The coefficient of variation of the POF estimator for a specified POF, ν_{POF} is given by Eq. (6) [7].

$$\nu_{POF} = \sqrt{\frac{1-POF}{n*POF}} \tag{6}$$

where *n* is the number of iterations. For problems with high reliability (i.e. low POF) *n* must be large to generate an accurate estimator of POF, requiring expensive computation which is time consuming. This limitation makes the application of MCS in some domains, such in stress analysis using Finite Element Analysis (FEA), expensive and difficult. Various variance reduction techniques have been proposed in the literature to speed up the standard MCS. Some of the most widely used variance reduction techniques in engineering reliability analysis are importance sampling [8] and Latin Hypercube Sampling (LHS) [9]. These techniques reduce the computation time over the standard MCS.

3.3 First Order Reliability Method (FORM)

In FORM, the limit state is linearised around the “design point”, the point on the limit state with the highest probability, also called the Most Probable Point (MPP). Owing to this linearisation, FORM may not be sufficiently accurate for highly nonlinear problems. In addition, as the degrees of freedom (DOF) or the number of variables of the problem increase, the error increases rapidly. It also fails to utilise the distributional information but has the advantage of simplicity. Transforming the original random vectors to standard uncorrelated normal vectors, the reliability index, β is then defined as the shortest distance from the origin to the MPP on the limit state function. In the standard normal space, the POF is given by Eq. (7) [10].

$$\text{POF} = \Phi(-\beta) \tag{7}$$

The Second Order Reliability Method (SORM) is constructed by fitting a parabolic surface, as opposed to a plane surface in FORM, to the limit state function at the design point. The information about the curvature of the limit state function is utilised in SORM, therefore improving results compared with FORM, but at the expense of greater computational complexity. Several methods have been proposed in the literature to solve for POF from the SORM approximation functions [11, 12].

3.4 Response Surface Methodology (RSM)

The RSM is an approximate mathematical function of the limit state, which avoids computations of the actual performance function. The principle behind the RSM is that a smooth surface and a suitable polynomial can be used as local approximations over some limited region of current interest. A second order polynomial is most common as it is sufficiently accurate in representing many engineering problems. Due to its experimental nature, RSM is very useful when a closed-form performance function is not available. Bucher suggests a response surface function, $g'(\mathbf{x})$ of the form Eq. (8) [13].

$$g'(\mathbf{x}) = a + \Sigma b_i x_i + \Sigma c_i x_i^2 \tag{8}$$

The points $\mathbf{x_i}$ are chosen to be mean values $\bar{x}_i$ and $x_i = \bar{x}_i \pm f_i \sigma_i$ where f_i is an arbitrary factor. This reduces the number of calculations of the actual performance function, $g(\mathbf{x})$, to $2k+1$, where k is the number of variables. A conventional MCS may then be applied, using $g'(\mathbf{x})$ as the performance function, often with considerable savings compared with using $g(\mathbf{x})$.

3.5 Advanced Mean Value (AMV)

AMV works on the basis of the assumption that if a limit state function $g(x_i)$ is smooth then a Taylor series expansion exists at the mean values, and the approximation given by Eq. (9) may be used [14].

$$g'(x) = g(x_i^*) + \sum \frac{\partial g}{\partial x_i}(x_i - x_i^*) \tag{9}$$

where x_i* is the co-ordinate of MPP. Using perturbation or finite difference techniques to evaluate the derivatives, $g'(\mathbf{x})$ is updated with information from the actual limit state function, $g(\mathbf{x})$ in each iteration, making it converge rapidly to the MPP and β can be determined from the shortest distance between the mean value point and the MPP. The number of evaluations of *g(**x**)* in AMV is reduced significantly compared to MCS, improving the efficiency for cases with complex performance function, *g(**x**)*.

3.6 Fast Probability Integration (FPI)

FPI is an approach which evaluates the actual performance function, *g(**x**)*, but approximates the MCS [15]. This approach differs from the RSM, which approximates *g(**x**)* and uses the actual MCS. The results from FPI are accurate and efficient, even for highly nonlinear performance functions, owing to the utilisation of information originating from *g(**x**)*. The FPI is in principle similar to AMV, but a nonlinear limit state function is first approximated with a quadratic function (Eq. 10) using a Taylor's series expansion around the design point. It is then transformed to a linear function (Eq. 11) via the transformation of x_i to y_i using Eq. (12). FPI is robust to variability in the input parameters, by effectively approximating non-normal distributions with equivalent normal distributions but this requires some extensive programming.

$$g'(x) = g(x_i^*) + \sum \frac{\partial g}{\partial x_i}\left(x_i - x_i^*\right) + \sum \frac{1}{2}\frac{\partial^2 g}{\partial x_i^2}\left(x_i - x_i^*\right)^2 \quad (10)$$

$$g(y)=c_0+\Sigma c_i y_i \quad (11)$$

$$y_i = \left[x_i - \left(x_i^* - \frac{a_i}{2b_i}\right)\right]^2 ;\ a_i = \frac{\partial g}{\partial x_i} \text{ and } b_i = \frac{1}{2}\frac{\partial^2 g}{\partial x_i^2} \quad (12)$$

4. CASE STUDY EVALUATION AND DISCUSSION

4.1 Development of Evaluation Criteria

In order to fully understand the obstacles of incorporating probabilistic design techniques into the design activity, a number of evaluation criteria are proposed, based on the issues listed below. It is suggested that these criteria may be used to compare the efficiency and effectiveness of each technique in solving problems in design analysis generally speaking:

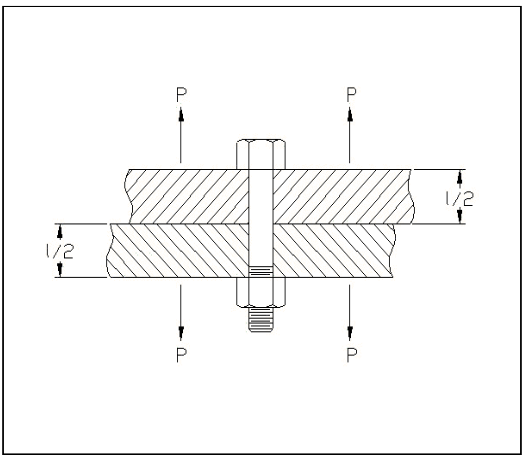
- **Criterion 1** concerns the **time** required for probabilistic analysis compared with the conventional deterministic approach. It is a combination of set-up (i.e. analysis) time and computation time. The set-up time is a “one-off” cost and includes time in learning the method and creating the model whereas computational time accounts for evaluating the model to identify POF.
- **Criterion 2** concerns the **cost** of implementing the methods in terms of the equipment and software required. This is the initial cost incurred in acquiring and implementing the techniques. Other costs involved are in learning and model creation, consolidated into the set-up time.
- **Criterion 3** concerns the **accuracy** of the probabilistic model. The accuracy required is dependent on the level of risk associated with the product and its failure consequences and the stage in the design process the technique is utilised.
- **Criterion 4** concerns **robustness** of the approach towards the complexity of the performance function. It comprises the ability of a method to cope with non-linearity in the performance functions, non-normal variable distributions and the number of DOF in the performance functions.

In the next sub-section, a case study will be introduced in which qualitative criteria are used in order to compare techniques in a simple engineering problem. In the next section, the implications of the criteria in a more complex example are discussed.

4.2 Bolted Joint Case Study

Avoiding bolted joint separation is a typical engineering problem faced by designers, a practical failure mode of which is steam escaping from a flanged connection. The performance function for this problem can be readily formulated and will be used to assess the merits and demerits of each technique based on the evaluation criteria discussed. The bolted joint diagram and derivation of the joint separation performance function are shown in Fig. 1. For a joint separation to occur, the forces in the clamped members must equal zero, indicating that the entire externally applied load is load, P is therefore greater than the reaction force in the bolt, R. Analytically, the condition of failure occurs when $P > R$ and the performance function can be written as in Eq. (1). Table 1 summarises the source data used to construct normal distributions to characterise the input variables in the final performance function [16, 17]. Where a range is given from the source, these limits are taken as $\pm 3\sigma$ from the mean values (which account for 99.73% of the normally distributed variables).

It is suggested that for this problem a useful indication of the relative merits of the different approaches can be obtained by using simple numerical or qualitative ratings for each criterion. For the time criterion, in this example the set-up time is rated on a scale ranging from 1 (no set-up effort) to 5 (extensive training and model development required) and the execution time is indicated in CPU seconds for different approaches. A qualitative judgment of time requirement is obtained by using the product of these two values. A qualitative judgment is again used for implementation cost, based on platform type. The values used were 1 (mental calculation only needed), 2 (calculator), 3 (spreadsheet), 4 (programming language) and 5 (full commercial package required). A scale of 1 to 5 was also used for judgement of robustness, where 1 = robust only to linear, normal and low DOF performance functions and 5 = robust to non-linear, non-normal and high DOF cases. In this example, accuracy was compared by taking a MCS with 10^6 evaluations as the accurate result.



$$p = 1.5$$
$$d_r = d - 1.226869p$$
$$d_m = d - 0.649519p$$
$$A_d = \frac{\pi d^2}{4}$$
$$A_T = \frac{\pi\left(d_r + d_m\right)^2}{16}$$
$$k_b = \frac{A_d A_T E}{\left(A_d l_T + A_T l_d\right)}$$
$$k_m = \frac{0.577\pi E d}{\left\{2\ln\left[5 * \frac{\left(0.577l + 0.5d\right)}{\left(0.577l + 2.5d\right)}\right]\right\}}$$
$$C = \frac{k_b}{k_b + k_m}$$
$$F_i = \frac{T}{Kd}$$
$$R = \frac{F_i}{\left(1 - C\right)}$$
$$g = R - P$$

Figure 1. Bolted joint arrangement and derivation of performance function.

Ideally, cost and total time should be low with a high robustness to the objective function. Since robustness is dependent on the complexity of performance functions, and cost and time factors are resource based, the combination of these suggest an improved rating for selection purposes. For this reason a combined rating based on cost, time and robustness is used to combine the overall traits of the techniques evaluated. Error, being inherent to each technique may be treated in isolation. The rating is given in Eq. (13).

$$\text{Combined\_Rating} = \frac{\text{Robustness}}{\text{Cost x Time}} \tag{13}$$

Table 1. . Source Data and Normal Distribution Parameters for Input Variables.

Variable	Symbol	Source Data/ Unit	Normal Distribution	
			Mean (μ)	Standard Deviation (σ)
Major diameter	d	(9.78 to 10.00) mm	9.89 mm	0.037 mm
Threaded length	l_T	Normal (12, 0.117) mm	12 mm	0.117 mm
Unthreaded length	l_d	Normal (18, 0.117) mm	18 mm	0.117 mm
Plate thickness	l	20 ± 0.07 mm	20 mm	0.023 mm
Modulus of Elasticity	E	208 MPa, C_v = 0.03 (worst)	208 MPa	6.24 MPa
Applied Torque	T	(16.272 to 16.724) Nm	16.498 Nm	0.075 Nm
Nut factor	K	(0.10 to 0.16)	0.13	0.01
Coefficient of Friction	μ	0.15 ± 0.03	0.15	0.01
Load	P	μ_P ± 30%	11.0 kN	1.1 kN

Table 2. Summary of Evaluation Criteria for Each Technique.

Technique (number of iterations)	Set-up Time (Ratings)	Comp. Time (s)	Comp. Time (Ratings)	Set-up Time x Comp. Time (Ratings)	Error (%)	Robustness	Cost
Coupling Formula	2	0	1	2	114.29	1	2
Commercial MCS (10^4) (spreadsheet package)	2	54	3	6	14.29	5	5
MCS (10^6)	2	440	5	10	0	5	4
FORM	3	0.05	1	3	114.29	2	4
RSM (10^5)	3	37.46	2	6	6.43	4	4
AMV	5	0.77	1	5	2.86	4	4
FPI	5	0.77	1	5	2.86	5	4
AWC	1	0	1	1	-	3	3
PSF	1	0	1	1	-	3	2

The results corresponding to the evaluation criteria are summarised in Table 2 for the bolted joint problem. By plotting these results a visual indication of techniques performance may be obtained. Fig. 2 shows an example – error vs an overall rating obtained using Eq. (13). As noted, error is based on the POF calculated using MCS with 10^6 iterations (POF = 0.0014). Error estimates of the POF are not available for the deterministic methods (PSF and AWC), however, POF = 0.05 relates to a design that is safe for a serviceability limit state [18]. In contrast, AWC, being more conservative, classifies it as an unsafe design. There exists an inverse relationship between the error and total time. MCS requires long execution time (for the given number of iterations) whereas analytical methods and

FORM are fast, but are the least accurate. The less accurate methods are more suitable in the initial stages where high accuracy is not needed as there is limited information about uncertainties in the design parameters. The sensitivity analysis results available from FORM can be used subsequently to guide more effective information collection. Methods such as RSM, AMV and FPI represent trade-offs between accuracy and time. MCS can be applied in later stages where accuracy becomes more important as the design proceeds. Fig. 2 shows that FPI, AMV and RSM are within 10% of MCS the result, and have a high overall rating, when robustness to the objective function is included. For problems with time constraints and complex performance functions, these methods are preferred, but the more robust the technique is, the more expensive it is.

4.3 Front Suspension Case Study

The bolted joint case study above mentioned uses an explicit, closed form, non-linear objective function with seven DOF and all variables are assumed to be normally distributed. It is also a static problem where loading is time-invariant. However, objective functions formulated for many real engineering problems are likely to be complex in nature. For instance, explicit performance function is not available in computational analyses such as FEA, Multi Body Systems (MBS) etc. A more realistic case study analyzing a suspension system for a sports utility vehicle is an example of a complex mathematical as well as engineering problem. In this situation, probabilistic design may not be so straightforward but probabilistic approaches may be used to draw correlations between analytical results for the loads on the chassis of the car (performance parameter) and experimental road load data, and to direct the collection of statistical information about the key parameters in the suspension system e.g. bush stiffness, damping coefficient, spring rate, jounce stiffness, etc. including the characteristics of the components (linear, non-linear). The aim is then to demonstrate the feasibility of probabilistic design and how it addresses uncertainty in such a complex analytical problem. For this purpose, a half vehicle model is developed in Automatic Dynamic Analysis of Mechanical Systems (ADAMS) and a simplified quarter car model is developed in Visual Basic.

- *Time*: As this is a pilot analysis of complex system, the difficulty of setting-up a probabilistic design within a computational package could be overwhelming, therefore more weight should be given to the set-up time over the computational time. The ability to understand and the ease of modeling with MCS are crucial to provide a better visualization of the problem. More advanced techniques could be justified once the probabilistic model is verified.

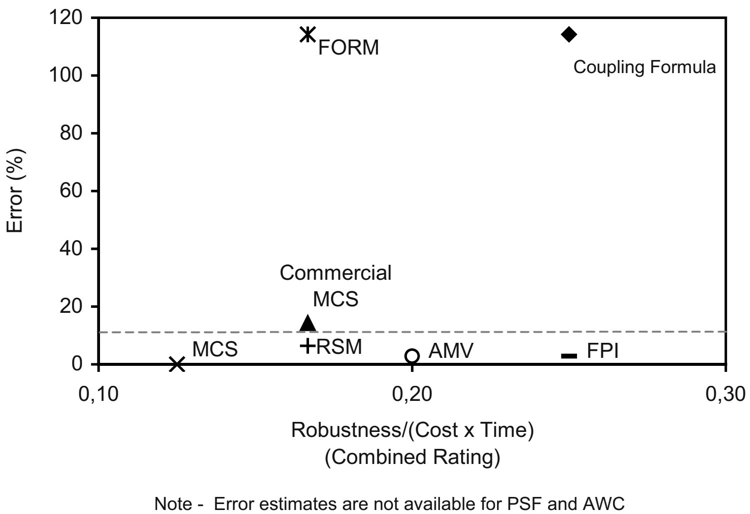


Figure 1. *Error versus combined rating for each technique.*

- *Cost*: The cost of implementation is not considered important in this problem.
- *Accuracy*: The ADAMS model needs to be verified to facilitate uncertainty analysis, therefore the accuracy of the probabilistic techniques incorporated is crucial.
- *Robustness*: The transfer functions in this dynamic problem are implicit as they require exhaustive numerical integration in solving them. The interface of probabilistic design and such complex computational analysis could be most easily achieved by executing the existing model repeatedly, updating the variables at each iteration. The random variables could be generated externally according to its characteristic probability distribution to form a matrix of realistic random systems, this is effectively a MCS approach.

It is demonstrated that the implementation of probabilistic design in such complex problem is feasible and it facilitates better understanding of uncertainty in design analyses but the inherent complexity of the problem governs the selection of probabilistic techniques. A response surface function obtained describes the relationship between the performance parameter and the critical design variables, which is valuable for a parametric study.

5. CONCLUDING REMARKS

A method for evaluating probabilistic design techniques to aid their selection based on the computational and modeling issues of time, accuracy, cost and robustness has been outlined. With the help of information presented, the probabilistic design techniques can be applied in more strategic contexts, so that design optimization procedures can be carried out more efficiently and judgement of when in the design process to apply them can be made to suit the nature of the problem. Some of the techniques also give sensitivity measures, which could augment the strategy for probabilistic design applications – providing insights into critical variables to aid further data collection. Case studies of both simple and complex design situations were adopted to provide the basis for discussion, but more cases of varying dimensions should be taken into consideration to aid a more comprehensive study taking into consideration opinions of many engineers. Selection of probabilistic and computational techniques is important to reproduce the physical problem for analysis.

With the availability of commercial software, learning time could be reduced drastically. This factor could be of less importance if commercial software is available to aid the probabilistic design process, the only software to date that includes the wide range of probabilistic techniques is called NESSUS [19]. However, the accuracy and validity of the probabilistic models still rely on the knowledge and experience of the engineers who develop them. Other less comprehensive probabilistic commercial packages are also available, mostly to provide the MCS and LHS on more user-friendly platforms. It can be argued that computational time is redundant given the increase in the speed of computers today, but is of concern when probabilistic design is to be applied in applications such as FEA and MBS, where probabilistic design, which typically involves thousands of iterations, could amplify the already time consuming simulation process, e.g. from hours to days. Performance functions that are implicit of the actual design variables and uncertain system parameters could only be evaluated in a numerical sense, i.e. FEA. Hence, numerically efficient probabilistic algorithms for evaluating this type of system response are required [20]. Although no extra modelling effort is required in the problem formulation in probabilistic design over a deterministic one, time to collect data has been recognized widely in literature as a contributing factor to the lack of probabilistic applications. Most published sources only quote deterministic values such as minimum expected values, nominal values etc. Making available statistical information such as standard deviation, coefficient of variation, etc. will certainly encourage more probabilistic design applications.

REFERENCES

1. K.V. Bury, Statistical Models in Applied Science (Wiley, NY, 1975).
2. B.H. Thacker, D.S. Riha, H.R. Millwater, M.P. Enright, Errors and uncertainties in probabilistic engineering analysis, Proc. 42nd SDM, Seattle, WA, 16-19 April 2001.
3. G.E. Dieter, Engineering Design, 3rd ed. (McGraw-Hill, Boston, 2000).
4. C. Arya, P.J. Wright, Design of Structural Elements: Concrete, Steelwork, Masonry and Timber Design to British Standards and Eurocodes, (Spon Press, GB, 1993).
5. D.A. Nethercot, Limit States Design of Structural Steelwork, 2nd ed. (Chapman & Hall, London, 1991).
6. E.B. Haugen, Probabilistic Mechanical Design (Wiley, USA, 1980).
7. M. Kleiber, J. Rojek, R. Stocki, Reliability assessment for sheet metal forming operations, Comp. Methods in App. Mech. and Eng., 191, 4511-4532 (2002).
8. R.E. Melchers, Structural Reliability Analysis and Prediction, 2nd ed. (John Wiley & Sons, UK, 1999).
9. M.D. McKay, R.J. Beckman, W.J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics, 21, 239-245 (1979).
10. A.M. Hasofer, N.C. Lind, Exact and invariant second-moment code format, J. of Eng. Mech. Div., ASCE, 100, 1227-1238 (1981).
11. L. Tvedt, Distribution of quadratic forms in normal space - application to structural reliability, J. of Eng. Mech., 116, 1183-1197 (1990).
12. A. Der Kiureghian, H.Z. Lin, S.J. Hwang, Second-order reliability approximations, J. of Eng. Mech., 113, 1208-1225 (1987).
13. C.G. Bucher, U. Bourgund, A fast and efficient response surface approach for structural reliability problems, Struct. Safety, 7, 57-66 (1990).
14. Y.T. Wu, H.R. Millwater, T.A. Cruse, Advanced probabilistic structural analysis method for implicit performance functions, AIAA Journal, 28, 1663-1669 (1990).
15. Y.T. Wu, P.H. Wirsching, New algorithm for structural reliability estimation, J. of Eng. Mech., 113, 1319-1336 (1987).
16. J.D. Booker, M. Raines, K.G. Swift, Designing Capable and Reliable Products (Butterworth-Heinemann, UK, 2001).
17. J.E. Shigley, C.R. Mischke, Mechanical Engineering Design, 6th ed. (McGraw-Hill, Boston, 2001).
18. ENV 1991-1, Eurocode 1 - Basis of Design and Actions on Structures - Part 1: Basis of Design, 1994.
19. D.S. Riha, B.H. Thacker, M.P. Enright, L. Huyse, S.H.K. Fitch, Recent advances of the NESSUS probabilistic analysis software for engineering applications, Proc. 43rd SDM, Denver, Colorado, 22-25 April, 2002.
20. H.A. Jensen, A.E. Sepulveda, Optimal design of uncertain systems under stochastic excitation, AIAA Journal, 38 (11), 2133-2141 (2000).

A FRAMEWORK SUPPORTING COLLABORATIVE OPTIMISATION FOR MULTI PROFESSIONAL DESIGN TEAMS

Vincent Riboulet[1], Philippe R. Marin[1], Stephen Gowers[2], and Frédéric Wurtz[3]

[1]*University of Grenoble, Institut National Polytechnique de Grenoble, Laboratoire « Sols Solides Structures », Domaine Universitaire BP 53 Grenoble Cedex 9 France, phone (33) 04 76 82 51 29, fax (33) 4 76 82 70 43; Philippe.Marin@inpg.fr;* [2]*University of Bath, UK;* [3]*Laboratoire d'Électrotechnique de Grenoble*

Abstract: Today in industry, product development is becoming an increasingly complicated activity and design teams can be composed of people from different cultures who are geographically dispersed. This project investigates the problems caused when required to work between engineers from different domains (differing 'professional cultures), in order to suggest new approaches and to develop a framework supporting collaborative global design. A general cultural study was made to analyse the product development processes within mechanical and electrical engineering that identified optimization as being a key stage in the development of electromagnetic products. Electrical engineers often use an analytical optimization environment to give dimensions to an 'overall' initial electromagnetic structure. Segregated analysis carried out by mechanical engineers at this stage can be very inefficient and ineffective. Analysing a specific example, the dimensioning of a spring in an electromagnetic plunger, has highlighted the advantages offered by improving collaboration. What has been proposed is an integrated approach to optimization, which involves mechanical and electrical engineers working together to jointly dimension a structure. An ICT (Internet Collaboration Tool) prototype supporting this type of work has been developed and optimization results can now be disseminated automatically to all of the designers in a global design team. This is a step in the development of a computer environment supporting collaborative design for multidisciplinary teams.

Key words: Distributed design, Collaboration, Optimiz ation, Electromagnetics

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 123-134.

1. INTRODUCTION

1.1 Economic context

Industrial Design is an activity where collaboration is fundamental. Indeed, many actors collaborate to achieve a common goal: The definition of a product that can be used and may be sold. Each one has its own knowledge and know-how, its own uses. If every one speaks about the same object, the product, the representation that each actor uses is not unique. "Designing industrial products is a collective, and distributed activity, implying cooperation, information sharing and coordination among several skilled actors who have got specific knowledge in their own domain" [1]

In this context collaboration must be organized to allow better communication between actors. Concurrent engineering methodologies are developed in order to involve the different points of view on the product, all along its life cycle. This trend is pushed by the evolution of the products: design of complex artifacts and systems requires the cooperation of multidisciplinary design teams providing optimization from any point of view. Several different people from different countries, companies and professions can now be required to work together during a product's development, which implies that different cultures have to be considered in the management of the product development process which can cause huge problems. Cultural studies carried out [2, 3] suggest four main types of culture that are apparent in industry today: National Culture, Corporate Culture, Branch Culture and Professional Culture. Professional culture, which relates to someone's profession, is at the heart of this study.

At the same time the globalization of the industrial activity and decentralization of many manufacturing processes leads companies to work in relation to very distant collaborators. Let us cite the example of Airbus planes built in several different European countries. Then, let us keep in mind the growing interaction with suppliers. B2B (Business to Business) needs strong support for integrating the design chain. A design which might, in the future, also integrate the consumer to allow full product customisation…

1.2 Collaborative design environments

The expansion of Internet-based tools has opened new opportunities for collaborative work improvement. This can explain the development of a new generation of tools designed to support this kind of activity: Internet/Web and CSCW (Computer Collaborative Work) in design [4]. They intend to offer to a team of engineers (design, production or sale offices, subcontractors...) the way to work together, sharing data about the product [5]. The functionalities of the bests tools of this kind are: the integration of several tools supporting communication, a PDM (Product Data Management) system, and the capacity to visualise a common 3 Dimensional model, and sometimes to annotate and modify it. The rigidity of these solutions seems to be a problem for the support of a satisfactory technical communication among a multidisciplinary design team [6]. The challenge for software developers is to provide users with solutions that allow the best compromise between communication among themselves, and ergonomics for each of them in his/her daily activities.

2. A DESIGN EXPERIMENT

2.1 Goals

To stress the communication exchanges, we decided to perform a distributed design with a team of engineers with different professional cultures [7], or say from different worlds, as defined in [8]. Mechanical and Electrical engineers have quite different ways to solve design problems, and different simulation tools that do not manipulate the same concepts, information or parameters. The problems caused by having to work between professional cultures are apparent in several areas of industry today and poor communications can sometimes prevent the advancement of innovative technology [9]. An interview with two engineers from Schneider Electric, suggested that when mechanical and electrical engineers are required to work together in industry there really are problems. In a lot of cases it seems that electromagnetic product development is a very segregated activity, designers adopting an 'over the wall' approach - an electrical engineer carrying out a study in isolation before passing the results onto a mechanical engineer to continue with product development. When the two engineers need to dialogue, it is not surprising that there are often misunderstandings and that the product development process can be very iterative and inefficient. What are the problems? How can we enhance collaborative work?

2.2 Teams

Eight Members of the “Laboratoire d’Electrotechnique de Grenoble” (LEG) and of the laboratory 3S "Sols-Solides-Structures”, (Grenoble, France), worked together in different physical places communicating over the Internet, trying to reproduce the working conditions of different professional groups collaborating over the Internet [10]. The different groups were often composed of one or two people depending on their professional competencies: mechanical or electrical engineering, structure definition, electro-mechanical sizing, manufacturing process definition, etc.

2.3 Product

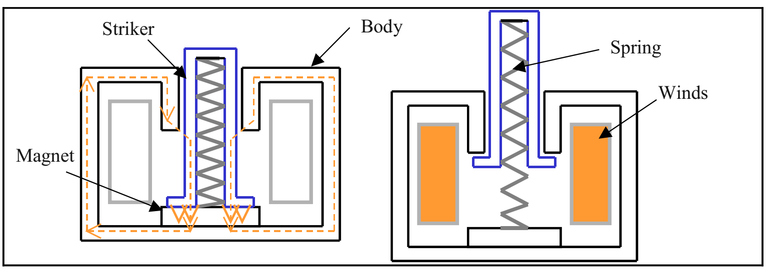


Figure 1. Stand-by position and released position

They designed an electromechanical plunger. (circuit-breaker) (see fig 1) which presents the particularity of being a problem involving numerous aspects: mechanical, magnetic, electrical, thermal which are moreover very coupled. This product was chosen to lay emphasis on collaboration and communication.

2.4 A Different Professional Culture : the electrical world

Different stages of the electromagnetic product development were analysed at the LEG and comparisons drawn with the approaches adopted by mechanical engineers. Electrical engineers use lots of variants and adaptive designs. As a result there is little support for upstream phases of innovative design projects. They put a lot of emphasis on the embodiment design stages with several computerized tools to support optimization and dimensioning [11]. An analytical (as opposed to a geometric) representation is usually at the heart of the product development process, which offers advantages and disadvantages when working with other engineers. The nature of electromagnetism restricts designers to an 'overall' approach, which makes the decomposition of products for isolated studies very difficult.

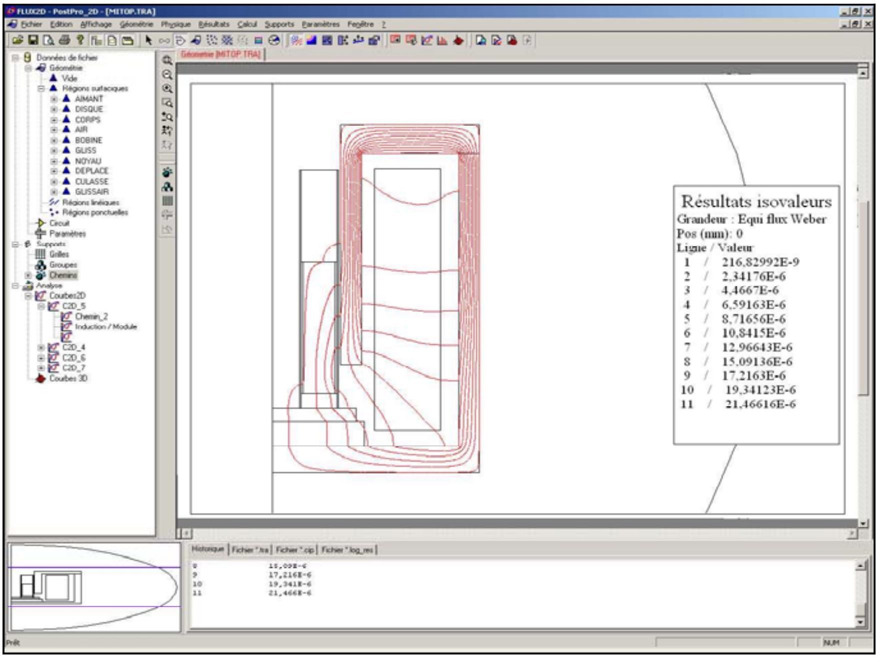


Figure 2. Flux 2D, magnetic F.E.M analysis

Nowadays, a specialist computer environment also contributes to the "specialist view" that has been discussed in the previous section. A study of computer applications used by electrical engineers has suggested that they use very specific software like Flux 2D (fig 2) and CDI Optimizer. In the LEG, optimization is often regarded as a method of dimensioning an initial structure with an objective function driving this embodiment stage of the design process. Note that this approach implies that most of the design parameters and optimization criteria can be analytically expressed, in a coherent manner.

2.5 Meetings

We observed that, from all the available tools, the team mainly used 2 kinds: highly specific technical tools such as a solver for magnetic flux, a

C.A.D (Computer Aided Design) geometric modeler, F.E.M (Finite Element Model) analysis or process planning… along with non specific communication tools like e-mail, forums, ftp, web space storage, Netmeeting, Paltalk (voice sharing). Special functionalities of the specific collaborative tools were not interesting in our case as for example, 3D model sharing is not adapted for the electrical engineers who do not use a 3D model as a working base! Different kinds of meetings were performed. Let us cite collaborative preliminary design, sizing, manufacturing, and review type meetings. This was completed by asynchronous exchanges when needed.

3. A SIGNIFICANT EXAMPLE – THE DIMENSIONING OF A SPRING

The problem of dimensioning a spring may look simple at first sight, but it is a very good example because the team of designers did not succeed in solving it! Finding an admissible set of values for the parameters in a distributed design context appeared to be very complicated.

3.1 Detailed view of the problem

The striker of the plunger is held in the low position by a magnetic force generated by the magnet. When an electrical current flows through the windings, a magnetic flux starts to counteract the magnet magnetic flux. Therefore the net magnetic force becomes weak. At this point the main force on the striker is the force of the spring pushing up. The striker begins to move.

Functional requirements:

- Holding an acceleration in the low position of ($a=2000 m.s^{-2}$)
- Maximum time to react (3,5 ms)
- Percussion energy (0,12 J)
- Residual force in high position (15N)
- Maximum size of the plunger 40*50*50mm

3.2 A complex and coupled problem

The behavior of the system is greatly influenced by the sizing of the structure. For a given geometrical configuration, the magnetic flux is influenced by aspects like the thickness of the walls, the gap between the different parts and the magnetic properties of the material.

The electrical engineers made an analytical optimization study in order to perform the structure sizing. Their objective was to minimize the overall volume of the plunger. At this point they asked the mechanical engineers for the spring specifications, more precisely the force equation (in order to know this value in the compressed state).

As the team did not know the size of the magnetic parts, it was not possible to size the overall structure. But it was possible to take into account the requirements which did not depend on the magnetic force. So the mechanical engineers started to develop a mechanical model of the spring compliant with these requirements.

$$F = k.X \quad (1)$$

$$k.Xlow + m.2000 < Fmagnetic + weight \quad (2)$$

$$k.Xhigh > 15 + weight \quad (3)$$

The problem is that the geometry of the striker is related to:

- its mass, hence the force corresponding to a $2000 m.s^{-2}$ acceleration, the time to move out, and the final kinetic energy. (Moreover the mass of the spring is not negligible).
- the magnetic flux which can flow through it.

The mechanical engineers gave their equations to the electrical ones, in order for them to try to find a structure satisfying the magnetic requirements. But it failed. First the optimization did not converge. On the second try with another spring, the magnet needed was too big for the structure.

Mechanical engineers tried to make new calculations, but were blocked because they did not know which parameter to move and in which direction. They faced the complexity of this very coupled problem. The whole procedure (exchanging the parameters and equations, looking for an admissible solution) took several hours of collaborative work between the two teams. The cultural issues led to misunderstandings that slowed the development process. But after cultural barriers were overcome, the technical difficulty of solving this very coupled problem appeared. Eventually, no satisfactory solution was found.

4. TOWARDS A COLLABORATIVE OPTIMISATION ENVIRONMENT

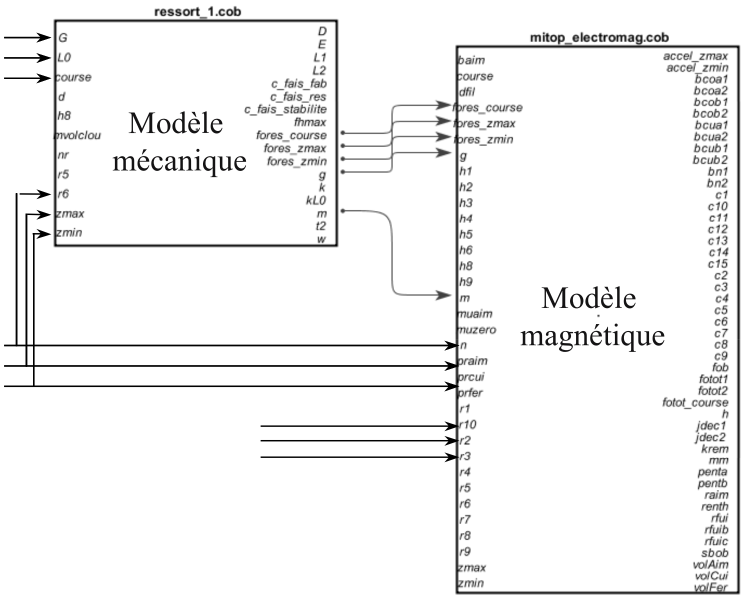


Figure 3. Coupled multidisciplinary optimisation

As opposed to handing over the results of an isolated study, an approach involving several elements of physics was required. In order to construct a coupled model to be computed in the existing optimization environment at the LEG, the mechanical team had to set up a group of equations describing the spring behaviour. Then the teams had to define common parameters between the two trade models, to define constraints to the parameters, and to create links between the common parameters. Once the overall analytical model is built (Fig. 3), the constraints and initial values have to be placed on the parameters. After this has been performed, the optimization engine carries out the iterative task of dimensioning the spring, whilst simultaneously optimizing the overall electromagnetic structure.

Table 1. Optimisation iterations and quality results

Optimisation	Number of iterations	Objective Function Value
No optimisation	6	No convergence
Continuous values	24	7931.98
Discrete values	422	8648,68

The first results were disappointing because the spring had but two coils, and the diameter of the wire was too thick. The mechanical result was good, but the technological one was not. So we added technological constraints to our model. Afterwards, realizing the importance of standard components in mechanical engineering [12], the optimization was carried out for discrete values corresponding to the real characteristics of springs available from one manufacturer [Tab. 1]. Although the discrete experiment did not incorporate the cost of the spring, this can be done in order to increase the supplier's involvement in the product development process. This approach was successful in providing a technical support to solve our very coupled problem.

4.1 A view of a collaborative optimization. Environment. Integration with CoDISS II

The study of the spring, linked with the electromagnetic plunger overall structure optimization, clearly demonstrates the advantages of developing a multi-trade optimization platform.

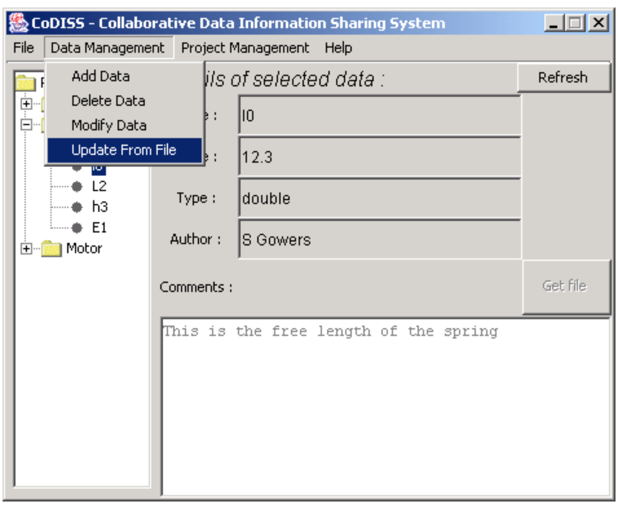


Figure 4. CoDiss2

In order to integrate this concept in collaborative distributed design, we have to support it in a distributed environment where you have to understand and comply with constraints imposed by other designers, more often than not from a different culture. In order to do this, we integrated the optimization process in our CoDiss2 (Collaborative Data and Information Sharing System) tool (see fig. 4).

The main objective of the CoDISS II [1] project is to create a framework supporting collaborative design with a common space of shared information, linked to several trades' interfaces. People use their ordinary software in conjunction with CoDISS II which supports the link between them. These links provide support for communication (both formal and informal) and also a mechanism to create "live" relations between the trade software's inner parameters. So the multi-trade optimization aspects studied in this paper are one of the aspects of this collaborative design environment

```
<Optimization>

      <Iterations Number="23">
         <Iteration Number="0">
            <Inputs Number="34">
               <Input Name="nr" Value="17.0"/>
               <Input Name="muzero" Value="1.2566E-6"/>
               ……… tous les autres 'inputs' pour la première itération
               </Inputs>
            <Outputs Number="59">
               <Output Name="c13" Value="24.0"/>
               <Output Name="c12" Value="2.9999999999999982"/>
               ……… tous les autres 'outputs' pour la première itération
            </Outputs>
         </Iteration>
         <Iteration Number="1">
            ………
         </Iteration>
         ………toutes les autres itérations
       </Iterations>

</Optimization>
```

Figure 5. XML formalization of the parameters exchanged between the trades

The implementation of a new function allows parameters within CoDISS II to be updated automatically using the XML results file generated by CDI Optimizer (Fig. 5). The designers are always aware of the last version of the value of the set of parameters they are interested in, and are able to submit a modification to all other designers. This creates a mutual awareness of the problem.

Furthermore, the CoDISS II mechanism will not only provide coherence in data, but will also allow constraints already exchanged during up-stream phases to be used for optimization. Eventually, it will provide a means of diffusing optimization results for individual analysis by a designer.

5. CONCLUSIONS

In carrying out this specific optimization study, it emerges that analytical optimization environments are often used as a method of giving dimensions to an initial structure. Segregated approaches between mechanical and electrical engineers during this stage may be ineffective. The very act of defining a component within a system puts constraints on the system from the component definition itself [13]. By adopting an integrated optimization approach, the optimization engine supports the iterative embodiment design tasks, thus avoiding long development loops between designers, and providing an effective way to find an acceptable set of values corresponding to the product requirements. It can also be quickly identified if common solutions do not exist, favouring trends towards dynamic construction of Product Development Specifications. As products become more complicated, the advantages offered by an integrated optimization approach still increase. Equally, if companies manufacture variants of the same products it is possible to capitalize on the time invested in the development of a composed analytical model.

Collaborative optimization represents an effective method of carrying out parts of the embodiment stages of an overall design project. The fact that designers are forced to discuss the optimization specifications is another benefit because this also means that collaborative optimization represents a platform for formal discussions of differing constraints imposed on design parameters by different engineers.

The design experiment has also shown the need to support collaborative problem formalization, constraints and interval negotiations. We have to enrich informal communication tools in a collaborative environment to better support these preliminary aspects of the optimization problem construction. Optimization only works on formalized models. Thus a new overall design experiment should try to answer the following questions:

- To what extent can mechanical considerations be incorporated into the analytical environment?
- Is it necessary to develop new tools in order to ensure data coherence?
- Once a composed analytical model is created, how easy would it be to adapt it to other variants of the product (thus capitalizing on the time invested at this stage)?
- Are there domains, other than mechanical and electrical, where this approach could be relevant? The development of Micro-Electro-Mechanical products (MEMS) is an area that also requires close co-operation between mechanical and electrical engineers, and may take benefit from this work.

ACKNOWLEDGEMENT

This article is based on the work carried out during the final year research project of Steven Gowers from the University of Bath, working under the supervision of Steve Culley and Daniel Brissaud. We would also like to thank Mark Nixon for his help.

REFERENCES

1. V. Riboulet, P. Marin, Towards a new set of tools for a collaborative design environment, Proc. of 7th Int. Conf. on Computer Supported Cooperative Work in Design CSCWD2002, pp.128-133, Sept. 25-27 2002, Rio de Janeiro, Brazil.
2. G. Hofstede, Culture and organization: The software of the mind. New York: McGraw Hill. 1991
3. J. Ulijn, and M. Weggeman, Towards an Innovation Culture: What are its national, corporate, marketing and engineering aspects, some experimental evidence, to be included in: C. Cooper, S. Cartwright and C. Early (London School of Economics and UMIST Business School, Manchester, Eds.), Handbook of Organisational Culture and Climate, London: Wiley. 2001
4. K. Schmidt, L. Bannon, Taking CSCW Seriously: Supporting articulation work, Computer Supported Cooperative Work, vol. 1, n^{os} 1-2, 1992, pp 7-40
5. Xiao, H-J Choi, R Kulkani, J K Allen, D Rosen, F Mistree, A web based distributed product realization environment, DETC2001/CIE-21766
6. V. Riboulet, P.R. Marin, J-C. Leon, Distributed Design – Current Approaches And Evolution Tendencies. Designing with Internet, in Methods and Tools for Co-operative and Integrated Design, S. TICHKIEWITCH and D. BRISSAUD Ed., 2003 Kuwler Academic Publishers, pp. 411-424
7. J. Ulijn, M. Weggeman. Towards an Innovation Culture: What are its national, corporate, marketing and engineering apsects, some experimental evidence, In: C. Cooper, S. Cartwright and C. Early (London School of Economics and UMIST Business School, Manchester, eds.), Handbook of Organisational Culture and Climate, London: Wiley
8. S. Mer, Les Mondes et les Outils de la Conception, pour une Approche socio-technique de la Conception de Produit, PhD thesis, INPG, 1998
9. T. Persun, (Contributing Editor). Online engineering collaboration. Design News. May 15, 2000.
10. B. Delinchant, V. Riboulet, L. Gerbaud, P. Marin, F. Noël, F. Wurtz, E-Cooperative Design Among Mechanical and Electrical Engineers: Implications for Communication Between Various Professional Cultures. IEEE Transactions on Professional Communication, vol.45, no. 4, Dec 2002.
11. E. Chapa, Outils et structure pour la coopération formelle et informelle dans un contexte de conception holonique. Thèse INPG , Oct 1997
12. B.J. Hicks, S.J. Culley, An integrated modelling environment for the embodiment of mechanical systems. CAD Journal. Elsevier Science Ltd. 2002
13. S. Pugh, Total Design - Integrated Methods for Successful Product Engineering. Addison-Wesley Publishers Ltd, Wokingham, England. 1991.

SPECIFICATION OF PRODUCT MODELLING CONCEPTS DEDICATED TO INFORMATION SHARING IN A COLLABORATIVE DESIGN CONTEXT

Frédéric Noël[1], Lionel Roucoules[2], and Denis Teissandier[3]
[1]3S, Université de Grenoble, Domaine Universitaire - BP n° 53 - 38041 Grenoble Cedex 9 – France ;
[2]LASMIS, Université Technologique de Troyes, 12 rue Marie Curie – BP 2060 - 10010 Troyes cedex, France
[3]LMP, Université Bordeaux 1, 351, cours de La libération 33405 - Talence Cedex, France

Abstract: Collaborative design efficiency expects to go further than concurrent engineering in breaking down sequential organizations: in concurrent engineering experts must act on the product as soon as possible but their actions constrain the next ones. To collaborate, the experts must share these constraints with a common language about the product. Here, a new product model dedicated to collaborative design is proposed. Some examples and usage sequences highlight the main benefits in a collaborative context. The model is used in the French IPPOP project (Integration of Product Process and Organization for Performance Enhancement in Engineering) where a framework dedicated to innovative collaborative design is under construction.

Key words: Concurrent engineering, Product model, Collaborative design, Multi-expertise

1. INTRODUCTION

CAD systems are mainly based on geometric models [1-2]. The current trend is to attach more and more technological information to form features, but these geometric models remain centralized and imply to be fully defined before any other expertise starts. Design activity remains sequential and is still led by geometric modelling. In the last decade, technological product

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 135-146.

models [3-8] were proposed to enlarge the scope of CAD system. The design of a new product is based on a very large set of computer services. Dedicated skilled services use a geometric model as input, complete this model with technological information, assist experts in the achievement of their task but do not help them to propose design evolution. Expertise mainly depends on expert knowledge. Sometimes, engineers configure generic services to encode their own knowledge. Spreadsheets, databases, mathematical solvers, etc, are used for this purpose but are hardly shared.

The model concepts may be the same for a couple of services. However, the model encoding is specific to each service. Each service has its own storage format, which is rarely compatible with others. Standards were developed to ease the translation between services but they are sources of data loss and of information duplication. PDM provides a way to handle the large amount of files created during the design process. It provides a repository where designers share files defining the product. However each file is understood by a few services. Moreover atoms of information inside a file are not available as sharable items. Engineers use PDM to publish information but not to collaborate. The STEP standard [9-10] was offered as an integrated model to support all the product information. However too few applications are proposed to share STEP objects in a dynamic and collaborative mode [11]. Moreover the evolution of STEP is not quick enough to deal with new concepts of new emerging technologies.

This paper proposes a new product model based on a small number of concepts to be easily used during collaborative design. The number of concepts is minimized to offer a very comprehensive language. This product model allows derivation of several views of the product and linking of these various definitions. Since few concepts are introduced, this product model can be extended with representations issued from other product models, like STEP representations. Thus it completes other product models to support collaborative design. The third section describes some examples of instantiation of this model while the final section describes how this product model can be shared in a collaborative design activity.

2. A NEW PRODUCT MODEL

2.1 A product model dedicated to collaboration in innovative design

A unique language for collaboration around product definition [12-13] is proposed. Many product models have already been defined in the literature.

They are more or less complex and specific to some design area. Even if STEP should be considered as the basis of a common product model, STEP Application Protocols are dedicated to specific fields or industries. Here, a first requirement is to provide a product model generic enough to support collaboration for any technical expertise. A second requirement expects this language to be simple enough for innovative design where the main structures, functions and behaviours of the product are not established at the beginning of the design activity. As a consequence, the product model must be defined on a small numbers of concepts.

2.2 Multiple axis definition of the product

Since the product model is dedicated to collaborative design, it must support the link between the various representations of the product handled by designers. Four axes to decline the product definition were integrated in the current product model:

1. **Service representations**: each computer service defines its own model encoding according to the managed theory but also of technical choices depending on the service developer.
2. **Various levels of detail** must be described in the model to allow the collaboration to focus on any detail respect to the undertaken problem.
3. **Various cultural representations** reflect various view points about the product. Within an electro-mechanical component the structure decomposition depends on the expert analysing the product: An electrician models the gaps between parts while mechanical analyst does not mind about these. They will not use the same decomposition.
4. The product life-cycle from the earliest design stage to the final recycling process involves **various states for its sub-systems**. Each sub-system may be defined at this various stages. Obviously, if several stages are considered during the design activity, then the various representations associated with the product evolution should be linked together to follow this view of the life cycle.

The product model proposed in the next sub-section intends to present and store in a same model every declination axis.

2.3 Main concepts

Since the product model must be used by many designers whatever their culture and know-how, we proposed to reduce its internal concepts to a minimal set of concepts presented in an UML class diagram on Fig. 1.

1. **Component**: a component is a partition of the product. A component can be subdivided into a set of sub-components. There are three kinds of components according to the semantic of their subdivision 1) Common Components (CC) that define a real partition of the part: – 2) Alternative Components (AC) are decomposed into optional sub-components: an alternative component can be substituted in any of its children – 3). View components (VC) are decomposed into several points of view: each sub-component is a correct representation of the product but it is attached to a specific culture shared by a sub-set of designers.

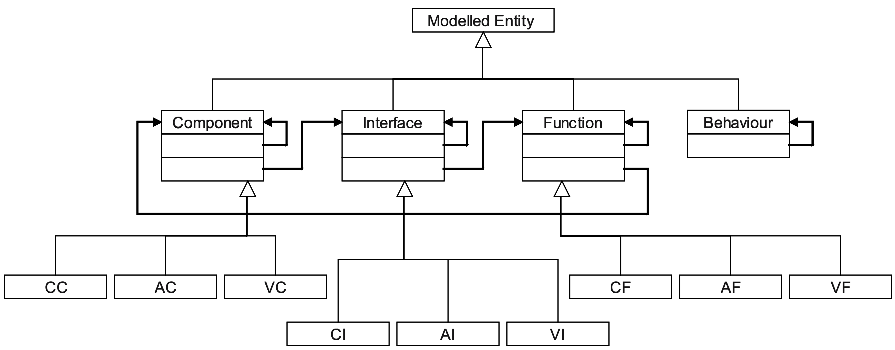
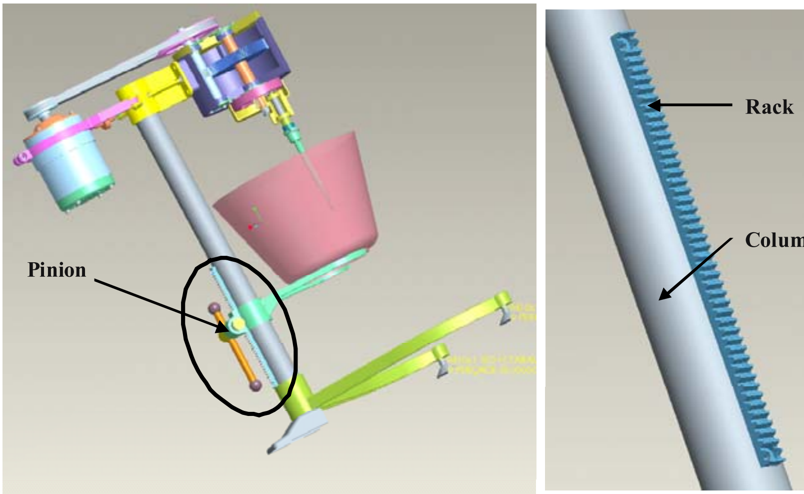


Figure 1. Product Model Class Diagram

2. **Interface**: an interface is a handler of a component, i.e.: a property by which the component may be linked to an interface of another component. A component may have mechanical, electrical, etc. handlers. Like components there are three kinds of interfaces 1) Common interfaces (CI) 2) Alternative interfaces (AI) 3) View interfaces (VI). Interfaces can also be subdivided into sub-interfaces to detail their definition.
3. **Function**: a function links at least two components through interfaces. It is a relation between the components. The function defines an objective to be achieved within some threshold defined by a criterion and a goal value for this criterion. Like components there are three kinds of functions 1) Common functions (CF) 2) Alternative functions (AF) 3) View functions (VF). For the definition of detail, functions are subdivided into sub-functions.
4. **Behaviour**: Behaviour defines a modal state inside the lifecycle of the product. Behaviour is defined by a set of components, interfaces and functions.

All these concepts define the main characteristics of the product. We assume that they are generic and simple enough to be understood and used by any designer. The model of the product is a data-base stored definition of these features and the relationships between them.



Mixing system **Rack-pinion detail**

Figure 2. A mixing machine and its rack-pinion detail.

Special representations may be associated to the previous feature. A large range of special types of value should be defined such as length, volume, images, and any other characteristic of interest as the location of a file url issued from any external service. Thus a part declared as a component in the product model can be associated with various external files modelling the part in heterogeneous software services.

3. EXAMPLE OF INSTANTIATION

3.1 A simple assembly: introduction

A mixing machine is used to illustrate the instantiation of the model. Fig. 2 shows the global assembly and focuses on a sub-system: the vertical column and the rack. Two alternative design solutions are considered: the column and the rack are bonded or screwed. Fig. 3 presents three trees corresponding to the structure, the interface, and the function tree. In the structure tree, the component named assembly is decomposed into three sub-components: the column, the fixed part and a component depending on a technological choice. If the parts are bonded together, the glue will be the third component whereas if the parts are screwed together two screws will be added in the system. The structure features are linked to their interface while interfaces are associated to the function. Finally, functions are linked with components fulfilling the function.

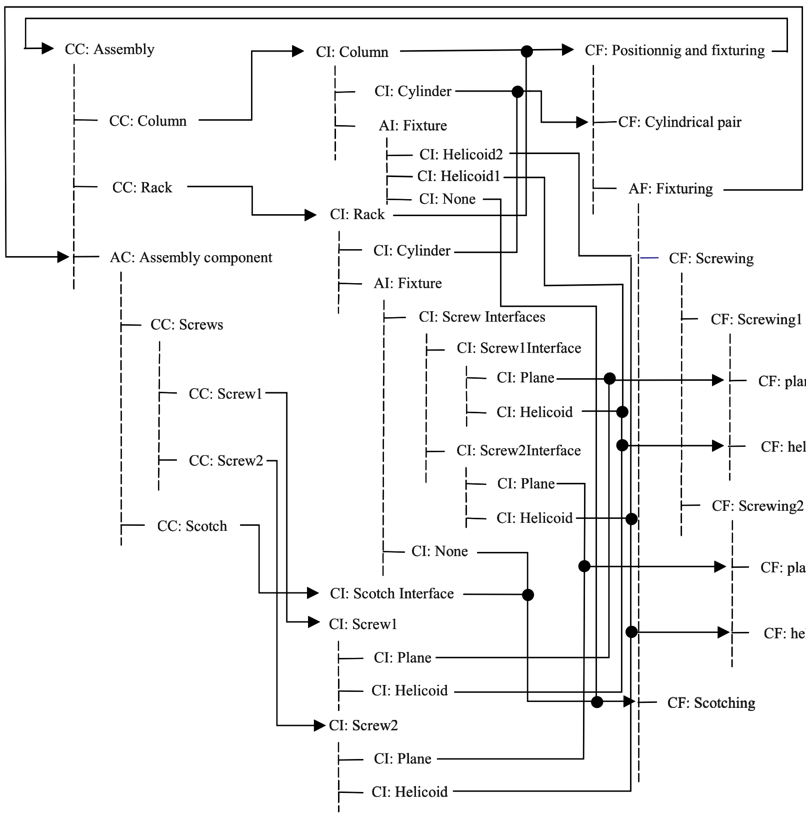


Figure 3. Model for the assembly of the column and the rack.

3.2 A geometric tolerancing cultural representation

In this section, only the bonded solution is considered. A geometric tolerancing tool manipulates three sources of essential information: 1) the standardized specifications. A standardized specification is a set of conditions on situation characteristics between elements of the same part and/or a set of conditions on intrinsic characteristics of an element (an element can be a face, a line or a point) 2) the contact specifications between the parts of an assembly. A contact specification is a set of mating relations between the parts of the product 3) the geometric condition of an assembly. A geometric condition is a set of conditions on situation characteristics between elements of the product and/or a set of conditions on intrinsic characteristics of the product.

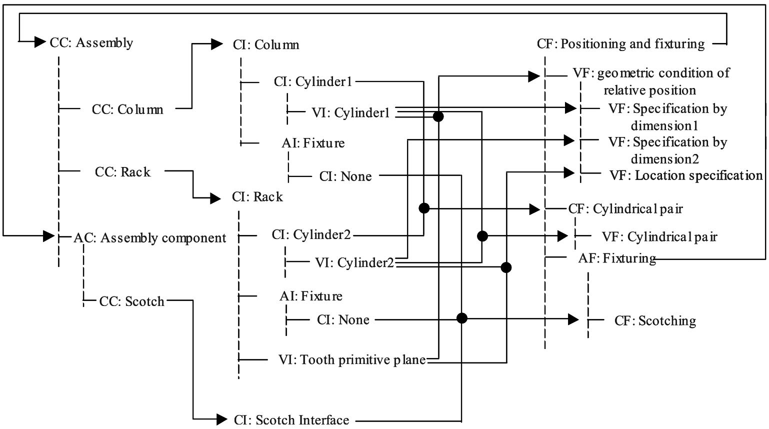


Figure 4. Geometric Tolerancing viewpoint of the assembly of the column and the rack

A view interface of the product model defines an element. In Fig. 3, a common function defines only the number of degrees of freedom between components. A view function (Fig. 3 and 4) may be a geometric specification by tolerance zone, a contact specification or a geometric condition. A specification by dimension defines a set of conditions on attributes of a view interface i.e. elements. Fig. 4 presents the 3D-dimension chains to ensure a geometric condition of the common function "Positioning and Fixturing". Fig. 5 gives an illustration of a surface graph as in [14] and [15] extracted from the instantiation of the product model from Fig. 4. The location specification and the dimension specifications according to envelope requirement give a result from the transfer of the common function "Positioning and fixturing" onto the parts column and rack: see Fig. 4 and 5. The 3D-dimensional dimension chains can be formalized in the surface graph [16] as well as in the product model. Current geometric tolerancing tools simulate the manufacturing defects of parts but they do not have a great influence on the design process. With the current product model, the trace of geometric specifications is available. If a geometric specification is not satisfied, the unsatisfied functions are identified. The integration of standard components (bearing, screw, nut, etc) and their impact on geometric specifications is taken into account and is linked with the choice of these components.

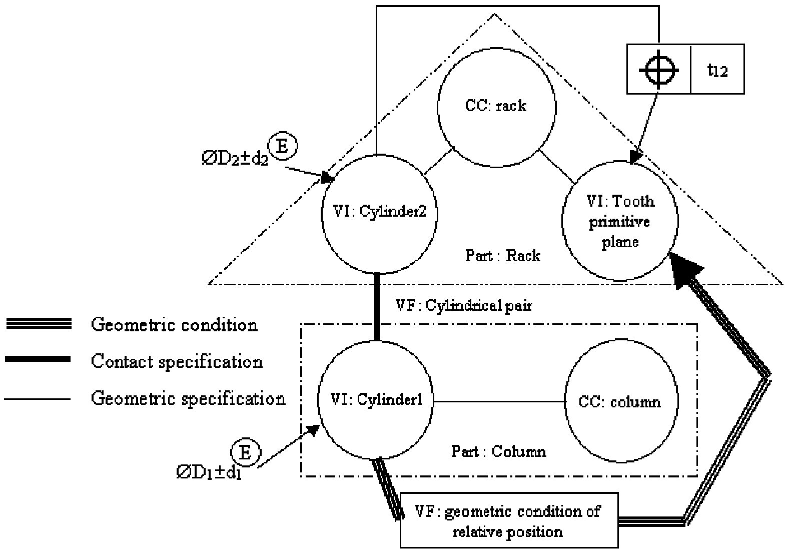


Figure 5. A 3D-dimensional chain in the product model for the assembly of the column and the rack according to a geometric tolerancing point of view.

4. SCENARIO OF COLLABORATION

4.1 General usage sequences

The next usage sequence presents the collaboration steps for several designers sharing the product model:

1. Each actor joins the project (product data). At this stage they all work separately: this involves a) Identification with a login and a password. b) Reception of the personal task list. Information required for each task (product data, objectives of the tasks, the human and material resources) is issued from the design process model [17].
2. The expert seeks if the product data already exists in the product model: if a) the data was already shared among the design group, then the expert finds the data in the shared database – or else if b) the data was not shared among project group. The expert will have to require information – or else c) the data exists but the expert does not find it. In fact, even ways to search for information (keywords, filters, dictionary, etc.) may

fail. The data could be defined from different points of view, different ontologies, etc.

3. The expert requires some data. So far, two ways of indicating requirement are proposed: 1) using a formal call mechanism developed in the IPPOP project, and 2) using a usual way such as e-mail.
4. Other members of the design group share the required information via the product model. Then the initial expert accesses this information.
5. The expert processes his product assessment. He uses his own model and services. At this stage he does not share his results that should only be declared in the product model when collaboration is needed (goto items 3 and 4): if expected by other experts, the expert shares his results via the product model.

The expert waits for new notifications: 1) Notification of a new task (goto item 2); 2) Notification of a modification on an already declared task (goto item 2); 3) Notification of a call for sharing product data (goto item 4).

4.2 Illustration on the rack-column sub-system

This example deals with a rack and pinion system. This system is part of the mixing machine shown in Fig. 2. Fig. 3 shows the initial stage of the product model in order to begin the illustration. At this stage, the shared information concerns the function and the mapped technologies that have been selected: 1) Function: to set the position and to fix the rack on the column; 2) Technology: two screws or bonded as an alternative; 3). Interfaces physically represent functional surfaces that transmit mechanical flows to fix the rack. Both functions and technologies are shared among the design group to provide a main common product definition to every expert.

Table 1. Various needs expected by mechanical analysisi and manufacturing expertise

Mechanical analysis needs	Manufacturing expertise needs
The form features of the rack	The form features of the rack
The material characteristics	The material characteristics
The loading values on the rack	

At this stage, this illustration presents how two experts connect the database in order to assess the product, to create new information and maybe to share this information with others. On the one hand the “mechanical analysis” expert carries out the structural analysis, on the other hand the “manufacturing” expert proposes what the manufacturing process for the rack could be.

Table 1 enumerates some basic needs required by the two areas of expertise. Currently, this information does not yet exist in the rack model. Each expert requires it (sequence item 4). Other members of the design group share geometrical, material and loading information by associating attributes to corresponding rack components: 1) A geometric representation is presented as a CATIA file. Only the path to this file is referenced in the component. 2) A vector that defines the loading of external components on the rack. 3) The material name that should be delivered by a database to reference all the characteristics of this material (Young's Modulus, elastic stress limit, etc.).

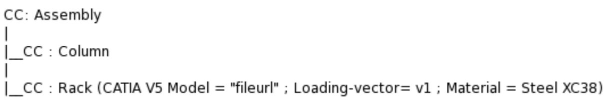


Figure 6. Shared information issued from the experts'call.

This information can be either totally specified or still free for modification or even incomplete. The two experts with this extra information assess the rack (sequence item 6): 1) The first one finds that the tensile stress (300 MPa) is higher than the elastic stress limit of the material (200 MPa). The rack will then fail – 2) The second one proposes to drill the screw holes using a conventional High Speed Steel drilling tool and thus conventional cutting speed (200 m/min)

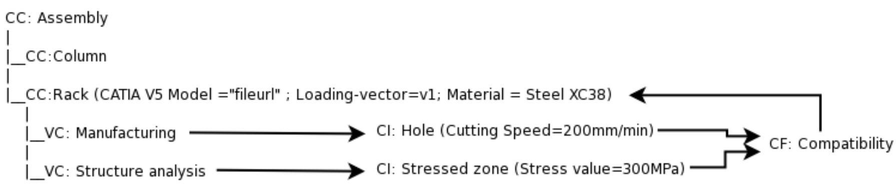


Figure 7. Rack model with new shared information

Information issued from both mechanical and manufacturing assessments are shared with already defined information, issued from the first call. The link between the two experts is not direct but specified via material compatibility identified by the corresponding function (Fig. 7). The notification of any information evolution (deletion, modification) can thus be propagated to each expert. In the current illustration, the "mechanical analysis" expert now specifies and justifies a material change to tempered to allow the loads to be borne. The "manufacturing" expert via interfaces and functions is then notified. The collaboration among them can be set.

5. CONCLUSION AND RECOMMENDATIONS

A new product model to share design information in a collaborative context was introduced. This model provides new concepts dedicated to information sharing and does not aim at competing with other existing product models. This product model was illustrated and demonstrated with examples. Usage sequences consolidate the information managed in the product model and the collaborative aspect via a shared environment. Sequences also assist the specification of the Graphic User's Interfaces that will be generated.

Presenting several product trees, the paper demonstrates on the one hand how a single expert can share his product data using multiple levels of representation. Those levels are necessary to manage data linked to the whole design process. The non shared information is kept in a local environment in order to avoid overload of the product database. On the other hand, data in the shared environment can be linked via Interfaces and Functions. Those links support the notification of changes (creation, deletion and modification) to respective experts. Further work will provide a computer-supported demonstrator to prove: 1). Multiple services connections to the model. A Direct Application Programming Interface would be developed. 2). Notifications mechanisms to experts. A list of messages could be tested to provide information in sequence 2.

ACKNOWLEDGEMENTS

The authors of this paper are participating to the French national project IPPOP (Integration of Product, Process and Organisation for Performance Enhancement in engineering). The product model proposed in this paper will be used in this project to support collaboration around the product definition. We would like to especially acknowledge G. Debarbouillé from Open-Cascade Company for his help in the discussions around this product model.

REFERENCES

1. T. De Martino, B. Falcidiano, F. Giannini, Feature-Based Modeling as a solution for Concurrent Engineering Design, CIMAT 96, Grenoble, 1996.
2. J. Shah., Assessment of Feature Technology, Computer Aided Design, vol. 23, n°5, June 1991.

3. S. Tichkiewitch, Specification on integrated design methodology using a multi-view product model, ESDA Proceedings of the 1996 ASME System Design and Analysis Conference, PD-Vol. 80, 1996.
4. L. Roucoules, S. Tichkiewitch, CODEMO: A Cooperative Design Environment - A new Generation of CAD Systems, Concurrent engineering: research and applications, vol:8 No 4 pp.263-280, 2000.
5. T. Kjelberg, H. Scmelkel, Product Modelling and Information Integrated Engineering Systems, Annals of the CIRP, vol. 41, n°1, pp 201-204, 1992.
6. F-L. Krause, F. Kimura, T. Kjelberg, S.C-Y. Lu, Product modelling, Annals of the CIRP, vol. 42, n°2, pp 695-706, 1993.
7. F.-L Krause,T. Kieswetter, S. Kramer, Distributed Product Design, Annals of the CIRP, vol 43-1-pp.149-152,1994"
8. R. Anderl, R. Mendgen, Modelling with constraints: theoretical foundation and applications, Computer Aided Design, Vol. 28, n°3, pp 155-166, 1996.
9. S. Arbouy, A. Bezos, A.-F Cutting-Decelle, P. Diakonoff, P. Germain Lacour, J-P. Letouzey and C. Viel, STEP, published by AFNOR, 1994.
10. T. Männisto, H. Peltonen, A. Martio, R. Sulonen, Modelling generic product structures in STEP, Computer Aided Design, Vol. 30, n°14, pp 1111-1118, 1998.
11. S.D. Urban, K. Ayyaswamy, L. Fu, J.J. Shah, L. Liang, Integrated product data environment: data sharing across diverse engineering applications, Int. J. Computer Integrated Manufacturing, Vol.12, n°6, 525-540, 1999.
12. F. Noël, D. Brissaud, Dynamic Data Sharing in a Collaborative Design Environment, Int. J. Computer Integrated Manufacturing, Vol. 7-8, pp. 546-556, 2003.
13. F. Noël, D. Brissaud, S. Tichkiewitch, Integrative Design Environment to improve collaboration between various experts, Annals of the CIRP, august 2003.
14. D. Teissandier, Y. Couétard, A. Gérard, A Computer Aided Tolerancing Model: Proportioned Assemblies Clearance Volume, Computer Aided Design, Vol. 31, pp. 805-817, 1999.
15. D. Teissandier, J. Dufaure; Tolerance synthesis with polytopes: Application to radial ball bearing arrangements; In: IDMME, Clermont ferrand, 2002
16. J. Dufaure, D. Teissandier; Geometric tolerancing from conceptual to detail design; 8th Computer Aided Tolerancing CIRP seminar, Charlotte (north carolina, USA), 2003.
17. P. Nowak, C. Merlo, B. Eynard, T. Gallet, From Design process specification towards PDM workflow configuration, in proc. of ICED 03, Stockholm, August 19-21, 2003.

ANALYSING DESIGN DECISIONS FROM A PRODUCT AND PROCESS PERSPECTIVE

Raymond Holt and Catherine Barnes
School of Mechanical Engineering, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT

Abstract: Most approaches to integrated product and process design focus on decision *support* – providing designers with information about the implications of their choices for different stages of the product life cycle. When the needs of different life cycle stages conflict, however, these techniques provide no basis for determining which should take precedence. This paper proposes a methodology for helping designers assess trade-offs between the competing needs to minimize cost and maximize quality, by drawing on principles from decision analysis. Research on decision analysis in design tends to focus on the mathematics of decision-making, but this methodology takes the view that the value of decision analysis is in the *insights,* not the numbers, it generates. It is argued that decision analysis should not be a normative tool for making optimal choices, but a framework for systematic discussion of important life cycle issues. This paper discusses the implications of manufacturing for design decisions, and proposes four conditions that a systematic approach to design decision-making should address to take these implications into account. The methodology is described, and illustrated with a case study. The paper ends by discussing the future work needed to validate and refine the methodology.

Key words: Design decision-making, Design for manufacture, Decision analysis

Every aspect of a product is determined by the decisions taken in its design, and those decisions have implications for every subsequent stage of the product life cycle. By failing to consider these implications, designers can inadvertently make choices that lead to expensive or low-quality products, or require subsequent alteration. Current Integrated Product and Process Design (IPPD) methods – design guidelines, software packages and cross-functional teams – provide information about manufacturing implications, but no basis for assessing trade-offs when functional performance and ease of manufacture conflict. These techniques provide *decision support,* concerned only with giving designers the information they

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 147-158.

need to make a decision. How the information is used is the domain of *decision analysis*. This paper presents a methodology for analysing design decisions from product *and* process perspectives, by combining conventional IPPD decision support with principles of decision analysis.

1. DECISION ANALYSIS AND INTEGRATED PRODUCT AND PROCESS DESIGN

Many researchers have investigated the use of decision analysis techniques in design. This section considers the approaches to decision analysis currently taken in design, and how decision analysis can be used to accommodate manufacturing concerns when taking design decisions.

1.1 Analysing Design Decisions

Decision analysis has long been used in design. Pahl and Beitz [1] and Pugh [2] both advocate simple techniques such as the weighted objectives tree, and controlled convergence. Other design researchers have shown interest in more formal techniques, such as Utility Theory [3, 4], Outranking [5] and the Analytical Hierarchy Process [6]. These reduce decision-making to a mathematical problem, with the emphasis on rational formulae for aggregating preferences and expectations numerically. This approach has two serious problems. The best known is Arrow's Impossibility Theorem [7], which demonstrates that there is no mathematically consistent way of aggregating the preferences of multiple individuals. To get around this, some researchers advocate the use of Game Theory [8, 9] – which models decision-making under conflict – as a way of reconciling the different needs of product and process design. This addresses the mathematical issue, but not the more fundamental problem: the importance of social interaction in design, highlighted by Bucciarelli [10]. Designers must negotiate a shared understanding of what they are trying to achieve, a process that cannot be substituted with a mathematical algorithm. This does not mean that mathematical decision analysis techniques are without value, but their value comes from the systematic basis they provide for analysing a decision situation, and the insights and discussions that arise from this process.

They function as *requisite decision models* [11]: instead of calculating an "optimum" compromise between product and process, designers collaborate to build a model they can all subscribe to. In fact, the *principle of the flat maximum* [12] suggests that regardless of the mathematical technique used, the "best" alternatives will tend to be rated above the "worst". It does not matter which technique is used, as they all encourage systematic thinking.

No methodology can provide the "correct" algorithm for considering both product and process perspectives when taking design decisions. Instead, the emphasis has to be on providing a *systematic* method for accommodating manufacturing concerns in design decisions. This means understanding how manufacturing relates to product performance.

1.2 Design Requirements and Manufacturing

Design is an intentional activity, undertaken to satisfy a perceived need, be it to overcome an existing problem, or to exploit a gap in the market. In design methodology, this need is interpreted as a formal requirements specification [1, 2], which forms the basis of all subsequent activities. The requirements specification, therefore, should be the basis of every design decision. The only manufacturing issue that designers need to consider is how manufacturing the product will imp act on its ability to satisfy the specified requirements. Requirements can be divided into two classes. *Functional* requirements measure a design's fitness for purpose: they define the conditions that the design must meet to satisfy the need that initiated the design process. There are also constraints upon the time and money that can be invested in making the product, represented by *Economic* requirements. Manufacturing impacts each class of requirements in a different way. A product's behaviour is determined entirely by its shape and material properties: two identical objects will function in exactly the same way, regardless of how they were made. So, when assessing a candidate design against functional requirements, *how* it was made is irrelevant. The only question is whether the specified design *can* be made. Conversely, manufacturing contributes directly to the cost and time of making a product.

To compare a design against economic requirements, designers need some idea of *how* it can be made. However, the range of processes available, and the development of solid freeform fabrication techniques in particular, mean that almost any design *could* be produced, given enough time and money. More pertinent is whether the proposed design can be made with existing production facilities. If not, the designers know that they must either abandon that design, or include the capital cost of investing in new processes when assessing the proposal's performance against economic requirements. This suggests that, to accommodate manufacturing concerns in their decisions, designers must: 1) identify the processes currently available for making that product; 2) consider what is feasible given those processes; 3) include the cost and time of manufacture when estimating a design's performance against economic requirements; and 4) make any trade-offs on the basis of the requirements specification. The next section proposes a methodology that takes designers through a series of activities that systematically satisfy these conditions, using decision analysis as a framework for exploring the decision.

2. METHODOLOGY

The methodology presented here only addresses choices about the form and substance of a product: it does not address process selection. Rather, it encourages designers to reflect on which issues are important for the product, and how its manufacture affects them. The methodology has four stages, as illustrated in Fig. 1. The *preliminaries* and *framing* set the decision context, preparing the information in the requirements specification for the subsequent stages. *Validation* compares candidate designs against the available processes, to help designers understand which are feasible, and how they can be made. *Evaluation* uses existing decision analysis techniques to compare feasible candidates against the requirements specification.

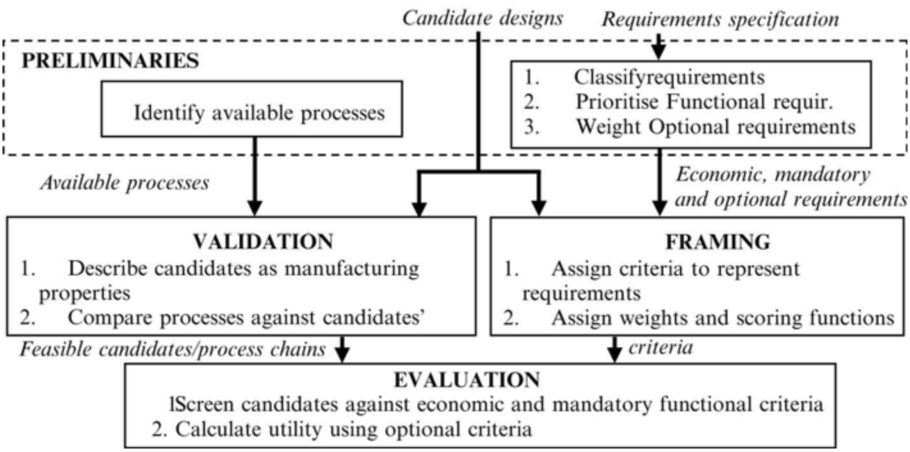


Figure 1. Main stages of the methodology

2.1 Preliminaries

The preliminaries lay the groundwork for all subsequent decisions. These activities encourage designers to think about the manufacturing processes currently at their disposal, the relative importance of each requirement, and how manufacturing affects them. The designers must first determine which requirements are *economic*, and which are *functional*. Economic requirements represent constraints, and *must* be satisfied. For functional requirements, designers must identify which are *mandatory*, and must be satisfied, and which are *optional* and can be traded-off (often referred to as *demands* and *wishes*, respectively [1, 2]). To quantify trade-offs between the optional requirements in evaluation, numeric weights must be assigned to indicate their relative importance. The exact method for assigning weights depends upon the decision analysis technique used for ranking (see §2.4), but larger weights always indicate greater importance. This provides the basic information needed for evaluation, but it is often difficult to assess whether an incomplete design will satisfy a given requirement, as this depends on subsequent choices [13]. For example, the maximum load of a structural member cannot be established from the choice of material: it also depends upon choices in relation to its shape. Measures that meaningfully compare the candidates in the current decision are needed, but can only be established once those candidates have been generated.

2.2 Framing

Framing develops surrogate *criteria*, relating the requirements to the candidates being considered. Where requirements apply to every decision about a product, each criterion is unique to just one. Each criterion relates to an attribute – a measurable property common to all the candidates, so that they can be compared – which influences the requirement's satisfaction. Sometimes a single requirement may be represented by several criteria; sometimes, the requirement itself will be suitable for comparing the current candidates, and no additional criteria are necessary. Criteria for optional requirements are used for ranking, and must be assigned weights to indicate their relative importance, and scoring functions, to relate each attribute level to a utility. Weights are assigned by dividing the requirement's weight among the criteria representing it. This done, the requirements specification can be used as a basis for evaluation, but the designers must first establish *which* candidates are feasible and *how* they can be made.

2.3 Validation

Validation analyses candidates from a process perspective, getting designers to think about the demands they are placing on the available processes. Where evaluation draws upon existing decision analysis techniques, Validation uses a new method called the feasibility test, described in §3. This draws on ideas from decision analysis to provide the same structured approach to comparing available solutions (the process) against requirements (the properties of the candidate designs). First, the designers must describe each candidate as a set of properties, representing the demands it places upon the processes used to make it. This stage only considers candidates at the component level, and how each component will be made – how the components are assembled is beyond the scope of this work. It does not matter how detailed the candidate is – the designers simply define the properties established at the time of the decision. No CAD model or formal drawings are needed. By identifying which processes can satisfy each property, the designers can construct chains of processes capable of making each candidate. If a design has at least one feasible process chain, then it is deemed feasible and passed on for evaluation. If not, the designers know that they must abandon the candidate; amend it; or investigate the cost of obtaining new processes.

2.4 Evaluation

Evaluation analyses candidates from a product perspective, encouraging designers to think about how well each candidate satisfies the requirements specification. This involves two stages. *Screening* eliminates any candidates that do not satisfy all the economic and mandatory functional requirements. When comparing candidates against economic requirements, the designers need to take account of the process chains identified in validation. If several process chains can make the candidate, the designers can either choose the most promising for economic estimates, or consider the implications of using different process chains. If *every* process chain for a candidate violates the economic requirements, the designers will know that the candidate cannot be produced given the available resources. If only one candidate survives screening, then there is no need to go any further – that candidate can be adopted. If several candidates survive, then the designers need to rank them. Any compensatory decision algorithm can be used for this purpose, in conjunction with the optional criteria, weights and scoring functions assigned in framing. In this research, SMART [13] has been adopted for its simplicity, but Multi-Attribute Utility Theory, or the Analytical Hierarchy Process, among others, could also be used without changing the methodology. The designers are under no obligation to base their choices on the numbers generated: the insights that emerge from applying these techniques are much more important.

3. THE FEASIBILITY TEST

The decision analysis techniques used in evaluation are well-established, and do not need to be examined here. Validation, however, draws upon a new method known as the Feasibility test, which this section describes. The feasibility test is *not* a method for process selection. Nor does it require detailed descriptions or CAD models – designers can describe each candidate in as much or as little detail as they have available. Instead, it provides an "early warning" so they can avoid choosing designs that, on reflection, they believe will be infeasible. A candidate deemed feasible now may prove infeasible when further details are added, but designers can only go off the information they have available. A candidate whose general design is deemed infeasible will not suddenly become feasible when more details are added. The feasibility test has three stages: assigning properties; testing compatibility between candidates and processes; and establishing feasible process chains for each candidate.

3.1 Assigning Properties

To assess a candidate's feasibility, it must be described in terms of five basic properties: material, shape, surface roughness, dimensions and tolerances. The candidate is first decomposed into its constituent components: as each component is manufactured separately, their feasibility can be assessed independently. Each component comprises a *basic component*, and a set of *features*. Features are not strictly defined: any area where properties deviate from those of the basic component - a hole, for example, or a face with a different surface finish – can be considered a feature. The aim is to get designers thinking about the candidate's properties, not to provide a perfect calculation of manufacturability. Both the basic component and its features are described in terms of the same five properties, with the exception of material: a feature cannot be a different material from the component it is part of. Fig. 2 illustrates an EXPRESS-G data model for capturing this information. Designers can assign as many, or as few of these properties as they are able: there is no need for the candidate to be fully defined. This done, they can begin comparing these properties against the capabilities of the available processes.

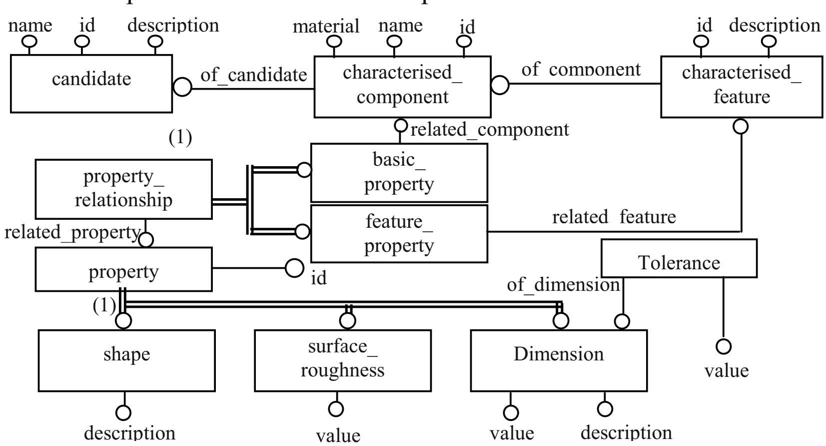


Figure 2. Data structure of design candidate properties used in validation

3.2 Process-Property Compatibility

Process capabilities are compared with candidate properties using a *compatibility matrix*, one for each component. This lists the available processes across the top, and the component properties down the side. Each cell in the matrix indicates whether the given process can satisfy the given property. If it can, the cell is left blank; if not, the cell is marked with a cross, to denote a violation. An example compatibility matrix can be found in §4. To determine whether a process satisfies or violates a property, designers must consult sources of decision support for process selection – PRIMAs [14] for example, or the engineers responsible for process planning. This often requires qualitative judgement. For each property, the designers have to ask themselves if there is any reason to believe that the process in question won't satisfy that property.

As well as the basic capabilities of the process – the materials and shapes it can handle, the surface roughness and accuracy it can provide – designers must consider two complicating factors. The first is process specific limitations: problems such as cold shuts, or hot tears, or thin walls crumbling in powder metallurgy. The other is dependency between properties – if a process violates one property, then it may violate others automatically. For example, if a process cannot produce a given shape, then it cannot produce any of that shape's dimensions. The dimension is meaningless without a shape to be attached to. Equally, if a process cannot make a given dimension, then it clearly cannot satisfy the associated tolerance. Fig. 3 illustrates the dependencies between properties. Machining and other material removal processes can be used to cut features into parts that have a basic shape that is beyond the process' capabilities. Where a process violates a property because of this precedence, the relevant cell can be shaded to highlight this. Once the designers have worked through each process in turn, and identified which properties it satisfies, they can then turn their attention to establishing which candidates are feasible, and how they can be made.

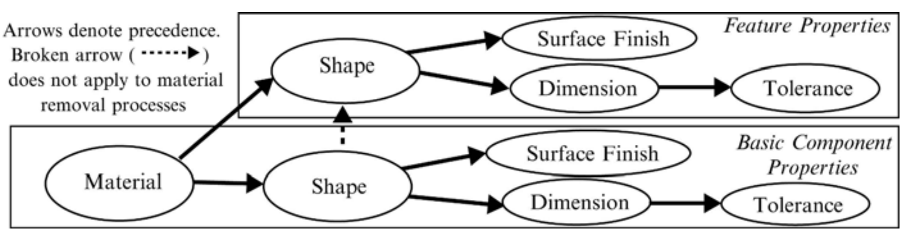


Figure 3. Dependencies between component properties in validation

3.3 Establishing Feasible Candidates and Process Chains

A product is normally made using a combination of processes known as a *process chain*. A process chain can make a given component if, for every property of the component and its features, there is at least one process in the chain that can satisfy it. A component is considered *feasible* if there is at least one process chain that satisfies all of its properties. A candidate is considered feasible if all of its components are feasible. The next section gives an example to illustrate the methods described here.

4. ILLUSTRATIVE EXAMPLE

The decision analysis techniques used in evaluation are already well-established, and can already be investigated more thoroughly in the available literature. This section aims to clarify the methods presented for the feasibility test, by demonstrating their use. The example given is deliberately simple – the purpose is to highlight and illustrate the mechanisms of the feasibility test, not to prove the feasibility of the component in question.

4.1 Scenario

The component considered is a counter flange for the inlet/outlet of a pump assembly. A simple candidate design for the component is shown in Fig. 4. Notice that not all aspects of the component have been determined at this stage. Some features are pre-determined – the size and location of bolt holes, and the diameter of the outlet hole, are determined by the equivalent characteristics of the pump. Four processes are available: sand casting, powder metallurgy, drills and lathes.

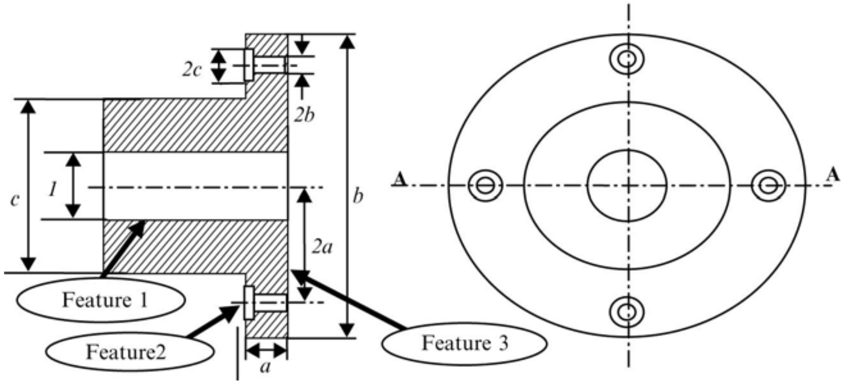


Figure 4. The Component design to be analysed

4.2 Assigning Properties

Three main features can be identified: the main inlet/outlet hole (Feature 1), a set of four bolt holes (Feature 2), and the mating face of the counter flange, which needs a smoother surface finish than the rest of the component. All other properties can be deemed part of the "basic component". This is not the only possible division of features for the part – the flange might be treated as a separate feature, for example, or the bolt holes might be treated as one feature each, but there is little to be gained by doing so. The properties identified for the basic component are shown in the Compatibility Matrix illustrated in Fig. 5.

4.3 Process-Property Compatibility

The design is assessed using Booker and Swift's PRIMAs [14]. Sand casting cannot produce the surface finishes and tighter tolerances needed for the features. The lathe cannot produce the bolt holes (feature 2), as they are not in line with the axis of rotation, and drilling is unable to handle the basic shape of the component, or the mating face (feature 3). On the strength of basic capabilities, powder metallurgy satisfies all the properties. In terms of process-specific limitations, sand casting requires a draft angle of 1 to 5°, but this can be added without interfering with the part's function. The sharp change in section between body and flange make it unsuitable for powder metallurgy, where marked changes in section should be avoided [14].

PROPERTY	VALUE/ DESCRIPTION	Sand casting	Powder Metallurgy	Lathe	Drill
BASIC PROPERTIES					
Material	Cast iron				
Shape	See Fig 4		X		X
Surface Roughness	12.5µmRa		X		X
Dimension *a*	13mm		X		X
Tolerance (on *a*)	±0.8mm		X		X
Dimension *b*	90mm		X		X
Dimension *c*	40mm		X		X
FEATURE 1: Thru' Hole					
Shape	See Fig 4		X		
Dimension *1a*	20mm		X		
Surface Roughness	1.6µmRa	X	X		
FEATURE 2: Bolt Holes (x 4)					
Shape	See Fig 4		X	X	
Dimension *2a*	29mm		X	X	
Dimension *2b*	7mm		X	X	
Tolerance (on *2b*)	+0.4mm/-0mm	X	X	X	
Dimension *2c*	15mm		X	X	
Dimension *2d*	12mm		X	X	
FEATURE 3: Mating Face					
Shape	See Fig 4		X		X
Surface Roughness	1.6µmRa	X	X		X

Figure 5. Compatibility matrix for illustrative example

In terms of dependencies, the lathe violates all properties of feature 2, as it cannot produce the required shape. For the same reason, drilling violates all properties of feature 3, and all but the material for the basic component. Because drilling is a material removal process, this does not prevent it being used for features 1 and 2. Powder metallurgy, however, violates the basic shape of the component, and therefore *automatically* violates all the properties of its features, as it cannot be used for post-processing. Fig. 5 summarizes these as a compatibility matrix. None of the available processes satisfy all the properties of the component, but this does not make it *infeasible* – as long as a feasible process chain can be formulated.

4.4 Process Chains

By referring to the compatibility matrix (Fig. 5), the designers can identify suitable processes for each feature. Sand Casting or machining on the lathe can be used to produce the basic component. Feature 1 can be produced by either the lathe or drilling. Feature 2 can only be produced by drilling, while features 3 and 4 can only be produced with the lathe. This offers four feasible process chains for making the component, as shown in Table 1. The component is deemed feasible, and passed on for evaluation, where the process chains identified can be used to assess its performance against economic requirements.

Table 1. Feasible process chains for illustrative example

Chain	Basic Shape	Feature 1	Feature 2	Feature 3	Feature 4
1	Sand Casting	Lathe	Drill	Lathe	Lathe
2	Sand Casting	Drill	Drill	Lathe	Lathe
3	Lathe	Lathe	Drill	Lathe	Lathe
4	Lathe	Drill	Drill	Lathe	Lathe

5. CONCLUSIONS AND FUTURE WORK

The needs of different product life-cycle stages often conflict, and designers must sacrifice one to satisfy another. This paper has argued that decision analysis can help designers assess these trade-offs in terms of what the product is meant to achieve. Not by calculating optimum trade-offs, but by helping designers reflect on the implications of their choices from different perspectives, in terms of the specified requirements. A methodology has been proposed for analysing design decisions from a product perspective – using existing decision analysis techniques – and from a process perspective, using a new feasibility test. This helps designers relate the information provided by conventional IPPD methods, so they can decide when manufacturability or functionality should take precedence.

Further work is needed in two areas. Firstly, success or failure in applying a methodology depends upon *how* it is used by the human beings who put it into practice. Testing is needed to determine the practicality of the methodology in design decisions taken by a group of designers; initially under controlled lab conditions, and subsequently in actual design projects. Regardless of any theoretical strength, a methodology is of no value if it is never used – or never used correctly – in practice. Also, manufacturing is only one aspect of the product life cycle that designers need to consider. The methodology presented here could readily be expanded to accommodate issues such as Design for Assembly, or Design for the Environment. The structured framework provided by decision analysis has the potential to help designers explore their decisions from many points of view, and not just manufacturing.

REFERENCES

1. G. Pahl and W. Beitz, Engineering Design: A Systematic Approach, 2nd Ed., Springer-Verlag, ISBN 3540199179, 1996
2. S. Pugh, Total Design: Integrated Methods for Successful Product Engineering, Addison Wesley, ISBN 0201416395, 1990
3. D.L. Thurston, J.V. Carnahan and T. Liu, Optimisation of Design Utility, Journal of Mechanical Design, N° 116, 1994, pp801-808
4. G.A. Hazelrigg, A Framework for Decision-Based Design, Journal of Mechanical Design, N°120, 1998, pp653-658
5. J. Wang, Ranking Engineering Design Concepts Using a Fuzzy Outranking Preference Model, Fuzzy Sets and Systems, N°119, 2001, pp161-170
6. S-W.Hsiao, Concurrent Design Method for Developing a New Product, International Journal of Industrial Ergonomics, N°29, 2002, pp41-55
7. K.J. Arrow, Social Choice and Individual Values, 2nd Ed., Yale University Press, ISBN 0300013639 m, 1973
8. T.L. Vincent, Game Theory as a Design Tool, Journal of Mechanisms, Transmissions, and Automation in Design, N°105, 1983, pp165–170
9. L. Chen and S. Li, A Computerized Team Approach for Concurrent Product and Process Design Optimisation, Computer-Aided Design, N°34, 2002, pp 57-69
10. L.L. Bucciarelli, Designing Engineers, MIT Press, ISBN 0262023776 m, 1994
11. L. Phillips, A Theory of Requisite Decision Models, Acta Psychologica, N°56, 1984, pp29-48
12. W. Edwards and D. von Winterfeldt, Decision Analysis and Behavioral Research, Cambridge University Press, ISBN 0521273048, 1986
13. W. Edwards, How to Use Multi-Attribute Utility Measurement for Social Decision-Making, IEEE Transactions on Systems, Man and Cybernetics, N°7, 1977, pp326-340
14. K.G. Swift and J.D. Booker, Process Selection: From Design to Manufacture, 2nd Edition, Butterworth-Heinemann, ISBN 0750654376, 2003

AIDED DECISION-MAKING FOR AN EMBODIMENT DESIGN PROBLEM

Dominique Scaravetti, Jérome Pailhes, Jean-Pierre Nadeau, and Patrick Sebastian
TREFLE-ENSAM,Esplanade des Arts & Métiers, 33400 Talence, France
dominique.scaravetti@bordeaux.ensam.fr; Tel.: +33 5 56 84 54 22; Fax: +33 5 56 84 54 36

Abstract: In the early phases of a traditional design process, many decisions are made by designers. For that purpose, they take advantage of their experience and company knowledge. These decisions are necessary in a sequential design process but may hide many embodiment solutions. Moreover, designers often use a trial-and-error mode to find a working combination of standard elements. To overcome these difficulties, a decision support system based on constraint programming is proposed. The object of the design process set out in this paper is to facilitate the embodiment design phase by avoiding a-priori decision making and searching for feasible architectures in which all the points of view of the various participants in the project are taken into account. A preliminary search of structuring characteristics of the design problem is necessary and is detailed here. This methodology is applied to embodiment design of the automatic weight-winding system of a monumental clock. To determine the most interesting solutions, objectives and performance indicators are entered. Finally, the benefits of our approach are discussed.

Key words: Embodiment design, Design problem analysis, Decision support system, Structuring characteristics, Constraint specification

1. INTRODUCTION

In a traditional design process, the conceptual design phase ends with a selection among several relevant concepts. Even if several methods and tools may be helpful to designers, concept selection is a critical step because all subsequent design activities are based on the decisions made at this stage [1],

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 159-172.

trade-offs must be made [2]. Then, embodiment design consists in rough arrangements and selections of structural dimensions, materials, components and technologies [3]. Designers have to identify relevant design parameters to start the design process based on the concept retained. Thus, the embodiment design phase is often performed according to professional habits and designers' experience [2]: designers must make assumptions [4] and make irreversible decisions which restrict the solution space. All these decisions may hide interesting solutions and lead to a non-optimal solution. Moreover, life-cycle cost can be influenced up to 70% by decisions taken during conceptual and embodiment design stages [5], as design knowledge is still low [6].

A decision support system (DSS) is proposed in this paper to assist designers in the embodiment design phase. The design problem is defined as a numeric Constraint Solving Problem (CSP). Inverted Integrated Design [7, 8] proposes to express functional requirement criteria, technical skill rules, physical behaviours, variables ranges, by using constraints. The domains of design parameters are finite, but they are set to the widest limits so that no solutions are dismissed. Using a CSP solver, these domains can be narrowed and the designer can focus on feasible embodiments while observing the objectives already expressed at this stage of design. A difficulty is to correctly formalise knowledge to make this base of knowledge coherent, non redundant and complete. A preliminary design problem analysis proves necessary. A four-level approach is proposed in order to express the structuring characteristics of the design problem [9]. This methodology also assists in the transcription of the design problem into a set of constraints.

In this paper, embodiment design is applied to a mechanical device to illustrate the selection between two relevant working structures, and the search for feasible embodiments. Two design processes are compared: The difficulties met in a traditional sequential method are highlighted. The benefits of a design process based on this decision support system are explained: the possible use of standard elements is checked out and standard elements as well as structuring dimensions are selected. *A priori* decisions are avoided.

2. EMBODIMENT DESIGN PROBLEM

The decision support system is used in embodiment design for the automatic weight-winding mechanism of a monumental clock. The design requirements specify that drums must be driven periodically by the mechanism. So, the drums must rotate freely in the opposite direction. A maximum overall dimension requirement must be met. The mechanical loads applied to the clock must be equivalent to those generated by a manual operation. The general concept retained, as a result of conceptual design, is to drive each drum with a geared motor located on a mounting above the clock (Fig. 5). Transmission is via pulleys and a flat belt. Engaging is achieved by the tension of the belt.

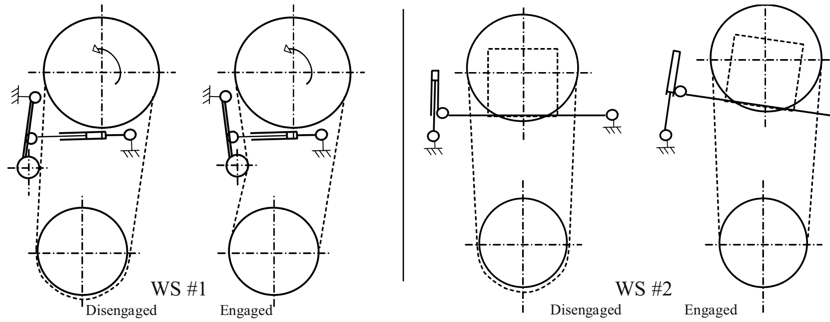


Figure 1. The two selected working structures (WS), in disengaged and engaged positions

In order to control the tension of the belt, two relevant working structures (WS) have been retained and are based on the use of a jack. The belt is stretched by a roller (Fig. 1, WS #1) or by moving the motor (Fig. 1, WS #2). Our decision support system assists in making a choice among these two working structures.

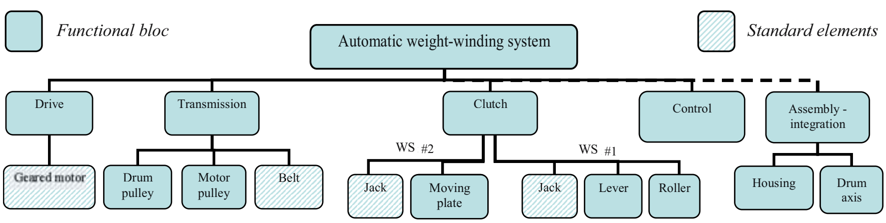


Figure 2. System decomposition (two alternative working structures for belt tensioning)

The technical organisation chart shown in Fig. 2 describes the two working structures: functional blocks are derived into components and standard elements (hatched), bespoke elements and interfacing constraints are identified. Indeed, at this stage of concept description, some choices are already made (using a geared motor for example). Some standard elements (concept already controlled or component defined in a catalogue) have to be selected in databases and others have to be dimensioned.

2.1 Difficulties experienced in traditional design processes

Several concepts are achieved as a result of conceptual design. Designers have to make a first selection among relevant concepts and they tend to focus early on solutions they can handle. Thus, interesting alternatives may be discarded for they are unknown or unusual.

In the early phases of the design process, designers usually validate a principle solution after setting many parameters and making various calculations. Thus, iterations cannot be avoided before design solutions are achieved [10-12] and detail design is started. Embodiment solutions are strongly dependent on all these initial decisions [11]. Moreover, all the elements of the design problem are not taken into account simultaneously in the choices made: qualification criteria, physical behaviours, design and manufacturing rules, etc. Designers often use their technical expertise and skills in pre-dimensioning processes. Sub-assemblies are simulated only once they are definite because most of the simulation tools require an exhaustive definition of the product.

Despite the low number of parameters and standard elements of the automatic winding mechanism dealt with in this paper (Fig. 5), a great number of combinations is possible. The main embodiment problem is to select standard elements in catalogues (geared motor, belt, jack) which match the geometrical parameters (relative positions of pulleys and roller, lengths of lever or moving plate). In a sequential design process, some parameters have to be set to make it possible to determine the other parameters, as all these parameters interact with one another. Designers use a trial-and-error mode.

A group of designers has been observed. They prove to make important decisions: First, they selected the working structure (WS #1). Then, they had to determine diameters and positions for pulleys. They also selected a motor compatible with the required rotational speed and torque. Then, they performed iterations in order to find a standard belt length. In every loop, they made calculations (roller offset and force) to meet the mechanism requirements. This design required many iterations and the proposed solution could have been enhanced: the radial load is about 123 N; the approach based on the decision support system leads to better solutions.

2.2 Constraint programming in design

A strategy to deal with conceptual design problems in a constraint modelling environment is proposed by [13]. Starting from a list of design requirements, design objectives and critical factors in a successful design process are collected in an unstructured manner. Each requirement is formulated into a testable constraint rule. Even if the list of rules is not completed, the authors specify that it is sufficient to allow various configurations to be developed. Even if inter-element compatibilities are checked, conflicting inter-component requirements must be solved [11, 13].

Embodiment design is a constraint specification and satisfaction task [14]. The author proposes a software ("Cadet") with a database of components. It facilitates the writing of the product model and automates the generation of component constraints for a single design concept. Many concepts can be successively tested by reducing the time to generate embodiments. However, an advanced concept is necessary to build the product model and to find a feasible embodiment [15]. Moreover, the studied mechanisms have to be described using the listed components.

Our decision support system assists in the structuring of design problems and in the generation of feasible embodiments, starting from different relevant concepts and working structures. All the conditions that determine the characteristics of the mechanism (geometry, physical behaviours, standard elements, technical skill rules, manufacturing constraint, cost-related elements, etc.) are translated into constraints (equality, inequality or logical rules). Variables may belong to continuous domains; thus designers do not have to perform a-priori selections and the potentialities of each concept are preserved. Variables may also be discrete, enumerated or tabulated. Each relevant working structure is described by a set of specific constraints.

The use of a CSP approach actually avoids a-priori decisions, by preventing causality and sequentiality in calculations. The CSP solver[1] searches for combinations of variables that satisfy all the listed constraints.

3. ANALYSIS OF THE DESIGN PROBLEM AND STRUCTURING ELEMENTS

When a designer initiates the design process, he only chooses the necessary and sufficient relevant functions and characteristics. The size of the design problem is limited. These essential elements structure the design activities when a first architecture is being searched for.

A method for analysing conceptual design data is proposed in [2]. Parameters for evaluating a concept and a set of most relevant components

[1] The solver used is "Constraint Explorer" (CE), developed by Dassault Aviation, within the framework of the "CO2" project (French national network of software technologies) [16], involving the LINA, LIP6, TREFLE and LIPSI laboratories.

are obtained. Design knowledge is extracted from a given set of previous designs and necessitates such a database. An analysis and structuring methodology is proposed in [9] in order to facilitate the writing of constraints by designers: a four-level (need, function, structure, physical behaviour analysis) approach is used to identify the structuring characteristics and functions which are then translated into constraints. The different stages and job status of our design approach are illustrated in Fig. 3. This process takes place between the conceptual and detail design phases. It starts from several relevant concepts and leads to a feasible embodiment solution.

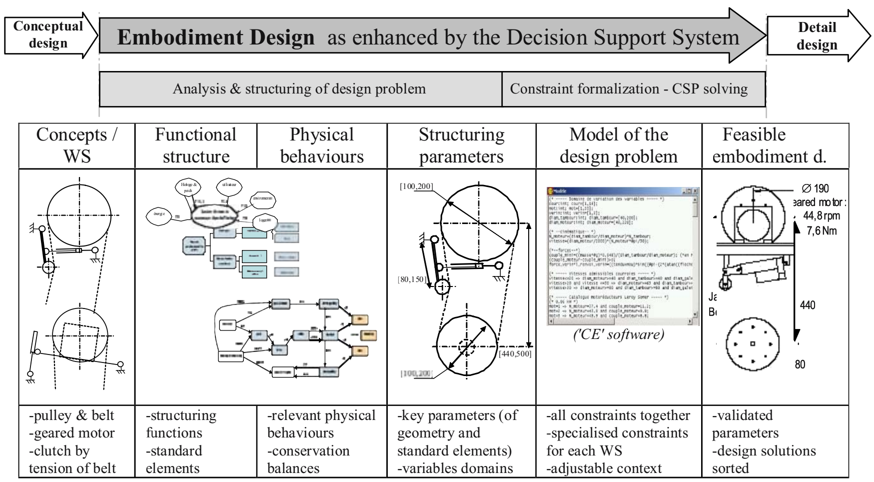


Figure 3. Different job status in the embodiment design phase, as enhanced by DSS

3.1 Need, function, structure analysis

With this methodology including tools for needs, functional and structural analysis, the following structuring characteristics and functions (Table 1) that have to be converted into constraints can be obtained.

Table 1. Structuring characteristics and functions.

Customer qualification criteria	- loads on the clock : radial force, rotational speed - service life and resistance to external stress - cost
Function under no flexibility	- maximum overall dimensions (x,y,z) - no damage to the clock (maximum loads)
Imperative constraint functions	- use of electrical power - weight to be wind up

Manufacturing constraint	- maximum pulley diameter
Function with risk	- clock undisrupted (residual friction in disengaged position)
Standard elements	- geared motor, belt, jack
Elements to be defined (geometry)	- pulleys- lever and roller / moving plate
Functional interfaces	- pulley-belt- pulley-drum axis , etc.

3.2 Physical behaviours involved in product operation

3.2.1 Analysis

For the most important life cycle stages ("automatic winding", "clock operation"), block diagrams of the mechanism are written. Contact and energy fluxes are thus identified. We propose to add "substances" (materials carried by fluxes) to these diagrams, to identify the effects of the fluxes with verbs of action between blocks, and to characterise these actions (useful, harmful, insufficient). The physical phenomena involved in the operation of the product which have to be taken into account in the design problem model can be defined using substance-field diagrams based on TRIZ theory [17]. Modelled effects can be described more or less precisely according to the analysis of the designer.

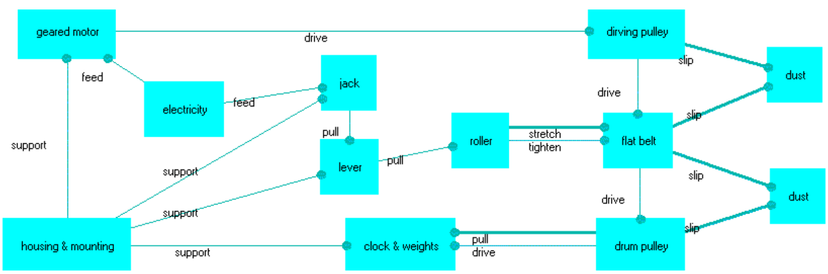


Figure 4. Substance-field diagram – working structure #1.

Figure 4 describes the physical behaviours of the elements of the WS#1, for the life-cycle stage "automatic winding". This graph has been determined using the software MAL'IN [18].

3.2.2 Interpretation

In the previous graphs, each action identifies one or more physical models and the relevant variables to be taken into account (Table 2). This graph is usually an innovation tool; it is used to identify technical or physical contradictions to be solved. Harmful actions must be eliminated. In our case, the working structures described are selected. The harmful effects are undergone and taken into account. To lead to acceptable consequences, the harmful effects induce limitations which are often translated into constraints by inequalities, whereas physical laws are usually defined by equalities [1].

Table 2. Identification of relevant parameters and physical behaviours (working structure #1).

Component	Action	Component	Variables	Relationship
Pulley	Drive / *Slip*	Belt	C, ω, ∅, belt tensions, coef. of friction	Adhesion conditions, Torque transmission
Roller	Tighten	Belt	jack force, ∅, belt tensions & lengths	Balance of forces, Geometry
Jack	Pull	Lever	lever lengths, forces, jack stroke	Balance of forces & torques, Geometry
Pulley	Drive	Drum	C, ω, ∅, clock weight	Torque calculation
Pulley	*Pull*	Drum	belt tensions, radial force	Balance of forces
Roller	Stretch	Belt	torques, % lengthening	Belt manufacturer recommendations
etc.	etc.	etc.	etc.	

3.3 Structuring parameters

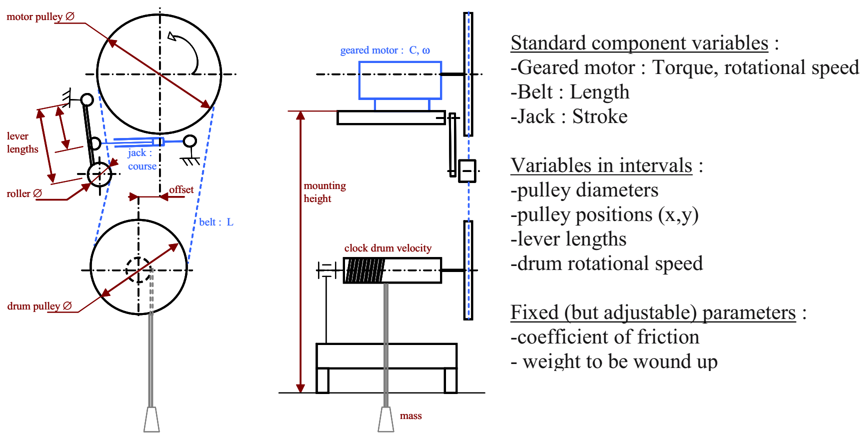


Figure 5. Key design parameters, for the working structure #1.

When possible, ranges of variation have to be proposed for variables. Their limits are as wide as possible so as not to limit the search space and to avoid dismissing solutions. For example, even if a rotational speed corresponding to a manual action is required for the clock drum (`N_drum`), the variable must belong to a given interval. Indeed, constraints imposed by standard motors (discrete rotational speed) cannot generate a value of `N_drum` exactly equal to an integer. Intervals translate either the flexibility linked to variables, or the uncertainty. Finally, the design problem takes the variables indicated in Fig. 5 into account. The clearance provided by the housing determines the maximum dimension ranges. As regards the standard components defined in industrial catalogues, the variables are discrete, enumerated or tabulated.

3.4 Translation into constraints

Some of the structuring functions have already been taken into account in the selections made in the first level of design. The other functions are translated into constraints, with the formalism of 'CE' software as shown below. Some constraints only relate to a single working structure.

▪ Key parameters : weight : 28kg (bell drum), 12kg (clock drum) ; maximum diameters : bell drum pulley=∅110mm, clock drum pulley=∅200 mm ; coefficient of friction (depending of conditions of use).

▪ Variable ranges: `∅_drum=[40,200];` (manufacturing capacities)
`mounting_height=[590,650];` (overall dimensions)
`N_drum=[55,61];` (requirements) etc .

motor_ ref	N_motor	torque_ motor	motor_ weight	motor_ cost
1	37.4	11.2	6.3	253
2	43.8	9.8	6.3	253
3	48.9	8.8	6.3	253
…	…	…	…	…
20	64.4	24.1	10.3	319

belt_ ref	belt_ length	belt_ cost
1	750	15
2	800	16
3	850	17
…	…	…
18	1700	34

jack_ ref	jack_ stroke	jack_ cost
1	50	166
2	100	182
3	150	198

Figure 6. Tables of key parameters for standard components.

▪ Catalogues : Lists of the available components and the corresponding parameters : the manufacturers' references are stored in a spreadsheet program (Fig. 6). For instance, the motors can only be selected among a defined list (motors in the range from 0.06 kW to 0.18 kW). Each motor reference is linked to relevant data. An embodiment solution (Table 3) can only use one of these existing components. The conditions of use of the standards (imposed by manufacturers) are also constraints; for example, for belts, the admissible rotational speed is translated into logical rules:

`speed<=20 → ∅_motor >= 40 and ∅_drum >= 40 and ∅_roller=40;` etc.

▪ Physical behaviours :

`T/t=e^(f*θ);` *(T&t:belt tensions; f:coef. of friction; θ:angle of belt winding)*

`torque=(T-t)*∅_drum/2; torque_motor > torque;` etc.

4. RESULTS

When the model and context have been completed, a first filtering operation is performed very quickly, which reduces the domains of variables or shows their incompatibilities. There may be numerous solutions due to the wide variable domains. Performance indicators are used to sort the solutions.

Table 3. First solutions for the 12kg drum weight, sorted according to the objective function

Sol. No	N motor	N drum	coef of friction	WS	min torque	motor torque	jack stroke	cost	offset	roller ∅	motor pulley ∅	drum pulley ∅	roller offset	jack force	mounting height	pulley length	jack length	belt length	weight	motor weight	low belt tension	high belt tension	radial force	cost indic.	r. force indic.	OBJ. funct.
	rpm	rpm			N/m	N/m	mm		mm	mm	mm	mm	mm	N	mm	mm	mm	mm	kg	kg	N	N	N			λ=0,5
12	56.3	60.80	0.5	1	6.10	7.6	50	516.0	0	40	216	200	100.82	168.03	590	111	51	1550	12	6.3	14.82	71.31	86.13	0.829	0.9	0.865
6	48.9	55.17	0.5	1	6.37	8.8	50	515.9	0	40	220	195	100.25	177.28	590	115	51	1550	12	6.3	15.20	73.14	88.34	0.829	0.923	0.876
22	57.1	60.81	0.5	1	6.02	12.4	50	532.0	0	40	213	200	107.04	173.81	590	109	48	1550	12	7.3	14.82	71.31	86.13	0.855	0.9	0.877
11	56.3	60.97	0.5	1	6.12	7.6	50	512.7	0	40	209	193	60.29	116.43	590	112	82	1500	12	6.3	15.36	73.90	89.26	0.824	0.933	0.878
10	56.3	60.13	0.5	1	6.03	7.6	50	511.5	0	40	204	191	78.57	143.13	590	114	67	1500	12	6.3	15.52	74.67	90.19	0.822	0.942	0.882
16	49.6	55.17	0.5	1	6.28	14.2	50	532.2	0	40	218	196	102.83	177.72	590	107	47	1550	12	7.3	15.13	72.77	87.89	0.855	0.918	0.887
137	56.3	59.13	0.5	2	5.93	7.6	100	541.1	0	40	209	199	0	146.14	590	90	209	1550	12	6.3	14.90	71.67	86.57	0.87	0.905	0.887
17	49.6	55.20	0.5	1	6.29	14.2	50	531.9	0	40	217	195	106.35	182.94	590	115	49	1550	12	7.3	15.20	73.14	88.34	0.855	0.923	0.889
20	57.1	60.93	0.5	1	6.03	12.4	50	529.0	0	40	207	194	63.87	119.03	590	114	81	1500	12	7.3	15.28	73.52	88.80	0.85	0.928	0.889
136	56.3	59.44	0.5	2	5.96	7.6	100	539.1	0	40	208	197	0	164.98	590	90	180	1550	12	6.3	15.05	72.40	87.45	0.866	0.914	0.89
3	48.9	55.14	0.5	1	6.37	8.8	50	512.4	0	40	212	188	63.35	130.64	590	107	71	1500	12	6.3	15.77	75.86	91.63	0.823	0.957	0.89

4.1 Objective function

Solutions are sorted out using the objective function (equation 1) defined by using relevance indicators (I_n) derived from customer qualification criteria. The objectives that can be expressed at this stage of the design process relate to cost elements, geometrical requirements, expected physical behaviours. These indicators are weighted by factors λ [10].

$$objective\_function = \left(\lambda_1 . \frac{I_1}{I_{1\max}}\right) + \left(\lambda_2 . \frac{I_2}{I_{2\max}}\right) + ... + \left(\lambda_n . \frac{I_n}{I_{n\_\max}}\right) \quad \text{with } \Sigma\lambda_i = 1 \quad (1)$$

In our study, relevance indicators are radial force to drum and cost. Cost evaluation: standard elements are determined using our decision support system; their costs are known and depend on characteristics such as power, size, etc. Moreover, the costs of raw materials are easy to calculate. As for the parts to be manufactured, economic indicators rather than real costs are expressed. The only costs retained are those relating to the parts' geometrical parameters. For the same manufacturing conditions, the differences between part costs correspond to differences between machining time. Nevertheless, embodied solutions can be compared using these economic indicators.

$$objective\_function = \left(\lambda . \frac{Cost}{Cost\_\max}\right) + \left((1-\lambda) . \frac{Radial\_force}{Radial\_force\_\max}\right) \quad , \lambda \in [0,1] \quad (2)$$

The objective function only necessitates two performance indicators (I/I_{max}). The objective function defined in equation (2) must be minimised. A comparison between solutions can be either absolute (each solution is compared to objectives by using criteria) or relative (solutions are compared by using criterion-defined measurements) as with performance indicators. A negotiation factor, λ, varies the weights of performance indicators: if the designer wants to give more importance to cost, he will set λ to a value closer to 1. Table 3 shows the first solutions, as sorted according to the objective function, for indicators of equal weights (λ=0.5). In this table, each row corresponds to a solution (feasible embodiment).

4.2 Results analysis

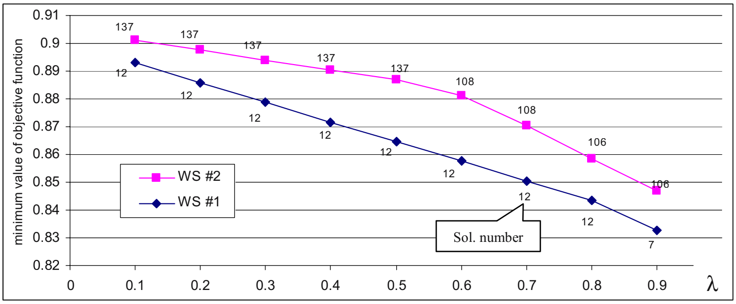


Figure 7. Minimum values of the objective function according to indicator weights, and corresponding solution.

Fig. 7 shows the minimum values of the objective function for each working structure, plotted against λ. The best solutions appear with WS #1 for each value of λ. Solution No. 12 is the most interesting for λ∈[0.1 , 0.8] and solution No. 7 emerges for λ=0.9. The use of this weight-factor method raises the problem of value assignment for the factors [10]. Moreover, the interest of such a weighting process does not appear clearly here. Indeed, whatever the value of λ, the solution No. 12 is almost always the most relevant one (Fig. 7). In fact, cost is influenced by each criterion. A more relevant method would be to write a global objective function only based on costs. The radial force can be linked to the service life of drum bearings, and thus to an amortisation period according to the cost. This work is under development. The solutions, as sorted according to increasing indicators (Fig. 8), shows that the WS #1 appears to be more relevant than # 2. The "staircases" observed in this figure correspond to the jumps existing in the enumerated values of the standard elements. The differences between the working structures mainly influence cost (Fig. 8b). However, no solution displays both indicators (radial force and cost) at their minimum values. The first solutions obtained (Table 3) display interesting indicators. Both working structures may be appropriate, depending on context. Thus, a solution using WS # 2 (No.137, table 3) appears among the first solutions. Another improvement of results analysis could be obtained by searching for the Pareto front in the whole solutions space. It highlights the non-dominated solutions, according to the minimisation of the performance indicators.

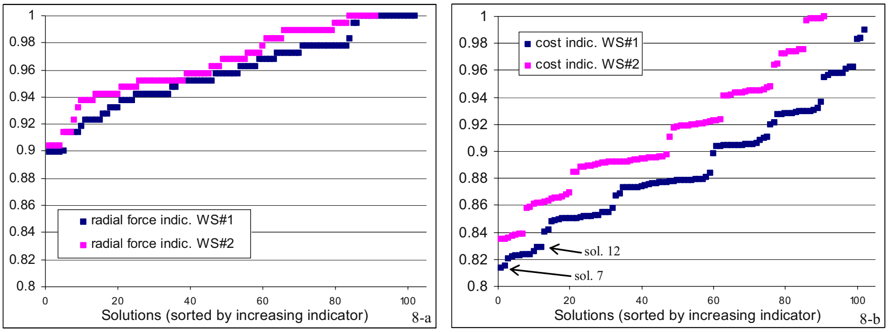


Figure 8. Performance indicators for each working structure.

5. BENEFITS OF THE PROCESS BASED ON A DECISION SUPPORT SYSTEM

5.1 Decision-making

The main difficulties result from decisions to be taken during the early phases of a classical design process. The proposed decision support system avoids the a-priori decision-making, which limits the field of design investigations. In our example, decision making is assisted in the selection of values for structuring characteristics and in the selection of standard elements. The designer is also helped when selecting a working structure that best meets all the requirements and objectives.

5.2 Detail design and time saving

The detail design phase has been shortened because the designer has a validated embodiment solution as well as dimensioning elements available. The results of the embodiment design phase have been used directly. Since all the structural parameters – dimension and positioning – (Fig. 5) have been defined, and standard elements selected, the detail design phase (Fig. 10) is quick to perform. In this design case, a global time saving was observed compared to the traditional process (§ 2.1). Traditional design methods require an iterative search of standard elements but also repeated topological fitting and calculations. Our approach focuses on an architecture guaranteeing that geometrical and mechanical constraints and technical skill rules are satisfied using only available standard components. It avoids *a-*

posteriori backtracking on the concept and limits iterations. Writing the design model is time consuming, but the model could be used afterward in order to adapt or change the design. This process capitalises knowledge, which is useful for the company. Standard component catalogues (spreadsheets) are supplied by the manufacturer. A new design often starts from a previous design and formalised knowledge appears to be useful then. However, this method must be applied to sub-problems that designers are able to describe, and for which many different configurations are possible.

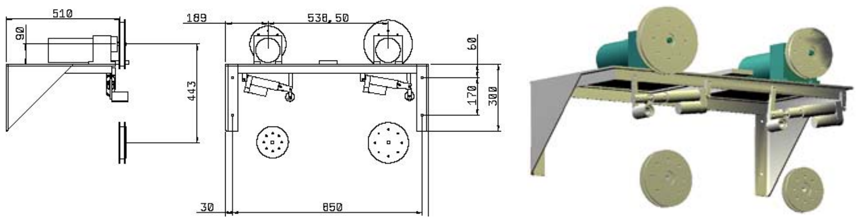


Figure 9. Assembling drawing of the automatic clock weight winding system

5.3 Minimisation of objectives

Sorting solutions can minimise performance indicators. In our case, the radial force to the clock is, for example, about 86 N for the solution No.12. Compared to the traditional method, radial force is reduced by 30%. The designer and his customer have to agree on the relative importance of the relevance indicators. There are no more decisions to be made and no more design parameters to be set.

6. CONCLUSION

The object of the design process set out in this paper is to facilitate the embodiment design phase by avoiding *a-priori* decision making, integrating many points of view, and searching for feasible architectures. The proposed methodology structures knowledge into structuring characteristics that are translated into constraints. Then, the CSP solver reduces the domains of problem variables. Starting from different relevant concepts or working structures, a field of feasible embodiment solutions is obtained. Optimal combinations of structuring parameters and standard components are worked out. The solutions obtained satisfy the functional, physical and technical skill-rule constraints expressed at this level of the design process. Then, solutions are sorted according to objectives. The total duration of the design process is low compared to the traditional method because the trial-and-error mode is avoided. We are currently using this methodology for an industrial design problem in the aeronautical field. The problem model requires the translation of complex mechanical and thermal behaviours. The evolution of the software should lead to the identification of non-compatible constraints, in the event of a conflict between several constraints.

REFERENCES

1. L.C. Chen, L. Li, Optimization of product configuration design using functional requirements and constraints, Research in Engineering Design, 167-182, n°13 (2002).
2. P.C. Matthews, L.T.M. Blessing, K.M. Wallace, The introduction of a design heuristics extraction method, Advanced Engineering Informatics, 3-19, vol. 16 (2002).
3. G. Pahl, W. Beitz, Engineering design; A systematic approach, 2nd edition (Springer-Verlag, London, 1996).
4. A. Chakrabarti, T. P. Bligh, T. Holden, Towards a decision-support framework for the embodiment phase of mechanical design, AI in Engineering, 21-36, n°7 (1992).
5. L.Zimmer, P.Zablit, Global aircraft predesign based on constraint propagation and interval analysis, in: Proceedings of CEAS Conference on Multidisciplinary Aircraft design and Optimisation, Köln, Germany (June 2001).
6. C.B. Chapman, M. Pinfold, Design engineering- a need to rethink the solution using knowledge based engineering, Knowledge-Based Systems, 257-267, n°12 (1999).
7. X. Fischer, Stratégie de conduite du calcul pour l'aide à la décision en conception mécanique intégrée; application aux appareils à pression, Ph.D thesis, ENSAM (2000).
8. X. Fischer, J.P. Nadeau, P. Sebastian, P. Joyot, Qualitative constraints in integrated design, in: Integrated Design and Manufacturing in Mechanical Engineering, P. Chedmail et al. (eds.), Kluwer Academic Publisher, 35-42, (2002).
9. D. Scaravetti, J.P. Nadeau, P. Sebastian, Structuring functions and constraints formulation; for an enhanced embodiment design, in: Proceedings of International CIRP Design Seminar, Grenoble, France (May 2003).
10. M.F. Ashby, Choix des matériaux en conception mécanique, Dunod, Paris (2000).
11. B.J. Hicks, S.J. Culley, An integrated modelling environment for the embodiment of mechanical systems, Computer-Aided Design, 435-451, n°34 (2002).
12. D. Tate, M. Nordlund, A design process roadmap as a general tool for structuring and supporting design activities, in: Proceedings of the 2nd world conference on integrated design and process technology, Austin, USA (1996).
13. A.J. Medland, G. Mullineux, A decomposition strategy for conceptual design, Journal of Engineering Design, vol. 11, 3-16, n°11 (2000).
14. A.C. Thornton, Constraint specification and satisfaction in embodiment design, Ph.D thesis, Cambridge University, UK (1993).
15. A.C. Thornton, A.L. Johnson, CADET a software support tool for constraint processes in embodiment design, Research in Engineering Design, 1-13, vol. 8 (1996).
16. L. Zimmer, CO2 Project, in: Proceedings of S3P seminar, Bayonne, (November 2001).
17. S.D. Savransky, Engineering of creativity: Introduction to TRIZ Methodology of Inventive Problem Solving (CRC Press, 2000).
18. TREFLE-ENSAM, MAL'IN: Méthodes d'Aide à L'INnovation (Methods for Innovation aid) (2003); http://www.lept-ensam.u-bordeaux.fr/malin.

SHARING MANUFACTURING INFORMATION AND KNOWLEDGE IN DESIGN DECISION AND SUPPORT

Bob Young[1], Anne-Francoise Cutting-Decelle[2], David Guerra[1], George Gunendran[1], Bishnu Das[1], and Sean Cochrane[1]
[1]*Loughborough University, Loughborough, LE11 3TU, UK,*
[2]*University of Evry, Evry, France.*

Abstract: The integration of product design and manufacture has been pursued and advanced for over 20 years. Further advances continue to be pursued in the knowledge that success can lead to significant competitive advantage. This paper provides a contribution to the view that integration through information and knowledge sharing has the potential to offer designers and manufacturing engineers improved decision support based on the provision of high quality information. Our current research into both the definition of manufacturing information and knowledge models and knowledge sharing techniques is explained. Further, the potential for process specification languages to provide improved knowledge sharing is discussed.

Key words: Information sharing, Knowledge sharing, Languages, Inter-operability, Manufacturing

1. INTRODUCTION

Since the 1980s, when CAD/CAM systems became the first integrated tools to support designers and manufacturing engineers, the capability of software tools have advanced substantially, with advanced parametric and knowledge based CAD tools linked to design analysis, manufacturing planning and code generation. In addition, the introduction of Product Data Management systems, Enterprise Requirements Planning systems and web-based tools are starting to provide a basis for information sharing and workflow analysis[1].

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 173-185.

However, the level to which information and knowledge can be shared is still very limited. Information sharing in today's systems tends to be at a human comprehension level and the recognition that clearly structured meta-data approaches are needed for computational comprehension has led to significant efforts in the area of ontologies and languages for information sharing and knowledge interchange [2]. The information sharing problem is further complicated by inter-operability problems when multiple system solutions are required which is typical of the situation in global development teams and supply chains. The work of the STEP standards community, [3], has made some progress with product data sharing, starting with geometry interchange and now providing further standards for product life cycle systems [4] and for manufacturing applications such as process planning [5]. Similarly the work of MANDATE [6] is providing a foundation for sharing manufacturing resource information. However, the breadth of information and knowledge types involved in the support of design decisions means that there is a significant amount of research still needed. While the currently accepted approaches of defining information infrastructures to provide comprehensive support to decision making can still provide valuable advances in understanding there is now a need to investigate approaches that can offer more flexible solutions [7]. This is especially true in design, which typically involves many members of a design team, each taking a different perspective on the problem.

This paper explores the issues involved in providing information and knowledge which can be shared across teams of designers, focusing particularly on manufacturing support. The approach taken attempts to ensure flexibility in system development, system maintenance, and the provision of high quality information to aid the decisions of design team members. The current status of our work on manufacturing information and knowledge modelling is explained and a route to providing flexible information sharing is discussed. The limitations of syntactic infrastructure development are also discussed and our current research exploration, to overcome these through the use of languages to ensure meaningful communication of information, is explained.

2. FUNDAMENTAL ISSUES IN MANUFACTURING INFORMATION AND KNOWLEDGE SHARING

Manufacturing businesses are becoming more and more globally dispersed and companies are becoming more willing to work closely together in order to remain competitive, It is therefore essential that information and knowledge sharing systems are able to support the global nature of business interactions. While Product Lifecycle Management (PLM) and Enterprise Resource Planning (ERP) systems are starting to offer effective support for communication, it is important to recognise that this support is typically in document and process management tools which provide a valuable aid to human interaction. However, for substantial additional benefit to be gained there is a need to share information at a level where computational sharing is possible. The resolution of this issue, illustrated in Fig. 1, requires an improved definition of the information and knowledge to be shared as no human intelligence is involved in its interpretation.

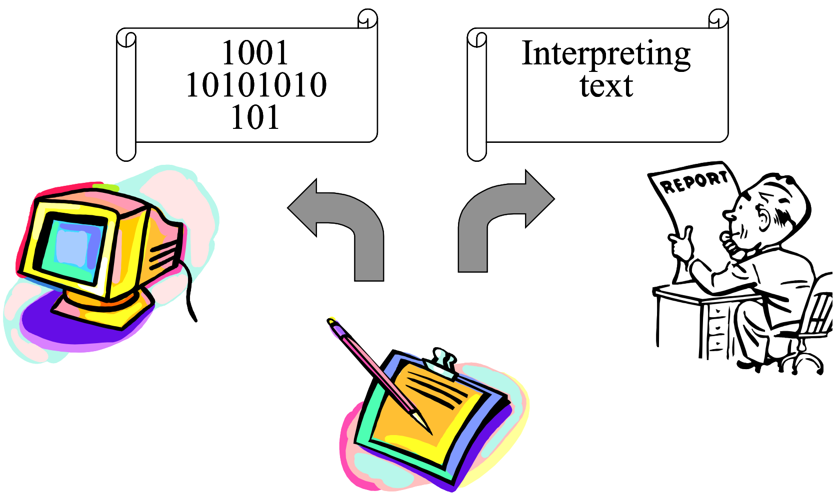


Figure 1: Computer or human interpretation of information

Even at the human interaction level, successful implementations of PLM and ERP systems require a huge business commitment to information classification and to change management. Nonetheless the potential value of successful computer based information sharing is substantial. Even at current levels of information communication between CAD systems it has recently been estimated that the costs of inter-operability problems in the US automotive sector run at around $1 billion per annum [8]. Add to this the broad range of additional information which could usefully be communicated from successful computer based information sharing and it can readily be seen that vast industrial benefit could result.

A further point, which can bring substantial confusion in the area of information and knowledge sharing, is the broad range of meanings used for the terms data, information and knowledge. Often data and information are used as interchangeable terms and no substantial distinction is drawn between the terms information and knowledge. In this work we use the following definitions [9]: Data relates to words or numbers, the meaning of which may vary, dependent on the context; Information is data which is structured so that it has a particular meaning; Knowledge is information with added detail relating to how it may be used. We believe that these are in line with the definitions arising from a recent discussion paper [10]: Data are simply symbols with no context and no relationships; Information is data within a specific context; Knowledge is an information relationship within or across contexts.

The range of information to be shared is not only wide, but needs to be viewed from different, multiple, perspectives. This is because each team member is likely to be interested in different aspects of the information, such that the significant attributes involved will be different and have different levels of significance. For example the designer of an injection moulded product may have a requirement for a particular set of dimensions and tolerances on a plastic component. The tool designer needs to have an understanding of the plastic to be used before he/she can specify the appropriate dimensions for the mould and whether the required tolerances can be achieved. Similarly the need for a location fit in a mechanical product has implications for the assembly dimensions and tolerances, which in turn will have implications for the manufacture of the components required. A simple illustration of this is provide in Fig. 2. In addition to viewpoint dependency, there is also a time dependency on information. The existence, and relevance, of particular information will vary with time through the design process.

A basic requirement of any effective information support tool is to provide the user with information that is useful without being swamped by a significant volume of irrelevant information. This type of information overload problem is common with Internet searches. If future information support systems are to be valued by their users it is important that a level of knowledge is encapsulated within the system to ensure that only high quality information is offered.

The provision of information, which can be communicated at a computer interpretable level, generally requires the definition of common infrastructures which can be shared by the systems which need to interact. However, on large scale interactions this tends to lead to monolithic data structures which are inflexible and difficult to change. While these may work for a specific organisation for a short period of time they are unlikely to be well suited to a variety of businesses which tend to develop their own organisational structures and functionality. Also as business requirements change rapidly the information and knowledge requirements of the business also change. It is therefore essential that methods for information and knowledge sharing are flexible and extensible if they are to be worthy of the investment which will be needed in their implementation.

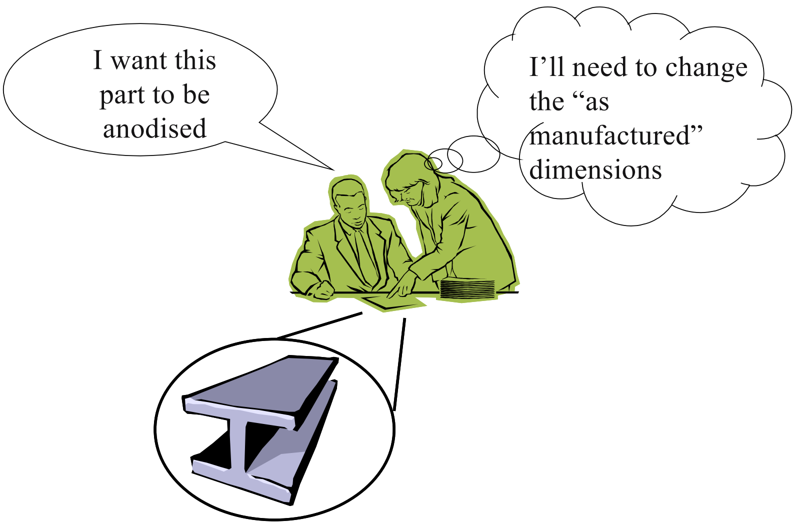


Figure 2: Viewing information from different perspectives

An issue, which is typically given inadequate treatment in the generation of information systems, is that of the meaning of the entities defined within information infrastructures. Typically definitions are made using simple textual statements which can then be mis-interpreted by system developers leading to confusion and misunderstanding of the levels of communication which can sensibly be achieved. For this reason, although difficult to achieve, a semantic definition of entities to be shared should be provided.

Information and knowledge sharing can be considered to be an issue within a specific support system. However the more extensive issue of inter-operability between systems should not be overlooked. While some progress may be made at a single system level of sharing, it is important to recognise that the costs mentioned earlier in this section cannot be substantially reduced without due consideration to inter-operability between systems. The unacceptable alternative would be to tie all users to a common system.

The overall concept which we have been exploring to provide an appropriate information and knowledge sharing environment is illustrated in Fig. 3.

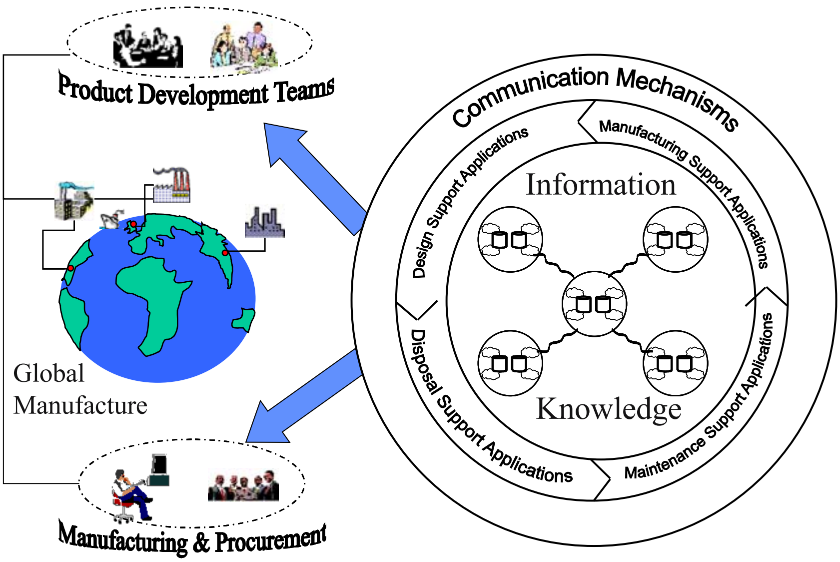


Figure 3: An Information & Knowledge Sharing Environment

3. KNOWLEDGE TRANSFORMATION LAYERS FOR INFORMATION SHARING

The commonly accepted approach to sharing product information has been through the construction of Product Models, which provide a source and repository for all product related information, utilized during product development. This approach can be extended to support information sharing by defining multi-viewpoint product data structures and adding a knowledge transformation layer which captures the understanding of the relationships between layers.

A multi-viewpoint product data structure is illustrated in Fig. 4. This takes the approach that each life-cycle perspective has a view and each sub-area of significance to the team has a sub-view. Each of these then provides access to the information needed from that viewpoint.

The problem then is how to deal with the inter-relationships between these multiple views. The relationships between views are of two types: (a) access to identical data, but from different perspectives; (b) access to information which changes from one view to another based on an understanding of the relationships between the views.

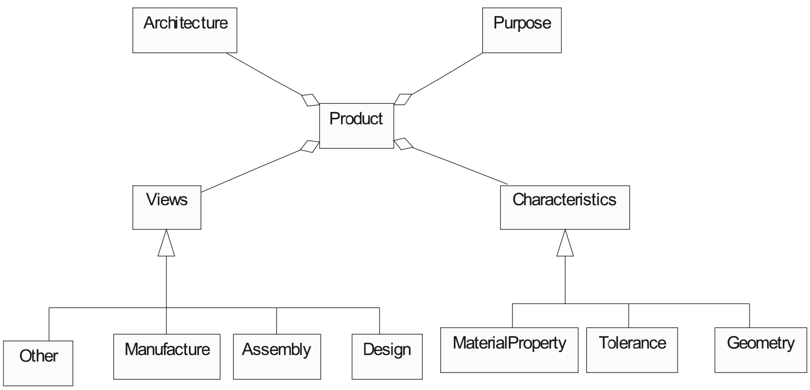


Figure 4: A multi-viewpoint product structure

The first of these is typical of feature technology approaches including shape, function, machining, and assembly features. These views may be related e.g. a functional requirement for a mechanical product can influence an assembly relationship between component parts, which in turn will effect the tolerance requirements on the components. More complex relationships exist where alternative manufacturing methods are considered. For example the choice of casting or injection moulding as a process has implications for the shape of the part as it will require taper angles in order to eject it from a mould. The choice of machining as a process would imply a regular shape for the part in order to ease fixturing requirements.

The approach being pursued in this research is to define a knowledge layer, which lies below the information views, to support the interaction and translation of information. This concept is illustrated in Fig. 5. The knowledge layer captures knowledge which support transformation between views and knowledge which is specific to a particular view.

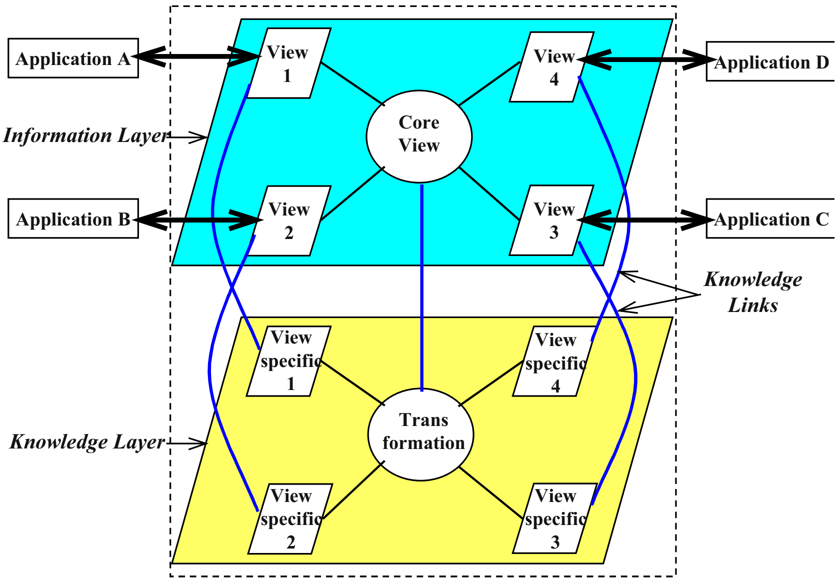


Figure 5: Supporting view relationships through a knowledge relationship layer

4. A MANUFACTURING INFORMATION AND KNOWLEDGE MODEL

The need for a model of manufacturing information, to capture the manufacturing capability of a manufacturing enterprise, has been accepted for some time now [11] as an appropriate means by which manufacturing information can be exchanged and shared across activities which require manufacturing inputs. The typical representations of such information has been to build structures which enable the capture of manufacturing resource information and in some cases manufacturing process information and the relationships between these two.

The significant issue in the work which has been pursued here has been to consider how to extend such an information model into the realms of a knowledge model, such that only high quality manufacturing information is provided to decision makers. A second requirement has been to have a clearly defined structure that will enable manufacturing information and knowledge to be readily maintained as a business changes and develops its capability.

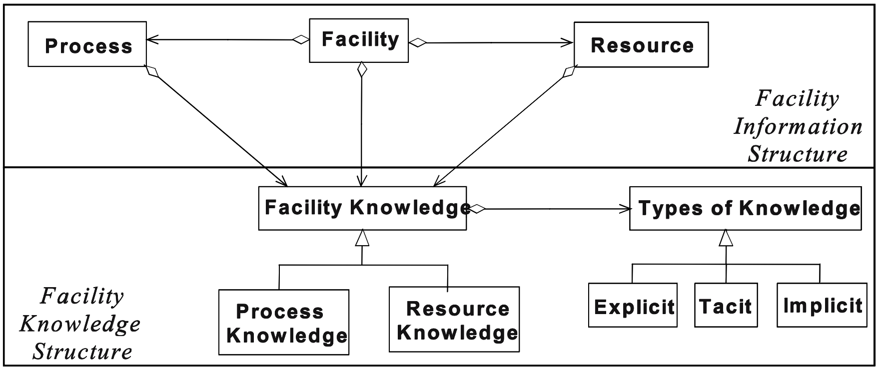


Figure 6: A structure for manufacturing information and knowledge

The range of manufacturing knowledge is extensive and the types of knowledge is equally wide ranging. While we have focused largely on machining knowledge in our work in this area, we have attempted to utilise an accepted structure of knowledge types, classifying knowledge as explicit, tacit or implicit. This has enabled us to identify the potential routes to the use of manufacturing knowledge. For example, explicit knowledge in the form of tables, procedures and graphs can all be captured in such a way that the knowledge can be computer processed. This is not the case with tacit knowledge, such as that captured in a video clip. The use of this structure and the understanding of the relationships between information and knowledge is enabling us to pursue an improved route to the capture, maintenance and use of manufacturing knowledge. An illustration of the top level structure of our manufacturing facility model, along with examples of information and knowledge, is illustrated in Fig. 6.

5. SHARING MEANING THROUGH THE USE OF DEDICATED LANGUAGES

The work described above provides significant syntactic improvement in the representation and capture of information and knowledge for information and knowledge sharing than is typically available in present systems. Further it provides the necessary first steps towards system definitions that can hope to provide a basis for sharing across system boundaries. That is, it provides a number of schema against which systems can interoperate.

However, as the sharing requirement between systems becomes more complex the need to avoid confusion in the definition of the meaning of the classes involved becomes significant. For example the sharing of geometric information is well understood if a boundary representation model is agreed as the basis for sharing. However, manufacture does not have the same focused level of understanding of terms. A manufacturing operation in one case may be all the machining to be done on a specific machine tool or, from another perspective, it may a single piece of geometry to be machined with a single cutter. If effective sharing is to be assured then there is a need to provide a means by which the terms shared can be assumed to share the same semantic definition. The use of ontologies and dedicated languages has the potential to provide this level of definition.

Our work is starting to explore the use of the Process Specification Language (PSL), [12] whose development started at the National Institute of Standards and Technology (NIST). This is a formal language aimed at creating a neutral, standard language for process specification. At this stage the scope of PSL is limited to the realm of discrete processes related to manufacturing, including all processes in the design/manufacturing life cycle. The language is being standardized through the ISO TC 184 SC4/SC5 to produce ISO 18629.

5.1 The PSL language

PSL is based on a lexicon, or a set of logical and non-logical symbols, and a grammar i.e. a specification of how these symbols can be combined to make well-formed formulae. All are chosen to represent the basic concepts in the PSL ontology. The underlying grammar used for PSL is based on the Knowledge Interchange Format (KIF).

Ontology can be defined as a lexicon of specialised terminology along with some specification of the meaning of terms in the lexicon [11]. The foundation of PSL is its ontology, which provides rigorous and unambiguous definitions, based on first order logic, of the concepts necessary for specifying manufacturing processes to enable the exchange of process information. The PSL ontology is essentially two-tiered. The foundation of the ontology is a set of process-related concepts that are common to all manufacturing applications. These concepts constitute the core of the PSL ontology and include concepts such as objects, activities, activity occurrences, and timepoints. Since these concepts, alone, only allow for the exchange of very simple process specifications, the ontology includes a mechanism to allow for extensions to these core concepts (the second tier) to ensure the robustness of the ontology.

The PSL core consists of core definitions of classes, relations and functions. These are object, activity, activity_occurrence, timepoint, before, occurrence_of, participates_in, between, beforeEq, betweenEq, exists_at, is_occurring_at, beginof, endof, object, activity, activity_occurrence and timepoints. Significantly, these are each supported by a set of axioms written in the basic language of PSL which provide a mathematical foundation for each definition.

The addition of outer core extensions is an essential part of PSL as these provide the basis for identifying terms that are more specific to a particular manufacturing purpose. The current extensions deal with activity, temporal and state, activity ordering and duration, resource roles, resource sets, and processor activities.

5.2 Exploring the use of PSL

In our most recent work we are committed to the need for well-defined semantics as a basis for effective knowledge sharing. We are exploring the application of the PSL approach to two areas of manufacturing. The first is in knowledge sharing between design and manufacture. The second is in supply chain interoperability.

Sharing knowledge across design activities requires multiple product, manufacturing and business factors to be considered. PSL potentially enables betters sharing and reuse of manufacturing knowledge, but does not solve the wider problem. We therefore anticipate that while one aspect of the research will involve exploring the applicability of PSL to manufacturing knowledge sharing, there will be a need to explore its relationship to other approaches which may be more appropriate to broader design requirements. The "Standard Upper Ontology" (SUO) [13], currently under development, may provide a framework which can be used to meet this need. This appears to be particularly appropriate as SUO already incorporates the PSL core.

Supply chain interoperability requires multiple businesses to share manufacturing knowledge if supply chain decisions are to be made effectively. Here we are exploring the interface between the construction and the manufacturing sectors, as companies in these two areas are likely to have interestingly different views of manufacture. For example a construction project is typically a one off project with multiple sub-contractors coming together for the duration of the project, while manufactured products are likely to have an on-going infrastructure of businesses involved in the supply chain.

The approach we are taking to these two diverse areas of manufacture is to attempt to understand the processes involved and their relationships, through the development of IDEF3 process models. These can then be evaluated against the existing lexicon available with PSL in order to identify the limitations of the existing language and the requirements for its extension. A further issue here is the need for flexibility and extensibility in a knowledge-sharing environment. The extent to which the use of this approach limits this will also be assessed.

6. DISCUSSION AND CONCLUSIONS

This paper has highlighted the importance of software tools that can enable the sharing of information and knowledge. It has highlighted that while communication between individuals may be problematic, the problems of communication between the software tools which people use are even more complex. However, the potential benefits to business of achieving information and knowledge sharing between these tools is immense.

The levels of information exchange and information sharing has been advancing over recent years, especially through the work of international standards bodies such as STEP and MANDATE. However, there is a further need for an understanding of how to share knowledge as well as information.

This paper has explored an extension to traditional information sharing approaches to enable knowledge sharing across multiple contexts and has focused particularly on manufacturing knowledge sharing. This has shown that clear classifications of information, knowledge and their relationships has the potential to advance current levels of system capability.

The paper has also identified the significance of semantic definition in manufacturing knowledge sharing and described two areas of research were the use of languages with mathematically defined semantics are being explored. It is believed that this approach has significant potential for the future if flexibility in the implementation of knowledge sharing tools can be achieved.

7. ACKNOWLEDGEMENTS

The authors would like to acknowledge the useful interactions with NIST personnel in building an understanding of PSL and especially Michael Gruninger. We would also like to acknowledge the Scholarships provided by Loughborough University and the EPSRC funded IMCRC at Loughborough University which is funding the current work of two related projects; “Knowledge Representation and Re-use for Predictive Design and

Manufacturing Evaluation" and "Cross-Disciplinary Communication Languages for Supply Chain Management".

8. REFERENCES

1. M. Abramovici, O.C. Sieg, Status and development trends of product lifecycle management systems, International conference on integrated product and process development. Wroclaw, Poland, 21-22 November 2002, ISBN 83-7085-667-5.
2. M. Uschold, M. Gruninger, Ontologies: Principles, Methods and Applications. Knowledge Engineering Review 11 (2)
3. J. Fowler, STEP for data management, exchange and sharing. Technology Appraisals UK, ISBN 1871802369, 1995.
4. T. King, Early practical realization of product lifecycle support. Product Data Technology Conference Turin, 7-9 May 2002.
5. C. Sakamoto, L. Slovensky, Product data representation and exchange: application protocol; process plans for machined products. ISO/CD 10303-240, 2003.
6. ISO TC 184/SC4. Industrial automation system and integration – Industrial manufacturing management data: Part 1: General overview, ISO FDIS 15531-1, 2002
7. R.I.M. Young, O. Canciglieri-Jnr, C.A. Costa, M. Dorador, J. Zhao, W. M. Cheung, Information Support in an Integrated Product Development System, in integrated Design and Manufacturing in Mechanical Engineering, Chedmail, Cognet, Fortin, Mascle, Penga (eds) Kluwer Academic Publishers, ISBN 1-4020-0979-8, 2002, pp 139 – 146
8. S.R. Ray, A.T. Jones, Manufacturing Interoperability. Proceedings of the Concurrent Engineering: Research and Applications. Enhanced Interoperability Systems, Ed. R. Jardim-Goncalves, J. Cha, A. Steiger-Garcao, Pub. Balkema, Netherlands 2003, ISBN 9058096238. pp 535- 540
9. J. Harding, A knowledge representation model to support concurrent engineering team working. PhD Thesis, Loughborough University, Loughborough, LE11 3TU, UK. 1996
10. J.J. Mills, J. Goossenaerts, Towards Information and knowledge in product realization infrastructures. Global Engineering, Manufacturing and enterprise networks ed MO J., Nemes L., pub Kluwer academic, Boston, ISBN 0-7923-7358-8, 2001, pp 245-254
11. R.I.M. Young, J.M. Dorador, J.Zhao, W.M.Cheung, A manufacturing infrastructure to link team based design to global manufacture, Global Engineering, Manufacturing and Enterprise Networks. Mo J., and Nemes L. (eds) Kluwer Academic Publishers, 2001,. ISBN 0-7923-7358-8, pp283-290
12. M. Gruninger, A.F. Cutting-Decelle, Industrial automation system and integration – Process specification language: Part 1: Overview and basic principles, ISO TC184/SC4/JWG8, ISO DIS 18629-1, 2001.
13. I. Niles, A. Pease, Towards a Standard Upper Ontology. Teknowledge Corporation, Palo Alto, CA 94303. [Web Accessed October 2003]. http://suo.ieee.org

Part 2

INTEGRATED DESIGN OF MANUFACTURING PROCESSES

MANUFACTURING PROCESS SIMULATION FOR TOLERANCE ANALYSIS AND SYNTHESIS

François Villeneuve and Frédéric Vignat
University of Grenoble, Laboratory 3S, BP53, Domaine universitaire, 38041, Grenoble cedex 9, France

Abstract: The choice of a manufacturing process (machine tools, cutting tools, part-holders, process plan) which allows the production of parts conform with the objective in terms of quality and price depends on many criteria. One of these is the compliance with the functional part tolerances. To check this point, the manufacturing process is simulated (positioning and machining operations). This simulation results in a Model of Manufactured Part (MMP). The deviations of all the MMP surfaces are described with regards to their nominal position (nominal part) using a small displacement torsor. Then, for all the functional tolerances to be checked, a virtual gauge is designed and developed and assembled with the MMP and the gap between the tolerance zone and the toleranced surface is calculated. The verification of each tolerance is performed via a system of inequalities representing the non-penetration conditions between the toleranced surface and the tolerance zone limits. The ability to machine a series of parts close to the functional tolerances being verified, the ISO or non-ISO manufacturing tolerances to specify the workpiece at the end of each set-up is determined. These tolerances are useful to follow the conformity of the workpiece at the end of each step of the machining process.

Key words: Manufacturing process, Manufacturing defects, Tolerance analysis, Simulation, Tolerance synthesis.

1. INTRODUCTION

Designing a new product in an industrial context supposes the certainty to be able to produce it with an improved quality and at a controlled cost. To achieve these goals, it is necessary to design the product and the process together. It is also necessary to be able to simulate the process designed to get information about the likely productivity and quality to be expected. The

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 189-200.

process will be considered suitable if it verifies the criteria imposed (by the client) in the field of productivity and quality. Complying with the functional tolerances for each manufactured part is in fact one of the quality criteria.

The tools most widely used (tolerances chain) [1] assume that the form and orientation defects are negligible in comparison with the dimensional defects. Each dimension can thus be modelled by a vector. Projecting these vectors, algebraic relations between various dimensions are calculated. These relations can be used to determine the manufacturing tolerances that verify the functional tolerances imposed. It is also used to distinguish the functional tolerances from the functional requirements. It can also be used to simulate the feasibility of a part starting from manufacturing process defects (dispersions) as in the Δl method. Several methods were then developed to enlarge transfer or simulation to 2D or 3D [2]. The works of Ballot and Bourdet [3] model the interactions between the parts of a mechanism, in order to predict the position and orientation variations of these parts in a three dimensional space. The variations are supposed to be small enough to use the concept of the Small Displacement Torsor (SDT) [4] except for the expected movements of parts. This method has been enlarged since 1999 to simulate the manufacturing process. F. Villeneuve and O. Legoff have developed a method using the SDT to simulate a machining process and to predict its defects. They have applied this method to an example of milling process plan [5]. The clearance and deviation spaces method [6] allows, just like the previous one, to simulate the influence of defects. The deviation spaces due to the defects and the clearance spaces due to the clearance between parts of the assembly are determined. The calculation of these spaces is achieved by the sum of Minkowski of the spaces due to the initial defects or clearances. To check the functioning of the assembly, it is then sufficient to check that the deviation space is contained within the clearance space. The applications presented only concern functional tolerance problems. Morse and Zou use quite a close method for tolerance transfer from functional to 1D manufacturing using a statistical approach [7]. The ΔL method has also been improved taking into account a three dimensional analysis, particularly the orientation defects [8]. This method proposes a process plan graph, a tool that includes both the process and positioning deviations and the functional tolerances. The defects description is carried out by a SDT. Checking the functional tolerances is done by the position analysis of the surface under consideration into a tolerance zone. The defects description and the analysis of each functional tolerance are briefly presented and seem to be somehow difficult to implement.

For all these approaches, the main problem consists in determining which surfaces and which set-up are implied in complying with a functional condition. The method presented hereafter proposes a solution to this problem.

The method hereby presented simulates the process and determines its effect onto the manufactured part in terms of surfaces deviations. To that aim, a Model of Manufactured Part (MMP) is virtually manufactured and the deviations of its surfaces are expressed with regards to their nominal position (nominal part). The nominal part is fixed in the MMP and allows to stack the defects generated at each set-up of the process. The MMP is one of the manufactured parts model . Knowing the 3D capabilities of the production means, it is possible to determine the probability density function or the extremes of the deviations of all the parts produced.

The MMP is then used to analyse the functional tolerances. Each tolerance is interpreted and a virtual gauge is created and assembled with the MMP. The next point is to check if the toleranced surface of the MMP is in the tolerance zone of the virtual gauge. This task is performed for the extremes MMP or a probability density can be calculated.

Once the process feasibility has been checked, there remains to find the manufacturing tolerances that will permit to follow the conformity of the workpiece at the end of each set-up. For each tolerance, starting from the verification result, the active (positioning and manufactured surfaces of each set-up) surfaces are identified and a set of ISO or non-ISO tolerances is proposed. The value of each tolerance is then calculated.

For this paper, to simplify the presentation, a 2D example of a manufacturing part has been chosen. However the method presented here has been developed to simulate 3D manufacturing processes.

2. THE METHOD

The method presented here is based on the determination and the analysis of the deviation of the surfaces of a part with regards to their nominal position (nominal part). The method is developed in two steps (Fig. 1):

- The first step consists in determining the effects of the process in terms of part surface deviation. At the end of this step, a model of the manufactured part with deviation defects is generated (it will be called MMP for Model of Manufactured Part).
- The second one consists in analysing the incidence of the process over the compliance with the functional tolerances (analysis) using the MMP and transferring these tolerances into manufacturing tolerances.

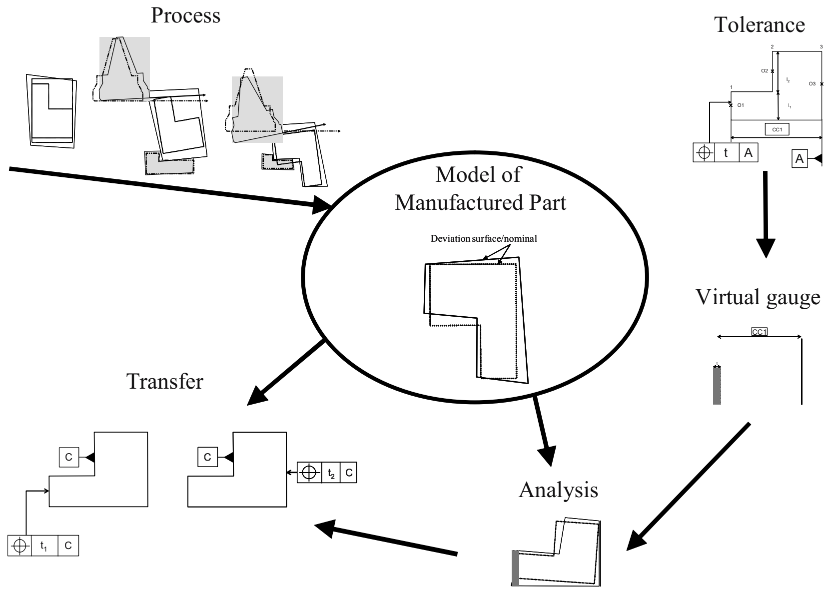


Figure 1. Overview of the method

The deviations of the MMP surfaces are expressed with regards to their nominal position in the part (nominal part). This is the key point of the method and this article will show that:

- The existence of a nominal part allows to collect throughout the process the defects generated by the successive operations.
- The position of the nominal part in the real part has no importance because the deviations are always at least grouped by 2 when a tolerance is analysed or transferred.

These deviations can be expressed by a small displacement torsor and the variations of the intrinsic characteristics of the surfaces (diameter in the case of a cylinder or a sphere). The variation domain of the parameters characterising these deviations represent the 3D capabilities of the machine-tools, the cutting tools and the part holders used during the machining process.

The MMP is composed of ideal surfaces that have to be associated with the real surfaces using an optimisation criterion (e.g least square criterion).

3. FROM THE PROCESS TO THE MANUFACTURED PART MODEL

The manufacturing process chosen for the part in Fig. 1 includes 2 set-ups. Starting from a block, the process consists in 2 successive milling operations. This process is described in Fig. 2. The surface positioning system is hyperstatic but organized in a hierarchical way (primary surface, secondary surface, ...)

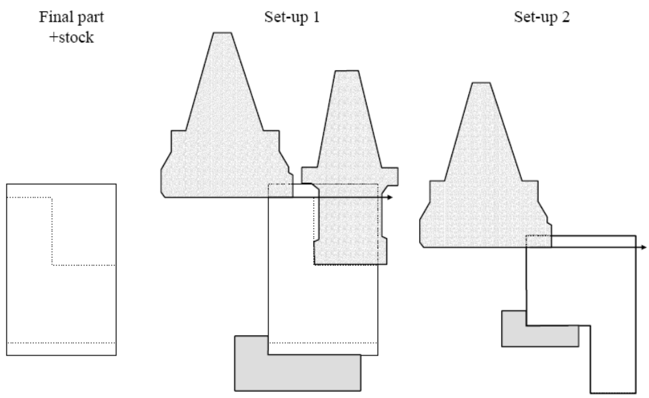


Figure 2. Nominal process

A process generates positioning deviations (due to the deviations of the part-holder surfaces and of the MMP surfaces) and machining deviations (deviation of a surface manufactured with regards to its nominal position in the machine) as represented in Fig. 3.

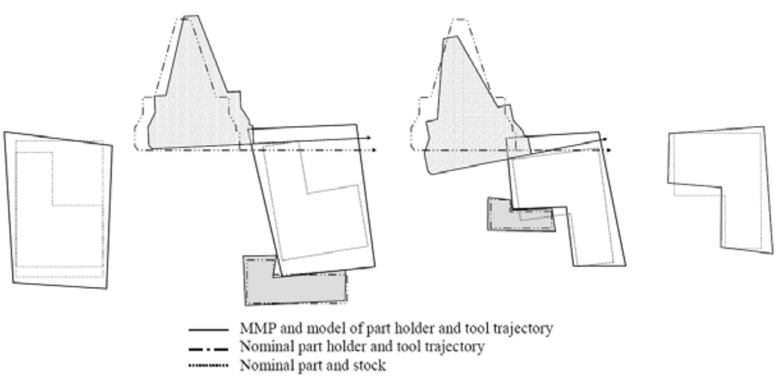


Figure 3. Process deviations

The deviation of each surface machined can be determined with regards to the nominal part by expressing for each set-up:

$$\text{Deviation}_{\text{realised surface/nominal part}} = \text{Deviation}_{\text{positioning}} + \text{Deviation}_{\text{machining}} \quad (1)$$

or, using a more compact expression

$$\text{Deviation}_{\text{part,surface}} = -\ \text{Deviation}_{\text{machine,part}} + \text{Deviation}_{\text{machine,surface}} \quad (2)$$

This operation must be repeated for each set-up until the final part is obtained.

For the process described in this paper, the stock is a parallelipipedic imperfect block with deviations of its surfaces with regards to the nominal stock. This stock is assembled with the part-holder in set-up 1. The part-holder surfaces have a deviation with regards to their nominal position in the machine. The determination of the position of the part due to this assembly operation is processed using the method developed by Ballot and Bourdet [3] and Thiebaud [10] for a standard mechanical system. The position of the part with regards to its nominal position in the machine is function of the deviation of the part-holder and part positioning surfaces and of the gap between these surfaces. The values of these gaps are determined by the contact properties (primary, secondary, fixed, floating). The part is then "machined" and the surfaces thus obtained have deviations with regards to their nominal position in the machine. Using equation (2) the deviation of each surface with regards to the nominal part can be determined. This task is repeated for each set-up until the deviation of all the surfaces of the part is obtained. Developments of this method have already been presented [5, 9].

The deviations of each surface depend only on the process chosen and can be expressed by a small displacement torsor. For example for surface 3:

$$T_{P,P3} = \begin{Bmatrix} u_3 + u_2 - u_{2S2} + \frac{l_1}{2}(\gamma_2 - \gamma_{2S2}) & U \\ U & 0 \\ U & \gamma_3 + \gamma_2 - \gamma_{2S2} \end{Bmatrix}_{O_3,\vec{x},\vec{y},\vec{z}}$$

This torsor describes in a generic way the deviations generated by the process on a part. Each defect parameter (u_1, u_2, …, γ_1, …) can be identified on the manufacturing means and can be expressed by a variation domain (or a probability density function). Thus it is possible to know the domain of variation of the surfaces of all the manufactured parts.

4. TOLERANCE ANALYSIS

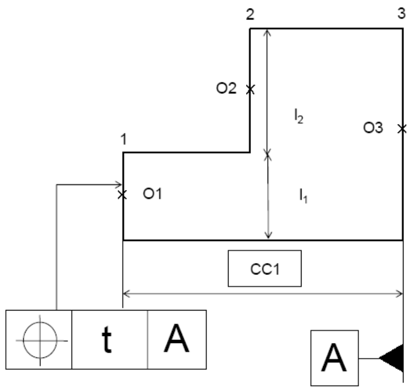


Figure 4. Tolerance of the part

In ordre to carry out a functional tolerance analysis, the first step consists in creating a virtual gauge corresponding to this tolerance in compliance with the ISO standard [10]. The virtual gauge corresponding to the position tolerance Fig. 4 is a datum plane and a tolerance zone (TZ) limited by 2 planes distant of t (Fig. 5). The second step consists in assembling the MMP and the virtual gauge abiding by the constraints imposed by the ISO standards (hierarchy of the surfaces, type of contact). For the example under study we have a fixed planar contact between surface 3 of the MMP and the datum plane of the virtual gauge. In the third step, the gap between the toleranced surface and the tolerance zone is determined using:

$$Gap_{TZ,TolerancedSurface} = Deviation_{TZ,Datum} + Deviation_{Datum,Part} + Deviation_{Part,TolerancedSurface}$$

where the assembly operation gauge part can be expressed by

$$Deviation_{D,P} = Gap_{D,PartDatumSurface} + Deviation_{PDS,P}$$

and the virtual gauge being a perfect part

$$Deviation_{TZ,D} = 0$$

so this gap can be calculated by

$$Gap_{TZ,TS} = Gap_{D,PDS} + Deviation_{PDS,TS} \quad (3)$$

For the present example there is a primary planar fixed contact between the datum plane of the gauge and the datum surface of the part. So we can write:

$$Gap_{D,PDS} = 0$$

and (3) become

$$Gap_{TZ,TS} = Deviation_{PDS,TS}$$

This deviation can be expressed by the following small displacement torsor

$$T_{P3,P1} = \begin{Bmatrix} u_1 - u_3 - u_2 + u_{2S2} - l_1/2(\gamma_2 - \gamma_{2S2}) \\ -l_2/2(\gamma_3 + \gamma_2 - \gamma_{2S2}) & U \\ U & 0 \\ U & \gamma_1 - \gamma_3 \\ & -\gamma_2 + \gamma_{2S2} \end{Bmatrix}_{O_1,\vec{x},\vec{y},\vec{z}}$$

The last step is the verification of the ability of the process to produce parts conform with the functional tolerance. This is done by verifying that the $Gap_{TZ,toleranced\ surface}$ comply with the non-penetration condition between the limits of the tolerance zone and the toleranced surface.

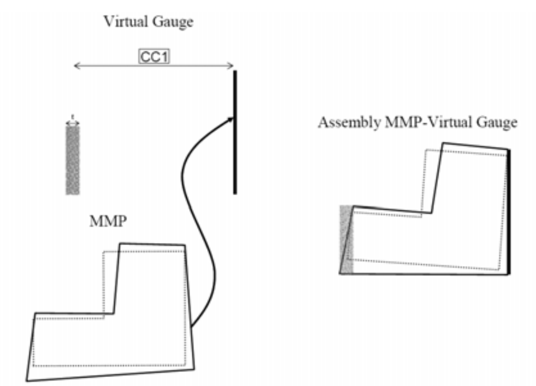


Figure 5. Tolerance analysis using a virtual gauge

This verification is expressed by a system of inequalities to check. For the example, this system of inequalities expresses that the upper and lower point of the ideal toleranced plane of the MMP do not intrude into the limits of the tolerance zone. So, 4 inequalities are obtained:

$$\begin{cases} u_1 - u_3 - u_2 + u_{2S2} - \frac{l_1}{2}(\gamma_2 - \gamma_{2S2}) - \frac{l_2}{2}(\gamma_3 + \gamma_2 - \gamma_{2S2}) + \frac{l_1}{2}(\gamma_1 - \gamma_3 - \gamma_2 + \gamma_{2S2}) < t/2 \\ u_1 - u_3 - u_2 + u_{2S2} - \frac{l_1}{2}(\gamma_2 - \gamma_{2S2}) - \frac{l_2}{2}(\gamma_3 + \gamma_2 - \gamma_{2S2}) - \frac{l_1}{2}(\gamma_1 - \gamma_3 - \gamma_2 + \gamma_{2S2}) < t/2 \\ u_1 - u_3 - u_2 + u_{2S2} - \frac{l_1}{2}(\gamma_2 - \gamma_{2S2}) - \frac{l_2}{2}(\gamma_3 + \gamma_2 - \gamma_{2S2}) + \frac{l_1}{2}(\gamma_1 - \gamma_3 - \gamma_2 + \gamma_{2S2}) > -t/2 \\ u_1 - u_3 - u_2 + u_{2S2} - \frac{l_1}{2}(\gamma_2 - \gamma_{2S2}) - \frac{l_2}{2}(\gamma_3 + \gamma_2 - \gamma_{2S2}) - \frac{l_1}{2}(\gamma_1 - \gamma_3 - \gamma_2 + \gamma_{2S2}) > -t/2 \end{cases}$$

This system, called functional inequalities system, has to be verified whatever the value of the parameters within their domain of variation in order to claim that the process chosen is able to produce conform parts.

5. TOLERANCE SYNTHESIS

The starting point of the tolerance synthesis is the inequalities system determined by tolerance analysis (functional inequalities system). It is also necessary to know the source of the defects (which set-up is concerned) and the portion (whole, lines, points, ...) of the surfaces used for positioning in each set-up.

From the functional inequalities system, a manufacturing inequalities system is determined for each set-up. These new systems limit the defect allowed for each set-up to produce a final part conform with the functional tolerance. In order to abide by this rule, the limits of these inequalities must guarantee the respect of the limits of the functional inequalities system [11].

In the present example, the functional inequality hereafter:

$$u_1 - u_3 - u_2 + u_{2S2} - \frac{l_1}{2}(\gamma_2 - \gamma_{2S2})$$
$$-\frac{l_2}{2}(\gamma_3 + \gamma_2 - \gamma_{2S2}) + \frac{l_1}{2}(\gamma_1 - \gamma_3 - \gamma_2 + \gamma_{2S2}) < \mathit{limit}$$

becomes

$$u_1 - u_2 - \frac{l_1}{2}(\gamma_2) - \frac{l_2}{2}(\gamma_2) + \frac{l_1}{2}(\gamma_1 - \gamma_2) < \mathit{limit1} \text{ for set-up 1}$$

$$-u_3 + u_{2S2} - \frac{l_1}{2}(-\gamma_{2S2}) - \frac{l_1 + l_2}{2}(\gamma_3 - \gamma_{2S2}) < limit2 \text{ for set-up 2}$$

In this case, if *limit1* + *limit2* = *limit*, the functional tolerance is complied with. *Limit1* and *limit2* represent the functional tolerance transfer. These manufacturing inequalities systems can be considered non-ISO manufacturing tolerances for each set-up. It may be sufficient to express the needs for the manufacturing process.

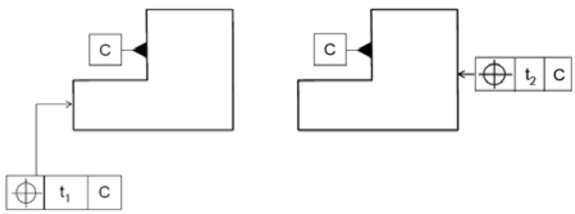


Figure 6. A solution of manufacturing tolerances

If one considers that it is better to determine ISO manufacturing tolerances, the first step is to identify the surfaces to be specified for each set-up. These surfaces are indicated by the parameters limited by the manufacturing system of inequalities. Second step, a set of manufacturing ISO tolerances has to be proposed for each set-up (Fig. 6). The manufacturing inequalities provide information on the type of tolerances likely to be convenient for the set-up. Some methods can be proposed to automatically propose a suitable type of tolerances starting from the DOF to constrain [12]. Third step, for each tolerance, a virtual gauge is assembled with the Model of Workpiece (MWP, model of the part at the end of each set-up) to generate a ISO manufacturing inequalities system.

For example

$$-t_1/2 < u_1 - u_2 - \frac{l_1}{2}(\gamma_2) - \frac{l_2}{2}(\gamma_2) \pm \frac{l_1}{2}(\gamma_1 - \gamma_2) < t_1/2 \text{ for set-up 1}$$

$$-t_2/2 < u_3 - u_{S2} + \frac{l_1}{2}(-\gamma_{2S2}) \pm \frac{l_1 + l_2}{2}(\gamma_3 - \gamma_{2S2}) < t_2/2 \text{ for set-up 2}$$

The limits chosen for these ISO inequalities (here t_1 and t_2) must guarantee the compliance with the functional inequalities system and thus with the functional tolerance. The last point is to determine the best set of manufacturing tolerances constraining each set-up as less as possible.

6. CONCLUSION

The method presented simulates a manufacturing process and determines its effect on the manufactured part in term of surfaces deviations. A model of manufactured part (MMP) was introduced. It is virtually manufactured and the deviations of its surfaces are expressed with regards to their nominal position. The nominal part is fixed in the MMP and allows to stack the defects generated at each set-up of the process. Knowing the 3D capabilities of the production means, it is then possible to determine the probability of the deviations of all the parts produced. The MMP is then used to analyse the functional tolerances. Then, for each tolerance, starting from the analysis results, the active surfaces of each set-up are identified and a set of manufacturing ISO or non-ISO tolerances is proposed.

Most of the similar approaches in bibliography are confronted with the problem of determining which surfaces and which set-up are involved in the compliance with a functional condition. The method presented proposes a 3D solution for this problem. It is important to keep in mind that the nominal part is the link between the different set-ups and that the position of this nominal part in the real part has no importance because the deviations are always at least grouped by 2 when a tolerance is analysed or transferred.

We showed, in previous papers [5, 9, 11] that it is possible to express our method using Small Displacements Torsors. We deliberately did not develop the mathematical expressions in this paper to propose an overall view of the method. However, some important points are still to be developed:

- For the moment, the tolerance analysis is proposed in the worst case. It seems to be more interesting to consider a statistical approach. But we have not defined yet how to use the SDT parameters in a statistical approach.
- The parameters of such a model need to be identified in terms of manufacturing tools. Some experiments have already been conducted [13], but it seems to be necessary to go further into this work direction.
- To automate the choice of ISO manufacturing tolerances should be a very important help for process planner. We have only proposed today some heuristics to solve that problem. On the other hand, we are able to identify which surfaces and which set-up influence the compliance with a functional tolerance.
- Proposing some help for the designers is one of our goals. We hope that the solutions presented here contribute to that goal, proposing a link between functional needs and manufacturing means. However, a graphical approach combined with a PDM should be necessary to express the process plans and to help define the product functional needs.

REFERENCES

1. B. Anselmetti, P., Bourdet, Optimisation of a workpiece considering production requirements, Computers in Industry, 21, pp. 23-34, 1993.
2. A. Desrochers, Modelling three-dimensional tolerance zones using screw parameters; In: CD-ROM proc. of 25th ASME design automation conference; Las Vegas 1999
3. E. Ballot, P. Bourdet, Geometrical Behaviour Laws for Computer Aided Tolerancing, Proceedings of the 4th CIRP Seminar on Computer Aided Tolerancing, Tokyo, Japan, pp. 143-154, 1995.
4. P. Bourdet, L. Mathieu, C. Lartigue, A. Ballu, The Concept of The Small Displacement Torsor in Metrology, In: Advanced Mathematical Tools in Metrology II, Series on Advances in Mathematics for applied Sciences, 40; pp. 110-122, 1996.
5. F. Villeneuve, O. Legoff, Y.Landon, Y., Tolerancing for manufacturing: a three dimensional model; In: International Journal of Production Research, Vol. 39, N°8; pp. 1625-1648, 2001
6. M. Giordano, B. Kataya, E. Pairel, Tolerance analysis and synthesis by means of clearance and deviation space, Proceedings of the 7th CIRP seminar on Computer Aided Tolerancing, pp. 345-354, 2001.
7. Z.H. Zou, E.P. Morse, Statistical Tolerance Analysis Using GapSpace, 7th CAT CIRP Seminar, ENS de Cachan, pp. 313-322, France, 2001
8. R. Benea, G. Cloutier, C. Fortin, Process plan validation including process deviations and machine tools errors; In: Proceedings of the 7th CIRP Seminar on Computer Aided Tolerancing; pp. 191-200; Cachan 2001
9. F. Vignat, F. Villeneuve, 3D transfer of tolerances using a SDT approach: application to turning process; In: International Journal of Computing and Information Science in Engineering, Vol. 3, pp. 45-53, 2003
10. F. Thiebaut, Contribution à la définition d'un moyen unifié de gestion de la géométrie réaliste basée sur le calcul des lois de comportement des mécanismes, thèse de Doctorat de l'ENS de Cachan, 2001
11. F. Villeneuve, F. Vignat, 3D synthesis of manufacturing tolerances using a SDT approach, in: Proceedings of the 8th CIRP Computer Aided Tolerancing Seminar, CAT2003, Charlotte, NC, USA, 29-30/04/2003
12. W. Yanyan, J.J. Shah, J.K. Davidson, Computer modeling of geometric variations in mechanical parts and assemblies, In: International Journal of Computing and Information Science in Engineering, Vol. 3, pp. 54-63, 2003
13. O. Legoff, F. Villeneuve, Three-dimensional geometrical tolerancing: quantification of machining defects, in: Geometric Product Specification and Verification, selected papers book of 7th CIRP CAT, Kluwer Academic Publishers, pp. 185-196, 2003

3D GEOMETRICAL MANUFACTURING SIMULATION

Compared approaches between integrated CAD/CAM systems and small displacement torsor models

Stephane Tichadou, Olivier Legoff and Jean-Yves Hascoet

IRCCyN, Institut de Recherche en Communications et en Cybernetique de Nantes, UMR CNRS 6597, 1 Rue de la Noe BP 92101, 44321 Nantes Cedex 3, France.

Abstract: In the context of tolerance allocation in process planning, our interest has focused onto manufacturing defects 3D modeling . This paper presents two approaches towards geometrical manufacturing simulation. The concept of small displacement torsor is used in order to model geometrical defects of part surfaces. The first approach uses a CAD modeler in order to extract the geometrical data for mathematical processing using a symbolic computational software. The model that we developed for manufacturing considers each manufacturing set-up as a single mechanism. A chart representation has been defined to express the geometrical manufacturing conditions of the process plan as a torsor chain. The second approach uses the features of an industrial CAD/CAM system in order to reproduce and industrialize this model. Then we carry out a virtual numerical metrology of the model thus obtained. Through examples, we present comparisons of numerical results from both applications. The approaches converge towards the same results. Finally we propose a comparison table which enables to identify the advantages of each of the two approaches.

Key words: Process planning; Geometric modeling and simulation; Virtual manufacturing; Tolerance chart; Small displacement torsor; CAD/CAM.

1. INTRODUCTION

In the industrial context of integrated design and manufacturing, tolerance charting represents an essential step in process planning. This activity is used to optimize the process choices by dimensioning the

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 201-214.

acceptable geometrical variations on the part while matching the functional and manufacturing constraints. Knowing the process plan, the usual method proceeds in two stages: Simulation (appraisal and validation of a process by predicting and estimating the minimal manufacturable values for each tolerance) and optimization (for instance enlarge some tolerances in order to reduce manufacturing costs). This paper presents two approaches for the geometrical manufacturing simulation. The concept of Small Displacement Torsor (SDT) is used to model geometrical defects of part surfaces. The first approach uses a CAD modeler in order to extract the geometrical data then used to generate the equations needed for a mathematical processing (synthesis and analysis of manufacturing tolerances). The second one aims at using a CAD/CAM system's functionalities in order to reproduce this model.

2. PROCESS PLANNING 3D MODELING

2.1 SDT Model

Several research works show that the use of the small displacement torsors allows to model the defects of the parts in mechanisms [1]. The hypotheses associated with this model are: the parts are rigid; the displacements are small and the form defects of surfaces are negligible. A SDT is defined by two vectors representing the values of three rotations and three translations of a surface. The displacement of a plane P compared to its nominal position in its local coordinate system Rp (origin Op on the surface and vector z the normal of the plane) is defined by:

$$\{\mathbf{T}_P\}_{Op} = \{\boldsymbol{D}_{Op}, \boldsymbol{\Omega}\} \text{ with: } \Rightarrow \begin{array}{l} \boldsymbol{\Omega} = \mathrm{U}\boldsymbol{x}_{Rp} + \mathrm{U}\boldsymbol{y}_{Rp} + \mathrm{w}\boldsymbol{z}_{Rp} = (\mathrm{U}, \mathrm{U}, \mathrm{w})_{Rp} \\ \boldsymbol{D}_{Op} = \alpha\boldsymbol{x}_{Rp} + \beta\boldsymbol{y}_{Rp} + \mathrm{U}\boldsymbol{z}_{Rp} = (\alpha, \beta, \mathrm{U})_{Rp} \end{array}$$

where U represents the undetermined components. Thus, the displacement of any point P_i of the plan is obtained by: $\boldsymbol{D}_{Pi} = \boldsymbol{D}_{Op} + \boldsymbol{\Omega} \wedge \boldsymbol{OpP_i}$.

2.2 Application to manufacturing

The model developed for manufacturing [2] considers each manufacturing set-up as a mechanism comprising (Fig. 1): a machine-tool MT, a part-holder H and its surfaces Hh, several manufacturing operations Mm, their surfaces Mmj and the active surfaces Pi of the workpiece P (location and machined surfaces).

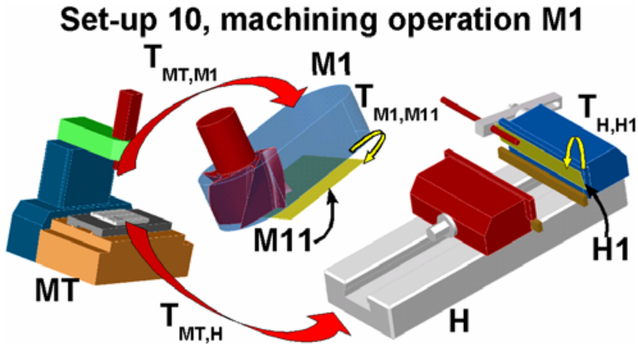


Figure 1. SDTs at set-up 10, machining M1.

On the charts of Fig. 2 and 3 each pole represents a geometrical entity and each arc a SDT characterized potential defect. If we want to define the defects of the part-holder surfaces relatively to the machine-tool using the *full model*, we need to be able to measure $\mathbf{T}_{MT,MTt}$, $\mathbf{T}_{MTt,Hh}$, $\mathbf{T}_{Hh,H}$, and $\mathbf{T}_{H,Hh}$ with $t \in \{1, 2\}$ and $h \in \{1, 2, 3, 4, 5\}$. In practice, it is easier and faster to measure $\mathbf{T}_{MT,Hh}$ directly on a machine-tool but the SDT model refers to the nominal geometry. Thus, we chose to use a *compacted model* (Fig. 2) with $\mathbf{T}_{MT,Hh} = \mathbf{T}_{MT,H} + \mathbf{T}_{H,Hh}$ for the part-holder and $\mathbf{T}_{MT,M} = \mathbf{T}_{MT,Mm} + \mathbf{T}_{Mm,Mmj}$ for the machining operations. Three types of torsors are used [3]:

- **SDT Deviation**: displacement of a surface relative to its nominal position: $\mathbf{T}_{H,Hh}$, $\mathbf{T}_{P,Pi}$, $\mathbf{T}_{Mm,Mmj}$ (undetermined components of this kind of torsor are not useful, so they are all set to 0).
- **SDT Connection**: displacement caused by the connection between two surfaces of two parts: $\mathbf{T}_{Hh,Pi}$, $\mathbf{T}_{Mmj,Pi}$.
- **Overall SDT**: displacement of a part: $\mathbf{T}_{MT,H}$, $\mathbf{T}_{MT,Mm}$.

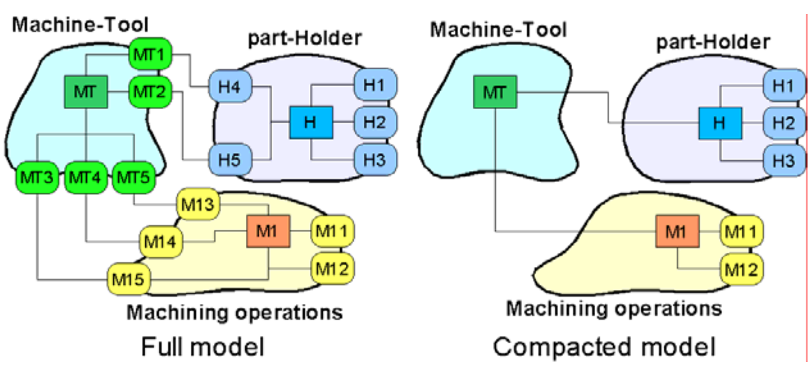


Figure 2. Full and compacted model of a set-up

In the machining context, the joints are organized hierarchically, i.e. the main contact $\mathbf{T}_{H1,Pi}$ is ensured having priority over the secondary one $\mathbf{T}_{H2,Pj}$ and so on. This is very useful during the unification operation which will be detailed further down in this paper. The complete graph of a manufacturing set-up (Fig. 3) considers the part only through its active surfaces because these are the only surfaces whose positions are controlled in the set-up. The data structure defined in those charts follow the following rules:

- Part-holder surface Hh: h numbers indicate the contact hierarchy (i.e. H1 is the main contact surface an so on).
- Machining operations Mm: m indicates the machining operation sequence (M1 is machined first).
- Machining operation surfaces Mmj: j is only a surface number, a different one for each surface.

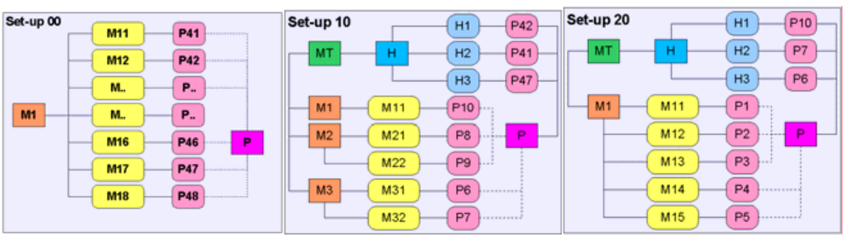


Figure 3. Charts of three set-ups.

We assume that each **Geometric Manufacturing Condition** can be specified with constraints on torsor components [4-5] (GMCs being made up with stock removal allowances and design tolerances). So, in order to check the feasibility of the process plan (tolerance analysis) or to optimize it (tolerance synthesis), the SDT related to each GMC has to be validated. Using the charts associated to the set-ups each one of these SDTs can be automatically written as a torsor chain (like a dimension chain by using a traditional 1D tolerance chart modeling). The expression of the defect between two surfaces of the finished part $\mathbf{T}_{Pi,Pj}$ is obtained by following the path (sequence of arcs and poles on charts) starting from the last manufactured surface Pi up to the first one Pj. If the first manufactured surface is not on same chart as the other one, this means that the two surfaces have been machined in different set-ups. In this case, it is necessary to pass through the location faces and to restart the same process. During the tolerance analysis of process planning, two classes of GMCs are considered related to the set-up planning: *direct* GMCs (those for which the two surfaces involved are active within the same set-up) and *indirect* GMCs (between faces that are not active in the same set-up, i.e. using an intermediate set-up datum to machine the two surfaces in different set-ups).

2.3 Example

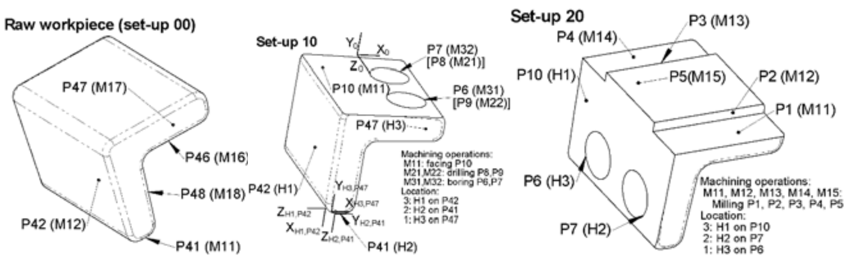


Figure 4. Manufacturing process.

The example used to illustrate our approach the part of the process plan shown on Fig. 4. The charts of the three machining set-ups 00 (in this set-up all surfaces are manufactured simultaneously), 10 and 20 are shown on Fig. 3. Moreover, we consider thereafter that the location of the workpiece on the part-holder is hyperstatic. All contact surfaces between the workpiece and part-holder are of the same type, i.e. in set-up 10: H1, H2 and H3 are considered as planes.

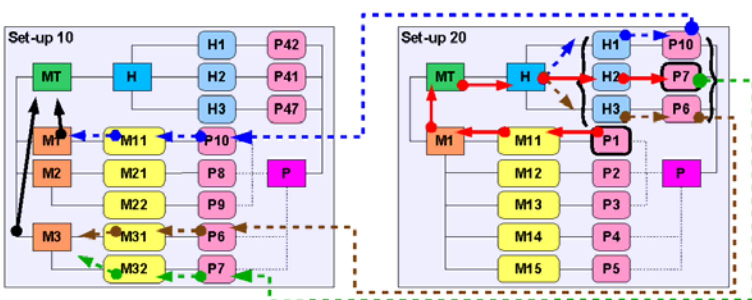


Figure 5. Path from P1 to P7.

Let us consider an example of a *direct* GMC between machined faces and location faces in the same set-up: $\mathbf{T}_{P1,P7} = -\mathbf{T}_{20M11,P1} - \mathbf{T}_{20M1,M11} - \mathbf{T}_{20MT,M1} + \mathbf{T}_{20MT,H} + \mathbf{T}_{20H,H2} + \mathbf{T}_{20H2,P7}$ is obtained by following the path from P1 to P7 on the chart of set-up 20 (Fig. 5). But, as H2/P7 is the secondary contact in set-up 20, a unification operation (this operation is detailed further) is performed in order to consider the primary contact H1/P10 and the relationships between P10 and P7 in the previous set-up (Fig. 5). So, our 3D model shows that the GMC is not so *direct* as the traditional 1D tolerance chart modeling would have initially let us thought. Let us focus our study onto the angular component about z_0 (Fig. 4): $\boldsymbol{\Omega}_{P1,P7}\cdot z_0 = \alpha_{20M11} - \beta_{20H1} - \beta_{20H1,P10} - \beta_{10M32} + \beta_{10M11}$.

H1,P10 is the main contact, thus the expression for this contact is: $-\beta_{20H1}-\beta_{20H1,P10}$; the manufacturing defect of P1: α_{20M11}. H2,P7 being the secondary contact, the location of P7 depends on the relative position of P7/P1 which are both manufactured in the preceding set-up: $-\beta_{10M32}+\beta_{10M11}$. This is why $\mathbf{T}_{P1,P7}$ is not really what we call a direct GMC because it depends on a previous set-up.

Let us consider an *indirect* GMC $\mathbf{T}_{P10,P48}$, following the paths on charts (Fig. 6), we obtain three torsor chains:

$$\begin{cases} \mathbf{T}_{P10,P48} = \mathbf{T}_{10P10,H} + \mathbf{T}_{10H,H1} + \mathbf{T}_{10H1,P42} - \mathbf{T}_{00M12,P42} - \mathbf{T}_{00M1,M12} - \mathbf{T}_{00M1,P48} \\ \mathbf{T}_{P10,P48} = \mathbf{T}_{10P10,H} + \mathbf{T}_{10H,H2} + \mathbf{T}_{10H2,P41} - \mathbf{T}_{00M11,P41} - \mathbf{T}_{00M1,M11} - \mathbf{T}_{00M1,P48} \\ \mathbf{T}_{P10,P48} = \mathbf{T}_{10P10,H} + \mathbf{T}_{10H,H3} + \mathbf{T}_{10H3,P47} - \mathbf{T}_{00M17,P47} - \mathbf{T}_{00M1,M17} - \mathbf{T}_{00M1,P48} \end{cases}$$

with:

$$\mathbf{T}_{P10,H} = -\mathbf{T}_{10M11,P10} - \mathbf{T}_{10M1,M11} - \mathbf{T}_{10MT,M1} + \mathbf{T}_{10MT,H}$$

$$\mathbf{T}_{00M1,P48} = \mathbf{T}_{00M1,M18} - \mathbf{T}_{00M18,P48}$$

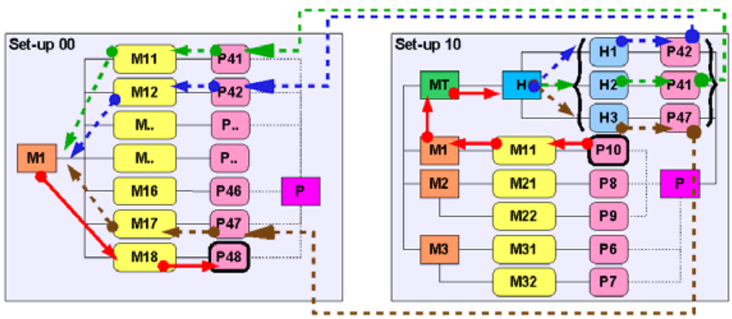


Figure 6. Path from P10 to P48.

The resolution of the 18 linear equations gives the needed $\mathbf{T}_{P10,P48}$ components. Let us consider condition $\mathbf{T}_{P5,P46}$: P5 is manufactured in set-up 20 and P46 in set-up 00 and P46 is never a location face. So, we have to pass through the location faces of the two machining sets-ups. This gives for $\mathbf{T}_{P5,P46}$: 9 torsor chains, i.e. a set of 54 linear equations. However, the resolution of all linear equation systems, related to all indirect GMCs, results in solving identical subsystems several times. These subsystems are those related to the unification of the workpiece location in a set-up. So, our strategy is to solve these subsystems before using their results in the initial equations. This process simplifies and speeds up computations during the resolution. In the example, two unifications have to be done, one for each set-up. To illustrate the method, the unification of the set-up 10 is developed. The location of the workpiece is defined by three similar SDTs:

$$\{\mathbf{T}_{10Hh,Pi}\}O_{10Hh,Pi} = \begin{cases} \boldsymbol{\Omega}_{10Hh,Pi} = (\alpha_{10Hh,Pi}, \beta_{10Hh,Pi}, U_{\gamma 10Hh,Pi})R_{10Hh,Pi} \\ \boldsymbol{DO}_{10Hh,Pi} = (U_{u10Hh,Pi}, U_{v10HhPi}, w_{10Hh,Pi})R_{10Hh,Pi} \end{cases}$$

with (h,i) ∈ {(1,42) , (2,41) , (3,47)}

The workpiece having a unique location on the part-holder and its location faces being not independent, they are defined by machining operations previously done. We can write that the relative position of two location faces in set-up 10 is the same as their relative position when manufactured: $\mathbf{T}_{10P42,P41} = \mathbf{T}_{00P42,P41}$ and $\mathbf{T}_{10P42,P47} = \mathbf{T}_{00P42,P47}$, this gives

$$\begin{cases} \mathbf{T}_{10P42,H}+\mathbf{T}_{10H,H2}+\mathbf{T}_{10H2,P41} = \mathbf{T}_{00P42,M12}+\mathbf{T}_{00M12,M1}+\mathbf{T}_{00M1,M11}+\mathbf{T}_{00M11,P41} \\ \mathbf{T}_{10P42,H}+\mathbf{T}_{10H,H3}+\mathbf{T}_{10H3,P47} = \mathbf{T}_{00P42,M12}+\mathbf{T}_{00M12,M1}+\mathbf{T}_{00M1,M17}+\mathbf{T}_{00M17,P47} \end{cases}$$

with $\mathbf{T}_{10P42,H} = \mathbf{T}_{10P42,H1}+\mathbf{T}_{10H1,H}$

Solving the system hierarchically (i.e. the equations related to the main location faces $\mathbf{T}_{10P42,P41} = \mathbf{T}_{00P42,P41}$ are solved before $\mathbf{T}_{10P42,P47} = \mathbf{T}_{00P42,P47}$ in which the first results are integrated) enables us to get: the expression of all undetermined components and the compatibility equations related to the hyperstatism of our machining fixture (3 equations in our case).

3. FIRST APPROACH: AUTOMATIC EXTRACTION OF THE GEOMETRICAL FEATURES

In order to solve the equations previously presented we use a numerical and symbolic calculations software (Mathematica). The resolution requires, as a preliminary step, the entry of all geometrical data (all SDTs related to the part-holder, the workpiece and the machining operations) while complying with the data structure we have defined with our charts. Thus, in the case of our example, if we consider only plane and cylindrical surfaces we obtain 47 surfaces. For each surface and for each contact workpiece/part-holder, it is necessary to define: a local coordinate system (an origin and at least two orthogonal 3D vectors) and the characteristic surface SDT (expressed in the local coordinate system at a chosen point, the origin, on the surface: two 3D vectors). We chose to extract all these data starting from a workpiece CAD model. First, it is necessary to re-design the workpiece by following the operation sequence of the process plan; then, all GMCs have to be defined. The objective of the feature extraction is to create a data structure identical to the one defined in our charts. We developed Visual Basic API within a CAD modeler based on an object-oriented approach. Here we do not detail all data structure of our VB API, the two main object classes are:

- *Geometrical feature*: the poles of our chart, composed by a nominal geometry and a set of surfaces.
- *Link*: the arc of our chart, defined by the SDTs (the SDT itself and its local coordinate system) linking the two *geometrical features* between it identifies a defect.

So, a process plan object integrates several set-up objects and a workpiece (*geometrical feature:* {P, Pi}). Each set-up can integrate a machine-tool, a part-holder and the machining operations (*geometrical features:* {MT}, {H, Hh}, {Mm, Mmj}). All these data and their related *links* are extracted directly by exploring the workpiece design tree and geometrical data. With regard to GMCs, they are computed according to the method we have presented previously after the extraction of the other data. They are plotted in tables of *links*. All the data extracted from the CAD model of the part into a file including the instantiation and the resolution of all the equations we need for the symbolic and numerical calculations. So, we only need to import this file into our computation software in order to use the model. In the case of our example, more than 300 equations are automatically generated.

4. SECOND APPROACH: CAD/CAM INTEGRATION

4.1 Description of the CAD/CAM integration

Usually the purpose of CAM systems is mainly to generate and validate tool paths. In our case it is used to create the solid resulting from the machining operations. Then we carry out measurements on this solid to note if it is in conformity. We use the CAM software like a “black box”. To simulate manufacture as well as possible, it is enough to integrate the geometrical and dimensional defects with an incidence in the set-up. Each element (MT,H, …) is modeled with its defects to lead to part P in its final geometrical state.

The necessary data are: the digital models of the rough part, the part-holder, the nominal finished part (P), the machine-tool and the tools. We start with the digital model of the rough part P(00) which we install in the part-holder of the first machining set-up 00, by complying with contact surfaces between part P(00) and part-holder with the defects (Fig. 7). After the definition of all machining operations for a set-up S, we can get the workpiece model in its final machined state P(S) from the CAD/CAM system. This new digital model then becomes an input datum for the following set-up.

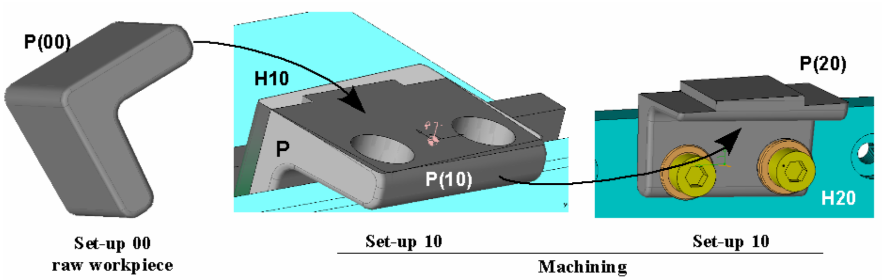


Figure 7. CAD/CAM integration.

The model of part P(S) obtained at the end of the set-up S depends on the machining operations of the set-up but also on part state P(S-1). The current CAD/CAM systems can update these dependencies dynamically. So, set-up after set-up (Fig. 7), by using this model as the stock of the next set-up P(S+1) we can finally get a geometric model of the finished part P(n) following the process plan. In this approach, the geometrical defects manufactured are integrated into each part's CAD model:

Stock:

- The geometrical variations of P(00) are modeled by adding geometrical and dimensional parameters to the CAD model of the rough part.

Machining:

- The positioning deviations (machine-tool/part-holder: $T_{MT,H}$ and part-holder/workpiece: $T_{Hh,Pi}$) are simulated by introducing a parameterized offset (rotation + translation): between P(s-1) and the part-holder H on the one hand, the machine-tool and H on the other hand.
- The machining defects are taken into account by parameterizing the tool shapes.

We compare P(n) with the nominal CAD model P in order to validate the manufacturing process. The whole set of the defects is described on a spreadsheet linked to the CAM files. Dimensions of the simulated workpiece are usable to check the conformity of the part. By using the CAD system, we carry out a virtual numerical metrology of the model obtained.

4.2 Examples of GMCs: distance line/plane

We can study the P7 axis parallelism and position compared with P1 (datum plane). We will analyze cylinder P7 with plane P1. Let us note that P7 is the secondary H2/P7 connection 20 (Fig. 8).

We saw $\mathbf{T}_{P1,P7} = -\mathbf{T}_{20M11,P1} - \mathbf{T}_{20M1,M11} - \mathbf{T}_{20MT,M1} + \mathbf{T}_{20MT,H} + \mathbf{T}_{20H,H2} + \mathbf{T}_{20H2,P7}$

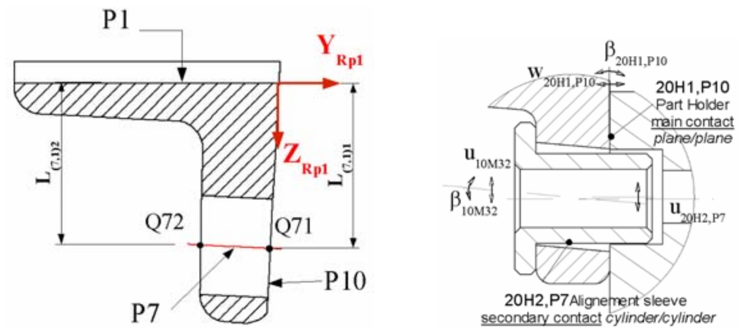


Figure 8. Parallelism and position (simplified two-dimensional view).

In view of the difficulties to quantify the deviations between the manufacturing operations [6] Mm and their surfaces Mmj during experiments, we initialize the SDT: $\mathbf{T}_{Mm,Mmj} = 0$. This cancellation does not simplify the model, but compacts it. Variations of P7 position compared with P1 are calculated by displacements:

$L_{(7,1)1} = \boxed{30} + \boldsymbol{DQ71}_{P1,P7} \cdot z_{Rp1}$ and $L_{(7,1)2} = \boxed{30} + \boldsymbol{DQ72}_{P1,P7} \cdot z_{Rp1}$ where $\boxed{30}$ is nominal size P1/P7.

After development, for $L_{(7,1)1}$ we find a linear equation with coefficients:

Coefficients		*Contribution*
β_{10M11}	–47.907	Influence variations generated at the time of the set-up 10 on the secondary contact of the set-up 20.
u_{10M31}	0.217	
u_{10M32}	–0.217	
β_{10M32}	47.907	
β_{20H1}	47.907	Main contact H1/P10.
$\beta_{20H1,P10}$	47.907	
u_{20H2}	1.217	Secondary contact H2/P7.
$u_{20H2,P7}$	1.217	
u_{20H3}	–0.217	Tertiary contact H3/P6.
$u_{20H3,P6}$	–0.217	
w_{20M11}	–1	M11 machining in set-up 20.
α_{20M11}	–46.354	
β_{20M11}	–14.721	

On the detail of Fig. 8, we notice the effect of some coefficients, in particular the defects 10M32 of the preceding set-up.

On the CAD/CAM system, we can introduce these deviations. w_{20M11} will be taken into account by parameterizing the machining stock allowances or dimensions tools. α_{20M11} mainly highlights a machining straightness defect related to the geometry of the machine tool. β_{20M11} corresponds to a tool orientation defect. Admittedly, the integration of these defects is not always simple and requires procedures which deviate from pure CAM activity. In practice, these angular variations α and β (associated usual dimensions of the parts) are negligible compared to w.

With the CAD/CAM system we simulate these distances function of $\beta_{20H1,P10}$. We propose a comparison $\Delta L_{(7,1)1}$ and $\Delta L_{(7,1)2}$ with SDT model (left of Fig. 9). It is noted that the errors are larger for point Q72 when the rotations are negative. The phenomenon is not symmetrical. The error is allotted to the CAD/CAM system because the model SDT is symmetrical. The error surfaces $\Delta L_{(7,1)1}$ according to the three values $u_{20H2,P7}$, $u_{20H3,P6}$ and $\beta_{20H1,P10}$ is plotted on Fig. 9 (right handside).

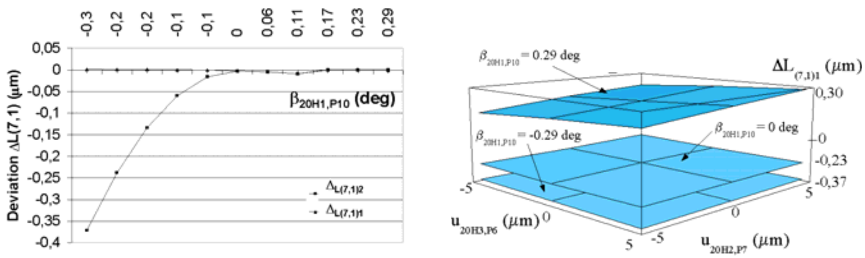


Figure 9. Comparison between STD Model and CAD/CAM simulation : Distance plane P1 / line (cylinder) P7.

In these two comparisons we note that the errors are acceptable compared to the necessary requirements in mechanics (0,37 µm, Fig.9 left).

4.3 Examples of GMCs : Perpendicularity between centerline and plane

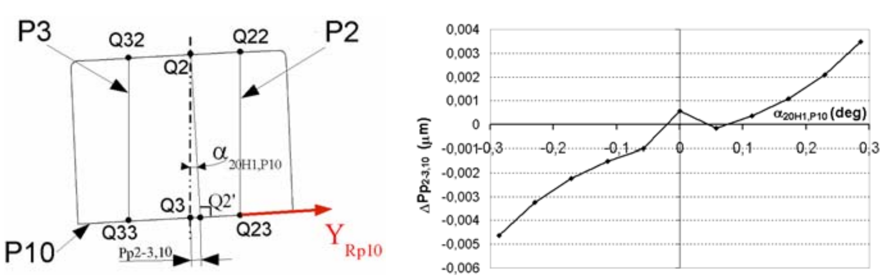


Figure 10. Perpendicularity between centerline of P2- P3 and the plane P10.

More complex specifications can be studied, e.g the perpendicularity between the centerline of faces P2 and P3 and plane P10 (Fig. 10). With the SDT model, perpendicularity is calculated :

$$\mathbf{Pp}_{2\text{-}3,10} = -(\boldsymbol{D}_{\boldsymbol{Q32}_{P3,P70}}.\boldsymbol{y}_{\boldsymbol{Rp10}} + \boldsymbol{D}_{\boldsymbol{Q22}_{P2,P70}}.\boldsymbol{y}_{\boldsymbol{Rp10}})/2 - (\boldsymbol{D}_{\boldsymbol{Q33}_{P3,P70}}.\boldsymbol{y}_{\boldsymbol{Rp10}} + \boldsymbol{D}_{\boldsymbol{Q23}_{P2,P70}}.\boldsymbol{y}_{\boldsymbol{Rp10}})/2$$

$$\mathbf{Pp}_{2\text{-}3,10} = 48.04[(\alpha_{20H1}-\alpha_{20H1,P10})+0.5(\alpha_{20M13}-\alpha_{20M12})]$$

with 48.04 is the part height.

With the CAD/CAD model, we define Q2, Q3 the mid-points of Q32Q22, Q33Q32 and Q2' the projected point Q2 on P10. Then in virtual metrology: $\mathbf{Pp}_{2\text{-}3,10}$ = Q3Q2'. $\boldsymbol{y}_{\boldsymbol{Rp10}}$. On the comparison graph of Fig. 10, we observe the difference between the approaches function of $\alpha_{20H1,P10}$: $\mathbf{\Delta Pp}_{\mathbf{2\text{-}3,10}} = \mathbf{Pp}_{\mathbf{2\text{-}3,10/CAM}} - \mathbf{Pp}_{\mathbf{2\text{-}3,10/SDT}}$. We again note a negligible difference between the two approaches.

From these two examples of GMCs, simulated dimensions and geometries are comparable by model SDT or CAM integration. This report also applies to other GMCs (*direct, indirect....*) which are not to be presented in this paper.

5. COMPARISON OF THE TWO APPROACHES

We first define the input data that the two approaches share: the part definition (a CAD model and a part draft with GD&T Geometric Dimensioning and Tolerancing) and the process plan (raw workpiece definition, machining operation sequencing, location faces and GMCs). Starting from these data, the stages for each approach are:

- **SDT Model and feature extraction**

1. Re-design of the workpiece (by following the operation sequencing of the process plan)
2. Geometric data extraction: Face features (deviation SDTs); Connection between components (H/MT P/H) in each set-up (connection SDT); Torsor chains related to each GMC build from the chart model
3. Mathematical resolution (symbolic and numerical) unification and computation of torsor chains
4. Representation of part defects (torsor components) to check GD&T
5. Tolerance analysis and synthesis using linear equation systems

- **CAM/Integration**

1. Creation and sequencing of all the set-ups (into CAM system)
2. Setting parameters of the variations in each set-up

3. Virtual metrology of the simulated workpiece
4. GD&T analysis

The construction of the comparison table 1 has enabled us to identify the advantages of each of the two approaches.

Table 1. Comparision of the two approaches.

Criteria	**SDT Model**	**CAM/Integration**
Tolerance analysis	➚➚ Linear equations	➚ Numerical results without tendency analysis.
Tolerance synthesis	➚➚ Linear equations	➘ Difficult with current CAM systems
Process plan modification management	➘ Rebuild the workpiece model	➘ Recreate all set-ups
Management of defects value modifications	➚➚ Symbolic calculations	➘ Difficult, re-execute the process plan
Virtual metrology	➘ Difficult, translation of GT&D into GMCs.	➚➚ Immediate except for stock removal allowances
Taking into account of all variations types	➚➔ Yes but without form defects	➘ Difficult, requires procedures which deviate from CAM activity. No form defects
Final 3D CAD model of the part	➘➘ No (not yet)	➚➚ Yes (automatically created for each set-up, usable in CAD)
Automation	➚ Yes	➘ Slower, stability depends on the CAM system
Accuracy	➔ Limited by SDT hypothesis	➔ Limited by the CAM system accuracy

As we show in the example, the two approaches converge towards the same numerical results. Thus, the numerical accuracy is not a criterion retained among the criteria used to compare them in the Table 1.

6. CONCLUSION

We have presented two different approaches from the 3D geometrical manufacturing simulation. One is directed towards a mathematical modeling using the Small Displacements Torsor (SDT) concept. The other one is applied and implemented on a CAD/CAM system. Our mathematical model allows us to know and control the process plan but can become somehow bulky. The CAD/CAM integrated approach quickly provides a numerical result but all the parameters are not easy to simulate. If a sensitivity analysis is wished it will then be necessary to associate both approaches.

A 3D modeler with an integrated representation of the geometrical defects of surfaces would be really useful, not only for our research work but also in a more general context (assembly, clearance, …). However, as we saw in the numerical results, the precision of 3D CAD/CAM systems should not be neglected, although the results converge with the calculations linearized by the SDT modeling.

According to the comparison of the two approaches, our future research works will focus on: How to simulate a process plan and to obtain virtually manufactured 3D model; How to design a method and tools for tolerance analysis and synthesis; How to automate as much as possible all tasks and have accurate calculations.

REFERENCES

1. E. Ballot, P. Bourdet, A Computational Method for the Consequences of Geometric Errors in Mechanisms, Proceedings of the 5th CIRP Seminar on Computer Aided Tolerancing, Toronto, 1997, pp. 137-148.
2. O. Legoff, F. Villeneuve, Three-dimensional geometrical tolerancing: quantification of machining defects, Geometric product specification and verification: integration of functionality, P. Bourdet and L. Mathieu, Kluwer academic publishers, 2003, pp. 185-196.
3. E. Ballot, P. Bourdet, F. Thiebaut, 3D tolerance chains: present conclusions and further research, Proceedings of the 7th CIRP Seminar on Computer Aided Tolerancing, Cachan France, 2001, pp. 65-74.
4. F. Villeneuve, F. Vignat, 3D synthesis of manufacturing tolerances using a SDT approach, Proceedings of the 8th CIRP Seminar on Computer Aided Tolerancing, Charlotte USA, 2003, pp. 279-290.
5. J.J. Shah, Y. Yan, B.C. Zhang, Dimension and tolerance modeling and transformations in feature based design and manufacturing, Journal of Intelligent Manufacturing, Vol. 9, october, 1998, pp. 475-488.
6. E.A. Lehtihet, S. Ranade, P. Dewan, Comparative evaluation of tolerance control charts models, International Journal of Production Research, Vol. 38, No. 7, 2000, pp. 1539-1556.

SOME ASPECTS OF INTEGRATED PRODUCT AND MANUFACTURING PROCESS

Patrick Martin
LGIPM - ENSAM- 4 rue Augustin Fresnel 57078 METZ cedex 3, patrick.martin@metz.ensam.fr

Abstract: In the frame of concurrent engineering the design of the parts, the production planning, even the production system must be addressed quite simultaneously. In this way the design and development cycle is reduced and the manufacturing constraints are taken into account as soon as possible. The need of a short reconfiguration time leads to the concept of Reconfigurable Manufacturing Systems (RMS): using those systems, the manufacturer only uses elementary cells (such as milling cells, supports, tooling, etc ...) which can be easily and quickly replaced as soon as the demand changes. Alternatively, this higher reconfiguration speed can be improved by using Virtual Manufacturing: the manufacturer can design, draw and simulate the line and its evolution in terms of quality, flows and costs. Evaluating the best performance of several configurations before the system is built, results in a gain of time. In the frame of virtual manufacturing this paper will deal with creating and simulating machining process using a computer aided engineering software in order to address technical, economical and production constraints.

Key words: Manufacturing process, Concurrent engineering, Manufacturing reconfigurable systems, Virtual manufacturing.

1. INTRODUCTION

Nowadays market fluctuation and fashion ask for more and more new and different patterns, decreased batch size and the need to design and manufacture with reduced delays. So the concept of concurrent engineering [1, 2] must be used. The design of the parts, the production planning and the production system must be considered quite simultaneously. Thus the design

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 215-226.

and development cycle is reduced and the manufacturing constraints are taken into account as soon as possible, and the design phase must take account of manufacturing constraints but these must not restrict the creativity of the design.

Integration product-process concept is illustrated in Fig. 1: the quality of mechanical parts depends on the expression of specifications (shapes, functions, dimensions, surface quality and materials), the ability of shaping processes and resource capability (defects in machine-tools, tools, fixtures). To meet optimization objectives, models and processes are used to generate production processes or design production systems or to validate a product's manufacturability, as well as defining the product qualification procedure. By process, we mean processes (forging, casting, stamping, machining, assembling, fast prototyping..), production resources (machines, tools..), and conditions for implementation (setting and holding in position, operating conditions…), operation scheduling and the structure of the installation. Production constraints (process capability, producible shapes, precision...) must be taken into account at the same time as economic (cost...), logistics (lead-times, reactivity, size of production runs...) or legislative (recycling, safety...) constraints. On the one hand, knowledge and constraints must be structured, formalized and represented (data and processes), using experimental data (processes and resources) and models of industrial conditions so that different types of expertise can be integrated coherently using appropriate models, methods and tools so as to meet production optimization objectives (quality, reactivity, productivity, cost…).

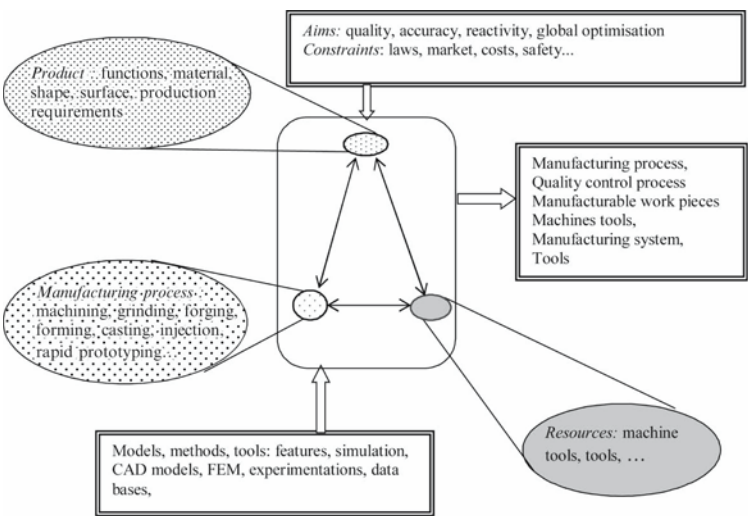


Figure 1. Reference diagram of our approach.

2. VIRTUAL MANUFACTURING OF RECONFIGURABLE MANUFACTURING SYSTEMS

In order to reduce design times and to give a help to choose the most suitable system for a specific process, Virtual Manufacturing appears helped by the developments of numerical methods for computing. We quote that Virtual Manufacturing is present all along the product cycle, hence its importance; the company must share its knowledge

Besides the concept, Virtual Manufacturing is based on numerical tools, such as:

- global database: all the ownerships within the company are listed and the staff can easily design or improve the factory with available stuff;

- 3D Manufacturing simulations: it enables the company to plan, detail, analyze and simulate the product life throughout the Manufacturing Lifecycle Engineering;

- logistic software (e.g. ERP and PDM): it matches all the actors of the product, from the provider of the provider to the customer of the customer.

Manufacturing global demand and rapid changes in process technology (e.g. the evolution of the product, of the quantities, of the variety of products) require creating production systems easily upgradable and into which new technologies and new functions can be integrated in the fields of manufacturing (mechanics, controls, flows and so on). Thus, there is a growing trend to produce smaller quantities with frequent small variations within a product family, so called mass customization. For example, if the customer orders a variation of the number of holes on an engine block and/or an angular variation of those latter's, the manufacturer must change his process as quickly as possible. So the concept of Reconfigurable Manufacturing Systems (RMS) appears [3, 4]. Previously we could only process with two extreme kinds of Machining Systems: Dedicated Systems (high productivity) and Flexible Systems (high flexibility). RMSs are placed somewhere in between those two boundaries, what we call customized flexibility, a solution is to use only elementary cells (machining, actuator and control cells) and to reconfigure them according to the demand of the customer (Fig. 2).

The idea is to design the production system around the concept of part family to achieve a lower cost and therefore a higher production rate with the required flexibility: basic process modules (hardware and software) are incorporated and can be rearranged or replaced quickly and reliably. So customized flexibility is provided for a particular part family. So rather than being replaced, they can be improved, upgraded, etc ... The appeal of reconfigurable manufacturing is to meet market demand with the right amount of the right product at the right time in an era where product development time is shrinking. So a given RMS configuration permits a reduction of lead time for launching and reconfiguring systems, and a rapid manufacturing modification and quick integration of new technology or specification. These systems will not run the risk of becoming obsolete, because they will enable the rapid changing of system components and the rapid changing of components and the rapid addition of application-specific software modules. They can be improved continuously by integrating new technology and be reconfigured rapidly to accommodate future products and changes in product demand rather than scrapped and replaced.

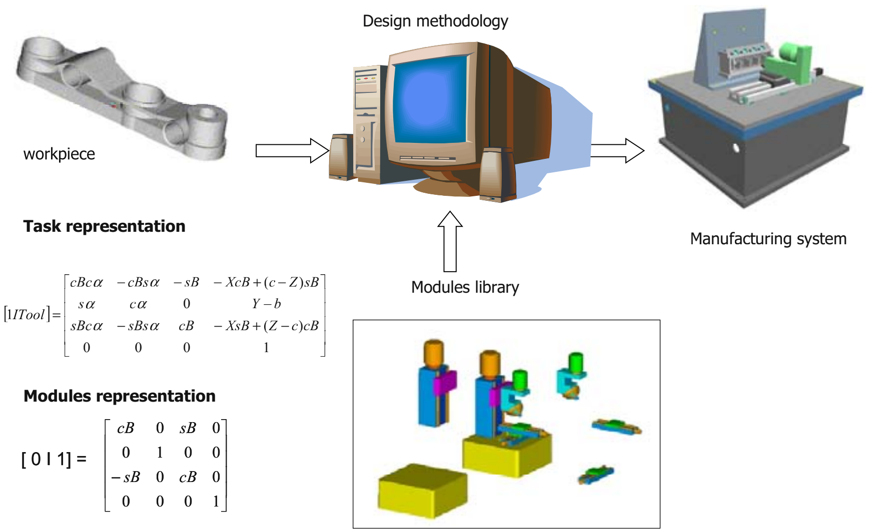


Figure 2. Geometrical and kinematical model of reconfigurable manufacturing system.

Developing a production system which constitutes a specific, but unique product necessitates some specific methods and tools. A great deal of work is being done on production system design, but is mainly concerned with logistics aspects (installation, scheduling) based on technical and temporal data (sequences, times, production resources) provided by the methods department. Work on designing the operative part of production systems is still limited [3, 5-7]. A methodology is therefore proposed to help in defining a flexible manufacturing line with the intention [8] of integrating the manufacturer's expertise based on a strategy using production features and in order to respect productivity, cost, flexibility and quality constraints.

The selected architecture must meet:

- technological requirements needed to produce each machining feature (respecting surface condition, geometric and dimensional tolerances) and equipment capability (precision, torque, power, speed...);
- economic constraints;
- logistics requirements: quantity, series, lead-times;

in a coherent approach based on a methodology, models and tools.

3. TOOLS AND METHODS USED

3.1 The idea of a feature

Product manufacture must simultaneously meet design and production requirements (processes and resources). The description of the part and of the production process is defined via the concept of a feature [2, 9-11] (feature, characteristic).

A feature is a semantic group (modeling atom) characterized by a set of parameters, used to describe an object which cannot be broken down, used in reasoning relative to one or more activities linked to the design and use of products and production systems.

The objective of modeling in features is to facilitate:

- the formalization of expert assessment;
- the capitalization of expertise;
- to provide information about production activities very early in the design phase;
- to improve communications between people working on the product throughout its life cycle.

This concept uses a high level of semantics, but the functional-feature – machining-feature link is not necessarily one-to-one. An universal catalogue of features cannot be envisaged because it depends so heavily on the industrial context. Every company, depending on its needs and habits, needs more than one baseline catalogue.

Features are characterized by a set of information:

- intrinsic characteristics: dimensions, surface quality, tolerances of the feature's own shapes;
- geometric relations between features: dimensions, geometric orientation tolerances, geometric position tolerances;
- topological relations: proximity or interaction relations

3.2 Homogeneous coordinates matrix

Each feature and part of the kinematics chain has an associated local identification, the position of which, relative to the part position, is defined by the matrix with homogeneous coordinates (4x4). The 12 terms which are non null or equal 1 characterize rotations (3 X 3 terms) and translations (last column). This matrix-linked notation is routinely used and found to be an efficient mathematical tool which can easily be manipulated in formal calculations tools. Positioning of the part feature on one hand and the tool on the other can be directly calculated by a product of matrices (open chain of solids), the terms for machining being written by the identity of the 12 terms of the two matrices (equation (1)) for passing from frame to feature and frame to tool:

[Base | Feature] ⇔[Base | Tool] with (1)

[Base | n] $=\Pi_{1,n}$ [i-1 | i]

where [i-1 | i] define the position (angular and linear) of the part i in relation with part i-1. So knowing the geometric and cinematic characteristics of each element the tool position in relation to the feature position can be computed.

This procedure is an efficient framework for kinematics modeling and systematic design of re- configurable machine tools. The common representation scheme used to capture the motion characteristics of both machining tasks and machine modules enables automatic selection and synthesis of machine tools. The module selection and synthesis process is reduced to simple computable operations such as matrix inversion and matrix multiplication (Fig. 2). The generalized kinematics modeling method we have shown above is only the first step towards development of an integrated tool for designing machine hardware and control schemes. The methodology can also be extended to include dynamic characteristics, error propagation and control logic by using small deviation matrices (paragraph 5.).

4. DESIGNING RECONFIGURABLE MANUFACTURING SYSTEMS

The methodology is defined in two steps:

- analyzing the part,
- designing the system,

which are split in several activities.

At first, the work piece is analyzed with the classical procedure known in process planning [12]; note simply that some machining operations, notably involving milling, can be performed on both casings during the same operation.

The part is fitted onto the pallet, we assume that all fitting features are sufficiently good qualities to help for fitting the part. Part fixturing is determined so as to remove any risk of the part overbalancing or sliding, while avoiding the risk of deformation due to the force applied. Of course it is necessary to analyze dimensional and geometrical tolerances surfaces roughness which constrain the part set-up and the cutting conditions.

The set of features to be machined, their types, intrinsic characteristics, topological relationships and associated basic processes are defined (Table 1). The location of each in the part baseline is also defined by using 4 X 4 matrixes.

For each feature, it is necessary to choose the operations, the tools and the cutting condition. For example every hole can be machined with a simple tool or a variable diameter tool, by using a multi-spindle or a single spindle head In a context of RMS, we should prefer to machine separately each hole, because a variable diameter tool is more expensive; however, if we are sure that the product will exceed an important number of parts, the variable diameter tool will become cheaper, because machining time will be reduced. Then the cutting conditions and the machining times are computed. It is also necessary to identify the directions of accessibility, for the most common case the characteristic polyhedral is a cube with a 2 or 3 axis CNC machine.

Table 1. Example of features definition

Features	Feature shape	Way of machining	Topology interaction	Tool axis	(x, y, z) axis motion	Acces sibility	Geometrical datas	Technological datas
P1	Plane	Surfacing	H1, H2, H3, H4, H5, H6	X/Y2/Z 2	PT, C, C	x-	Included in Ø130	Ra2.5 Localization 0.2 Parallelism 0.05 Planeity 0.05
P2	Plane	Surfacing	D	Z2/X/Y 2	C, C, PT	Z_2-	Included in 70(x1)x30(y 2)	Ra3.2 Planeity 0.05
P3	Plane	Surfacing	E	Z2/X/Y 2	C, C, PT	Z_2-	Included in 70(x1)x30(y 2)	Ra3.2 Planeity 0.05
C11	Emerging hole	Boring	C13, C21, P1	X	PX, PT, PT	X	Ø57.4 D7	Localization 0.04 Ra2.5
C13	One-eyed hole	Boring	C11	X	PX, PT, PT	x-	Ø22.3 JS7	Localization 0.04 Ra2.5
C21	Emerging hole	Boring	C23, C11, P1	X	PX, PT, PT	x	Ø57.4 D7	Localization 0.04 Ra2.5

Next machining times have to be determined. They are the sum of technological and auxiliary times. Technological times are calculated from cutting conditions defined using standard software (Couple tool material) and take into account the possibilities of simultaneous machining to obtain a baseline for technologically feasible minimal times. Auxiliary times are estimated by timing or analogy.

Then it is necessary to define the machining process with elementary machining cells in order to shape a RMS. So, how many holes have to be drilled simultaneously? How many cells does the machining process require? These issues are among the main challenges of RMS. The different elements that can be used are: drilling cells, multi-drilling heads, tooling, actuator cells, fitting tools, supports on which elements are attached. Thus a library of these elementary devices which allow the design of the virtual manufacturing system is created

From all this information and the precedence's 'conditions, several architectures which fulfil the technical requirements can be proposed. Position of each module on the line depend of technical and economical constraints such as: the available volume, the kinematics links to be created, the possible vibrations during milling, the module price, the flexibility required or foreseen. Several configurations (Fig. 3) can thus be built and their performances compared by multi-criteria analysis.

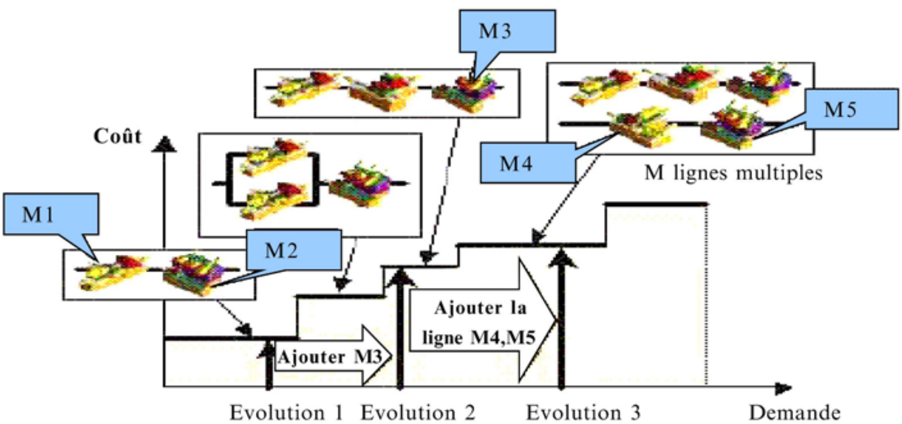


Figure 3. Several manufacturing system configurations [13]

After that it is easy to develop the digital model of these architectures by using a software dedicated for the design of robotized manufacturing systems (Technomatix or DELMIA for example). Translation movements are then added and several devices are attached in adequate position on the support all along the manufacturing system in order to build the manufacturing system (Fig. 4).

By defining relevant criteria, the performance of a system can be accurately assessed. It should be a part of the Product Data Management System and can be used by several persons of the company:

• process planner: from the parts that will be machined, he defines manufacturing feature, precedence constraints, machining time …;

• machine designer: he will choose the equipments in order to fulfil the technical requirements (size, accuracy, speed, power), type of energy needed (electrical, hydraulic, and pneumatic)…

• manager: he will see the cost (production , maintenance), production time, productivity, flexibility (type of part produced), flexibility reconfiguration average time for each product variation, integrability, maintenance, evolution ability, customization, human factors, safety, ergonomic …..

• software master: he has to develop the models, undertake the simulation in order to obtain relevant data.

So a lot of characteristics for multi-criteria analysis can be defined to evaluate the system some of then can be easily quantified, for others their quantification is mainly subjective.

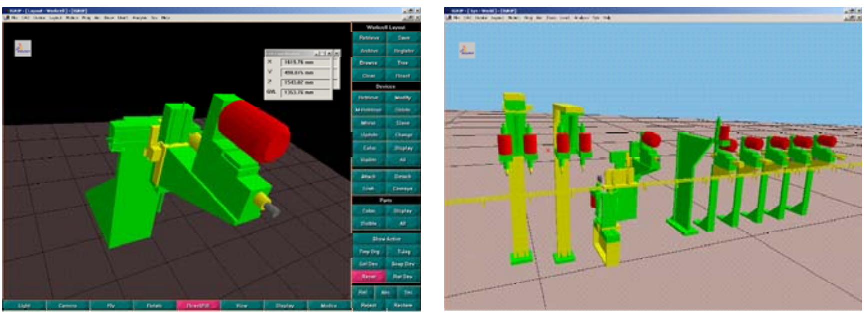


Figure 4. Module and manufacturing line digital models

5. WORKPIECE ACCURACY PREDICTION

The obtained surface quality is function of the machine tool movements quality during its functioning. The defaults of surfaces are the important characteristics to identify the causes and sources in order to predict work piece accuracy and/or undertake corrective tasks. The cutting process and the active forces, the resources capability, the process planning (set-up, various parts' states) are the main elements that exercise influence on the work piece quality, usually these elements interact, so simulation ensables the prediction of part quality.

For each part, it is necessary to describe the deviation resulting from the joint between the part and the holder in each set-up in a base part reference

(Fig. 5). At the end we obtain the maximum deviation on a manufactured surface so the values of the tolerance which can be obtain with the real manufacturing equipments. For this objective, a parameterized formal expression of the nominal position, the normal vector and the deviation of each manufactured point in a single base reference according to the manufacturing deviation causes, adjustment parameters and nominal lengths is determined by using small deviation homogeneous coordinates matrix.

The origins of defaults are also caused by the machine tool mechanical structure stemming from geometry, kinematics, dynamics, thermal, and vibrations. We can say the same thing for technical structure (vice, fixturing systems). The structure of actuators and sensors can lead to defaults of stability and motions errors. These errors are less important than the mechanical and technical structures, except for the control of high speed machining. The different factors influencing part accuracy, more to complete the description quasi static errors and dynamic errors [14] must be taken intoaccount.

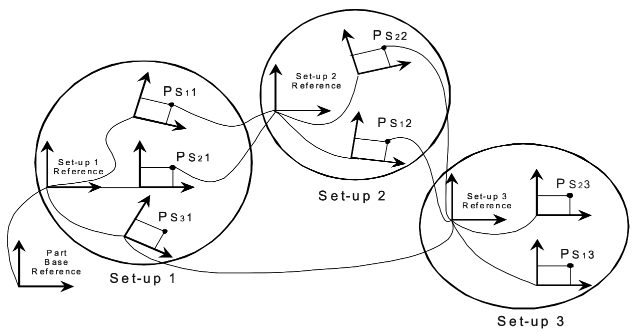


Figure 5. Deviation of a manufacturing surface estimation due to several set-up

In order to analyze the effects of the work cell defects (addition of machine, work piece, and set-up defects) on the work piece we present the result of the simulation based on a 3 axis NC machine tool of our workshop. We study each type of work cell defect separately and compare them with a work cell that has all the defects inside. Thus, we design five work cells each one composed by the same machine, tool, fixture and work piece but with five different types of defect. Each of them will machine the same designed work piece (they all have the same machining program). Consequently, the five final machined work pieces should all be different, four having the result of a particular work cell defect type and a fifth having the results of all the work cell defects. By the way, we can compare the effects of each work cell defect type on the work piece by analysing each piece (we do not take into account tool and work piece deflections). After taking the measures of each work piece, we compute all those measures to get valuable and comparable results. Fig. 6 gives the picture of the machined work piece compared to the perfect piece and the influence of work cell defect on the machined work piece accuracy (which work cell defect causes what type of work piece geometric error).

It seems that for the piece location errors, the most influential defects are the rotation and the piece location errors.

		Workcell defects			
		Orient.	Straight.	Rotation	Piece_pos
Form	A unevenness	Low	Very Low	Medium	High
	B unevenness	Very Low	Low	Medium	High
	Slots sides casing	Very Low	Low	Medium	High
	Slots bottom casing	Low	Very Low	Medium	High
Orientation	D & d parallelism	Low	Very Low	Medium	High
	D / d parallelism	Low	Very Low	Medium	High
	XS & YS parallelism	Very Low	Medium	Low	High
	XS/YS perpendicularity	Very Low	Medium	High	Low
Location	d locating	Medium	Very Low	High	Low
	D locating	Medium	Very Low	High	Low
	XS locating	Very Low	Low	High	Medium
	YS locating	Medium	Very Low	High	Low

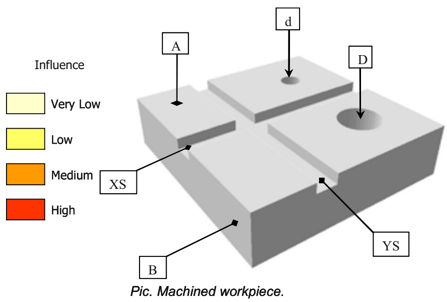


Figure 6. Work piece tested and influence of work cell defects on work piece geometric accuracy.

6. CONCLUSION

Virtual Manufacturing gives a new dimension to manufacturing design by dramatically reducing times at each step of product life cycle. This type of methodology, to which simulations of kinematics or dynamic performance of the equipment and flow can be added, provides a tool to aid in decision making for process engineering, thus helping reduce the time and cost of production engineering. But the use of these new tools necessitate a structured methodology we have described.

Simulation technique enables the prediction of the surface quality of a machine part function of the machine tool quality, the set-up and the cutting process. Then, the user can anticipate those errors and manage to reduce the influence of the work cell defects. But at this moment the models are so complex that it is not possible to take into account all the phenomena involved in particular tool and work piece deflections.

REFERENCES

1. G. Sohlenius, Concurrent engineering, Annals of the CIRP - 41/2, 1992, p645-655

2. M. Tollenaere (sous la direction de), Conception de produits mécaniques, Hermes, ISBN2-86601-694-7, 1998.
3. Y. Koren, Reconfigurable Manufacturing systems, Annals of the CIRP vol 48/2/1999.
4. M.G. Merhabi; A.G. Ulsoy, Y. Koren; P. Heytler, Trends and perspectives in flexible and reconfigurable manufacturing systems, Journal of Intelligent Manufacturing, 13, 2002, pp.135-146
5. O.Garro, P.Martin, Towards new architectures of Machine Tools, International Journal of Production Research, Vol. 31, n° 10, pp. 2403-2414, ISSN 0020-7543, October 1993.
6. O.Garro, P.Martin, M. Véron, Shiva A Multiarms Machine Tool, CIRP Annals 1993, Vol. 42, pp. 433-436, ISBN 3-905-277-19-0, 1993.
7. R. Schukschik, Requirements for structures of automated machine tool, Annals of the CIRP, vol.32/1/1983
8. P. Martin, A. D'Acunto, Design of a production system: an application of integration product-process, Int. J. Computer Integrated Manufacturing, Vol.16, N° 7-8, 2003, pp 509-516.
9. J.J. Shah, Y. Shen, A. Shirur, Determination of machining volumes from extensible sets of design features in Advanced in feature based manufacturing, Shah, Mantyla, Nau editors, Elsevier, 1994
10. O.W. Salomons, F.J. Van Houten, H.J.J. Kals, Review of research in feature-based design, Journal of Manufacturing Systems, vol.12 n°2, 1993
11. K. Case, J. Gao, Feature technology: an overview, Int. Journal of Computer Integrated Manufacturing, Vol.6 n°1&2, 1993
12. G. Halevi, R. Weill, Principles of process planning, Chapman and Hall 1995.
13. Y. Koren, U. Heisel, F. Jovane, T. Moriwaki, G. Pritchow, H. Van Brussel, A.G. Ulsoy, Reconfigurable Manufacturing Systems, CIRP Annals, v. 48, No. 2, 1999
14. S. Chatterjee, Spindle deflections in high speed machine tools - modelling and simulation, Int. J.of Manufacturing. Technology., 11, 1996, p. 232-239.

INTEGRATING DESIGN AND MANUFACTURING PROCESSES IN AUTOMOTIVE SUPPLIERS

Peter R. Frise

University of Windsor, 401 Sunset Ave. Windsor, ON Canada N9B 3P4, voice: (519) 253-3000 ext. 3888,

Abstract: As the global automotive sector strives to achieve enhanced levels of efficiency and unprecedented cost reductions in the product realization process, a number of key issues have emerged in the area of the integration of the design and manufacturing processes. These include new considerations in the decision-making processes that must be integral in determining the major parameters of a new product and the formation of a new technical team of product development personnel. The paper describes these issues and based on the experience of the author in working with companies, it relates how they are being addressed in the North American automotive sector with particular reference to auto parts suppliers who are under severe pressure to lower costs and decrease product development time. The paper poses questions for consideration in advancing the state of the art in these key topics of concern.

Key words: Tools for design and manufacturing decision making, Design personnel

1. INTRODUCTION – THE CONTEXT

The modern automobile is one of the most complex and pervasive products ever created for the mass consumer marketplace. The capital, physical infrastructure, logistics and skilled human resources required to produce even a simple vehicle are vast. For example, it is estimated that the development of a new vehicle platform can cost from \$1.5B USD to as much as \$6B USD and require from 500,000 to over a million person-hours of work. Even a relatively simple freshening of an existing product can cost

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 227-242.

upwards of $3-500M USD. Automakers are attempting to compress the design cycle into ever-smaller periods of time with the most advanced companies developing a new vehicle in 15-25 months as opposed to as much as 60-72 months taken to develop previous generations of vehicles. The investments and product development tasks must be borne by a number of companies ranging from the automakers themselves to a whole array of supplier companies ranging from small tool and die shops to large system integrators who may be responsible for entire vehicle systems (or chunks) such as cockpit modules, wiring harnesses or seating systems. All of the design decisions made in this process must be coordinated and communicated throughout the supply chain and production system and it must be done quickly and without error while still maintaining strict confidentiality of the new vehicle's design features. The fact that most of the decisions are made by the automakers who are only responsible for designing a portion (often less than half) of the vehicle parts means that suppliers must be agile and use foresight in planning their design and manufacturing processes.

1.1 The Automotive Supply Chain and Over-Capacity

The automotive supplier system is arranged in a series of layers or "tiers" and the tier level of a given company and a given product depends upon to what type of organization it is sold. Generally, the companies which supply directly to the automakers are known as Tier 1 suppliers while those who supply components to the Tier 1's are referred to as Tier 2's etc. Thus, a tool and die shop would be a Tier 2 supplier to a company that supplies interior mouldings for the cockpit of the vehicle and the plastic resin supply company would be a Tier 2 or 3 supplier to the same moulding company. Tier 1 companies tend to be global in structure while lower tier suppliers are sometime more local in nature. Most automakers are trying to decrease the number of supplier companies in the supply chain for reasons of cost and business complexity as will be detailed below. This trend has resulted in considerable displacement within the supplier community as companies jockey for position as one of the chosen suppliers to the companies in the tier level above. The global auto industry suffers from a significant amount of excess capacity in vehicle assembly volume and, to a certain extent in the attendant parts manufacturing operations of external supplier companies and within the in-house parts divisions of the automakers alike, especially in some markets for certain companies. The over-capacity issue has been estimated to represent the output of up to 60-80 assembly plants world-wide, a number that is basically equivalent to all of the vehicle assembly plants operating in North America at present.

Contributing to the issue of over-capacity is the trend toward common platforms and corporate consolidation that has been sweeping the automotive industry in the last five years. Automakers are increasingly using a given platform and powertrain for a wide array of vehicle types in order to amortize the cost of development over a larger number of sales. The term "mass-customization" has often been used to signify the fact that each customer wants a personally unique vehicle but economies of scale dictate that it is more cost effective to have the basic underpinnings of a number of vehicles be identical, resulting in large scale production of a majority of parts, especially those requiring expensive tooling. In fact, cost control and reduction is the major issue for virtually everyone in the automotive sector today and the pursuit of cost reductions while maintaining quality is the major task of all supplier companies. Fig. 1 below shows the results of survey taken at the Automotive World Congress of the Society of Automotive Engineers in March 2002 and 2003 in Detroit. he whole issue of cost reductions is clearly at the "top-of-mind" for most engineers working in the auto industry today [1].

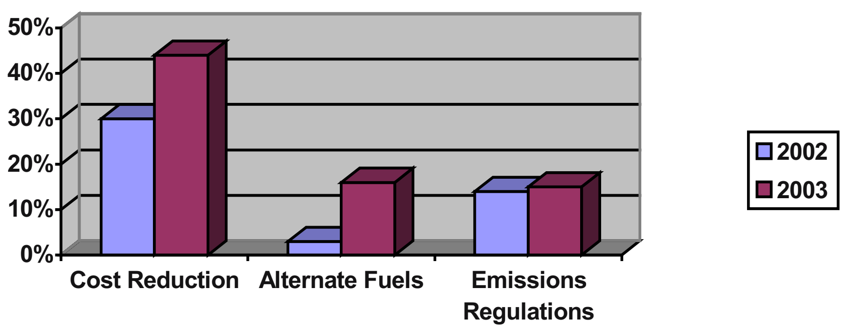


Figure 1. Survey of major concerns of automotive engineers taken at the SAE World Congress – Detroit 2002 and 2003 [1].

An example of mass customization and cost consciousness is seen in the number of recent models of light-duty sport utility vehicles which have all wheel drive but are being built on the same platform as passenger sedans which only have front-wheel drive. There is one platform used by a major European automaker across their entire stable of brands built in several countries that is used for well over a dozen distinct vehicles ranging from a high performance two seat sports coupe to an economy four-door hatchback sedan for family use. Given the practice of most automakers to assemble vehicles in, or close to the markets where they will be sold, it is likely that a significant number of new factories will be opened in certain parts of the world (chiefly in the emerging markets of China and potentially India) while other locales will likely see a number of plants close or be downsized (although the latter option is not attractive as there are economies of scale that larger plants can achieve). This trend of consolidation in more mature vehicle markets such North America and Europe has resulted in some plant closures in recent years, particularly for automakers that had excess assembly capacity in those markets. The resulting disruptions of labour and infrastructure are serious issues for the communities where these plant closures have occurred and represent very difficult adjustments for both supplier companies and individual workers.

1.2 The Modern Vehicle as an Integrated Product

The modern motor vehicle is built of a wide array of materials including sheet and wrought steel, cast, sheet and forged aluminium, cast iron, stainless steel, magnesium, many grades of both thermoplastic and thermoset materials, glass and fabrics including both synthetics and natural materials such as leather along with an increasing array of natural fibres (such as wood fibres or even agricultural crop waste fibres) for applications related to noise and thermal insulation and even some lightly loaded structural parts. In addition, there are significant amounts of electronic materials incorporated into the information and entertainment (often referred to as infotainment system) of the vehicle and small amounts of precious metals such as platinum in key components such as catalytic converters which provide improved environmental emissions performance. As well, advanced coatings are applied to the interior and exterior surfaces of the body structure to protect it from corrosion. In manufacturing motor vehicles virtually all traditional processes are used including die and sand casting, injection, blow and rotational moulding, stamping, machining, welding, bonding, bolting and even riveting (especially in the case of some of the latest aluminium body structures such as in some recent Audi, Jaguar and Rolls Royce models). Many relatively new processes such as hydroforming are used in certain components such as high strength truck frames and cross-car beams and new coatings processes such as dry powder coating, e-coating and water-borne paints are also in routine use. The introduction of new manufacturing processes has necessitated the development of a range of new materials such as hydroformable alloys of both steel and aluminium, more reliable alloys of magnesium for some very large die cast parts along with a large number of specialized engineering plastics for high temperature underhood applications including a vast array of specialized electrical connectors which must function at military standards of reliability in a very severe environment.

1.3 Auto Sector Market Realities and Integrated Design and Manufacturing

One of the other key issues in vehicle design is that of time-to-market and the need to respond very rapidly to the changing tastes of the automotive market. These needs and tastes are difficult to anticipate and so the lead time to develop new vehicles plays a key role in an automaker's ability to capture a market segment or respond to a new trend such as a shift in the price of fuel due to international events or the introduction of a new government safety regulation. In addition, various automakers have attempted to create new market segments by developing whole new vehicle types and these have sometimes been very successful (such as for example the family minivan) while others have been failures, or at best, limited successes. In North America, there have been a number of vehicles which were initially successful but have lost market appeal very rapidly and several have even been discontinued after only 2-4 years on the market versus a usual life of 7-15 years. The Ford Thunderbird (introduced as a 2002 model) has just been cancelled with the 2005 models likely being the last available. This is in-spite of the vehicle winning a number of prestigious design awards and being in great demand for a period of time following its introduction. The Volkswagen New Beetle has suffered a similar peak and fall in sales but it remains on the market at this point. The cost of creating a new vehicle design is obviously a major concern to car companies in the highly competitive world market as it plays a key role in determining the final selling price as well as in defining the resources needed to actually bring the vehicle to market. The design decisions made by the technical team plays a role in the materials and manufacturing of each new vehicle model and the increasingly competitive automotive industry is demanding new decision-making models and a new technical team to maximize efficiency and minimize product development time. A key factor in minimizing development time is the use of computer-based design tools and data management systems. These can allow a designer to evaluate a large number of design alternatives quickly to ensure that an optimum solution has been developed. A major pitfall of this type of virtual design, especially from the standpoint of a parts supplier, is that the designer must have a solid understanding of the manufacturing processes and typical materials selection options that are available to his or her company. This need for experience and knowledge has not been diminished by the introduction of computer design tools, in fact, the need for better educated technical people has actually increased because several roles are being combined and decisions must be made more rapidly by fewer people in order to meet demanding product development schedules. The ability of a given parts supply company to compete will be founded on their ability to provide effective and innovative products rapidly at the lowest possible cost while maintaining quality and the reliability of delivery logistics.

2. DECISIONS ABOUT MANUFACTURING AND MATERIALS – THE CAPABILITY INVENTORY

Among the key decisions to be made by any supplier are which materials to use and which manufacturing processes to employ in producing the product. Each company has a distinct range of design and manufacturing expertise and supply-chain capabilities. This could be thought of as the available "capability inventory" of the company and it forms a major factor in what type of project the company can realistically undertake and points the way to what sort of decisions would need to be made as the company tries to enter new markets. Fig. 2 below illustrates the capability inventory concept and lists what lies within a given company's inventory. Note that figure does not provide a comprehensive listing for all companies but shows typical items which must be considered in making decisions about design and manufacturing of a typical automotive component.

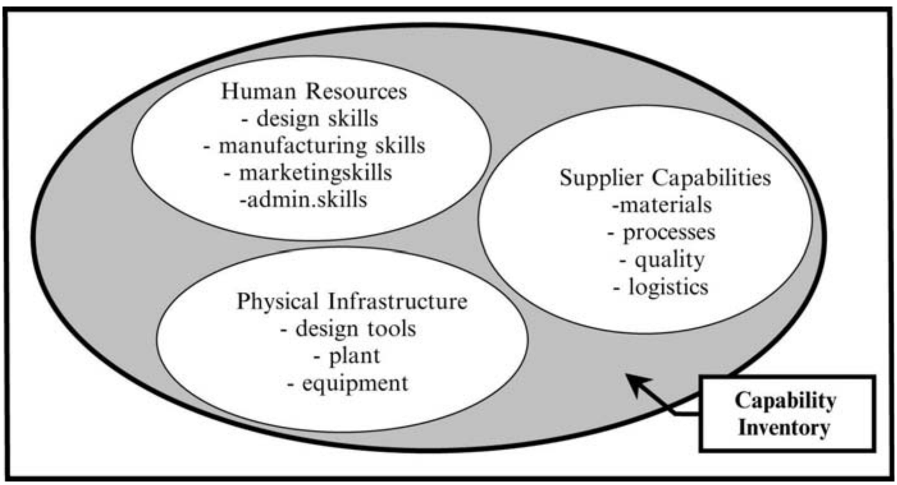


Figure 2. The capability inventory of a given company is made up of the human and physical resources infrastructure of the company, plus the capabilities of its suppliers.

In considering design decisions about new products, every company makes strategic judgements about new business areas and these decisions are usually made by senior level executives. This group of people can be considered the *primary* decision-makers.

2.1 The Primary Decision-Makers

In considering new projects or product lines, the primary decision-makers consider issues such as the availability of financing, staffing and major physical resources needed to carry out the project in question and what pricing and payment terms are being offered by the potential customer. The measure of effectiveness of the primary decision-makers is really a measure of how aware they are of the capability inventory of their company as well as the likelihood that a particular business opportunity will be successful. In that respect, most automotive supplier companies today are successful and make informed decisions at the macro-business level. If this were not the case, these companies would likely have gone bankrupt over the last few years of very demanding business competition. As was noted above in Fig. 1, the major concern of auto sector companies today is cost reduction and in that respect, most of the decisions that are within the typical purview of primary decision-makers have already been made. Typically, these types of cost reduction decisions include facility location, consolidation and rationalization, financial arrangements and employee staffing and compensation levels and supply and delivery logistics.

2.2 Design Engineers - The Secondary Decision-Makers

In working with a number of manufacturing companies both within and outside of the automotive sector, the author has noted the following facts:

- many key decisions are made at the drawing board (or CAD station) in the design office,
- these decisions are often made by design engineering personnel (the secondary decision-makers), who are often relatively junior people with little access to primary decision-makers or to the information and decision making tools they employ,
- these drawing board decisions are seldom reviewed by the top management of the company (i.e. the primary decision-makers).

These "drawing board" decisions are often as important, if not more important than those made by the primary decision-makers in terms of their impact on product cost and competitiveness. Typical decisions made within the above group of people include:

- choices for materials including new or non-traditional materials,
- choices for manufacturing methods including processes are new to the company,
- choices of stock items such as fasteners, coatings, joining methods and heat treatments.

These decisions will essentially determine the key characteristics of the product and thus they will have a direct impact on the cost and eventual competitiveness of the company in the market. In addition, these types of decisions will have a direct effect on the company's internal cost structure because they impact issues such as:

- the type of equipment and consumables needed to produce the product,
- the required floor space will be needed in the production facility,
- the amount of warehouse space required to stock parts and materials,
- the type of training needed by production workers to use the chosen manufacturing processes,
- the number of different suppliers required to assemble the resources to produce the new product,
- the likely product life-time in the competitive marketplace.

A typical example would be that of a manufacturer which has traditionally produced a line of products that are made of stamped metal components assembled in jigs and welded into an assembly. A new line of products is developed and instead of a stamped steel weldment, an aluminium structure is chosen by a designer and approved by his or her immediate superior (also a secondary decision-maker). The aluminium structure will require an entirely new welding process with different equipment, consumables and training requirements for the operators, new finishing and coating processes and consumables and likely a new supplier to provide the aluminium stock for the product. The seemingly simple decision to substitute aluminium for steel effectively results in a cascade of supply and manufacturing process decisions for the new product. The higher cost of aluminium will dictate a much higher material cost, resulting in a higher final product cost, although the lower weight may provide an off-setting benefit in certain market segments. In fact, decisions such as this are made every day by relatively junior people and yet they can result in significant extra costs for the company before the decision actually gets a final review by the primary decision-makers who must approve the capital purchases necessary to implement the new design. If the capital purchases are not approved by senior management, then the company may not have sufficient time to implement the design in the traditional materials before the close of bidding takes place and so the whole project may be lost as a result. Thus, the secondary decision-makers actually take on a role that can be as important as that of the primary decision-makers without having the knowledge or authority to make these judgements. This situation may be indicative of poor communications within the company, but nonetheless, these things sometimes occur even in well-managed companies. The magnitude of the decisions made by secondary decision-makers may not be as large as the one described above, but they may still have a significant impact on the corporate cost structure and thus the competitiveness of the whole product line. A prime example of the type of seemingly small decision that can have a large effect is a case where a designer calls for a fastener that is not presently in the company's stock fastener library. The new fastener may be a non-standard size or it may even duplicate an existing part number. To introduce the part into the production system will require that a purchasing agent put out a call for quotes, source and buy the part which will then require space in the company's stockroom and resources in the inventory system. All of these activities cost money and thus they impact the bottom line as well as affect the time required to bring the new product to market. The author has worked with a number of companies who were not aware of the number of different fasteners in their own internal parts supply system and there were several that had identical components that had been issued two or even three separate part numbers and were being sourced from different suppliers simultaneously. Even large automakers are susceptible to this type of problem. One major international carmaker admitted to having over 20 different windshield washer bottle lids simultaneously in their internal parts system while their competitors only had three or four different types of lid. Another example cited by an automaker was the case of interior ashtrays for cigarette smokers. One company had 28 different types of ashtrays while another only had four in their system for a comparably comprehensive line of vehicles. The cost of carrying the different part numbers in their respective systems was estimated by the company to be on the order of $8-15,000 USD *per part, per year* resulting in a total cost of as much more than $400K USD per year for ashtray variants. Note – that figure is only the cost of having the open part numbers – not the cost of actually producing, buying, stocking or installing any of these parts. One must of course consider the added value to the customer of having 28 different variants of ashtray, and there may indeed be some value, but clearly this type of duplication drives up cost, diminishes efficiency and impairs the company's ability to compete with leaner competitors. In the automotive sector, the scale and pace of the industry means that not caring for these small details costs so much money that companies cannot remain competitive for long, especially when automakers continually ask for cost reductions of as much as 8-15% from year to year.

2.3 The Way Forward

The fact that key decisions are often made at the drawing board or CAD tube by relatively junior people is not new, nor is the fact that sometimes these decisions never receive a thorough review by senior corporate management before it is too late to reverse them. The new fact is that in today's cost driven environment, even the largest automotive companies cannot afford to leave these seemingly minor decisions in the hands of people who do not understand or consider the full range of potential consequences. Given that the time of senior executives is at a premium and that no single person can ever be fully aware of all the ramifications of every set of technical decisions, there is a clear need for the development of a mechanism to ensure that all of the factors and components in the corporate capabilities inventory are communicated to those who are making the secondary decisions – the design engineering staff.

Such an awareness program can be relatively simple and it will seem familiar to many more senior members of the design community because most companies had similar programs in previous times. Examples of measures that have been found to be effective include:

- having all new designers spend a period of time (typically a month or more) on the shop floor learning first-hand what processes and equipment are in use at the company,
- having a display board of all standard fasteners and other stock components currently in the supply system of the company prominently located within the design office and then requiring all designers to defend choosing a non-standard component to their superiors.

The second of the above two items can actually have a benefit before even introducing it to the design staff because the mere act of assembling the display board will make obvious any duplication in the existing supply system of the company. It may also suggest opportunities to reduce the number of different components and thus achieve economies of scale for future supply purchases. The development of a corporate capabilities inventory tool is of key importance for all companies but most especially for smaller companies who are less able to add manufacturing processes or equipment and who are more sensitive to fluctuations in internal cost structures in today's competitive world of integrated design and manufacturing of modern products.

3. THE HUMAN RESOURCES SIDE OF INTEGRATED DESIGN AND MANUFACTURING

The other important factor in ensuring good technical decisions is the quality of technical human resources. To achieve good decisions that do not need to be reviewed within the top echelons of the company, it is necessary that the people who make those decisions be better equipped to understand all of the ramifications of their decisions. Given the drive to lower costs and speed-up the process of developing new products, it is necessary to concentrate more capability in a smaller number of people on the technical team by distilling it to only include higher capability people [2, 3].

3.1 The Automotive Technical Team – Past and Present

In past times and in many places today, the product development process in automotive companies is carried out by a team that practices in relatively narrow specialties as shown below in Fig. 3. The product engineer is responsible for the basic conceptual design of the part while the designer prepares the working drawings and details them. Analysis tasks are passed to a specialist analyst and the manufacturing plan for the part is devised by yet another specialist. Each of these groups in the technical team has had a typical career path and set of prerequisite qualifications for entry. For example, analysts usually had at least a Bachelor's degree in engineering while many designers are trained at technical colleges rather than at degree-granting institutions and some only had a secondary school education. Their responsibility for performing any engineering calculations is strictly limited [3]. This system has some distinct disadvantages including: difficulty in managing the work of several people developing the same component, limited career progression for some people due to the potential for "pigeon-holing" and overall higher costs due to the need to provide a salary and benefit package and physical infrastructure for each person on the product team.

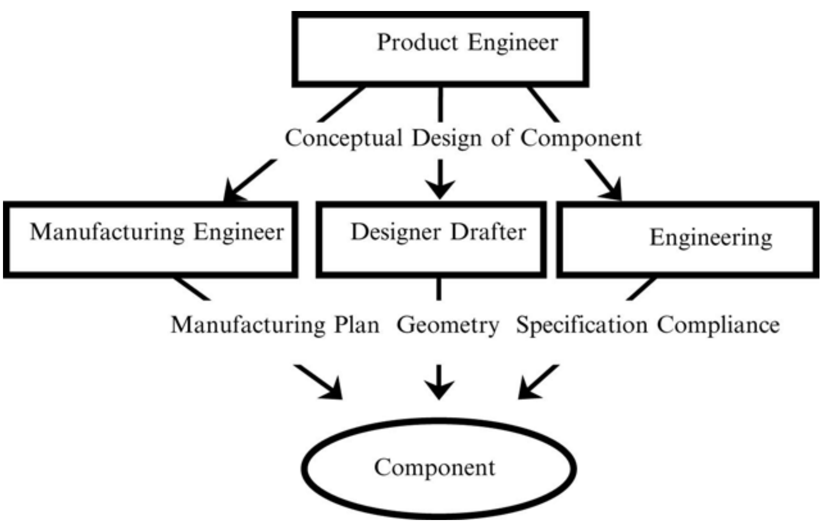


Figure 3. The existing automotive engineering product team – made up of several specialists each contributing to the development of the physical component.

The availability of powerful, reliable and easy to use integrated design/analysis packages (such as EDS Unigraphics, and CATIA) along with specialized tools for optimization of design alternatives has made the need to have separate people perform tasks such as stress analysis and detailing of drawings much less clear. In the hands of a sufficiently experienced and well-qualified person, the same computer running an integrated software suite can now be used to create the geometry of the part, quickly detail the working drawings and finally, carry out the engineering calculations to ensure compliance with specifications. Thus, it is becoming less cost-effective to perform these functions in separate parts of the company using separate groups of personnel. In fact, it is often the case that some of these tasks are being done in different countries and the results of the design effort are being shipped over the internet directly to the manufacturing facility. Intense competition has caused industry to require each employee to contribute as much value as possible to the product realization process. If the team can be consolidated without compromising quality, product performance or logistics and delivery schedules, then only the most capable employees, specifically those capable of performing multiple tasks, can be retained if the organization is to maximize profitability [4].

3.2 The Design Engineer

One of the key points on the issue of consolidating the technical team is that no amount of sophisticated computer hardware or software can replace engineering judgment or remove the necessity for a critical review of all product characteristics before a production decision is made. As was seen above, the secondary decision-makers must be fast and they must be correct or corporate competitiveness will be seriously impaired. As a result, a product development team is emerging that combines the roles of designer, analyst and, to a certain extent, manufacturing engineer into a highly capable individual known as a Design Engineer. Due to the need for the Design Engineer to be able to make complex decisions in several fields or subspecialties, the logical educational background to fill this role is an engineering degree, perhaps with formal post-graduate education but certainly with an on-going program of upgrading and professional development to update skills and build new capabilities. In comparison with the technical team shown in Fig. 3, the Design Engineer combines several roles and creates both the part geometry and does the analysis needed to ensure compliance with specifications. He or she may also develop the manufacturing plan for the part. In some cases, the Design Engineer will call upon senior specialists in a mode which uses them in consultative fashion as needed rather than routinely for every job as shown below in Fig. 4.

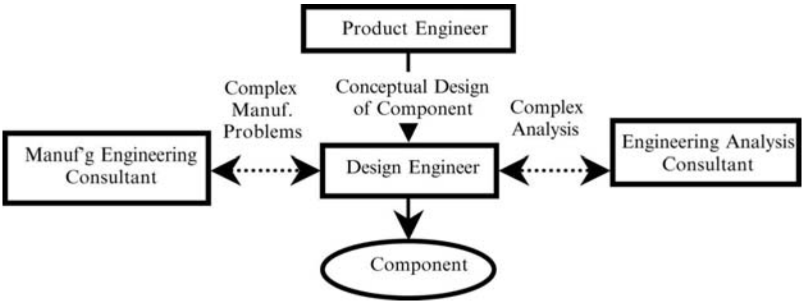


Figure 4. The future automotive engineering product team – the Design Engineer combines several of the previous roles and uses modern design technology to speed the product realization process. Advanced technical specialists are brought-in to assist with complex problems or to research new methods rather than on routine matters.

The term "consultant" does not necessarily imply an outside consultant. In the case of larger companies, these people are in-house but are only available in sufficient numbers to assist with projects on an occasional or as-needed basis rather than on every project. Also, while the roles of Product Engineer and Design Engineer may be combined, it is often the case that the Product Engineer is employed by the automaker while the Design Engineer works for the supplier company and so they would, of necessity, be separate roles.

3.3 Educating and Training the Design Engineer

The focus of the automotive supplier industry on cost reduction and speed of the product development process is driving companies to upgrade their human resources to include only high value-added employees. This imperative is made doubly clear when one considers the effect of the secondary decision-maker issue which highlights the need for a high level of sophistication even in people who work on CAD stations due to their role in making far-reaching decisions for the company. Modern product development companies simply cannot afford to have poorly educated and trained people designing their products because the cost of a senior management review of every single decision is too high and the consequences of poor decisions are disastrous. A large number of engineering personnel are therefore needed to actually design new products and operate the manufacturing plants. These people are mainly mechanical and manufacturing engineers, but there is a significant need for electrical, computer and materials engineering expertise as well. The educational foundation of all of these disciplines will likely remain the Bachelor's degree but people who want to progress or perform industrial research and development will need to consider doing post-graduate work as well, with a large majority being at the Master's level and some key positions requiring a Ph.D. In the context of educating people and preparing them for life, universities have a dual mission. The first is to provide instruction in specific material (training) and the other is to prepare the student for whatever changes and developments might occur in the course of their professional life (this is education). Training is concerned with the development of skills to meet the needs of employers for highly qualified personnel today. Education provides knowledge and preparation that will allow people to respond to future developments in an agile fashion. The evolution of the Design Engineer is not an easy or even a particularly rapid process because of the broad range of knowledge required to carry out the role effectively. In general, few recent engineering graduates possess all of the skills and knowledge needed to fill this role effectively. In Europe, the basic degree program is longer than in North America and many programs require a significant amount of industrial experience. This has been found to provide a much better level of skill and accounts for the fact that many successful Design Engineers presently practicing in Canada were educated in Europe [5-6]. Recently, hiring efforts in Europe have found fewer candidates available to fill these jobs suggesting that EU programs may be moving away from imparting the very characteristics sought by product development companies around the world and prompting Canadian schools to revamp their curricula to provide better preparation for the Design Engineer role [7]. One key problem that has emerged in the development of the Design Engineer and the curricula for teaching these people the necessary body of knowledge is that, while there is no shortage of positions available for people with the necessary skills, there is a reluctance on the part of many engineering graduates to accept a long term position working on a CAD station. Thus, the very inclusion of routine design tasks (such as geometry creation) in the Design Engineering profession is actually making it difficult in some instances to attract young engineers to take up this vital career path and as a result, some less qualified people are still doing this work to the detriment of the companies in question.

4. CONCLUSIONS AND QUESTIONS FOR THE FUTURE

The integrated design, development and manufacturing of a new product in the automotive industry is an extremely complex process involving billions of dollars in assets and the livelihoods of hundreds of thousands of people in many countries. The entire process is highly optimized with the goal of making the best possible product at the lowest possible cost and doing it in the shortest feasible time to respond to changing market trends. With trade barriers dropping around the world, higher cost manufacturing nations and areas such as Western Europe, Japan and North America are under intense pressure to optimize even further and so the consequences of either a bad business decision or a sub-optimal technical decision can be extremely serious.

The decision making processes involved in the product development process include both macro-level (or primary) issues such as corporate finance and plant location along with more immediate technical choices (the secondary decisions) such as whether or not to implement a new manufacturing process, introduce a new material or even call for a non-standard fastener in the next generation product. Each of these decisions, when considered separately may seem minor, but taken in the context of the scale of automotive product development and production, they can spell disaster for a given company or product line [5].

All of the foregoing suggests two questions which the design engineering community might wish to consider. These questions are intended to spark discussion and lead to further research efforts working closely with existing product development companies and in collaborative groups within the academic part of the community.

1. Can a tool be developed to assist design engineers to consider the capability library of the company in making design decisions?

Such a tool would have to be relatively fast and simple to use and it would need to encompass all of the issues raised above in Section 2 of this paper as well as others that are considered relevant. It could take the form of a simple checklist for design reviews or perhaps a more sophisticated tool should be devised. The author has worked with a number of companies in several manufacturing sectors, each one of which would benefit from such a simple and robust decision-making tool for design personnel, especially for younger people.

2. How do we achieve the level of capability in the future generation of design engineers that will enable them to use all of the tools available in modern CAD packages while maintaining their level of interest in doing the more routine tasks such as geometry creation and detailing of working drawings?

The second question is as much a local issue as a profession-wide concern since different locales have different employment outlooks for qualified people. The author looks forward to working with colleagues on these issues and making progress on them in the future.

REFERENCES

1. Oberg Industries, Survey of Top of Mind Issues for SAE Congress Attendees, www.oberg.com, 2003.
2. American Society of Mech. Eng., Integrating the Product Realization Process into the Undergraduate Curriculum, ISBN 0-7918-0126-8, ASME, New York, 1995.
3. P.R. Frise, The Challenge of Staying Current in Today's Automotive Industry for the Design Engineer, Eighth World Cong. on Continuing Eng'g Educ., Toronto, May 2001.
4. Advisory Council on Science and Technology, Expert Panel on Skills, Skills Profile of the Automotive Industry Sector, May 1999.
5. Industry Canada, Automotive and Transport Branch, The Automotive Competitiveness Review: Industry Identified Issues, ISBN 0-662-63613-9, Ottawa, 1998.
6. Industry Canada, Automotive and Transport Branch, The Automotive Competitiveness Review: A Report on the Canadian Auto Industry, ISBN 0-662-63585-X, Ottawa, 1998.
7. P.R. Frise, Identifying the Design Engineering Body of Knowledge, proc. International Conference on Engineering Design – ICED-03, Stockholm, August, 2003.

AGILE MANUFACTURING AND ITS FITNESS FOR INNOVATION

Kerry J Mashford
Interfacing Limited, 3 Hartfield Close, Milton Keynes, MK7 6HN, Tel: 01908 200585, e mail: kerry@interfacing.co.uk

Abstract: The current manufacturing climate is characterised by increasing frequency of new product introductions, shorter product lifecycles and the need for shortest possible time to market. High innovation rate companies today make more than 60% of their profit on products less than two years old. In this environment a key responsibility of any manufacturing company is to maximise its fitness for product innovation. The commercial benefits from doing this usually far outweigh the cost of manufacturing machinery, yet we are still very conservative about investing capital in manufacturing capability and, in this high innovation rate arena, equipment capabilities must be agile enough to change (rapidly) to follow the product portfolio. The issue is therefore how to decide on the right equipment capabilities for your particular product family and how to identify the untapped product innovation space available from your currently installed asset base. This paper presents a method for characterising both the agility requirements of the product(s) and the agility capabilities of the manufacturing equipment and illustrates how these can be used to support capital investment decisions and guide new product development by reference to a specific example.

Key words: Agile manufacturing; Mass customisation; Manufacturing fitness; Innovation; Responsiveness.

1. BACKGROUND

The current commercial climate for manufacturing enterprises is characterised by increasing frequency of new product introduction, shorter product lifecycles and the need for shortest possible time to market. These

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 243-254.

trends are most pronounced in consumer goods but are also apparent in OEM products and, in all cases, are propagated along the supply chain. High innovation rate companies today make more than 60% of their profit on products less than two years old. Product life cycles are typically halving every five years. In a competitive market, time to market for new products has a direct bearing on market share and hence on total sales volume, see Fig. 1. Whether first to market and seeking to secure market share or second (or nth) to market, seeking to catch up, the speed at which products can be introduced has a direct bearing on a company's bottom line.

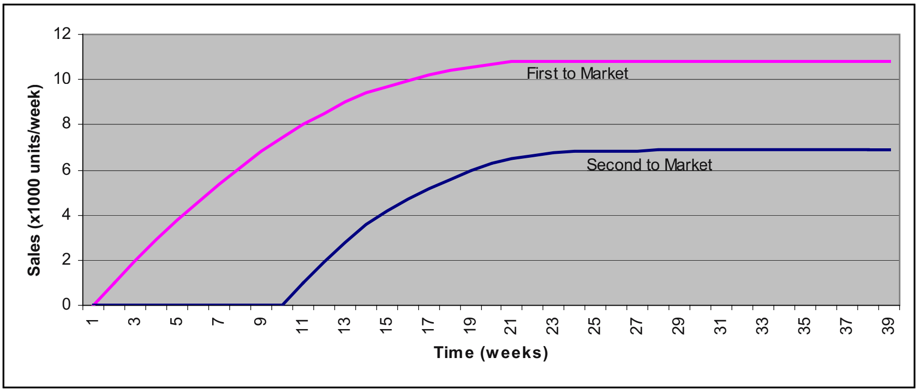


Figure 1. Time to market impacts total sales volume

1.1 Sequential Product and Process Development

In sequential product development where the nature of the manufacturing processes is well established, the product size, shape, etc. are determined with little involvement of the manufacturing function. Even if Manufacturing is closely involved in the product development, it is usually not possible to order change parts for production machinery until the product form has been decided - see Fig. 2. Take, for example, a bottle of bath lotion. Let's ignore, for a moment, the contents of the bottle! The size, shape and height of the bottle must all be known before suitable transport carriers can be sourced. The screw thread for the cap and the size and shape of the cap itself are needed before the assembly end effectors are ordered. Cartons, labels and their associated erection and application equipment depend in turn on the design decisions made on the base product. With lead times for production machinery change parts being as much as sixteen weeks (it used to be at least sixteen weeks) the time taken from deciding to launch a relatively simple new product to having it available to ship can be almost six months. In today's demanding and competitive market place time scales as long as these are unacceptable. For many years, industries have been striving to increase parallelism in their new product introduction activities – Fig. 2 again. Especially in the case of products of an established genus, such as the bath lotion, where the processes and operations used in their manufacture are understood but just require dimensional design, parallel product design is the norm. The container shape, bottle top, contents, labels and cartons are designed with a great deal of overlap in the programme. But there is still a significant delay in designing the change parts for the production equipment as these usually require the product attributes to be known before they can be made (Fig. 2). The reasons for this are both attitudinal and organisational. Providing engineers, product development professionals and business decision-makers with the tools to improve agility in product and equipment design and in investment can increase innovative capacity and improve time to market.

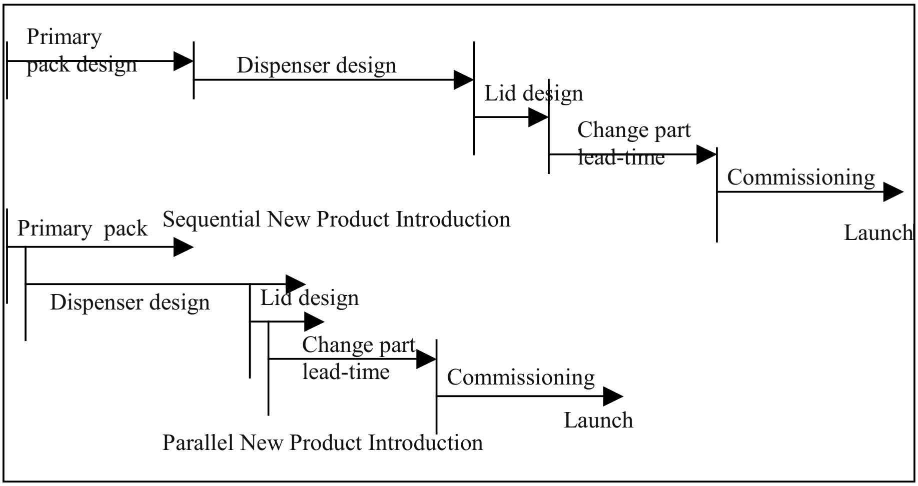


Figure 2. Delay in parallel design of products and processes

1.2 Investment Decisions

In an effort to devolve decision making and improve accountability, especially in larger companies, responsibility for investment in capital equipment and change parts is given to the manufacturing site or function. Their target is to spend as little as possible to manufacture the company's products. Capital expenditure is usually subject to severe scrutiny and requires watertight justification. The swiftest and least controversial way of making such justification is on the basis of it being essential in order to manufacture products currently marketed, e.g. if existing equipment breaks down catastrophically. If investment is tied to a new product introduction which itself is commercially risky, then the less capital expended the better and the more difficult the capital expenditure approval is and the riskier the investment is perceived. Under these circumstances, taking a decision to invest more heavily – even if it can be argued that the proposed new capital equipment, modules or change parts could be used on other future products, is very difficult. Manufacturing is criticised for being profligate and wanting better toys, the product innovation team is unhappy because the proposed additional expenditure is associated with its specific new product and thus affects the cost benefit analysis of the product, perhaps even threatening its launch. Thus, it is clear that, without a framework to view capital investments strategically in the light of the whole current and anticipated product portfolio, a company's manufacturing capability is destined to creep forwards incrementally, doing no more than satisfying the immediate need and placing a recurring brake on the launch of each and every new product.

So, while the declared strategy for the company may be for highly innovative commercial activity with frequent new product introductions, so long as the manufacturing function is being constrained to invest for the present and immediate future production, it will be constantly constrained because of its inherent lack of fitness for product innovation. The responsibility and funding to act to deliver this strategy falls on the marketing and new product development functions, but frequently insufficient thought is given to the role manufacturing might play and, most importantly from a strategic perspective, to fund and manage this function to improve the company's manufacturing fitness for innovation. This organisational disconnection, reinforced by entrenched financial processes and practices within companies (cost centre mentality), hampers their ability to capitalise on product innovative potential, which in turn has severe consequences for bottom line profitability.

2. THE AGILE MANUFACTURING ENTERPRISE

2.1 Future Market Place

For all companies, but especially for those with strategies for a high product innovation rate, having a manufacturing base only able to produce your current product portfolio is a severe limitation. Equipment capabilities must be agile enough to change rapidly to follow the developing product portfolio and lean enough to produce an expanding portfolio economically. Even more than this, we must anticipate that the product portfolio will not only change more frequently but will contain greater variety, as customers become more attracted to mass customised products. In time more products will be designed in terms of a product realisation space, see Fig. 3. Only a small proportion of the possible products may have been made so far but customers can order any product within the product realisation space. This space can be determined by market decisions and/or supply chain process capabilities and it is important that a company's manufacturing capability is sufficiently agile to deliver the required scope of products.

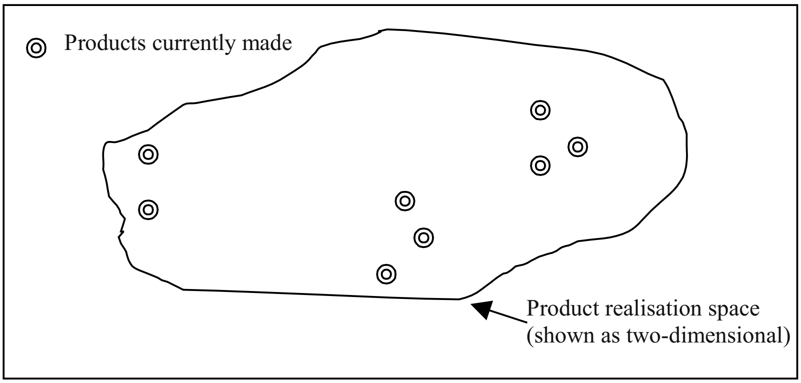


Figure 3. Two-dimensional representation of product realisation space

2.2 Product Design

From the product designers' perspective, new products fall into two broad categories:

i) products that can be made using established processes
ii) products dependent on entirely new processes

Products that fall into the second category are usually related to a significant strategic process development, undertaken to ensure product differentiation that is difficult to replicate and hence provide a competitive advantage – does anyone remember the early years of Viennetta ice cream deserts, Velcro, even sandpaper? Once a process becomes established, the close relationship created between product development and manufacturing during the development of the new process, fades away. Subsequent new products are designed that rely on the, now established, process, and fall into the first category.

In most industries, the design of new products based on established processes is carried out with little and usually badly structured knowledge of the process capabilities of existing equipment. Thus product designers have little basis for anticipating what might be easy or difficult to make, cheap or expensive, swift or time consuming to prepare for, and hence what is involved in time, cost and other resources to bring a particular product to market.

2.3 The Manufacturing Function

For manufacturing engineers their crucial importance to the company's ability to fulfill and profit from its strategy for product innovation is often overlooked. Frequently consulted too late in the product development process, it is too often the manufacturing engineer's depressing duty to inform the product developers that their new product cannot be made with existing equipment and change parts, and, even if the capital expenditure could be justified, it would take several months to source and fit the necessary machine modifications. Furthermore, if this was done, it would preclude manufacture of some established products, or the changeover time between the various machinery configurations would be prohibitive. The frustrating reality for the manufacturing function is frequently that if it had only known that new products were likely to be developed in the direction now taken, it might have recommended purchase of a different machine or line initially. While one accepts that there will always be some new product development in unanticipated direction, the majority of new products are within a product realisation space that was known intuitively to the marketeers for some considerable time – often several years. Good manufacturing departments hate to have to say that a proposed product cannot be made and always pull out all the stops to meet the target launch date, but the consequences are frequently detrimental to existing products. For companies in high innovation rate businesses the general mode of operation inside the company is frequently reactive, always striving to catch up, with departments being caught out by unanticipated demands.

2.4 The Need for New Tools

For a company, the need is for a combination of knowledge and communication: how do you decide on the right equipment capabilities for your particular product families? How do you characterise the agility requirements of the product(s) and the agility capabilities of the manufacturing equipment.[1] Once these are available, how do you communicate this information to those who can use it to make capital investment decisions, product design decisions, estimate time to market, undertake cost benefit assessments and so on and do all this in a strategic, rather than operational time scale.

In the following sections I present a way to view and characterise the dimensions of agility in manufacturing machinery and a simple method for characterising the agility requirements of products. These facilitate the

[1] This can also be extended to considering the agility characteristics of the whole supply chain and the internal company processes, but those are topics for another time.

dialogue between the manufacturing and product development functions, providing a basis for decision making on the design and introduction of new products and the purchase of capital equipment.

3. AGILITY CHARACTERISATION OF MANUFACTURING MACHINERY

Manufacturing machinery tends to be of a generic type, e.g. assembly, machining, filling, cartoning, etc., each of which carries out specific types of operation. The agility of a particular machine can be characterised by the range of specific parameters it can operate in or deliver, the speed at which it can be changed from one setting to another and the way in which those changes are achieved (e.g. change-parts, peg settings, continuous adjustment with jigs, which also gives an indication of the lead time to enable the machine for a new setting). A significant amount of work was undertaken by partners in the ARMMS Network on this topic, some of which is documented in reference [1], exploring alternative approaches to classifying manufacturing agility. For practical purposes it is useful for a particular company to create tables of the various equipment types that comprise their production capability, together with the parameters and parameter values that determine the products the machines are able to process – see Table 1 which is based on specifications of bottling plant in a personal care factory . In addition to this it is valuable to tabulate the time and cost associated with enabling the machinery for a new product (the manufacturing contribution to the product launch programme) and the time to change between products once the machines are enabled (informs decisions on frequency of manufacture, stock turns etc.), see Table 2, which is based on the manufacturing requirements for introducing a 50% extra free promotion for a personal care product, by increasing the height of the bottle.

Table 1. Comparison of alternative machines using key parameters

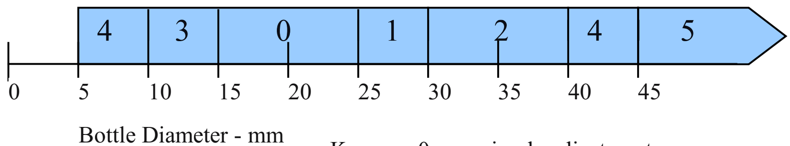
Candidate Bottling Machines	Speed (bpm)	Filling fluid viscosity range (kg/(m.s))	Bottle height range (mm)	Bottle diameter range (mm)
Dedicated Machine (with change parts)	120	$1.0x10^{-3} – 1.5$	120 / 140	50 / 65
Conventional Machine	130	$1.0x10^{-3} – 1.5$	100 - 140	40 - 75
High speed	200	$1x10^{-3} – 1x 10^{-1}$	80/100/120/140	40/45/50/55/60
Low speed / high flexibility	80	$1.0x10^{-3} – 1.5$	60 - 200	30 - 100

These tables can be used to enable decisions to be taken on which plant to use for a specific new product or when considering manufacturing options in early and late life of the product. They can also be used to support purchase decisions in conjunction with an analysis of historical and projected product innovation activity. Tables 1 and 2 are examples of such tables for bottling machines.

Table 2. Consequence of accommodating change of parameter in new product on alternative installed equipment

Bottling Machine	Consequences associated with bottle height change					
	Likely method / innovation category	Parts cost	Lead time to enable	C/O time	Waste cost / change over	Labour cost
Dedicated Machine	New module / 2	£12,000	16 weeks	2 hr	£100	£50
Conventional Machine	Change parts / 2	£2,000	4 weeks	1hr	£100	£25
High speed Machine	Programmable /1	£50 (code)	1 day	1 cycle	zero	zero
Low speed / high flexibility	Manual adjustment /0	zero	zero	4hr	£50	£100

For a specific change in a specific parameter value a machine will not always need the same work to enable it to deliver it. Some changes may be very easy, cheap and swift, while others prohibitively lengthy or expensive. Another way of visualising the capability space of specific machines is presented in Fig. 4 and 5, which depicts a mapping of the difficulty of achieving variable values in one or two-dimensional form. The representation depicted in Fig. 4 was generated by the author and colleagues as part of the ARMMS project [1]. For variables that are independent a series of one-dimensional mappings will suffice, one for each parameter, whereas if variables are dependent it may be necessary to represent them in two-dimensional space or even in n–dmensional theoretical space.



Key
0 – simple adjustment
1 - modifications to software
2 - change parts required
3 - new module required
4 - modifications required
5 - new plant / line required

Figure 4. Example of mapping of independent variables

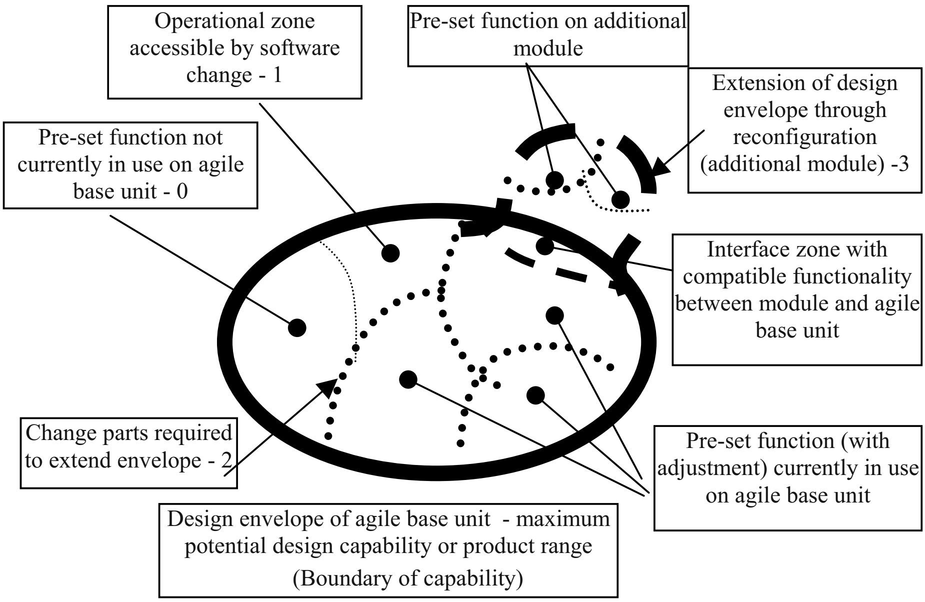


Figure 5. Example of mapping of independent variables

Since we are interested not only in the capability of the machines but also the agility we can represent this with respect to each key parameter by plotting its range against the speed to change from one value to another within the range, Fig. 6.

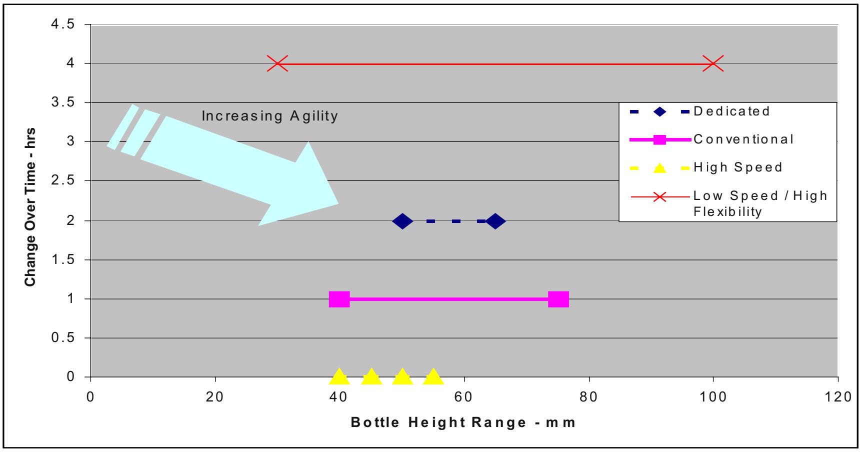


Figure 6. Representation of the height agility of various bottling machines

4. NEW PRODUCT CHARACTERISATION

While it is interesting to map out the process parameter values for a particular new product, the value of this becomes apparent when the next step is taken; that of determining how quickly and easily the available equipment alternatives can be configured to produce the new product. Take again the example of the bottling line and the example that a proposed new product is of a different height to existing products made on that machine, a different shape and requires the application of differently shaped and sized labels. The table summarises the difficulty encountered in making the new product on each of the plant types currently available to the manufacturer, using the same key. Thus a level 5 innovation requires whole new plant, whereas a level 1 innovation needs only a new machine control program to be generated. The innovation category level of a proposed new product can be used by product line management to inform their product portfolio management decisions. More quantitatively, using the tables created for each variable, and taking care not to double count cost or time, the total cost and time to introduce a proposed new product on each of the available machine types can be assembled. In this case, data is gathered from equipment manufacturers and internal production data.

Table 3. Example of ease / difficulty of equipment to process new product

Plant	New Bottle Shape	New Bottle Height	New Label
Dedicated Machine	5	4	2
Conventional Machine	3	2	2
High speed machine	1	1	1
Low speed / high flexibility	1	0	0

Table 4. Evaluation of time and cost associated with the introduction of a proposed new product on alternative plant

	Parts cost	Lead time to enable	C/O time	Waste cost / change over	Labour cost per C/O
Dedicated Machine	£150,000	16 weeks	2 hr	£100	£50
Conventional Machine	£20,000	6 weeks	1hr	£100	£25
High speed machine	£2000 (coding)	1 week	1 cycle	zero	zero
Low speed / high flexibility	zero	zero	4hr	£50	£100

5. HOW TO ACHIEVE MANUFACTURING FITNESS FOR INNOVATION

The key question each company must ask before embarking on a campaign to enhance its manufacturing fitness for innovation is, "How much product innovation and of what kind are we going to have and so, what level of agility do we need?" Agility costs and this cost must be commercially justifiable to ensure ongoing profitability of the company. The aim for all companies whatever their level of product innovation, should be to make the introduction of the vast majority of new products run-of-the-mill. These product introductions, many of which are straightforward variants of existing products, should be achieved with minimal cost and disruption to existing production and, perhaps most importantly, without undue demands on the valuable resources of manufacturing and product development personnel. It is important for these experts to be able to concentrate the bulk of their effort at a strategic level on the development of new categories of product and the new process developments to enable them.

5.1 Common New Product Introduction

Thus, for those run-of-the-mill new products:

1. Manufacturing teams document the capabilities of existing equipment using the agility mapping options given in Table 1 and Fig. 5,6 & 7.
2. Manufacturing teams communicate these capabilities to product design teams to enable rapid evaluation of potential new products.
3. New product development teams use the capability maps and charts to identify potential innovative space for new products that can be delivered to market with greatest ease.
4. New product development categorise new product ideas as in Table 3 to gain an understanding of the difficulty of bringing each concept to market.
5. For short-listed product concepts, product development teams, together with manufacturing build composite tables containing the time, enabling cost and operational cost, etc. of introducing a particular new product.
6. Teams use this information to guide decision-making and prepare cost / benefit proposals for specific product launches.

5.2 Strategic New Product Innovation

For the strategic new product families and associated new process developments:

1. New product development teams record the parameters that characterise the new product family and decide on a range of values for these, within which anticipated new products will reside, as in Fig. 4.
2. Product line managers decide target number of new products within this family over a two to five year period (this can usually be derived from company strategy), that will need to be made using the new plant or line.
3. Product line managers using market research to determine projected volumes and product life cycles of such new products and, with supply chain management, determine the stock turns and frequency of manufacture (even to manufacturing batches of one).
4. For commercially available equipment (or elements of the whole line that are commercially available), manufacturing teams map their characteristics as in Table 1 and Fig. 4, 5, & 6, and use this to select suitable candidates.
5. For completely new function / processes, process and machine development teams use the product and supply chain requirements to hone the final equipment design.

5.3 Complementary Approaches

There are other complementary ways in which various players can contribute to the delivery of higher product innovation rates. These are mentioned below for completeness but are not expanded here. Product design can focus on ways of using product customisation and delaying product differentiation to increase variety of final products while minimising the need for agility at the early stages of the supply chain. Machine designers can look to increasing modularity, turning design variables into configuration variables and configuration variables into operational variables in order to increase the agility of individual machines and machine types. Machine suppliers can look to new business relationships with their customers, for example leasing of machines and modules, to enable their customers to operate with greater apparent agility in their plant, while simultaneously shifting capital purchases into revenue spending.

REFERENCES

1. ARMMS final report, section D2.3, Prepared by AMTRI

This paper draws on the work of the author in industrial practice, much of which is confined to internal reports, and her role as Director of the DTI's Manufacturing 2020 Foresight Progamme, whose work is summarised in its final report "Manufacturing 2020: We Can Make it Better", DTI, 2000.

HOLISTIC APPROACH FOR PRODUCTION RAMP-UP IN AUTOMOTIVE INDUSTRY

Günther Schuh, Jens-Christian Desoi and Gregor Tücks
Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen, Germany Steinbachstrasse 53, D-52074 Aachen, +49 (0)241 80- 27387, +49 (0)241 80- 22293 g.schuh@wzl.rwth-aachen.de

Abstract: The product development and the production Ramp-Up are still major cost drivers in automotive industry. An outstanding new product performance can only be achieved by combining the influence of technical product design complexity and cost drivers in production as well as influences of market potentials. The approach presented in this paper serves as a key for the formation of the Ramp-Up fit to proactively support the Ramp-Up management by developing market/ design as well as production triggered Ramp-Up scenarios. Thus, it will be possible to detect margin-optimised and complexity-reduced Ramp-Up-strategies and the explicit planning of the Ramp-Up curve. The present approach is separated in the three stages - Strategies, Planning and Evaluation. At first, production-optimised Ramp-Up strategies will be developed, which guarantee the specific absorption of market potentials. Afterwards, from a production point of view, planning determinates for building up the Ramp-Up curve will be presented. Finally, a Ramp-Up benchmarking process for OEMs and suppliers during Ramp-Up phases will be derived. The benefit of the presented approach has been verified by industrial experts as a result of a nationally funded study to open up potentials within the Ramp-Up phase.

Key words: Production ramp-up; New product launch; Market oriented.

1. INTRODUCTION

In a global, dynamic and competitive environment the introduction of new products has become a focal point of attention for automotive OEMs and suppliers. Increasing speed of innovations, shortened product lifecycles and a continuous differentiation of product programmes characterise the current situation in the automotive industry. During the last four decades

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 255-268.

product lifecycles have decreased by 60 % and, at the same time, the number of new product innovations has significantly increased. The production Ramp-Up becomes a more complex business process and is of special interest as the lifecycle profit of a product is primarily determined in the early phases of product market introduction.

To increase their level of competitiveness European companies especially have to amplify product customisation. Ramp-Up specific individualisation potentials are mainly generated by the ability to cope with complexity and variety [1]. The short time of changeover form R&D to series production emerges to a strategic chance for real differentiation from competitors due to own product innovations. Lost sales profits due to production problems in the Ramp-Up phase can never be compensated because of decreased product life cycles [2]. Thus, the proper control of production Ramp-Ups advances to an eminent success factor in the automotive industry [3].

2. PROBLEM STATEMENT

From authors point of view the main Ramp-Up problems as well as existing chances can be described by four theses.

Thesis 1 - Everlasting problem: The Ramp-Up quantity and impacts within the whole supply-chain will increase. The quantity of Ramp-Ups will escalate because of decreasing innovation-cycles and rising model and derivative diversity as seen in Fig. 1. Moreover the trend of dispensing R&D and production responsibility to suppliers will continue. As a consequence, it will be on a significant challenge for OEMs and suppliers to be able to cope with the rising Ramp-Up quantity and complexity on each value-added stage. Reliable evaluation instruments do not exist but would generate auspicious benefit.

Thesis 2 - Complexity of products and technologies will further increase and strongly influence the Ramp-Up process effectiveness. The essential strength of the European automotive industry and differentiation against global competitors will still be achieved by characteristics like Mass-Customization, strength of innovations and quality orientation. Retaining on such key-characteristics, only such Ramp-Up strategy and planning approaches would be applicable, which support and keep the rising product- and process complexity manageable. Variance for diversity is necessary, the key of success however is a market oriented and temporal reduction of variance during the Ramp-Up phase.

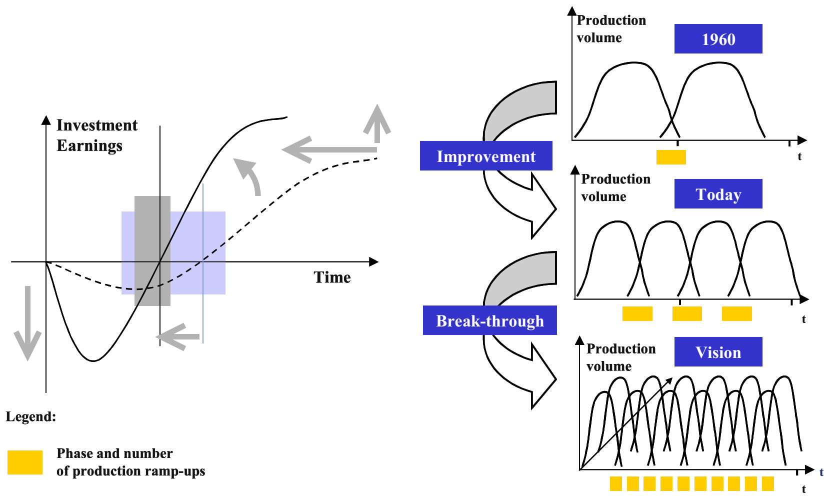


Figure 1. Current situation and potential in Ramp-Up management

Thesis 3: Skimming of Margins in early stages and prevention of Lost-Sales: If a company focuses on innovation leadership instead of price leadership, it is essential to gain profit and to guarantee the return on investment to compensate innovation efforts. Losses in time and money due to the need for additional installation and troubleshooting efforts significantly lower profit margins. To achieve notable profit in automotive industry vehicles have to be personally configured and to be highly diverse. In consequence, in future not only generating volume in early stages after launching new products will be of major importance, but also the ability to cope with the combination of margin focussed variants and according volumes for located customer segments is showing promise.

Thesis 4: Only Ramp-Up robust production systems avoid troubleshooting: Today's production Ramp-Up is very often still managed on a trial and experiment short-time basis and is not free of process failures. There is a significant lack of knowledge and on hour to react on critical disturbances production. This generates high additional costs for production downtime and for the ability to provide the right knowledge to the right people at the right time. Thus, the Ramp-Up phase is often working on a stability edge. As a consequence the Ramp-Up process has to be secured and evaluated to provide the correct reaction strategies due to plan variances.

3. STATE-OF THE-ART THEORIES AND APPROACHES IN RAMP-UP MANAGEMENT

In scientific literature no or only less meaningful information for the production phase “Ramp-Up” can be found. State-of-the-art approaches are mostly isolated applications and solutions of specific Ramp-Up problems. A complete overview of the “Ramp-Up” phase and the management of this complex phase does not exist. Nevertheless within the Scientific Community a couple of researchers focus on the field of production Ramp-Up. Frizelle presents a framework for Ramp-Up manufacturing where the three dimensions novelty, learning and performance in Ramp-Up manufacturing form the body of the framework [4]. Zeugträger argues that the essential potentials in effort reduction in this phase stay unused [5]. According to Bungard no automaker has developed a Best Practices Model to reduce losses in the change of product models because of adherence to tayloristic thinking patterns to handle this phase with sequent processes [6]. Zangwill and Kantor maintain that in the Ramp-Up of serial products a mostly over-fast transfer from research to a commercial product is taking place, which results in the enterprise suddenly facing outstanding problems [7].

Some studies pick up the problem of Ramp-Up management. The latest are “Word Class Vehicle Launch Timing” (1992-1996, USA), “An Exploratory Study of International Product Transfer and Production Ramp-Up in the Data Storage Industry” (2000, USA) [8], “Fast Ramp-Up – Schneller Anlauf von Serienprodukten” (Germany, 2002) [9] and “Developing the R&D to Production Ramp-Up Process 2001-2003” (Finland) [10]. In practice some companies introduce special Ramp-Up teams with participants of all departments involved during Ramp-Up activities for a smooth communication and coordination. For example at DaimlerChrysler commercial vehicle division the Ramp-Up management consists of nine partial projects with fixed project phases, responsibilities and dates. An interdisciplinary team measure the project status at defined quality gates. For achieving the economic goals within the project a single project manager is responsible [11]. To evaluate the productivity of Ramp-Ups BMW implemented a parameter based reporting system which identifies critical characteristics and certain measures within a production Ramp-Up. Actual data will be compared with knowledge-based nominal values. The parameters of the reporting system are separated into quantitative (technical) parameters and qualitative parameters such as quality of communication between the different departments involved. A reasonable usage of this reporting system can only be guaranteed by charging it with actual reviewed Ramp-Up data [12]. Another reporting system established at Ford is a lesson learned documentation based on a web platform. Target deviations, failures and their measures of worldwide production Ramp-Ups/ product launches can be archived systematically. The tool supports the improvement of future Ramp-Ups by using the evaluated data of different Ramp-Ups in global Ford plants.

There are completed studies and practical approaches for Ramp-Up management in the automotive industry. These existing methods include Simultaneous-Engineering, Project-Management, Total Quality Management, different Re-engineering approaches and Reporting Systems. However, each of these methods only provides a limited view of the issue of Ramp-Up management and neither present nor future problems have been addressed. A cutting edge reference business process for the Ramp-Up phase is required. Thus, a key characteristic of a new Ramp-Up approach will be to shift focus from the resource driven production process to the knowledge intensive Ramp-Up process.

4. CONCEPTUAL DESIGN OF THE HOLISTIC RAMP-UP APPROACH

Notable improvements can only be achieved by a holistic approach which emphasises three thematic directions.

1. A general Ramp-up strategy is need to orient Ramp-up preparation between partners towards concerted objective. But the chosen strategy has to be type and process specific adapted. The chosen strategy provides the essential input variables for the planning process to be initiated afterwards.
2. The Ramp-Up has also to be planned specifically. Meanwhile classical production planning parameters are insufficient to coordinate complex and frequent production run-ups and shutdowns in own and supplier plants. A detailed tomography with additional parameters could help to solve that problem. Furthermore the planning is not a singular process but has to be adjusted regularly on ever-changing ancillary conditions.
3. Because of high-grade interconnectedness of partners it is arguable whether actual strategy and planning can exceed the limits of future demands. It has to be ensured that all players within the partner network have necessary Ramp-Up competences and act conformable in kinds of strategy and planning declarations. A benchmark based evaluation audit enables the identification of deviations, points out optimisation potentials and allows to initiate remedial actions.

At first – within the strategy related thematic direction – parametric Ramp-Up strategies with focus variant and volume triggering will be developed. To raise profitability while concurrently lowering complexity will be the major advantage. With such parametric Ramp-Up strategies it will be possible to align the production e.g. for providing at first the high margin variants and related volumes for special customer segments. Hence it will be affordable to forego on high volumes in early launch phases but to gain customers willingness to pay for novelties. That means complexity could be actively managed by strategy and planning parameters by means of controlling volume and variety. Thus the parameters variance and volume are the main influencing factors in strategy design.

With a general Ramp-Up strategy partners are aligned towards the long term view. Based on this the tomography is used for the short-term planning procedure. Due to the fact that ordinary production planning parameters are insufficient, with the tomography appropriated parameters are presented to solve this problem. Furthermore, the Ramp-Up planning will not be accomplished only once, but has to be adopted according to actual and changing boundary conditions. The planning steps will feature more detailed time-based planning with extended planning determinants from two (variety and volume) to eight. This advanced planning procedure will help to avoid layoffs in times of restructuring of industries. To react on deviations to unexpected disorders also an active deviation management will rely on the developed tomography.

In Ramp-Up management considering only the strategy and planning aspects is not sufficient due to highly integrated partner networks in automotive industry. It has to be assured that suppliers are able to meet the Ramp-Up strategy and planning values. Therefore a benchmarking based evaluation instrument enables the identification of derivations and optimisation potentials and measures, as presented in the third section.

5. DETAILED PRESENTATION OF THE HOLISTIC RAMP-UP APPROACH

5.1 Ramp-Up Strategy

Within this chapter a strategy model will be presented. Beneath one quite familiar strategy three new Ramp-up strategies will be introduced. As seen in Fig. 3 the parameters expressing the strategies alignments are variety, decoupling level, Ramp-Up time and utilization.

At first the already existing strategy – with focus on high utilization degree within the Ramp-Up phase – will be introduced. Major aim of this strategy is a high volume production at once after new product launch. This kind of strategy is called “Volume-First” because it is dominated by thoughts of gaining profit with advantage of scale. Major leveraging parameters are

Ramp-Up time and Utilization. Though according to the author's expectations in future there has to be more intelligent Ramp-Up strategies. Complex Ramp-Ups and higher quantity can only be harnessed as a cash-cow when it will be possible to rule efforts and benefits. Ramp-Up efforts related to handling rising product- and technology complexity has to be adjusted on achievable market potentials. A time wise reduction of product variety, interworking with customers higher price attendance, will be necessary. The major goal is to make profit first with reduced complexity and afterwards to increase volumes on secured processes. The three new Ramp-up Strategies take this change of paradigm into account and will be called "Tuning-Profit-into-Scale Strategies". Major leveraging parameters are variety, decoupling level and Ramp-Up time.

Turning -Scale-into-Profit - Volume First

In industrial practice the question of the right Ramp-Up strategy is dominated by two basic procedures. The most suitable euphemism for both is "Volume-First-Strategy" because both focus particularly on short production downtime periods. The classical "Shut-Down" approach - practiced in recent years – facilitates certain production facility changes and new installations. However such a strategy can cause at least a few weeks of down-time and extensive losses. Efficiency requirements oblige OEMs and suppliers to disavow more and more from this procedure. Today a more fluent changeover from old to the new model production is preferred. The so called Back-To-Back-Launch offers these efficiency potentials. Here a total changeover from the old to the new model production within a few days is aspired. Even though flexibility potentials of today's modularised facility concepts make this procedure possible there are critical boundaries due to the resource rivalry in the Ramp-Up. Pre-Series production, while the old model is still produced, is necessary to provide vehicles for homologation, crash tests and marketing purposes. The Volume-First Strategy is most suitable for high volume and low variety production programmes.

Turning-Profit-into-Scale - High Margin Variants First

As a consequence of high product diversity balancing the essential bottleneck resources and hence managing the Ramp-up related complexity is an increasing challenge. It is quite obvious that for European OEMs product diversity is one of the major success factors. Managing the ramp-up related complexity by means of ramp-up equalisation and staggering provides an opportunity to combine a marketwise as well as production oriented long-term planning. Product variety adjustment in attractive customer segments and volumes will be one of the major leverages of the new Ramp-Up strategies. Related configuration parameters are Variety, Decoupling Level and Ramp-Up Time. With these configuration parameters an enhancement of the strategy portfolio with three new strategies and specific optimisation potentials is feasible as seen in Fig. 2. A solution to decrease the Ramp-Up specific system-complexity features the „Dedication-Strategy“ the “Slow-Motion-Strategy as well as the “Step-by-Step” strategy.

The “Slow-Motion-Strategy” is characterised by a parallel Ramp-Up of all vehicle variants whereby the whole volume keeps on a constant and low level for a longer time period. However, afterwards the volume will upswing conspicuously more steeply in comparison to the sequential Ramp-Up. In general this kind of Ramp-Up type offers advantages for run-up of highly automated manufacturing systems, because specific problems of automated production normally only appears on maximum system severity. Advantages of this strategy are expected especially in phases of pre-series where an overlapping of the old and new models production is quite familiar to prepare the new model launch. Ongoing production and distribution of the old model will be possible and resource rivalry will fail to appear.

The “Dedication-Strategy” is characterised by a sequential Ramp-Up of consecutive vehicle variants with accumulated volumes. Major target of this strategy is to produce and to launch variants on a gathered way. A relief of production and logistics will be achieved because of an optimised bounding of orders in charges. This Ramp-Up strategy is suitable especially for final assembly requirements and to support the Ramp-Up quality management. Reasons for inadequate quality and process-related disturbances could be allocated up to related variants. New variants will only be launched when all new assembly operations are learned. So assembly collaborators could be relieved during the Ramp-Up because of a well directed reduction of variance.

The “Step-by-Step Strategy” is especially suitable when an enormous technical complexity has to be accomplished. This Ramp-Up strategy focuses on process and area specific problem characteristics, the Ramp-Up curves are decoupled and the area and process related and optimised separately. It is not a matter of an isolated consideration of different areas in production, because the processes proceed in parallel or in sequence. The clue is in the holistic integration and optimisation of the controversial requirements of the different areas within the value-added chain. The step-by-step strategy requires a number of buffers within the areas, which is contrary to the lean-layout approach.

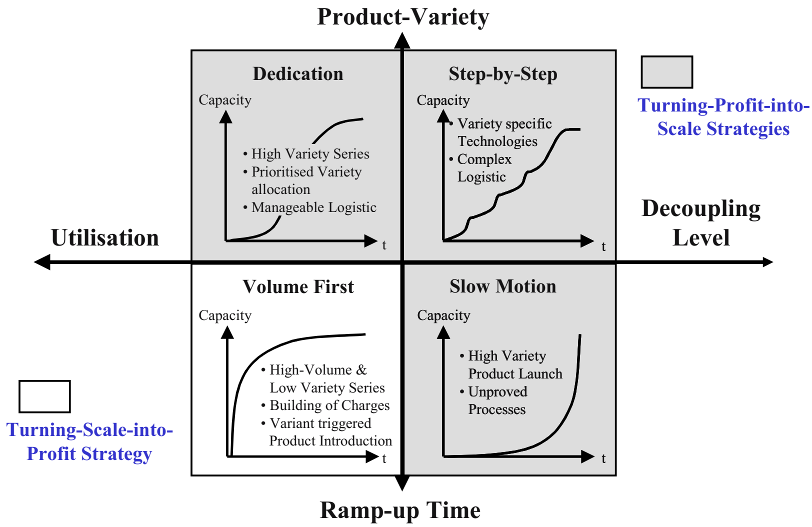


Figure 2. Different kinds of Ramp-Up Strategies

5.2 Ramp-Up Planning

The second part of the holistic approach deals with the realisation of Ramp-Up strategies in the Ramp-Up planning from production point of view. As described in chapter 5.1 the Ramp-Up strategies are characterised by the active planning determinants volume and product variety. These active planning determinants of influence in their entirety passive planning determinants, which are not directly or only in part controllable. Primarily the passive determinants process chain and quality are driven by the volume and product variety. The realisation of the Ramp-Up strategies in the Ramp-Up planning phase takes place by a determined control of these planning determinants.

The Ramp-Up planning phase goes in parallel with the product development process in sense of a concurrent engineering process. In early phases the planning reliability for Ramp-Up strategies is not given due to the poor forecasting stability, because the influences of the market is determined a long time before start of production. Thus, the Ramp-Up curve will be planned to certain point of time along the Ramp-Up phase. Due to the increasing planning and forecasting reliability along the planning phase the characteristics of the planning determinants will be more concrete and detailed. The changes of characteristics of the planning determinants are described in the so called “Ramp-Up Tomography” as seen in Fig. 3.

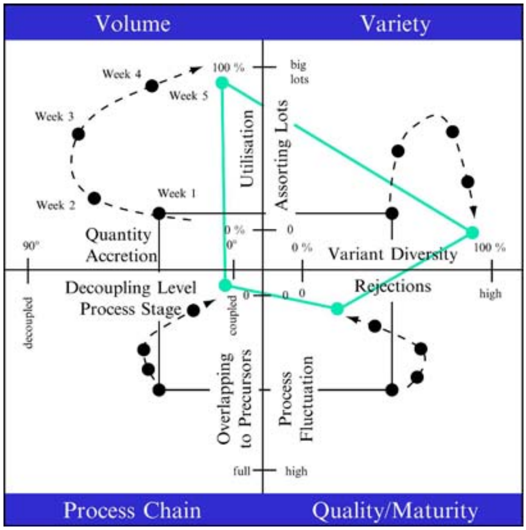


Figure 3. Configuration of the Ramp-Up Tomography

In the Ramp-Up tomography the four planning determinants are described by eight operative planning parameters. The volume is specified by the quantity accretion and utilisation. The planning parameters of the product variety are variant diversity in the sense of the model mix and the assorting lots. In addition to that, the planning parameters for the determined process chain and quality/ maturity are the decoupling level on process stage, the level of overlapping to precursors (Ramp-Up/ Ramp-Down), process fluctuation and rejections.

As a basis for the realisation of the Ramp-Up strategy in the planning of the production new and innovative design concepts for production equipment are necessary. These concepts have to fulfil the requirements concerning volume flexibility, product mix flexibility and change-over flexibility. Especially modular built production equipment can be easily configured to particular stages due to the different stages of the planning phase. But from author's point of view a generic and optimal Ramp-Up curve will not be possible for example due to unplanned changes in product. In consequences production Ramp-Ups in automotive industry are faced with plan deviations and failures in production. If deviations take place the Ramp-Up tomography will be adjusted to new management objectives. Thereby, modular flexible production equipment, fixtures and tools can be adjusted to the Ramp-Up planning parameters as well as for achieving the operative Ramp-Up targets.

In summary, by implementation of the Ramp-Up tomography in planning activities, aggressive launch targets can be planned for the long term and adjusted quite easily in the short term with operative Ramp-Up parameters. Furthermore the tomography serves as a basis for internal and external supplier performance measurement as described within the next chapter.

5.3 Evaluation/ Benchmarking

The third stage of the holistic Ramp-Up approach describes the Ramp-Up evaluation concept using established benchmarking theory. Because of increased number of partner and highly integrated supplier networks it is necessary that internal and external suppliers can guarantee their Ramp-Up fitness due to strategy and planning requirements. The characteristics of the Ramp-Up planning parameters have to be achieved. Due to that a Ramp-Up benchmarking serves as an useful evaluation instrument to measure the internal and external supply capability of suppliers for Ramp-Ups.

The benchmarking concept is divided in two areas "Benchmarking-Process" and "Fields of Action", where the benchmarking-process forms the backbone of the concept as seen in Fig. 4.

Within the benchmarking process four stages are proposed starting with the preparation of the benchmarking process, analyse in terms of actual comparison, implementation of measures and finally process maturity in terms of circulating realisation of benchmarking projects. In addition to that, fields of action are considered, which have to be investigated at the supplier to analyse the Ramp-Up ability due to the selected Ramp-Up strategy. The main field of action forms the supplier's business processes followed by organisational and production structure including equipment and process ability. The third field deals with the product degree of maturity followed by the supply chain and personnel. The Ramp-Up tomography gives the input for the benchmarking with its planning determinants volume, variety, process chain and quality. The benchmarking takes place during the Ramp-Up planning phase where the supplier has to provide its Ramp-Up feasibility, ability and stability. In the following the proceeding during the preparation phase will be explained in more detail, because this phase is of particular importance.

The preparation phase is divided in three sub-phases as shown in Fig. 4. In the beginning the Ramp-Up benchmarking project has to be defined. Within the second stage the benchmarking team is determined, that means, it has to be established who is involved in the benchmarking project. Because of the complexity a matter of particular interest is the determination of key parameters for performance assessment within the benchmarking project.

Within this phase all relevant quantitative and qualitative key parameters for the benchmarking are determined.

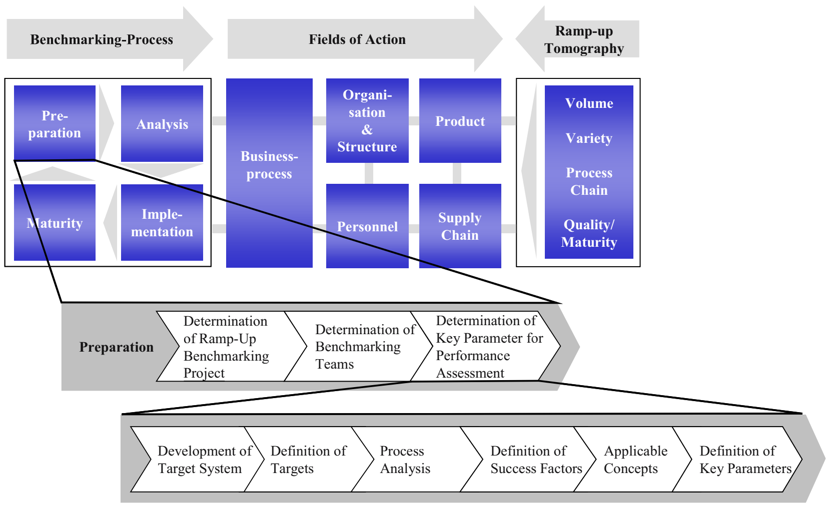


Figure 4. Detail of Preparation phase within the "Ramp-Up Benchmarking Concept"

Starting point of the proceeding is the development of a target system, which is defined especially for the requirements of a Ramp-Up benchmarking. After that, targets for the production Ramp-Up have to be defined and classified within the target system. The realisation of the defined targets is often afflicted with problems, which are under investigation during the supplier's process analysis within the third stage. Based on the identified problems in the next step success factors for the Ramp-Up have to be defined and classified into the target system. In the fifth step a usable concept is developed, where the success factors can be realised adequately in practice. In the last step the definition of the benchmarking parameters takes place, which represent the values for the performance measurement of the supplier and are derived from the success factors.

In summary, the Ramp-Up benchmarking serves as a suitable evaluation instrument for the Ramp-Up planning phase. From the results of the benchmarking process optimisation measures can be derived and implemented within the Ramp-Up process.

6. CASE STUDY

The above presented approach is a suitable instrument especially for the European automotive industry. The German Federal Ministry of Education and Research (BMBF) agreed on the potential chances and risks of production Ramp-Ups in serial production for OEMs and their suppliers and promoted a study with the topic "Fast Ramp-Up - schneller Anlauf von Serienprodukten". The aim was to detect the essential influencing factors and potentials within the complex phase of production Ramp-Up. On top of that, future research fields and initial solution approaches should be identified. With support of the Association of German Automotive Industry (VDA) the WZL identified the central and industry specific Ramp-Up problems in automotive industry. With several experts of OEMs and their supplier solution approaches were discussed. Based on that, the holistic approach was detailed exemplarily for a defined value chain in production by WZL in cooperation with an automotive OEM. As a result of industrial expert inquiries the potential benefits of the holistic Ramp-Up approach are enormous. This approach will be enhanced with different companies participants in an international project which will be applied for.

7. CONCLUSION AND FUTURE PROSPECTS

Based on the problem statement of production Ramp-Up in the automotive industry presented in this paper a hypothesis has been derived, that the increasing number of always complex production Ramp-Ups can only be faced by developing a holistic approach for market and value chain optimised production Ramp-Up. The approach presented is based on three stages Ramp-Up strategy, Ramp-Up planning and Ramp-Up evaluation using benchmarking technique.

REFERENCES

1. G. Schuh, H. Riedel, I. Abels, J. Desoi. Serienanlauf in branchenübergreifenden Netzwerken - eine komplexe Planungs- und Kontrollaufgabe, Werkstattstechnik online, Nr. 11/ 12, 2002
2. G. Wiesinger, G. Housein. Schneller Produktionsanlauf von Serienprodukten - Wettbewerbsvorteile durch ein anforderungsgerechtes Anlaufmanagement, Werkstattstechnik online, Nr. 10, 2002
3. H.-P. Wiendahl, M. Hegenscheidt, H. Winkler. Anlaufrobuste Produktionssysteme, Werkstattstechnik online, Nr. 11/12, 2002
4. G. Frizelle. A framework for manufacturing Ramp-Up, PICMET 2003 international conference, Portland, Oregon, USA

5. K. Zeugträger. Anlaufmanagement für Großanlagen, Fortschrittsbericht VDI Reihe 2, No. 470, VDI Düsseldorf, 1998
6. W. Bungard, K. Hofmann. Anlaufmanagement am Beispiel der Automobilindustrie, Neue Organisationsformen im Unternehmen, Bullinger & Warnecke, Springer, 1996.
7. W. Zangwill, P.B. Kantor. Toward a Theory of Continuous Improvement. Management Science, 1998
8. Study. Word Class Vehicle Launch Timing - An Exploratory Study of International Product Transfer and Production Ramp-Up in the Data Storage Industry, USA, 2002
9. A. Kuhn, H.-P. Wiendahl, G. Schuh. Fast Ramp-Up - Schneller Produktionsanlauf von Serienprodukten, Studie, Verlag Praxiswissen Dortmund, 2002
10. Study. Developing the R&D to Production Ramp-Up Process 2001-2003, Finland, 2003
11. H.D. Bürgel, M. Binder, R. Bayer. Das Projektmanagement zur Optimierung der Schnittstelle Entwicklung/ Serienanlauf im Nutzfahrzeug-Geschäft der Mercedes-Benz AG am Beispiel des Montagewerks Wörth, in P. Horváth, G. Fleig. Integrationsmanagement für neue Produkte, Schäffer-Poeschel, 1998
12. J. Weber. Den Serienanlauf sicher im Griff – ein erfahrungsbasiertes Kennzahlensystem, VDI-Z, No 1/2, 2001

INNOVATION TOWARDS IMPROVED LAMINATED STRUCTURE DESIGN AIDS

Daniel Coutellier

LAMIH-UMR CNRS 8530, University of Valenciennes and Hainaut-Cambrésis, le Mont Houy, F-59313 Valenciennes Cedex 9, France

Abstract: Industrial performance goes through an ongoing improvement of the whole process permitting the transition from R & D to product design, down to the industrialization or even the final recycling. Innovation plays a pivotal part in design, organisational changes and improvement : It makes it possible to define increasingly performing processes. However the first and foremost preliminary step quite obviously boils down to the improvement of the different tools used at one time or another throughout all the steps pertaining to the processes being used. This article hereby proposes an overall view of the innovating breakthrough in the field of improved design aid tools for laminated structures.

Key words: Composite, Damage, Design, Simulation

1. INTRODUCTION

Within the design of laminated structures (made up with multi-layer materials) numerous problems still occur, more particularly whenever the issue of the numerical study of those structures are submitted to extreme changes such as impacts or overall shocks (e.g. structure collision in the field of transportation) is addressed. Composites, more particularly laminated ones, are currently used in order to design wide ranging structures in the field of transportation (aeronautics, marine, automotive or railway technology…). Improved numerical tools and higher passive safety levels have therefore resulted in changed design specifications and increasingly rigorous standards.

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 269-280.

In the field of composites, the experimental approach is still too often resorted to for the validation of numerous steps in the design phases, but this has been made necessary on account of behaviour complexity. For instance, the material transition from the test-piece scale to the whole structure scale is difficult on account of all the differences not only in terms of geometry but also of manufacturing process. As an aid to improve the scaling-down of those experiments, the requirement is to improve the numerical simulation tools for the study of those structures. This study, innovatively goes through the performance of the codes which take into account the physical phenomena and translate them into computational models. It is all the more complex as those phenomena significantly vary with the constituents making up the laminates used. In order to consider the heterogeneous behaviour of the material and therefore of its constituents, the numerical tools have to provide a representation as accurate as possible of the damage modes, and consequently of the failure modes, encountered in most cases. The approach presented here, brings about the first tentative answers to this issue. It is part of the improvement of the numerical simulation tools aiming at an innovative approach to laminated structures design.

2. PHENOMENON PRESENTATION

A laminated structure is made up of the assembly of several different layers which may themselves in their turn be made up with so-called standard materials (featuring well-known behaviour laws as in the case of metallic materials), or alternatively composites made up with fibres in a matrix. As regards the latter, the phenomena usually encountered during the damage phase are primarily at the layer level such as fibre fracture, fibre/matrix debonding, and matrix cracking (Fig.1). This propagating damage then generates ensuing damage at the laminate level, particularly with delamination phenomena revealing splitting between the different material plies (Fig.2). All these phenomena will depend not only on the basic constituents but also on the loading being applied. More particularly, when issues of collision or impact phenomena are to be considered, strain rate effects can no longer be neglected within the laminated material configuration.

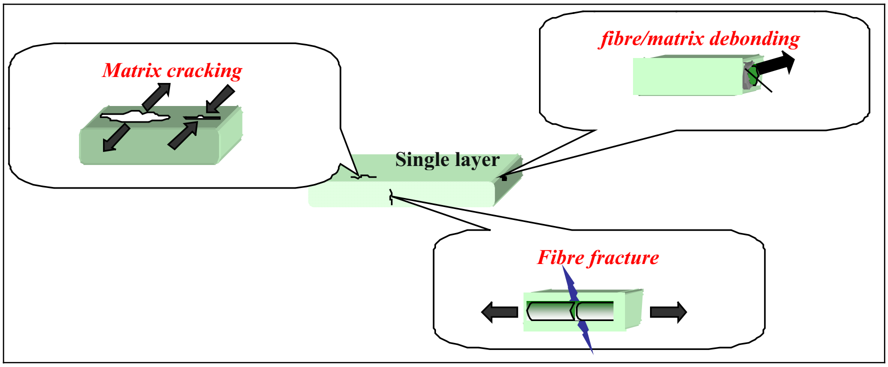


Figure 1. Damage modes within the elementary layer.

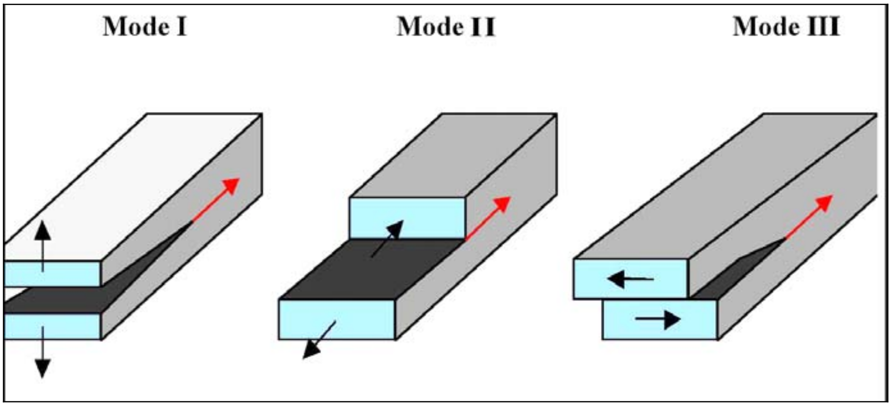


Figure 2. Different opening modes between two layers of a laminate.

If for example the case of the impact of a structure geometrically close to a helicopter energy absorption device configuration is to be considered [1], Fig. 3 shows the damage obtained after an axial impact onto the structure. This model type is fairly complicated to represent in a numerical model for the following reasons:

- The problems pertaining to behaviour changes at the fullscale manufacturing level in comparison with the characterization test-pieces.
- Strain rate effects are quite significant and thus generate important changes both in the behaviour and the deformation modes.

The fairly complex failure modes are still poorly represented in the methods of computational models such as Finite Element Method (FEM).

Figure 3. Energy absorption structure failure modes.

3. TOOL IMPROVEMENT CONTEXT

The experimental laboratory work started from the preliminary consideration of the modelling of GLARE (GLAss REinforced aluminium laminates) or ARALL (Aramid Reinforced ALuminium Laminates) type materials, used a few years ago for airplane fuselage design. More often than not, aeronautic structures are submitted to severe point loading which does not necessarily entail visible damage at the upper surface of the material but generates significant inner damage, thus jeopardizing both the component safety and lifespan. In order to partly address this problem, it is proposed to develop a multi-material multi-layer element [2] which enables the topic of impact modelling of laminates made up with quite different materials from the point of view of behaviour to be addressed. Numerous examples [3] have made it possible to validate this finite element integrated within the industrial PAMCRASH^{TM} software. Once this first step was completed, the main behaviours of a fibre and resin type composite was also improved and an in-depth basic model [4] was integrated within the computational code in order to reproduce behaviours close to reality. To that aim, the approach used was often identical in order to try and reproduce within a numerical model a phenomenon or set of phenomena pertaining to the damage previously presented (Fig. 1 & 2). This can be broken down as follows:

- Implementation of a very accurate experimental campaign allowing an in-depth, or at least as precise as possible, analysis of the phenomenon under study.
- Analysis of the existing bibliography possibly to be found in the "Benchmarking" field and in-depth study in order to try and define a

mathematical model or an innovative methodology permitting to tackle the phenomena under close scrutiny.

- Model or methodology integration within a computational code, industrially based whenever possible, thus permitting the testing of developments within discrete applications making final validation easier. The latter often being carried out in close correlation with an experimental campaign applied on a real structure.

4. PRESENTATION OF EXAMPLES OF PHENOMENA TAKEN INTO CONSIDERATION

The main objective does not consist of defining the models developed in different studies conducted within our research-lab, but in demonstrating the influence that either taking into account or ignore certain phenomena. Three examples are presented, thus allowing the effect of integration into a computational model, as applied to laminated composites (Fibre-matrix), to be defined.

4.1 Modelling resin ductability within a laminate

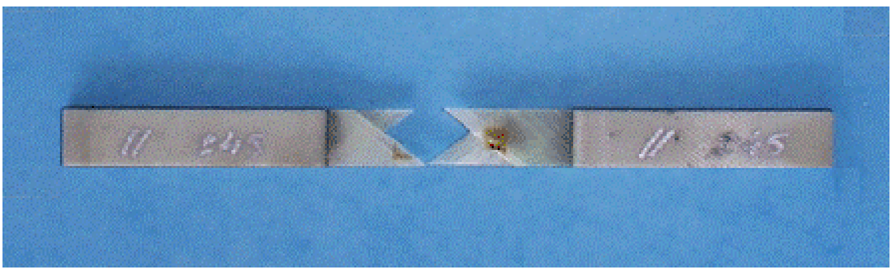


Figure 4. Tensile test on a $[45]_{2s}$ laminate.

Here, a numerical visualization of the likely influence matrix ductility may have over the overall behaviour of a laminated material [5] was proposed. A tensile test on a laminate $[45]_{2S}$, featuring a 60% E glass fibre rate and 40% epoxy resin rate (Fig.4), has been simulated. The experimental stress/strain curve was compared with a composite behaviour model featuring an elastic type law with damage and a plasticity integrating model (Fig.5). A 50% error rate on the shear deformation can be noticed if the inelastic phenomema are not taken into account. This study case is extreme since the stacking under study favours shear stress.

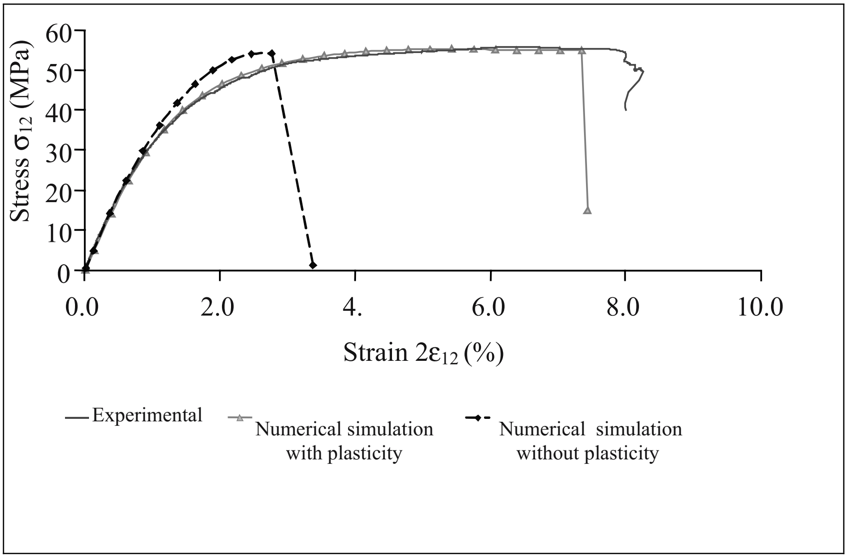


Figure 5. Tensile comparisons on $[+/-45]_{2s}$ laminate.

4.2 Study of strain rate influence over a laminate behaviour

4.2.1 Experimental approach

Tests [6] on the influence likely to be exerted by strain rate onto a composite in-built into a 60% E/Glass fibre volume rate and epoxy resin have been carried out: Tensile tests on $[0]_4$, $[\pm 45]_S$, $[\pm 67.5]_S$, $[+45]_4$ laminates, and finally a compression test onto a $[0]_6$ laminate. The experimental setup incorporated a hydraulic jack. For each type of test, four different loading velocities were mainly applied : 5 mm/min, 500 mm/min, 0.2 m/s and 4 m/s. The experimental tests brought forward two essential conclusions :

- As Fig. 6 & 7 clearly show, the elastic modulus, the initial plastic flow stress and finally the longitudinal fracture stress significantly increase from a certain strain rate threshhold. This increase ranges from 30% up to 300% depending upon the constants.
- In Fig. 6, a conservation of the behaviour type can be noticed, in the case of a tensile test on the 0° material, since it is still a brittle elastic type behaviour. However, in all the other cases, the behaviour tends to become "damageable elastic" under the influence of strain rate.

Consequently, the behaviour law may be defined as "viscous" damageable elasto-plastic, starting from a "reference" behaviour law.

The experimentally observed behaviour type was hard to identify: it was neither visco-elasticity nor visco-plasticity. The modelling approach hereby proposed was associated with the viscous fluids approach [6] : The "pseudo-viscoelasticity" of resin (and under the present circumstances of the composite) was taken into account by means of a viscous strain to be added to the computed one (Fig. 8). The clearly identified behaviour was considered at a reference velocity level.

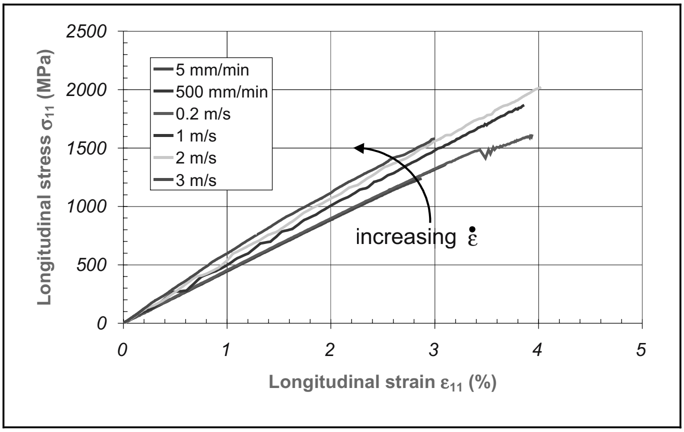


Figure 6. Dynamic tensile loading test on a $[0]_4$ laminate.

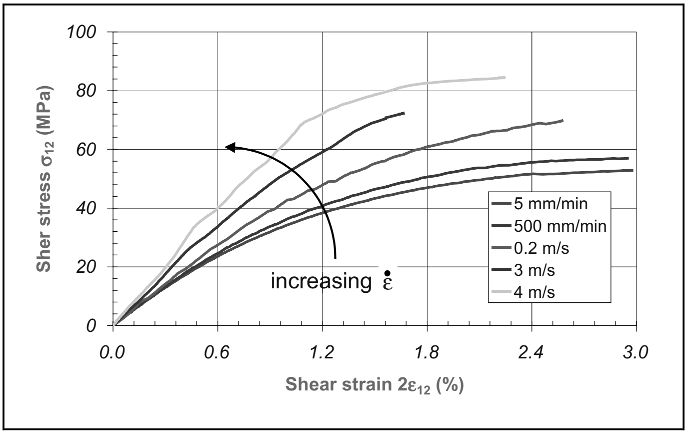


Figure 7. Dynamic tensile loading on a $[+/-45]_s$ laminate.

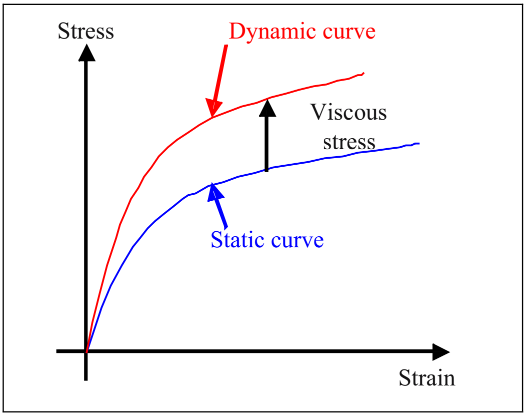


Figure 8. Viscous stress definition.

4.2.2 Numerical simulation

In order to make sure the theoretical “primary” behaviours were properly represented, elementary validations, corresponding to experimental tensile or compression tests and originating from a static and dynamic framework study, were investigated. Fig. 9 & 10 as compared with Fig. 6 & 7, visually demonstrate that the approach developed seemingly correlates satisfactorily with experiments.

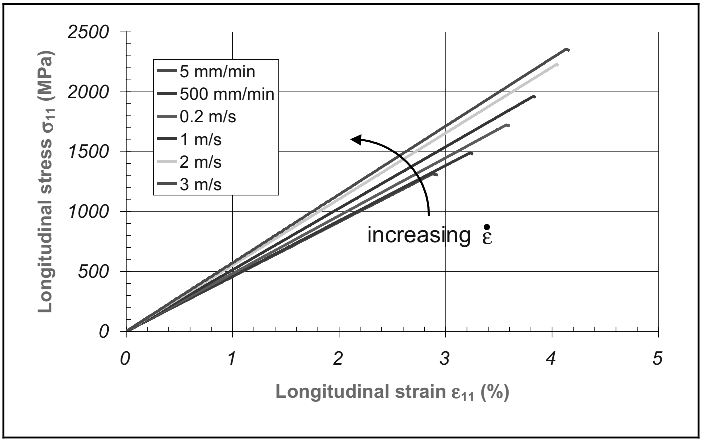


Figure 9. Numerial dynamic tensile loading on a $[0]_4$ laminate.

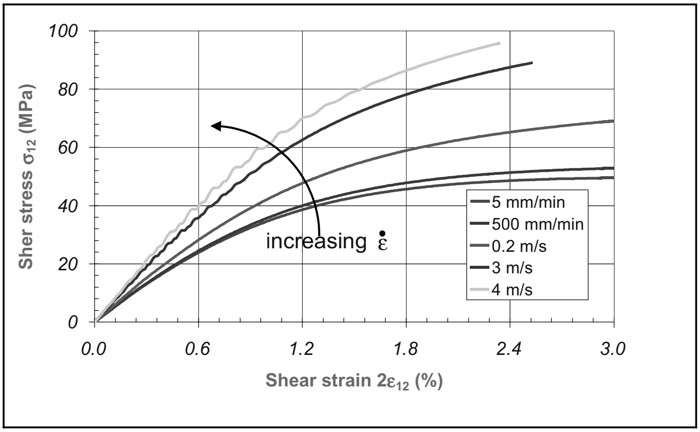


Figure 10. Numerical dynamic tensile loading on a $[+/-45]_s$ laminate.

4.3 Delamination phenomena into numerical models

Through a methodology [7] developed within the framework of experimental studies conducted in the laboratory, the delamination phenomena can be fully integrated into the finite element models and thus bring about improvements to the expected results. The presentation of this methodology in this publication is not the main objective but can be summarized as follows.

The first part concerns delamination detection within thin laminated structures. As a result of to the computational code used in rapid dynamics, the laminated materials were modelled by means of shell-type elements. The approach was quite ambitious although limited to modes II & III of Fig. 2. This methodology used post-processing criteria, based on fracture mechanics and related to damage mechanics by the effective stresses computed with the composite ply behaviour model. The second step allowed both the consideration and translation of delamination influence within the overall structure behaviour. This influence was introduced by a local change of the material mechanical characteristics. These integrated effects generated a change in the structure behaviour.

Tests have been conducted with a three point dynamic bending loading applied to laminated plates. Here a glass-epoxy $[90_2/0_4/90_2]$ plate is presented (Fig.11). The dynamic experimental results have been obtained using a velocity controlled jack. The damage mechanisms have been validated through observations made experimentally through the use of a high-speed video camera, and strain gauges and displacement sensors.

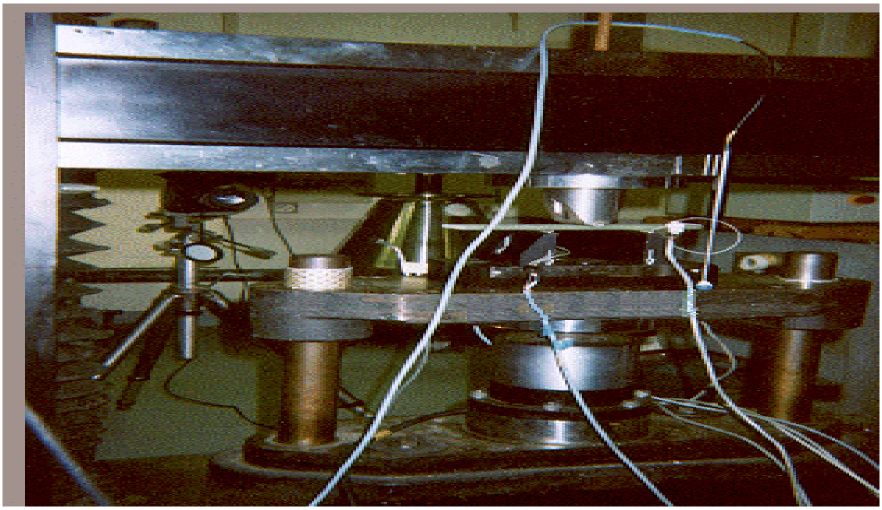


Figure 11. Bending test on a laminated plate

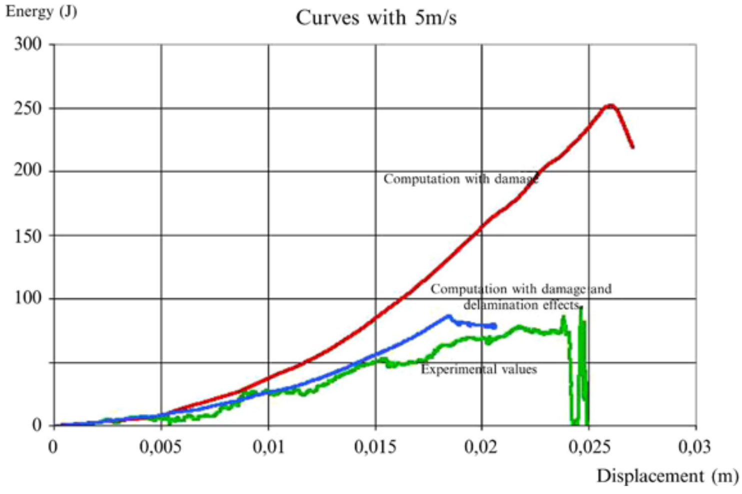


Figure 12. Comparison of the energy/displacement curves

In order to demonstrate the relevance of the developments, two different types of computations have been achieved: the first, referred to as “general”, does not take into account the post-processing of mechanisms detected by the methodology; the second, referred to as “sequential”, integrates the changes in the layer-intrinsic characteristics. The results presented here are given for a 5 m/s loading velocity. A complete set of numerical curves has been plotted versus the experiments. The gap between the overall computational approach and the experimental results was large . As regards the sequential computational method, it provided a considerable improvement in the assessment of the energy evolution (Fig.12). The sequential computation showed a significant contribution for the overall structure behaviour prediction at the end of the impact. Thus, it allows a relatively accurate result to be reached, particularly as regards the fundamental quantity of energy absorbed.

5. CONCLUSIONS AND FUTURE PROSPECTS

Within laminated structure design, the phenomena encountered during severe loading were far from being perfectly represented in the existing numerical models, even with the most advanced codes. Depending upon the designers' requirements and of their technical means, models specifically adapted to each particular situation have to be considered. The improvement of industrial performance results from the development of numerical tools allowing, for some given application cases, development of precise models which will then facilitate the adjustment of the design parameters aiming at structure optimization.

Within the field studied in this paper, the models developed provided valuable material. However, further studies still need to be conducted for an in-depth understanding of the domain of composite material damage.

ACKNOWLEDGMENTS

"The Région Nord-Pas-de-Calais", le GRRT, le FEDER and the CNRS for their financial support of this work is gratefully acknowledged. The ONERA institute and ESI Group have also to be thanked for their contribution.

REFERENCES

1. D. Dormegnie, Contribution à l'étude des lois de similitude applicables au crash de structures stratifiées du type absorbeur d'énergie, thèse de l'Université de Valenciennes et du Hainaut-Cambrésis, Valenciennes, 2001.
2. C. Gauthier, Contribution à la modélisation du comportement en crash des structures stratifiées métal/composite. Développement d'un élément de coque multi-couches multi-matériaux, thèse de l'Université de Valenciennes et du Hainaut-Cambrésis, 1996.

3. D. Coutellier, C. Gauthier, Y. Ravalard, X. Ni, E. Haug, Un élément multi-matériaux pour l'étude du comportement en crash des structures stratifiées : synthèse et perspectives, Revue Européenne des Eléments Finis, Vol-7 n°1/2/3, 1998, pp.177-192.
4. P. Ladeveze, Sur la mécanique de l'endommagement des composites, JNC 5, Ed. Pluralis, Paris, 1986, pp. 667-683
5. D. Coutellier, P. Rozycki, New functionalities for PAMCRASH™ multi-layered multi-material element, Composites Part A – Applied Science and Manufacturing, 2000, Vol-31, pp. 841-851.
6. P. Rozycki, Contribution au développement de lois de comportement pour matériaux composites soumis à l'impact, thèse de l'Université de Valenciennes et du Hainaut-Cambrésis. Valenciennes, 2000.
7. J.C. Walrick, Contribution au développement d'une nouvelle méthodologie pour l'étude du délaminage dans les structures stratifiées composites : application à l'impact basse vitesse, thèse de l'Université de Valenciennes et du Hainaut-Cambrésis, Valenciennes, 1999.

MICRO EDM DRILLING: ACCURACY STUDY

Duc-Truong Pham, Stefan Dimov, Samuel Bigot, Atanas Ivanov, and Krastimir Popov
Manufacturing Engineering Centre, School of Engineering, Cardiff University,CF24 0YF

Abstract: This paper examines the influence of various factors onto the final accuracy in EDM (Electro Discharge Machining) micro drilling. In particular, the main parameters affecting the size and position of the hole are discussed and techniques for minimising errors are proposed. The paper analyses the technological capabilities of different methods of setting up and dressing the drill (electrode) on the machine. The paper also evaluates the limitations of the EDM drilling process.

Key words: Micro EDM, Micro-EDM accuracy, Micro holes

	Notations
Δd	d variation
ΔD	D variation
Δg_d	g_d variation
Δg_e	g_e variation
Δg_{meas}	g_{meas} variation
ΔH	H variation
ΔX_{meas}	X_{meas} variation
ΔX_{pos}	X_{pos} variation
Δy_d	y_d variation
ΔY_{meas}	Y_{meas} variation
ΔY_{pos}	Y_{pos} variation
Δy_{unit}	y_{unit} variation
d	effective dressed diameter of electrode
D	electrode initial effective diameter of electrode

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 281-294.

	Notations
D_{guide}	diameter of the ceramic guide
D_{init}	initial diameter of the electrode
D_{init_min}	minimum initial diameter of the electrode
g_d	spark gap during dressing of the electrode
g_e	spark gap during erosion of the workpiece
g_{meas}	spark gap during measuring cycle
H	achieved hole diameter
L_{guide}	length of the guide
t_{meas}	time interval between each contact signal check
V_{meas}	speed of measuring cycle movement
X_H	X coordinate of the hole relative to the workpiece reference
X_{meas}	X coordinate of the measured point of contact between electrode and workpiece surface
X_{pos}	X coordinate of the hole relative to the machine zero point
y_d	target dressing position in Y axis
y_{d_init}	an initial target dressing position
Y_H	Y coordinate of the hole relative to the workpiece reference
Y_{meas}	Y coordinate of the measured point of contact between electrode and workpiece surface
Y_{pos}	Y coordinate of the hole relative to the machine zero point
y_{unit}	Y coordinate of the dressing point relative to the machine zero point
z_{guide}	length of electrode protruding from the ceramic guide

1. INTRODUCTION

Electrical Discharge Machining (EDM) is a non-contact machining process for conductive materials that has been applied for more than 40 years and has proved particularly useful in the tool making industry. Due to its high precision and the good surface quality it can ensure, EDM is potentially very suitable for micro-fabrication [1, 2]. The EDM process utilises the thermo-electric energy released between a workpiece and a highly charged electrode submerged in a dielectric fluid. A pulsed electrical discharge across the small gap (known as the “spark” gap) between the workpiece and the electrode removes material from the workpiece through melting and evaporation. Clearly, due to the contactless nature of EDM, there are only very small process forces at work. This, combined with the availability in recent years of advanced computer controlled spark generators that help improve machined surface roughness, promises to make EDM the preferred method for producing micro features.

This paper focuses on the use of EDM to create one type of micro feature – holes. The paper examines the influence of various factors on the final accuracy in EDM micro drilling. In particular, the main parameters affecting the size and position of the hole are discussed and techniques for minimising errors are proposed. The paper analyses the technological capabilities of different methods of setting up and dressing the drill (electrode) on the machine. It also evaluates the limitations of the EDM drilling process.

2. PROCESS OVERVIEW

The electrodes usually used for micro EDM drilling are tungsten (W) or tungsten –carbide (WC) rods or tubes 0.1mm to 0.2mm in diameter and 150 mm to 300mm in length. One end of the electrode is clamped in a high-speed spindle and the other end goes through a fixed ceramic guide positioned a few millimetres above the workpiece. By rotating the electrode improves the achievable aspect ratio and surface finish due to the fact that a rotating electrode helps to remove debris from the work zone and makes the final roughness less dependent on the initial electrode roughness. During the EDM drilling process, the electrode wears and this can be compensated for with *Z* movement while the ceramic guide remains at the same fixed position. When an electrode diameter smaller than the smallest rod available is needed, the electrode section protruding from the ceramic guide can be ground (dressed) on the machine using either a WC block [3], a spinning WC wheel or a special wire grinding device [4, 5].

3. ACCUMULATION OF ERRORS

3.1 Process Definition

This work investigates the accuracy achievable with the micro EDM drilling process. Drilling a small hole in a block of material by single pass machining and using a single dressed electrode will be adopted as an example.

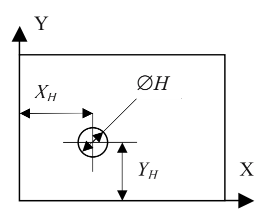


Figure 1. Dimensions of Interest

The main dimensions of interest are (Fig. 1):
H – Diameter of the hole
X_H and Y_H – position of the hole with respect to the co-ordinate system.

3.2 Factors affecting hole diameter

The achieved diameter *H* depends on the diameter of the effective dressed electrode *d* and the spark gap g_e (Fig. 2).

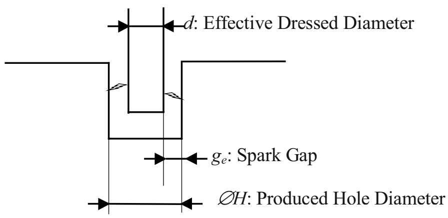


Figure 2. Achieved diameter H

$$H = 2 \cdot g_e + d \quad (1)$$

The deviation from the nominal *H* (ΔH) will depend on variations of the spark gap Δg_e and of the effective dressed electrode diameter Δd.

$$\Delta H = 2 * \Delta g_e + \Delta d \quad (2)$$

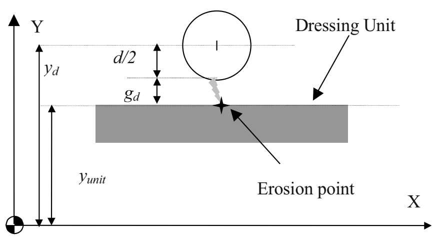


Figure 3. Dressing process

In order to reduce the initial effective diameter *D* down to a micro effective diameter *d*, an electrical-discharge-grinding unit is used as shown in Fig. 3. For this study, movement in the dressing process is performed along the Y axis. Distance y_{unit} gives the position of the eroding point in the work area of the machine relative to the machine reference point. The electrode is eroded until the centre of the spindle reaches a target position y_d

resulting in an effective dressed electrode of diameter *d*. Taking into account the spark gap g_d between the electrode and the dressing unit, the effective diameter obtained *d* is defined by equation 3.

$$d = 2 * (y_d - y_{unit} - g_d) \tag{3}$$

The variation in the effective dressed diameter *d* (*Δd*) will depend on the variation in the position of the grinding device Δy_{unit}, the variation in the positioning of the centre of the electrode Δy_d and the variation of the spark gap when grinding Δg_d.

$$\Delta d = 2 * (\Delta y_d + \Delta y_{unit} + \Delta g_d) \tag{4}$$

Finally the variation in the diameter of the hole drilled by a single dressed electrode will be determined by equation 6:

$$H = 2 \cdot (g_e + y_d - y_{unit} - g_d) \tag{5}$$

$$\Delta H = 2 \cdot (\Delta g_e + \Delta y_d + \Delta y_{unit} + \Delta g_d) \tag{6}$$

3.3 Factors affecting the position of the hole

The position of the hole is described by the following equations directly derived from Fig. 4.

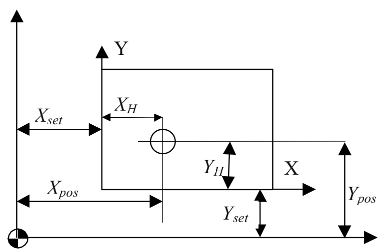


Figure 4. Hole position description

$$X_H = X_{pos} - X_{set} \tag{7}$$

$$Y_H = Y_{pos} - Y_{set} \tag{8}$$

In order to set up the workpiece position in the work area, an electrode of nominal effective diameter (*D*) is used. It should be noted that the use of probes or other setting devices is ruled out because it would require reattachment of the high speed spindle and readjusting the ceramic guide and is therefore likely to introduce more errors. The set-up process is represented in Fig. 5.

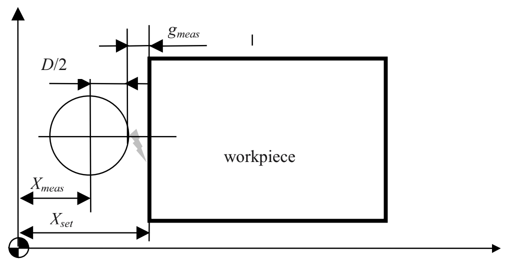


Figure 5. Set-up process

$$X_H = X_{pos} - (X_{meas} + D/2 + g_{meas}) \quad (9)$$

$$Y_H = Y_{pos} - (Y_{meas} + D/2 + g_{meas}) \quad (10)$$

The deviations are respectively:

$$\Delta X_H = \Delta X_{pos} + \Delta X_{meas} + \Delta D/2 + \Delta g_{meas} \quad (11)$$

$$\Delta Y_H = \Delta Y_{pos} + \Delta Y_{meas} + \Delta D/2 + \Delta g_{meas} \quad (12)$$

The accuracy of the hole position will depend on the accuracy of machine positioning (ΔX_{pos}, ΔY_{pos}), the detection contact signal accuracy with surface (ΔX_{meas}, ΔY_{meas}, Δg_{meas}) and the variation in the initial effective diameter of the electrode (ΔD).

4. FACTORS AFFECTING THE ACCURACY

4.1 Accuracy and repeatability of positioning

The accuracy and repeatability of positioning (Δyd and ΔXpos, ΔYpos) of the machine were measured according to BS ISO 230-2:1997 and the results are have been plotted in Table 1 for the three axes of the machine. The accuracy of machine movement was measured using a laser interferometer. The parameters affected by the accuracy of machine movement (Δy_d and ΔX_{pos}, ΔY_{pos}) are discussed in the following sections.

Table 1. Positioning results

Repeatability of positioning		Repeatability of positioning		Repeatability of positioning	
Unidirectional A↑	3.76μm	Unidirectional B↑	4.49μm	Unidirectional C↑	1.48μm
Unidirectional A↓	2.95μm	Unidirectional B↓	4.92μm	Unidirectional C↓	1.27μm
Bidirectional	5.33μm	Bidirectional	7.83μm	Bidirectional	1.9μm
Accuracy of positioning		Accuracy of positioning		Accuracy of positioning	
Unidirectional A↑	14.03μm	Unidirectional B↑	5.03μm	Unidirectional C↑	2.39μm
Unidirectional A↓	11.85μm	Unidirectional B↓	5.18μm	Unidirectional C↓	1.82μm
Bidirectional	15.7μm	Bidirectional	7.91μm	Bidirectional	2.73μm
X axis		Y axis		Z axis	

4.1.1 Dressing position (y_d)

y_d is the target dressing position to be reached by the centre of rotation of the electrode relative to the machine reference point during the dressing process. It is defined by the operator (or by a program) in order to obtain a specific effective dressed diameter *d* and therefore a specific hole diameter. Variation Δy_d will arise due to the machine accuracy and repeatability of positioning. An obvious way of reducing Δy_d during the dressing process is always to approach the position from the same direction (unidirectional approach). Another way of limiting the error is to identify an area on the machine and to fix the dressing unit wherever the repeatability of positioning is at its best. For instance, when focusing on one single spot in the machine where the dressing unit is placed, the calculated repeatability of positioning according BS ISO 230-2:1997 in this case is 2.2μm, which is much better as compared to values given in Table 1. The measured Δy_d is 1.98 μm.

4.1.2 Hole position (X_{pos} and Y_{pos})

X_{pos} and Y_{pos} are the coordinates of the hole in the machine coordinate system. The position of the hole depends on the workpiece and its position in the work area. Thus the only way to improve ΔX_{pos} and ΔY_{pos} is to adopt a unidirectional approach to the hole. To machine micro holes, multiple dressed electrodes might be required and therefore the machine positioning accuracy will mainly affect the hole position while positioning repeatability will impact on the hole size and shape.

4.2 Spark gaps and effective electrode diameter

4.2.1 Gap between electrode and workpiece (g_e)

g_e is defined as the spark gap between the electrode and the workpiece. Its nominal value is fixed by the chosen pulse of the generator and the dielectric used. In conventional EDM, the selection of a pulse is directly linked with the removal rate and surface roughness required. In micro EDM, electrode wear is another important criterion which also needs to be carefully considered. In addition, in order to achieve the micro features, the spark gap should be minimised. Variations in g_e (Δg_e) generate random errors which can appear due to flushing conditions and lack of surface/material integrity [5].

4.2.2 Gap between electrode and dressing unit (g_d)

g_d is defined as the spark gap between the electrode and the dressing unit.

As in the case of g_e, the value of g_d is fixed by the chosen pulse parameters and dielectric material, and variations in g_d (Δg_d) can arise due to flushing conditions and lack of surface/material integrity. The pulse parameters are selected depending on the surface roughness required on the electrode and on the dressing speed. Because the electrode is rotating, its surface roughness should not significantly affect the roughness of the machined surfaces. However, due to the small dimensions involved, a high roughness will affect the strength of the dressed electrodes, which could break during the process. Estimating Δg_d is difficult but it was assumed that it would not exceed Δg_e in the worst case. This is because, during dressing, sparking conditions are better than during drilling itself as dressing involves single point sparking with better flushing conditions.

4.2.3 Effective diameter (D)

The effective diameter of the electrode D is determined by the initial diameter D_{init} of the electrode (WC rod) and the assembly conditions between the electrode and the ceramic guide. The difference between the diameter D_{init} and the diameter of the ceramic guide D_{guide} generates a gap that introduces potential errors as shown in Fig. 6.

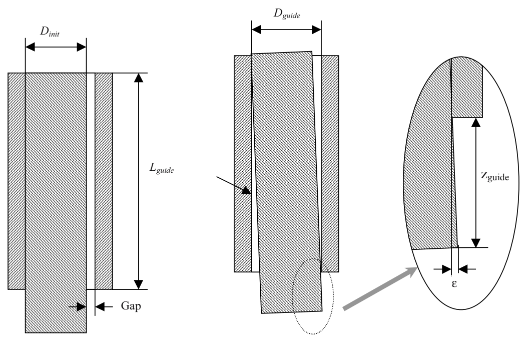


Figure 6. Effect of gap between guide and electrode

Thus, variations in the effective diameter (*ΔD*) can occur, which reflect the tolerance of the electrode and the assembly conditions between the electrode and the ceramic guide. Based on the parameters shown in Fig. 7, the maximum variation in effective electrode diameter is defined by equation 15:

$$\frac{\left(D_{guide} - D_{init\_\min}\right)}{L_{guide}} = \frac{\varepsilon}{z_{guide}} \tag{13}$$

and $\Delta D = D_{guide} + 2 \cdot \varepsilon - D_{init\_\min}$ (14)

$$\text{Thus, } \Delta D = D_{guide} + 2 \cdot \frac{\left(D_{guide} - D_{init\_\min}\right) \cdot z_{guide}}{L_{guide}} - D_{init\_\min} \tag{15}$$

In the equations above, D_{guide} is the guide diameter, D_{init_min} is the initial electrode minimum diameter according to the manufactured tolerance and z_{guide} is the length of the electrode protruding from the ceramic guide. In the case of experiments discussed in this paper, the diameter of the electrode was D_{init}= 0.146±0.002mm, thus D_{init_min}= 0.144mm, and the measured diameter of the ceramic guide was D_{guide}=0.154mm. L_{guide} was 12mm, and z_{guide} was within 2mm. Based on those values, the calculated maximum deviation *ΔD* was 13.3µm. However, this maximum variation only occurs when the electrode position within the guide is modified to a number of extreme positions. This is only possible when there is a significant movement of the electrode along the X and Y axes relative to the guide. This is highly unlikely. To support this point, two cases are considered. First, when using a non-rotating electrode, the position between guide and electrode should not change because the only movement between guide and electrode is in the Z axis and no force acts along the X or Y axis that would change this position. Significant change should only occur when altering the position of the guide with respect to the head holding the electrode, for instance, when replacing an electrode or after a significant movement in the Z direction because this would affect the angle at which the electrode enters the guide and therefore might create a significant movement along the X and Y axes. Also, changes could arise due to the tolerance of the manufactured electrode, but the diameter variation along the length of a single manufactured rod is considered negligible. Therefore, it can be assumed that, for a small feed and without change of electrode, the variations in effective diameter when using a non-rotating electrode are negligible, $\Delta D \approx 0$. Secondly,, when the electrode rotates, X-Y movement can occur due to the rotation and friction between guide and electrode. However, it can be assumed that this movement follows a cycle in phase with the rotation, resulting in a small increase in effective diameter but with negligible variations between two rotation cycle periods, therefore $\Delta D \approx 0$. It should be noted that the length of the dressed section of the electrode should be smaller than z_{guide}. This is to avoid the dressed part of the electrode touching the guide, as this would increase the potential error and the dressed electrode could be damaged. However, this introduces another limitation in the depth achievable using a dressed electrode.

4.2.4 Estimation of Δg_e and Δg_e

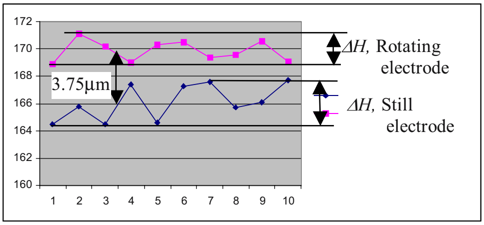


Figure 7. Hole produced with rotating and stationary electrodes

An experiment was conducted to estimate Δg_e and ΔD. A ∅150µm WC electrode was used to drill two series of 50µm deep holes. The experiment was conducted with rotating and stationary electrodes and the results are given in Fig. 7. The erosion process of a hole is represented by the following equations:

$$H = 2 \cdot g_e + D \quad (16)$$

$$\Delta H = 2 \cdot \Delta g_e + \Delta D \quad (17)$$

All holes were machined with the same electrode/guide assembly to a depth of 50µm, resulting in a total Z feed of only 1mm for the total process. Thus, according to the previous section, it can be assumed that for both series of holes $\Delta D \approx 0$. Therefore, $\Delta H \approx 2 \cdot \Delta g_e \Leftrightarrow \Delta g_e \approx \Delta H/2$. The difference between the mean diameter of the two series is 3.75µm, which represents the mean increase of effective diameter due to the rotation. From the experiment, it can be concluded that $\Delta g_e = 1.1$ µm when the electrode is rotating and $\Delta g_e = 1.6$µm when the electrode is not rotating. As seen in Fig. 8, when the electrode is rotating Δg_e is smaller, which might be due to the better flushing conditions and a better way of removing debris from the cutting zone.

4.3 Temperature instability error (Δy_{unit})

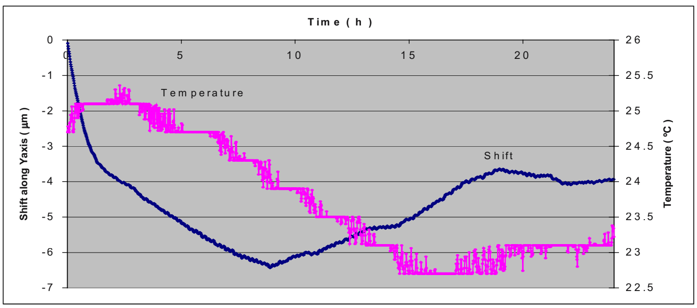


Figure 8. Shift of the relative position between the rotating head and the table along the Y axis due to temperature variation

y_{unit} is the position of the erosion point on the dressing unit in the machine co-ordinate system. Changes in the temperature in the room and in the machine structure create variations in the relative position between the rotating head and the machine table and therefore affect the position of the dressing unit with respect to the machine zero point, y_{unit}, and to the electrode. An example of that error is shown in Fig. 8. Displacement measurements along the Y axis were taken using a laser interferometer during the machine warming-up cycle. The obvious way to minimise the variation is to work in a temperature-controlled room and to ensure thermal-stability of the machine structure. Each machine should be tested to find out the time for the temperature of the machine to stabilise for certain ambient conditions and the temperature relative deviation of each axis should be measured in order to plan electrode dressing with minimum error. Under certain conditions (temperature-controlled room and minimum dielectric temperature variance, short working hours) it can be assumed that $\Delta y_{unit} \approx 0\mu m$.

4.4 Measurements errors

4.4.1 Workpiece surface detection error (ΔX_{meas}, ΔY_{meas})

During the setting up of the workpiece, when an electrical contact occurs between the electrode and the workpiece, a contact signal is registered by the machine system processor. The processor has set priorities in checking each machine status signal, which means that the checking of the contact signal is not carried out continuously. There is a time interval (t_{meas}) between each signal check. This causes an error when detecting the position of the measured surface when measuring X_{meas} and Y_{meas}. If the speed of approaching the surface is V_{meas}, the variation will be:

$$\Delta X_{meas} \text{ or } \Delta Y_{meas} = V_{meas} * t_{meas} \tag{18}$$

Usually the contact signal is checked every 2- 5msec (depending on the controller). Obviously, to minimise the error the speed should be as low as possible but high enough to avoid stick-slip. In this case, t_{meas} is 3msec and the measuring speed is from 20mm/min to 1mm/min. The calculated variation is 0.05 to 1µm.

4.4.2 Surface position detection error (Δg_{meas}, ΔD)

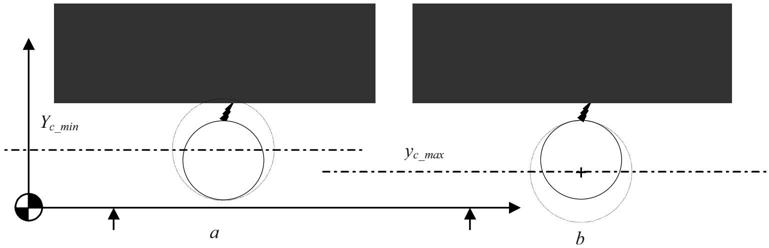


Figure 9. Surface position detection with ΔD influence

During the measuring cycle, voltage is applied between the table and the spindle. The machine moves until an electrical contact is reached. As the surfaces tend to oxidize, a different gap, or different pressure is needed for the spark to break through. All these factors contribute to a surface detection error introducing variation in the measuring spark gap Δg_{meas}. In order to minimise the effect of surface detection error, a measuring probe can be used. The variation using such a probe on a WC block approaching the surface with a very low speed of 1mm/min was measured as Δg_{meas} = 1µm. As explained earlier in this paper, it will be difficult to remove the high-speed spindle and the ceramic guide to perform the measurement with the probe and then replace them on the machine, because more errors will be introduced. The electrode itself can be used to do the measurement instead of the probe. In this case ΔD will be included in the measurement as well. Unlike the erosion or dressing process, ΔD is not negligible. This is because as mentioned in section 4.2.4, the rotation of the electrode generates a cyclic X/ Y movement . Therefore, contact between electrode and the surface might occur at different positions in the cycle, as shown in Fig. 9. The measurement accuracy will depend on the measurement speed. The lower the speed of the approach in relation with the electrode rotational speed, the smaller the error . This is comfirmed by an experiment, where the variations of surface detection on a WC block with ∅150 µm WC electrode using different speeds were measured. For speeds of 20mm/min, 5mm/min and 1mm/min (the lowest speed on the machine), (Δg_{meas}+$\Delta D/2$) is respectively equal to 5.7µm, 3.9µm and 3µm.

5. EXPERIMENTAL SET UP

Table 2. Experimental ΔH

	y_{d_init}	y_{d_init} -10µm	y_{d_init} -20µm	y_{d_init} -30µm	y_{d_init} --40µm
ΔH exp 1	19.4µm	3.6µm	6.0µm	13.9µm	11.9µm
ΔH exp 2	12.2µm	12.2µm	9.4µm	13.9µm	8.6µm

The experiment consisted in drilling two series of 10 holes, for 5 different dressing positions. The first target dressing position was y_{d_init} and the 4 others were respectively (y_{d_init} -10µm), (y_{d_init} -20µm), (y_{d_init} -30µm), (y_{d_init} --40µm). The ΔH values measured have been plotted in Table 2. A 10µm reduction in target dressing position y_d should result in a 20µm diameter reduction for the hole drilled. In the experiment, the mean differences between each series of diameters were 20.49, 19.55, 21.31, 19.08, which gives a 2.23 µm variation. This shows the potential for an accurate dressing process using EDM grinding.

6. CONCLUSIONS

This paper has given an overview of the factors affecting the accuracy during micro EDM drilling and the results of the conducted experiments can help to plan the process within the expected tolerances. The following conclusions may be drawn. The machine used for micro EDM drilling should be placed in a temperature-controlled environment with a constant ambient and dielectric temperature. If a grinding device is used, the type of device should be justified and its position should be selected after careful investigation of the geometrical accuracy of the machine. Tests should be made in advance for the preferred sparking conditions and the spark gap deviation should be measured. Speed is the main factor contributing to errors when using measuring cycles. When assigning process tolerances for micro EDM drilling all activities during the process, such as electrode grinding type, positioning type and operation duration, should be considered. All these activities will accumulate errors, which should be taken into account. This paper has discussed the level of accuracy achievable with micro EDM drilling which has been studied as the first step into investigating micro EDM milling strategies.

ACKNOWLEDGMENTS

This work was carried out as part of the ERDF project "Supporting Innovative Product Engineering and Responsive Manufacturing" (SUPERMAN), the ERDF project 'Micro Tooling Centre', the EC Network of Excellence '4M' and the EC Network of Exellence 'I*PROMS'.

REFERENCES:

1. K.H. Ho, S.T. Newman, State of the art electrical discharge machining (EDM). International Journal of Machine Tools&Manufacture, 43, 2003, pp. 1287-1300.
2. W. Ehrfeld, H. Lehr, F. Michel, A. Wolf, H.P. Gruber, A. Bertholds, Micro Electro Discharge Machining as a Technology in Micromachining. SPIE, 2879, 1996.
3. N. Ravi, S.X. Chuan, The effects of electro-discharge machining block electrode method for microelectrode machining. Journal of Micromechanics and Microengineering, 2002, pp. 532-540.
4. T. Masuzawa, M. Fujino, K. Kobayashi, T. Suzuki, Wire electro-discharge grinding for micro machining. Annals of CIRP 1985, 34(1).
5. H.S. Lim, Y.S. Wong, M. Rehman, M.K. Edwin Lee, A study on the machining of high-aspect ratio micro-structures using micro-EDM. Journals of Materials Processing Technology, 140, 2003, pp. 318-325.
6. K.P. Radjurkar, W.M. Wang, Improvement of EDM Performance with advanced monitoring and control system. Transactions of the ASME, 119, 1997, pp.770-775.

ADAPTATIVE CONTROL OF THE ERROR IN THE VON MISES STRESS IN 2D ELASTICITY

Eric Florentin, Laurent Gallimard and Jean-Pierre Pelle
LMT-Cachan ; ENS de CACHAN/CNRS/Université P. et M. Curie ; 61, avenue du Président Wilson, 94235 Cachan Cedex ; FRANCE

Abstract: In current industrial situations, it is necessary to have reliable evaluations of local quantities such as Von Mises stress. These quantities are evaluated using F.E. code. Even if the mechanical model chosen is adequate, the mesh used in F.E. analysis introduces errors on the quantities being computed. For the engineer, it is essential to study and, if possible, to improve the quality of the computations carried out. In this work, we focus on the quality of a 2D elastic finite element analysis. Our objective is to control the discretization parameters in order to achieve a prescribed local quality level over a dimensioning zone. The method is illustrated through 2D test examples.

Key words: F.E., Local error, Quality, Dimensioning, Adaptation

1. INTRODUCTION

A major concern in the mechanical field has always been to control the quality of FEM computations. However, the information thus obtained (evaluation of the overall error, contribution of each element to this error) permits to develop procedures to adapt the calculation parameters in order to achieve the overall level of accuracy aimed at while reducing calculation costs. Naturally, such procedures have been developed mainly for linear analysis and, in this context, they have led to very robust mesh adaptation techniques [1-2]. Parameter adaptation techniques for nonlinear and dynamic problems can be found in [3-4]. In recent years, estimators allowing the control of the quality of local quantities (stresses, displacements, contour integrals) have been proposed by different teams [5-11].

Naturally, parameter adaptation techniques have been associated with these estimators [5, 7, 12]. The objective of this paper is to propose a mesh

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 295-304.

adaptation technique for linear problems permitting to guarantee the quality of the stresses calculated in a dimensioning zone while reducing the cost of the finite element analyses. This technique is based on the local estimator developed at LMT-Cachan [9, 13-14]. In the first part of this paper, after having reviewed the standard formulation of a linear elasticity problem, we outline the basic principles of the error in constitutive relation and its application to the estimation of the local quality of the calculated finite element stresses. In the second part, we review the mesh adaptation technique permitting to control the overall quality of a finite element analysis. In the third part we present the modifications that have been made to this technique in order to guarantee a given quality level in a zone specified by the user (*a priori* a dimensioning zone) while reducing the computational cost of the finite element analysis. First examples of applications to two-dimensional problems are presented. They clearly show the interest and the effectiveness of the method proposed.

2. ERROR IN CONSTITUTIVE EQUATION

2.1 Error in constitutive relation

Considering an elastic structure occupying a domain Ω subjected to :

- a prescribed displacement,
- a prescribed volume force density,
- a prescribed surface force density,

Hooke's material operator is denoted **K**. Then, the problem can be formulated as follows:

Find a displacement field $\underline{U}_h$ and a stress field σ_h defined over Ω verifying:

- the kinematic constraints
- the equilibrium equation with external forces
- the constitutive relation

2.2 Discretization error

This problem is solved approximately using the finite element displacement method. We want to know the quality of the couple $(\underline{U}_h, \sigma_h)$ as an approximation of the solution $(\underline{U}_{ex}, \sigma_{ex})$ of the corresponding continuous problem. Classically, a measure of the solution error over the structure is defined as:

$$e_h = \| \underline{e}_h \|_{U,\Omega} = \| \underline{U}_{ex} - \underline{U}_h \|_{U,\Omega} = \| \sigma_{ex} - \sigma_h \|_{\sigma,\Omega} \qquad (1)$$

where $\| \cdot \|$ represents the energy norm of $\cdot$ over Ω.

e_h provides only some overall scalar energy information on the quality of the finite element calculation. It could be interesting to decompose e_h taking into account the finite element partition introduced, and to break down the error into contributions from each element *E* of the mesh τ_h :

$$e_h^2 = \sum_{E \in \tau_h} e_{h,E}^2 \tag{2}$$

with : $e_{h,E} = \| \underline{e}_h \|_{U,E} = \| \underline{U}_{ex} - \underline{U}_h \|_{U,E} = \| \underline{\sigma}_{ex} - \underline{\sigma}_h \|_{\sigma,E}$

Remark: $e_{h,E}$ is a local error measure over element E: $e_{h,E} = 0 \Leftrightarrow \sigma_h = \sigma_{ex}$ in E

2.3 Error in constitutive relation

The discretization error is estimated through the error in constitutive relation. The concept of error in constitutive relation relies on splitting the equations into two groups:

* the admissibility equations: constraints and equilibrium,
* the constitutive relation.

The constitutive relation has a particular status. In practice, this is often the least reliable equation. An admissible pair $(\underline{U}_{KA}, \sigma_{SA})$ verifying the first group of equations is built up. Then, the non-verification of the constitutive relations permits to define e_{CR}. This estimate boils down to a sum of elementary contributions:

$$e_{CR}^2 = \sum_{E \in \tau_h} e_{CR,E}^2 = \sum_{E \in \tau_h} \| \sigma_{SA} - \boldsymbol{K}\varepsilon(\underline{U}_{KA}) \|_{\sigma,E}^2 \tag{3}$$

2.4 Construction of an admissible couple

Since the pair $(\underline{U}_h, \sigma_h)$ is not *admissible*, in order to use an error in constitutive relation it is necessary to construct an *admissible* pair $(\underline{U}_{KA}, \sigma_{SA})$ starting from the finite element solution and the problem data.

- In the framework of a displacement finite element method of the conforming type, the displacement field $\underline{U}_h$ is admissible. For the sake of simplicity, $\underline{U}_{KA} = \underline{U}_h$ in Ω is chosen
- However, the stress field σ_h is not admissible (i.e. it does not verify the equilibrium exactly). The techniques used for the construction of an

equilibrated stress field σ_{SA} starting from $\hat{\sigma}_h$ and the problem data are detailed in [1, 13, 14].

2.5 Solution error overall upper bound

With this choice of admissible fields, the error in constitutive relation can be connected to the solution error through Prager-Synge's hypercircle theorem [15]:

$$e_{CR}^2 = \| \sigma_{SA} - \sigma_h \|_{\sigma,E}^2 + \| \sigma_{ex} - \sigma_h \|_{\sigma,E}^2 \tag{4}$$

In particular, this theorem yields the following inequalities:

$$e_h \leq e_{CR} \quad \text{and} \quad \| \sigma_{SA} - \sigma_h \|_{\sigma,E} \leq e_{CR} \tag{5}$$

2.6 Local quality of the finite element solution

In [14, 16], quantity $e_{CR,E}$ is shown to be an upper bound of the error actually made $e_{h,E}$. The idea is that if a quality field σ_{SA} better than σ_h is built then, there exists a constant C such that:

$$\| \sigma_h - \sigma_{ex} \|_{\sigma,E} \leq C\, e_{CR,E} \tag{6}$$

The construction of σ_h, which in practice leads to (6) with C close to 1, is detailed in [13, 14, 16].

2.7 Local quality of Von-Mises equivalent F.E. stress

Using orthogonality property between deviatoric part and spheric part of the stress, equation (6) leads to an upper bound of the error made on Von Mises' equivalent stress $J(\sigma)$ [17] :

$$\left| J(\sigma_{ex}) - J(\sigma_h) \right| \leq K_{VM}\, e_{CR,E} \tag{7}$$

where K_{VM} is a constant depending on material properties.

3. ADAPTED MESH TO CONTROL THE FINITE ELEMENT STRESSES QUALITY

In this section, methods are presented that permit to choose the discretization parameters in order to achieve a given quality level. This quality can be overall (but, in this case, the practical benefits are limited) or local, in which case the discretization obtained leads to a result of direct value to the user. It is to notice that the prescribed quality can be given by an absolute error (energy norm or Von-Mises equivalent stress using K_{VM}) or a relative error, which is of a more pratical use.

3.1 Adapted mesh for controlling the global quality

The idea consists in utilizing the results of an initial finite element analysis and the associated error estimates to determine an optimum mesh, i.e. a mesh which provides the desired overall accuracy while minimizing the computational costs.

In short, the procedure principle is the following:

* An initial calculation is performed on a relatively coarse mesh $\boldsymbol{\tau}$;
* On this mesh $\boldsymbol{\tau}$, the relative overall error ε is calculated as well as the contributions ε_E;
* This information is then used to determine the characteristics of the optimum mesh $\boldsymbol{\tau}^*$.

3.1.1 The optimum mesh

The optimum mesh $\boldsymbol{\tau}^*$ whose concept was introduced in [18, 19], is such that :

- $\varepsilon = \varepsilon_d$ (prescribed accuracy)
- ε_E^* uniform over $\boldsymbol{\tau}^*$.

This definition is equivalent to having an evenly distributed error. This criterion does not necessarily correspond to minimum computational costs. A criterion without this drawback, introduced in [20], consists in taking the following definition:

- $\varepsilon = \varepsilon_d$ (prescribed accuracy)
- minimize the number N^* of element for $\boldsymbol{\tau}^*$.

This is the criterion chose, since it naturally leads to minimum computational costs.

3.1.2 Determination of the optimum mesh

The idea, in order to determine the characteristics of the optimum mesh

$\boldsymbol{\tau}^*$ is to calculate for each element E of the mesh $\boldsymbol{\tau}$ a size modification coefficient $r_E = h_E^* / h_E$.

h_E^* is the size of element E in $\boldsymbol{\tau}^*$ (unknown) and h_E is the size of E in $\boldsymbol{\tau}$. The calculation of the coefficients r_E is based on the convergence rate of the error $\varepsilon = O(h^q)$, constant q depends on the type of element being used, but also on the regularity of the exact solution of the problem being considered. One can show [21] that the resolution of such a problem yields tsize map r_E :

$$r_E = \left(\varepsilon_d^2\right)^{1/q} \varepsilon_E^{-2(q+1)^{-1}} \left(\sum_E \varepsilon_E^{2(q+1)^{-1}}\right)^{-1/2q} \quad \text{for E in } \Omega \tag{8}$$

3.2 Adapted mesh to control the local quality

The mesh obtained by adaptation guarantees a given overall error over the structure. This allows to obtain a better-quality finite element analysis by setting a lower target error. Here, an approach is proposed, which still aims at achieving an overall error level, but, in addition, also guarantees a given quality level in a given zone ω. Thus, the user can obtain an optimized mesh which provides, at a minimum cost, local information with a given accuracy. This is achieved by first setting the element sizes in zone ω in order to achieve the desired local quality, then minimizing the number of elements in the whole structure.

The principle of the procedure is the following:

- An initial calculation is performed on a relatively coarse mesh $\boldsymbol{\tau}$;
- on this mesh $\boldsymbol{\tau}$, the relative overall error e is then calculated as well as the relative local error $e_{loc}(\omega)$ over zone ω;
- this information is used to determine the characteristics of the local-optimum mesh $\boldsymbol{\tau}_{loc}^*$.

3.2.1 Local-optimum mesh

A local-optimum mesh $\boldsymbol{\tau}_{loc}^*$ is such that:

- $\varepsilon = \varepsilon_d$ (prescribed accuracy)
- $e_{loc}^*(\omega) = e_{loc,d}$ (prescribed local accuracy)
- N^* number of elements of $\boldsymbol{\tau}_{loc}^*$ is minimum.
- $e_{loc}^*(\omega)$ is uniform over ω

This criterion naturally leads to minimum cost while guaranteeing good-quality information for all elements in zone ω. In a mesh thus optimized, each element within zone w has a local error equal to $e_{loc,d}$.

3.2.2 Determination of the local-optimum mesh

Thus, the determination of the local-optimum mesh τ^*_{loc} is equivalent to the determination over the initial mesh of a size modification map with size r_E. In fact, this is a simple modification of the overall case. Indeed, the optimization is performed in two steps:

- the sizes are set up in ω, using the convergence ratio

$$r_E = \left({e_{loc,d}}/{e_{loc(E)}} \right)^{1/q} \text{ for E in } \omega \tag{9}$$

- one minimizes the number of elements over Ω-ω, using the same methodology as for (8):

$$r_E = \left(\varepsilon_d^2 - \sum_{E \in \omega} r_E^{2q} \varepsilon_E^2 \right)^{1/2q} \varepsilon_E^{2(q+1)^{-1}} \left(\sum_{\text{E}} \varepsilon_E^{2(q+1)^{-1}} \right)^{-1/2q} \text{ for E in } \Omega\text{-}\omega \tag{10}$$

3.3 Illustration of the method to a simple test case

This example illustrates the method on a beam subjected to bending. The part is dimensioned using the yield stress in the most highly solicited zone. An optimum mesh is being sought after in order to obtain good-quality stresses in the zone where the mechanical loading is maximum. Both the geometry and the loading are given in Fig. 1.

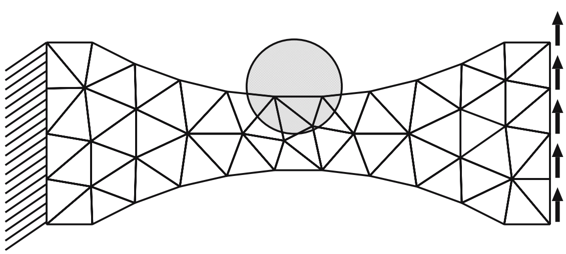


Figure 1. The initial mesh

A coarse and regular initial mesh is constructed in order to get an idea of the most highly solicited zones at a low cost. The zone with the highest stresses, as expected, is the central zone. We are more particularly interested in the compression zone (see Fig. 1). For this initial mesh, the local (for the zone of interest) and overall quality levels are given in Table 1.

The first step consists in seeking the mesh which yields 2 % global error with the minimum number of elements. The corresponding mesh is shown on Fig. 2.

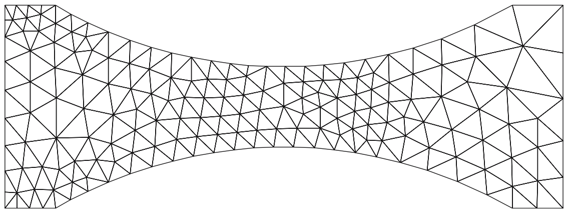


Figure 2. The global-optimum mesh

The local and overall quality levels obtained via this mesh are given on *Table 1*. Now, a mesh is considered with the same local quality level over the zone of interest, i.e. 1%.A mesh is sought, guaranteeing this local quality level for all elements in the zone and, at the same time, yields 5% global error while, of course, minimizing the number of elements used. The corresponding mesh is shown in Fig. 3.

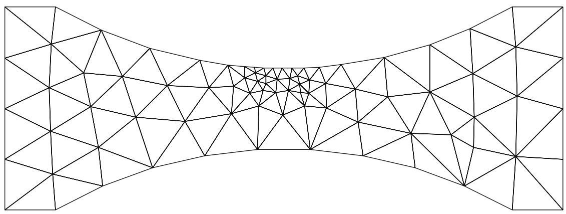


Figure 3. The Local-Optimum Mesh

Table 1. The corresponding quality levels

	Initial Mesh	Global-optimum Mesh	Local-optimum Mesh
Global error	6.60%	1.58% (2% prescribed)	4.61% (5% prescribed)
No. of elements	60	299	130
No. of nodes	151	666	299
Local error	3.99%	0.90% (for information)	0.80% (1% prescribed)

4. CONCLUSIONS

The test cases which were carried out show several things. First of all, a new tool has been presented that guarantees a local quality level prescribed by the user on the stress (Von Mises equivalent stress). This tool is directly usable in the design stage. Then, in terms of cost, this technique provides an additional benefit as compared to previous methods. Indeed, remeshing methods based on a global objective, by creating an adapted mesh, result in considerable time savings ; remeshing methods based on a local objective, as we have seen, result in a saving in terms of the number of elements. This time saving is significant, particularly in the design phase, where it allows several constructive solutions to be tested. The saving would be even more significant if this mesh adaptation technique were applied in 3D. Indeed, in 3D problems, the computational cost is vital and, in general, it sets the limits.

REFERENCES

1. P. Ladevèze, J.P. Pelle, and P. Rougeot. Error estimation and mesh optimization for classical finite elements. Engrg. Computation, 8:69–80, 1991.
2. P. Coorevits, J.P. Dumeau, and J.P. Pelle. Analyses éléments finis adaptatives pour les structures tridimentionnelles en élasticité. Revue européenne des éléments finis, 5 (n°3):341–374, 1996.
3. In P. Ladevèze and J.T. Oden, editors, Advances in adaptive computational methods in mechanics. Studies in Applied Mechanics 47, Elsevier, 1998.
4. L. Gallimard and J.P. Pelle. Adaptive analysis in elastoplasticity based on an enhanced constitutive relation error. Engineering Computations, 18:974–995, 2002.
5. R. Rannacher and F.T. Suttmeier. A feedback approach to error control in finite element methods: Application to linear elasticity. Comp. Mech., 19:434–446, 1997.
6. J. Peraire and A. Patera. Bounds for linear-functional outputs of coercive partial differential equations: local indicators and adaptive refinement. In P. Ladevèze and J.T. Oden, editors, Advances in Adaptive Computational Methods, pages 199–216. Elsevier, 1998.
7. S. Prudhomme and J.T. Oden. On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors. Comp. Meth. in Applied Mech. and Engrg., 176:313–331, 1999.
8. Cirak F. and Ramm E. A posteriori error estimation and adaptivity for linear elasticity using the reciprocal theorem. Comp. Meth. in Applied Mech. and Engrg., 156:351–362, 1998.
9. P. Ladevèze, P. Rougeot, P. Blanchard, and J.P. Moreau. Local error estimators for finite element analysis. Comp. Meth. in Applied Mech. and Engrg., 176:231–246, 1999.
10. T. Strouboulis, I. Babuska, D.K. Datta, K. Copps, and S.K. Gangaraj. A posteriori estimation and adaptative control of the error in the quantity of interest. part 1: A posteriori estimation of the error in the von mises stress and the stress intensity factor. Comp. Meth. in Applied Mech. and Engrg., 180:261–274, 2000.
11. Stein E. Ohmnimus S. and Walhorn E. Local error estimates of fem for displacements and stresses in linear elasticity by solving local neumann problems. Int. J. for Num. Meth. in Engrg., 52:727–746, 2001.
12. S. Ohnimus, E. Stein, and E.Walhorn. Local error estimates of fem for displacements and stresses in linear elasticity by solving local neumann problems. Int. J. Numerical Methods Eng., 52(7):727–746, 2001.
13. P. Ladevèze and P. Rougeot. New advances on a posteriori error on constitutive relationin f.e. analysis. Comp. Meth. in Applied Mech. and Engrg., 150:239–249, 1997.

14. E. Florentin, L. Gallimard, and J.P. Pelle. Evaluation of the local quality of stresses in 3D finite element analysis. Comp. Meth. in Applied Mech. and Engrg., 191:4441–4457, 2002.
15. W. Prager and J.L. Synge. Approximation in elasticity based on the concept of functions space. Quart. Appl. Math., 5:261–269, 1947.
16. E. Florentin, L. Gallimard, P. Ladevèze, and J.P. Pelle. Local error estimator for stressesin 3D structural analysis. Computers and Structures, 81:1751–1757, 2003.
17. E. Florentin, L. Gallimard, P. Ladevèze, and J.P. Pelle. A tool for verifyiong the local quality of a 3D F.E. analysis in design. In International Conference on Integrated Design and Manufacturing in Mechanical Engineering'02, 2002.
18. P. Ladevèze. Nouvelles procédures d'estimations d'erreur relative à la méthode des éléments finis. In proc. Journées Eléments finis, Rennes, 1977.
19. P. Ladevèze and D. Leguillon. Error computation in finite element method and application. Proc. GAMNI, Dunod, 1981.
20. P. Ladevèze, G. Coffignal, and J.P. Pelle. Accuracy of elastoplastic and dynamic analysis. In I. Babuska, J. Gago, E.R. de A. Oliveira, and O.C. Zienkiewicz, editors, Accuracy estimates and adaptive refinements in Finite Element computations, pages 181–203. J.Wiley, 1986.
21. P. Ladevèze and J.P. Pelle. La maîtrise du calcul en mécanique linéaire et non linéaire. Hermes, 2001.
22. P. Coorevits, J.P. Pelle, and P. Rougeot. Sur l'adaptativité des maillages à base de quadrilatères. application à l'automatisation de calculs par éléments finis en élasticité 2D. Revue européenne des éléments finis, 3(3):379–409, 1994.

INCREMENTAL SHEET METAL FORMING: A DIE-LESS RAPID PROTOTYPING PROCESS FOR SHEETMETAL

Eric S. Masuku, Alan N. Bramley, Anthony R. Mileham, and Geraint W. Owen

Innovative Manufacturing Research Centre University of Bath, Claverton Down, Bath BA2 7AY, UK ++44 1225 385366, E.S.Masuku@bath.ac.uk

Abstract: This paper presents an account of some investigations into a new die-less rapid prototyping process for sheetmetal forming, namely Incremental Sheetmetal Forming (ISF). The process involves the use of a single smooth tool to carry out local sheet metal deformation progressively on a CNC milling machine. The controlled movement of the tool enables a 3-Dimensional profile to be made. The process can offer rapid prototyping advantages for sheetmetal parts to be made directly from a 3D CAD model to the finished product without the traditional intermediate stage of tool design and manufacture. The paper describes an investigation into the capabilities of the ISF process by using experimental data to develop a tool-path generation methodology for generic materials and shapes. This is achieved by looking at a number of process variables, such as tool rotational speed and tool feed rates, on surface finish, as well as dimensional accuracy. The ISF process can offer the advantages of short lead times, high flexibility and low cost for volume applications.

Key words: CAD/CAM, Incremental sheetmetal forming, Die-less forming

1. INTRODUCTION

Manufacturing new products brings new challenges to the product introduction and commissioning process. In traditional sheetmetal forming the tooling costs are a drawback especially for bespoke and small batch manufacturing. The ISF process is a sheetmetal forming method that does

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 305-314.

not require a die or a unique punch to form complex shapes and therefore is much suited for low volume and high flexible manufacturing of sheetmetal products. It involves the use of a single smooth tool to carry out local sheet deformation progressively on a CNC milling machine. The deformation path is generated directly from CAD/CAM software application and downloaded to the online CNC milling machine. No specific tooling is required other than the sheetmetal holder fixed to the machine bed. This process can benefit sheetmetal rapid prototyping manufacturing as a die-less sheetmetal process. Process variables and process limitations have been identified and are discussed in this paper. Other approaches to incremental sheet forming have been reported including Hirt [1], Hagan [2],Tanaka [3] and Kitazawa [4] and Sawada [5].

2. THE ISF PROCESS

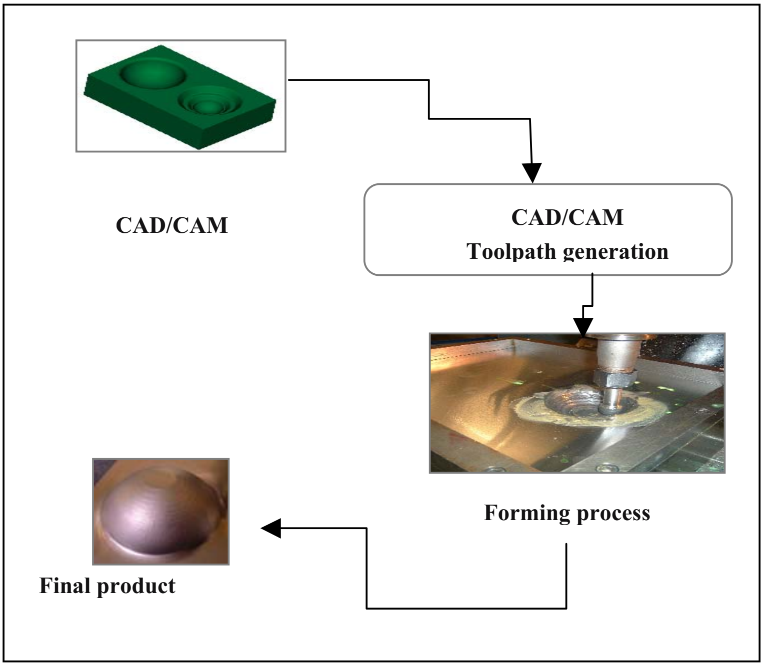


Figure 1. Schematic representation of the ISF process

Fig. 1 shows a schematic diagram of the manufacture of sheetmetal products using the ISF process. The local sheetmetal deformation is achieved by the tool following a given path, generated by the CAD/CAM software.

The sheetmetal is clamped around its periphery and fixed onto the CNC machine using the holder as shown in Fig. 2. Different size tools can be used for a typical part, Fig. 3, or different sub-part of the part.

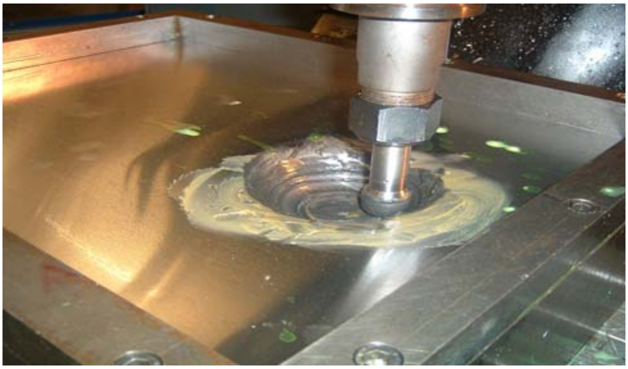


Figure 2. Sheetmetal holding mechanism

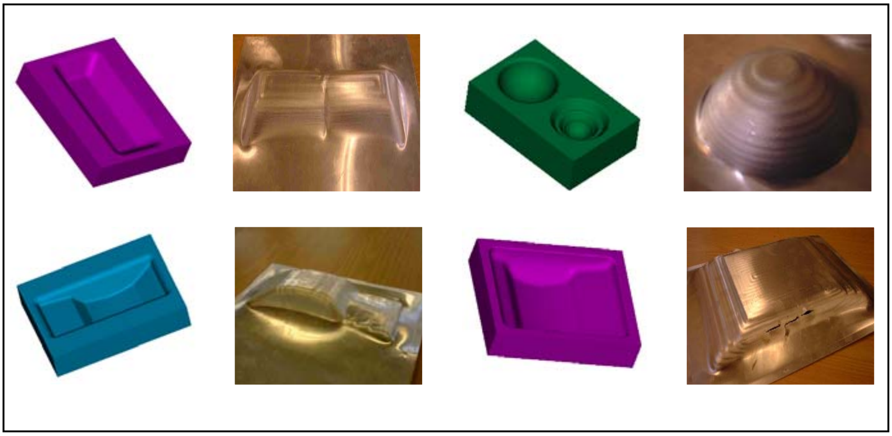


Figure 3. Typical products made with the ISF process and their respective CAD models

3. PROCESS ACCURACY

Unlike a material removal process the ISF process is dependent on the elastic-plastic behaviour of the material. The sheetmetal tends to spring back during the initial stages of deformation. This causes variation of the intended dimension of the formed part as shown in Fig. 4, as measured on a

Coordinate Measuring Machine. The main variation though is more pronounced in the depth of the shape than the width.

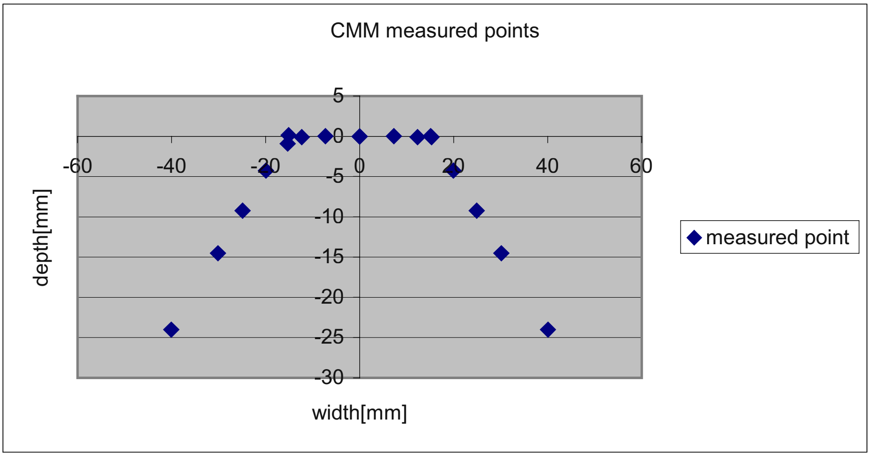


Figure 4. Intended ISF toolpath and actual measured points

Therefore the toolpath can be adjusted in the depth direction to account for this spring back effect. Toolpath strategies can be adopted to overcome these effects. Some of these are discussed in the next section together with shortcomings of the process.

4. TOOLPATH STRATEGIES

Generating an accurate toolpath is one of the main challenges of the ISF process. This is dependent on the mechanical properties of the sheetmetal, the holding mechanism, tool speed, tool feed rates and tool sizes. Experimental studies have been carried out to investigate the different tool-path strategies and process accuracy. Tool-paths for a variety of generic shapes have been generated and used to determine process accuracy capabilities. A selection of these are discussed in this section.

4.1 Initial deformation and toolpath strategies

Due to the elastic-plastic properties of the sheetmetal, the toolpath of a given shape will vary from the final shape as shown in Fig. 5. This is obvious at the beginning of the process where there is evidence of both elastic and plastic deformation. The final shape of the formed part has been found to be dependent upon a number of factors including the tool-path, the material properties of the sheetmetal, the tool material, tool speed and the tool feed rates.

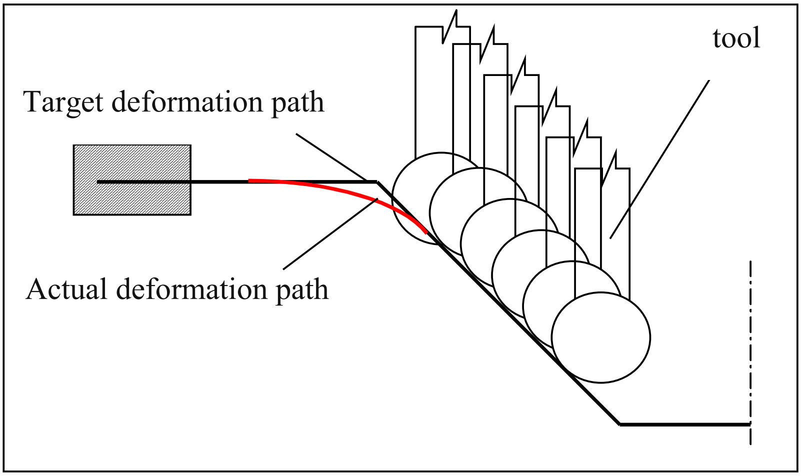


Figure 5. Target and actual ISF toolpath difference

Toolpath strategies to improve this have been developed for the selected case studies. Fig. 6 shows the various possible toolpath solutions to correct the difference in the targeted and actual deformation paths.

Option (d) was seen to be the most effective toolpath over the variation in the toolpath. Fig. 7 shows how this variation overcomes this problem.

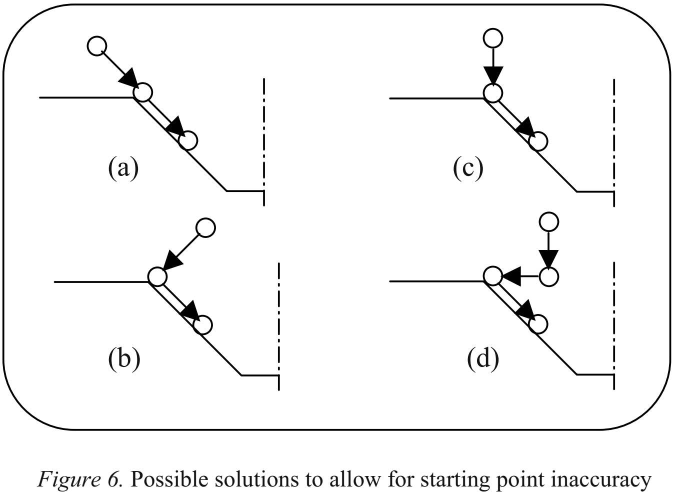


Figure 6. Possible solutions to allow for starting point inaccuracy

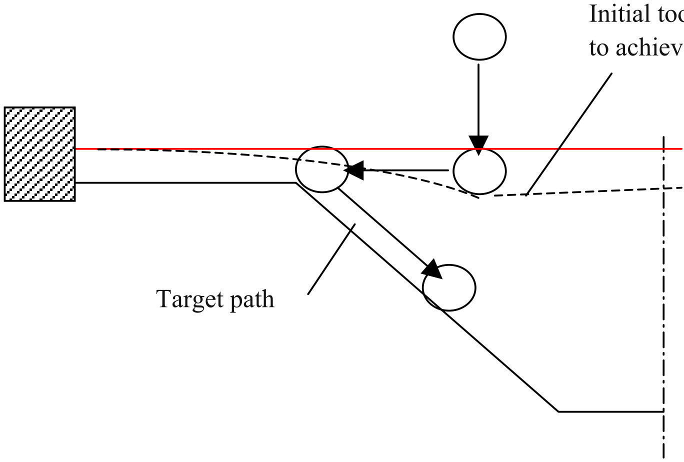


Figure 7. Initial deformation approach strategy

4.2 Maximum slope achievable using ISF

The incremental deformation of the sheetmetal is limited by the amount of depth to which the sheetmetal can be deformed. This can be verified by determining the limiting angle α as shown in Fig. 8. In order to compare different parts with various forming angles, the thickness of the sheet metal must be uniform for all parts. The only parameter that must change is the angle.

The part chosen is a cone with dimension 100 mm for the top diameter, 40 mm for the bottom diameter

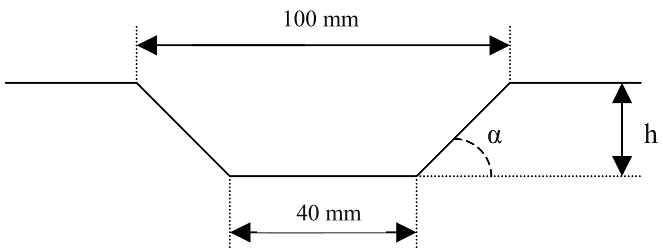


Figure 8. Dimensions of the cone

The tool has a diameter of 17.5 mm. The rectangular holder was used for this experiment

Table 1. Achieved angle of deformation

Angle [α]	Height [h][mm]	feed ratio [ΔX/ ΔZ]	Pass/fail
45	30	2 / 2	pass
56.31	45	2 / 3	pass
63.43	60	2 / 4	pass
71.57	90	0.5 / 1.5	pass

4.2.1 Results

The first three angles tested were successful with a relatively coarse feed ratio. The fourth case was achieved with small increments, whilst maintaining the feed ratios for the feed. These results, table 1, show that high deformation angles are achievable with the ISF process but the feed rates are important in achieving higher deformation angles.

5. FORMING LIMIT DIAGRAMS (FLD)

Figure 9. Formed gridded material

In traditional sheetmetal forming, processes are strains induced to form it into a given required shape. Typical strains achievable are for example in the process of deep drawing in the range of up to 30% extension. A gridded Aluminium sheet of 1.5mm thickness was formed into a conical shape (Fig. 9), and the strains were analysed Fig. 10.

The strain distribution along section A-A from Fig. 9 is shown in Fig.10. This is characteristic of a stretching forming process and it also shows that higher levels of strain are achievable without the onset of the necking of the sheet. These high strains can be attributed to the incremental deformation of the process as compared to more traditional processes like deep drawing. This characteristic is an advantage of the ISF processes as there can be more deformation induced on the sheet to generate complex shapes using the process.

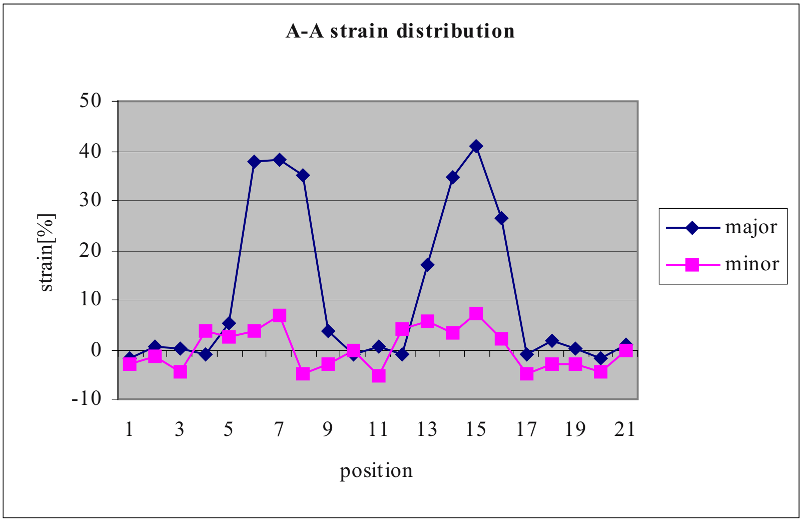


Figure 10. Strain distributions in the section A-A

The measured strains, in both major and minor directions, were used to construct a Forming Limit Diagram. The strain distribution of the major and minor strains indicate that very high strains are achievable which are in excess of traditional limiting boundaries of sheetmetal forming processes. Fig. 11 shows this FLD for the cone.

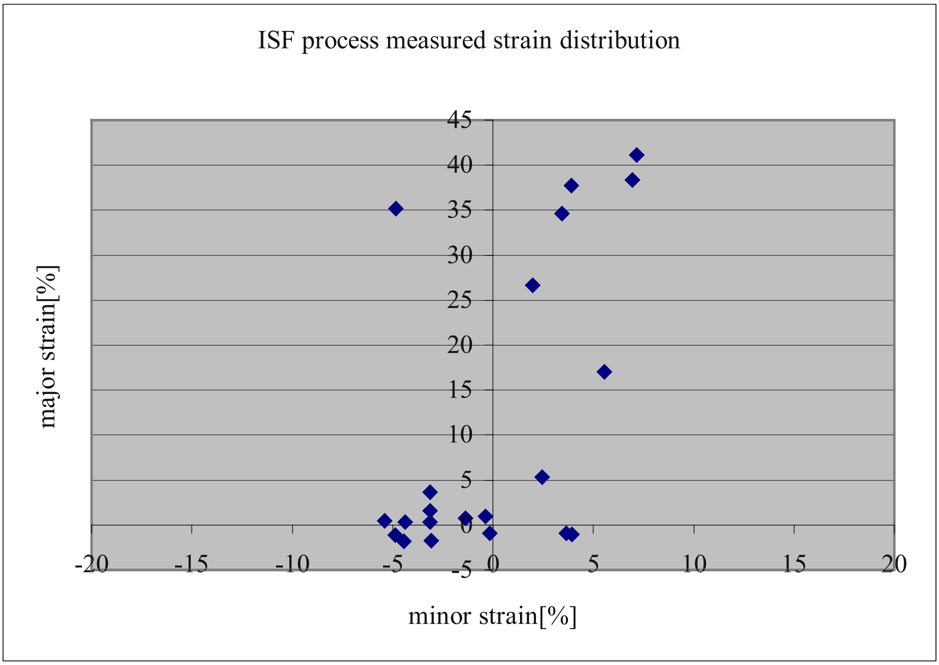


Figure 11. Forming limit diagram for cone

6. CONCLUSION

This paper has introduced the ISF process that uses CAD/CAM technology to manufacture sheetmetal products without the use of die and punches as in traditional processes. The process is therefore suited for bespoke, low volume and highly flexible manufacturing. Different process parameters have been identified which include tool speeds, feedrates, and toolpath strategies. It has also identified areas that need further investigation in order to improve the process.

Future work is looking at numerically modelling the process and comparing the results with FLD from the experimental work. Also further work will be carried to construct the FLD for the process as current techniques are directed at the traditional processes. The limits for these are shown to be below that of the ISF process.

ACKNOWLEDGEMENTS

The authors wish to thank the EPSRC for their financial support for this project. They also want thank Corus RD&T in Rotherham, UK for the assistance in the FLD experiments and data capture and all the Industrial collaborators in this project.

REFERENCES

1. G. Hirt, S. Junk, N. Witulski, Incremental Sheetmetal Forming: Quality Evalution and Process Simulation, 7th ITCP, Yokohama, 2002, pp. 925-930.
2. E. Hagan, J. Jeswiet, A review of conventional and modern single point sheet metal forming methods, IMECHE part B, Journal of Engineering Manufacture, vol.217 No 2. 2003, pp. 213-225.
3. S. Tanaka, T. Makamura, K. Hayakawa, Advanced Technology of Plasticity, Vol. II, 1999, pp. 1495-1500.
4. K. Kitizawa, A. Nakajima, Cylindrical incremental drawing of sheet metals by CNC incremental forming process, Advanced Technology of Plasticity, Vol. II, 1999, pp. 1495-1500.
5. T. Sawada, S. Matsubara, M. Sakamoto, G. Fukuhara, Deformation analysis for stretch forming of sheet metal with a CNC machine tool, Advanced Technology of Plasticity, Vol. II, 1999.pp. 1501-1504.

CONSTITUTIVE EQUATIONS FOR JOINTS WITH CLEARANCE

Charles Forster

Department of Mechanical Systems, Roberval Laboratory, University of Technology of Compiègne, BP 20 529, 60 205 Compiègne cedex, France

Abstract: The small displacements notion is widely used in tolerancing, particularly to describe the kinematics of joints with clearance. But this purely geometrical approach does not take into account the joint forces description. Our paper aims at examining the consequences of such an approach on the forces representation. After a brief reminder of the constitutive equations of perfect joints, our interest will focus on joints with clearance. We demonstrate that the forces in such joints have the same expression as the forces in the corresponding perfect joints. But their components do not have any value any longer, they must match some inequalities depending on the displacements actual value. The main interest of the proposed approach is to determine the joint forces properties by considering only the geometrical study, well mastered by now.

Key words: Mechanism theory; Tolerancing; Perfect joint; Small displacements.

1. INTRODUCTION

The mechanism theory is quite useful in the mechanical product design process. Particularly, the notion of perfect joint between two rigid bodies permits to analyse the product's dynamic behaviour independently from its technological development. This is the first step of a dimensioning process, leading to preliminary choices as regards geometry, dimensions and materials. In the following steps, additional mechanical phenomena can be taken into account, for instance by introducing stiffness in joints, or compliance of the parts (see for instance [1]).

Another extension of this theory takes into account the clearances between joints. Particularly, in the field of computed aided tolerancing, this

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 315-326.

approach has permitted the construction of 3D tolerance chains [2]. Here, the fundamental notion is that of the small displacements of a rigid body relative to another one. This tool is also used to describe the mechanical parts' geometric defects and to determine their influence on the assembling and the geometric behaviour of mechanical systems (see for instance [3-5]).

Even though these approaches have permitted to formalise, and partially solve the problems of tolerance analysis and synthesis, they are purely geometric and do not take into account the forces at work within the mechanical system. Now the consideration of the forces would permit to establish a complete and consistent theory, liable to help the solving of such problems and to determine the influence of the geometric defects on the mechanical system's behaviour .

Recent pioneering works [6] have just started to take into account clearance and stiffness in particular case studies. The present work aims at establishing the theoretical fundamentals of such an approach, i.e at determining the constitutive equations of joints featuring clearance.

In section 2, we recall the general theory of the constitutive equations for perfect joints. In the third section, we focus our interest to the particular case of joints with clearance. After developing the changes in kinematics induced by the presence of the clearance, we apply the general method in order to exhibit the relations between displacements and joint forces. Finally, in section 4, we illustrate the efficiency of the method proposed by studying a particular joint.

2. CONSTITUTIVE EQUATIONS FOR PERFECT JOINTS

In this section the theory of the constitutive equations for perfect joints, developed by Germain [7], and based on the virtual power method, is briefly recalled.

2.1 Geometry and kinematics of a rigid bodies system

We consider the motion of a system of rigid bodies S within a frame $\Re$. Let $\underline{q}$ be the set of independent time-dependent parameters describing completely the position of the bodies in $\Re$. Moreover the parameters $\underline{q}$ are subjected to some additional constraints (due to unilateral joints) of type:

$$g_l\left\{ \underline{q}\right\} \geq 0 \qquad l = 1,...,u \tag{1}$$

For each point M belonging to the system S, the velocity vector relative to the frame $\Re$ at time t can be written:

$$\underline{V}_M = \sum_i \frac{\partial \underline{M}}{\partial q^i} \dot{q}^i + \frac{\partial \underline{M}}{\partial t} \quad with \quad \dot{q}^i = \frac{d}{dt} q^i(t) \tag{2}$$

Due to Eq. (1), the time derivatives cannot be assigned any value, and must comply with the following conditions:

$$\sum_i \frac{\partial g_l}{\partial q^i} \dot{q}^i + \frac{\partial g_l}{\partial t} \geq 0 \quad if \quad g_l(\underline{q}) = 0, \quad l = 1,...,v \quad and \quad v \leq u \tag{3}$$

expressing the fact that, if $g_l = 0$ at time t, g_l will keep this value or become positive immediately afterwards (so its variation must be positive).

2.2 Virtual motions under consideration

By definition, a frozen virtual motion associated with parameters $\underline{q}$ (FVMA) is, at time t, the field of virtual velocity vectors defined by:

$$\underline{\hat{V}}_M = \sum_i \frac{\partial \underline{M}}{\partial q^i} \hat{q}^i \tag{4}$$

and obtained by replacing the derivatives by virtual parameters $\underline{\hat{q}}$ [1] in the expression of $\underline{V}_M$. The word "frozen" expresses the fact that such a virtual motion is instantaneous, so it cannot explicitly depend on time t.

A FVMA defines a frozen virtual motion compatible with all the constraints (FVMC) at time t, if the virtual parameters fulfil the relations:

$$\sum_i \frac{\partial g_l}{\partial q^i} \hat{q}^i \geq 0 \quad if \quad g_l(\underline{q}) = 0, \quad l = 1,...,v \tag{5}$$

The FVMC represent all the possible motions respecting the unilateral properties of the joints considered.

[1] We do not write $\underline{\hat{q}}$ for simplicity's sake.

2.3 Associated representation of forces

According to the virtual power method, the virtual power of the joint forces within the S system is a linear function of the virtual parameters $\underline{\hat{q}}$:

$$\hat{\mathcal{P}}_S = \sum_i \pi_i \, \hat{q}^i \tag{6}$$

the dual variables π_i defining the components of the forces. Now the constitutive equations of perfect joints can be formulated:

"Joints constraining the motion of the system S are said to satisfy the constitutive equations of perfect joints if, for each arbitrary FVMC, the virtual power of the forces is non-negative".

It can be proved that this definition implies that the components π_i necessarily satisfy:

$$\pi_i = \sum_{l=1}^{u} \lambda_l \frac{\partial g_l}{\partial q^i} \quad with \quad \begin{cases} \lambda_l = 0 & if \quad g_l(\underline{q}) > 0 \\ \lambda_l \geq 0 & if \quad g_l(\underline{q}) = 0 \end{cases} \tag{7}$$

3. CASE OF A JOINT WITH CLEARANCE

3.1 Nominal joint

We now consider two rigid bodies S and Σ mobile within a frame $\mathcal{R}$ (with $\mathcal{E}$ its associated three-dimensional Euclidean affine space). The motion of Σ (resp. S) relative to $\mathcal{R}$ is represented in the space $\mathcal{E}$ by a one-parameter family of Cartesian coordinates systems, $\mathcal{R}_\Sigma(t) = \{O_\Sigma(t), \mathcal{B}_\Sigma(t)\}$[2] (resp. $\mathcal{R}_S(t) = \{O_S(t), \mathcal{B}_S(t)\}$).

Suppose that a joint of movability m exists between both rigid bodies, i. e. a passive, time-independent and bilateral joint. This joint will be qualified as "nominal" in what follows.

The kinematics associated to this joint is defined at each time t by the velocity distributor $\{C_{S/\Sigma}\}_{nom}$ (or equivalently by $\{C_{\Sigma/S}\}_{nom}$) :

$$\{C_{S/\Sigma}\}_{nom} = \{\underline{\Omega}_{S/\Sigma}, \underline{V}_{Q\in S/\Sigma}\} = -\{C_{\Sigma/S}\}_{nom} \tag{8}$$

[2] In the following, and for simplicity's sake, the variable t will be omitted.

where $\underline{\Omega}_{S/\Sigma}$ is the rotation's rate pseudo-vector of S relative to Σ. $\underline{V}_{Q\in S/\Sigma}$ denotes the velocity vector of the particle of S coinciding with point Q.

Joint forces are represented at time t by the torsor $\left[T_{\Sigma\to S} \right]_{nom}$:

$$\left[T_{\Sigma\to S} \right]_{nom} = \left[\underline{R}_{\Sigma\to S}, \underline{M}_{Q,\Sigma\to S} \right] = -\left[T_{S\to\Sigma} \right]_{nom} \tag{9}$$

where $\underline{R}_{\Sigma\to S}$ and $\underline{M}_{Q,\Sigma\to S}$ are respectively the force vector and the moment pseudo-vector at point Q, exerted by Σ on S.

The nominal joint matches the perfect joint constitutive equations. So $\left\{ C_{S/\Sigma} \right\}_{nom}$ and $\left[T_{\Sigma\to S} \right]_{nom}$ comply with Q for each point:

$$\left[T_{\Sigma\to S} \right]_{nom} \otimes \left\{ C_{S/\Sigma} \right\}_{nom} = \underline{R}_{\Sigma\to S} \cdot \underline{V}_{Q\in S/\Sigma} + \underline{M}_{Q,\Sigma\to S} \cdot \underline{\Omega}_{S/\Sigma} = 0 \tag{10}$$

Standard NF EN 23952 - ISO 3952 defines a set of reference joints. For each of these reference joints, there exists a Cartesian coordinate system $\mathcal{R}_{nom}(t) = \{O_{nom}(t), \mathcal{B}_{nom}(t)\}$ in which the vectors of $\left\{ C_{S/\Sigma} \right\}_{nom}$ and $\left[T_{\Sigma\to S} \right]_{nom}$ have the simplest expression, i. e. most of their components are equal to zero.

3.2 Geometry and kinematics of a joint with clearance

The joint being in practice manufactured with a finite precision, there exists a free space between the surfaces of the two bodies after assembling, commonly called "clearance". Due to this space, the motion of S relative to Σ is no more constrained, and six time-dependent parameters $\underline{q} = \left\{ q^1, ..., q^6 \right\}$ are required to completely determine the position of each particle of S relative to Σ.

These parameters are chosen as: the three coordinates (u, v, w) of particle O_S in $\mathcal{R}_\Sigma$, and three angles (α, β, γ) representing three sequential rotations around the base vectors of $\mathcal{B}_\Sigma$ (rotation sequence "123").

In general, the clearance is small as compared to the other dimensions of the solids. So additional motions (relative to the nominal joint) due to the clearance remain fairly limited.

It makes sense introducing a small perturbation assumption. Coordinate systems $\mathcal{R}_S$ et $\mathcal{R}_\Sigma$ are supposed to coincide at initial time t_o with $\mathcal{R}_{nom}$ (reference configuration); and the parameters $\underline{q}$ have little values compared to the unity.

Consequently $\mathcal{B}_{nom}$, $\mathcal{B}_\Sigma$ and $\mathcal{B}_S$ are equal to first order. And it can be proved [1, 7] that the displacement field of S relative to Σ is a distributor, the small displacement distributor $\left\{ D_{S/\Sigma} \right\}$, defined at time t by:

$$\{ D_{S/\Sigma} \} = \{ \underline{\omega}_{S/\Sigma}, \underline{U}_{O_S/\Sigma} \} = -\{ D_{\Sigma/S} \} \tag{11}$$

$\underline{\omega}_{S/\Sigma}$ is the pseudo-vector of small rotations of S relative to Σ, its components in the basis $\mathcal{B}_{nom}$ are equal to (α, β, γ). $\underline{U}_{O_S/\Sigma}$ is the small displacement vector of particle O_S (coinciding in the reference configuration with O_Σ and O_{nom}); its components are (u, v, w) in $\mathcal{B}_{nom}$.

Moreover it can be established that:

$$\underline{\Omega}_{S/\Sigma} = \frac{d}{dt}\underline{\omega}_{S/\Sigma} \quad \underline{V}_{O_S/\Sigma} = \frac{d}{dt}\underline{U}_{O_S/\Sigma} \tag{12}$$

the time-derivatives being computed relative to base $\mathcal{B}_{nom}$. So the components in $\mathcal{B}_{nom}$ of the vectors $\underline{\omega}_{S/\Sigma}$ and $\underline{U}_{O_S/\Sigma}$ of the velocity distributor $\{ C_{S/\Sigma} \}$ are the time-derivatives of the parameters $\underline{q}$.

As proposed in Germain [7], $\{ C_{S/\Sigma} \}$ can be split into two parts:

$$\{ C_{S/\Sigma} \} = \{ C_{S/\Sigma} \}_{nom} + \{ C_{S/\Sigma} \}_{cle} \tag{13}$$

where the m non-zero components of $\{ C_{S/\Sigma} \}_{nom}$ represent the degrees of freedom of the nominal joint, and where the n = 6 - m non-zero components of $\{ C_{S/\Sigma} \}_{cle}$ represent the supplementary degrees of freedom due to the clearance presence. Equation (13) is nothing but the canonical decomposition of a distributor, by projection on two orthogonal sub-spaces of the distributors' linear space. It permits to separate the set of parameters $\underline{q}$ into two sub-sets: a first set of m parameters $\underline{q}_{nom}$ associated with the nominal joint, and a second one of n parameters $\underline{q}_{cle}$ due to the clearance.

In the same manner, following Giordano [5], $\{ D_{S/\Sigma} \}$ can be split in:

$$\{ D_{S/\Sigma} \} = \{ D_{S/\Sigma} \}_{nom} + \{ D_{S/\Sigma} \}_{cle} \tag{14}$$

where $\{ D_{S/\Sigma} \}_{nom}$ $\left(\text{resp.} \{ D_{S/\Sigma} \}_{cle}\right)$ depends only on m parameters $\underline{q}_{nom}$ (resp. the n parameters $\underline{q}_{cle}$).

3.3 Additional inequalities due to the clearance

If S and Σ enter into contact at time t, there exists at least one particle M_S belonging to the surface of S and one point M_Σ of the surface of Σ such that: $M_S = M_\Sigma$. Moreover, it is assumed that in this point, at least one of the two surfaces is regular, of unit outward normal $\underline{n}_S$ and/or $\underline{n}_\Sigma$. In what follows, $\underline{n}$ will denote the unit outward normal $\underline{n}_S$ to S if it exists, or the opposite of $\underline{n}_\Sigma$.

Contact can be expressed by:

$$\underline{OM_\Sigma} = \underline{OM_S(t)} = \underline{OM_S(t_o)} + \left(\underline{U}_{M_S/\Sigma} \right)_{nom} + \left(\underline{U}_{M_S/\Sigma} \right)_{cle} \tag{15}$$

$\left(\underline{U}_{M_S/\Sigma} \right)_{nom}$ is the displacement vector of particle M_S associated to the nominal joint (parameters $\underline{q}_{nom}$). This vector is necessarily orthogonal to $\underline{n}$, because the technological development of the nominal joint is such that the allowed (small) displacement is tangent to the surfaces. $\left(\underline{U}_{M_S/\Sigma} \right)_{cle}$ represents the displacement vector of M_S due to the clearance (parameters $\underline{q}_{cle}$), necessarily parallel to $\underline{n}$.

Then, Eq. (15) implies:

$$\left(\underline{U}_{M_S/\Sigma} \right)_{cle} . \underline{n} = \left(\underline{OM_\Sigma} - \underline{OM_S(t_o)} \right) . \underline{n} = j \qquad j \geq 0 \tag{16}$$

Configurations without contact in M_Σ correspond to:

$$\left(\underline{U}_{M_S/\Sigma} \right)_{cle} . \underline{n} = \left(\underline{OM_\Sigma} - \underline{OM_S(t_o)} \right) . \underline{n} < j \tag{17}$$

We do not develop the effective search for Eqs. (16)-(17). This would require a detailed study of the geometry of the bodies S and Σ. In practice, due to the technological joint development, unique or multiple contacts between the surfaces of S and Σ can be expressed by u independent[3] inequalities:

$$g_l \left\{ q^1_{cle}, \ldots, q^n_{cle} \right\} = -\left(\underline{U}_{M_S/\Sigma} \right)_{cle} . \underline{n} + j \geq 0 \quad l = 1, \ldots, u \tag{18}$$

where $\underline{n}$ and j may depend on the parameters q_{cle}.

3.4 Associated virtual motions

The velocity field of body S relative to Σ being defined by the distributor $\left\{ C_{S/\Sigma} \right\}$, FVMA are represented at time t by distributor $\left\{ \hat{C}_S \right\}$, obtained from $\left\{ C_{S/\Sigma} \right\}$ by substituting the time-derivatives of the parameters $\underline{q}$ by virtual parameters $\hat{\underline{q}}$.

Due to the definition of the parameters $\underline{q}_{nom}$ and $\underline{q}_{cle}$, and taking into account the canonical decomposition of $\left\{ C_{S/\Sigma} \right\}$ given in Eq. (13), virtual

[3] When these inequalities transform into equalities, if the parameters $\underline{q}_{cle}$ are solutions for some of them, they are not solutions for the others.

parameters $\underline{\hat{q}}$ are separated into two sets $\underline{\hat{q}}_{nom}$ and $\underline{\hat{q}}_{cle}$ of natural interpretation. $\left\{ \hat{C}_S \right\}$ can be written:

$$\left\{ \hat{C}_S \right\} = \left\{ \hat{C}_S \right\}_{nom} + \left\{ \hat{C}_S \right\}_{cle} \tag{19}$$

where $\left\{ \hat{C}_S \right\}_{nom}$ $\left(\text{resp.} \left\{ \hat{C}_S \right\}_{cle}\right)$ is built from $\left\{ C_{S/\Sigma} \right\}_{nom}$ $\left(\text{resp.} \left\{ C_{S/\Sigma} \right\}_{cle}\right)$ by replacing the parameters $\underline{q}_{\text{nom}}$ (resp. $\underline{q}_{\text{cle}}$) by the parameters $\underline{\hat{q}}_{nom}$ $\left(\text{resp. } \underline{\hat{q}}_{cle}\right)$.

In the same way, a FVMC will be defined by the two distributors $\left\{ \hat{C}_S \right\}_{nom}$ and $\left\{ \hat{C}_S \right\}_{cle}$ such that:

$$\sum_{i=1}^{n} \frac{\partial g_l}{\partial q^i_{cle}} \hat{q}^i_{cle} \geq 0 \quad if \quad g_l(q^1_{cle}, ..., q^n_{cle}) = 0, \quad l = 1, ..., v \tag{20}$$

3.5 Representation of the forces

Because the virtual power is in the present case a linear function on the MVFC vector space, the joint forces are represented by a torsor $\left[T_{\Sigma \to S} \right]$, and the virtual power of these forces becomes:

$$\hat{\mathcal{P}}_{\Sigma \to S} = \left[T_{\Sigma \to S} \right] \otimes \left\{ \hat{C}_S \right\} \tag{21}$$

Combining Eqs. (21) and (19) leads to:

$$\hat{\mathcal{P}}_{\Sigma \to S} = \left[T_{\Sigma \to S} \right] \otimes \left\{ \hat{C}_S \right\}_{nom} + \left[T_{\Sigma \to S} \right] \otimes \left\{ \hat{C}_S \right\}_{cle} \tag{22}$$

According to the virtual power method, the consequences of Eq. (22) are explored by considering different particular cases.

If one chooses $\left\{ \hat{C}_S \right\}_{cle} = 0$, i. e. $\underline{\hat{q}}_{cle} = 0$, one easily finds:

$$\forall \left\{ \hat{C}_S \right\}_{nom} \quad \hat{\mathcal{P}}_{\Sigma \to S} = \left[T_{\Sigma \to S} \right] \otimes \left\{ \hat{C}_S \right\}_{nom} \geq 0 \quad \Rightarrow \quad \left[T_{\Sigma \to S} \right] = \left[T_{\Sigma \to S} \right]_{nom} \tag{23}$$

We thus prove that the torsor of the joint with clearance has necessarily the same shape as the torsor relative to the nominal joint. The dual variables $\underline{\pi}$ are the components of this torsor: π_i is a component of the force $\underline{R}_{\Sigma \to S}$ (resp. of the moment) if $\underline{\hat{q}}^i_{cle}$ is a dimension (resp. an angle). Moreover, from Eq. (10), we can derive the new virtual power expression:

$$\hat{\mathcal{P}}_{\Sigma\to S} = \left[T_{\Sigma\to S}\right]_{nom} \otimes \left\{\hat{C}\right\}_{cle} \tag{24}$$

If $\left\{\hat{C}_S\right\}_{cle} \neq 0$ and all the functions g_l satisfy: $g_l > 0$ (there is no contact between the bodies), all the λ_l are equal to zero, $\underline{\pi}$ is equal to zero and consequently $\left[T_{\Sigma\to S}\right]_{nom}$. This is logical: when there is no contact because of the clearance, forces must be null.

In the general case where $\left\{\hat{C}_S\right\}_{cle} \neq 0$ and one or more functions g_l satisfy: $g_l = 0$, the components $\underline{\pi}$ of the force torsor are computed from Eqs. (7).

4. EXAMPLE

We consider a composite joint with four degrees of freedom, the association of a planar joint of normal $\underline{z}_{nom}$ and a cylindrical one of axis (O_{nom}, $\underline{x}_{nom}$).

The nominal Cartesian coordinate system is (O_{nom}, x_{nom}, y_{nom}, z_{nom}), and the components (plotted in columns) of the vectors of the velocity distributor $\left\{C_{S/\Sigma}\right\}_{nom}$ and the force torsor $\left[T_{\Sigma\to S}\right]_{nom}$ are:

$$\left\{C_{S/\Sigma}\right\}_{nom} = \begin{Bmatrix} A & U \\ 0 & V \\ G & 0 \end{Bmatrix} \qquad \left[T_{\Sigma\to S}\right]_{nom} = \begin{bmatrix} 0 & 0 \\ 0 & M \\ F & 0 \end{bmatrix} \tag{25}$$

It is worth noting that F et M may have any value, because the nominal joint is bilateral.

The joint is technologically built up with a cylinder (body S) of diameter 2r, of length 2l, lying between two parallel upper and lower planes $\mathcal{P}_u$ and $\mathcal{P}_l$ belonging to the body Σ and distant from 2R (see Fig. 1). In the initial configuration, S is half way from both planes. Consequently there exists a gap j = R – r between the upper and lower linear generators $\mathcal{G}_u$ and $\mathcal{G}_l$ of S and the planes of Σ.

To take into account the symmetry of this configuration, we change the definition of the coordinate system $\mathcal{R}_{nom}$. But it is worth noting that $\left\{C_{S/\Sigma}\right\}_{nom}$ and $\left[T_{\Sigma\to S}\right]_{nom}$ remain unchanged.

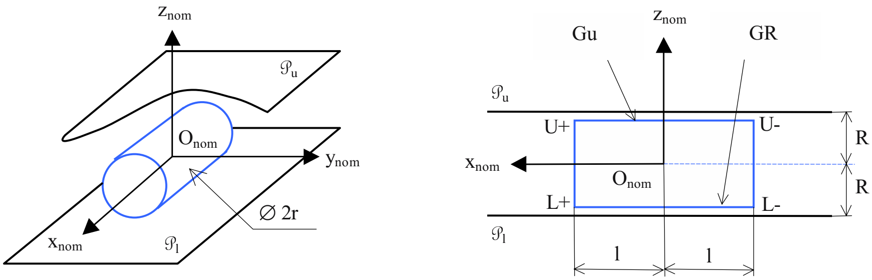


Figure 1. Technological realisation of the joint and definitions.

Small displacements of S relative to Σ involve the three coordinates u, v, w of particle O_S, and three angles relative to the axis x_{nom}, y_{nom}, z_{nom}. The distributor $\{ C_{S/\Sigma} \}$ can be split into:

$$\{ C_{S/\Sigma} \} = \begin{Bmatrix} \dot{\alpha} & \dot{u} \\ \dot{\beta} & \dot{v} \\ \dot{\gamma} & \dot{w} \end{Bmatrix} = \begin{Bmatrix} \dot{\alpha} & \dot{u} \\ 0 & \dot{v} \\ \dot{\gamma} & 0 \end{Bmatrix}_{nom} + \begin{Bmatrix} 0 & 0 \\ \dot{\beta} & 0 \\ 0 & \dot{w} \end{Bmatrix}_{cle} \tag{26}$$

Expressions of $\{ D_{S/\Sigma} \}_{nom}$ and $\{ D_{S/\Sigma} \}_{cle}$ are:

$$\{ D_{S/\Sigma} \}_{nom} = \begin{Bmatrix} \alpha & u \\ 0 & v \\ \gamma & 0 \end{Bmatrix}_{nom} \qquad \{ D_{S/\Sigma} \}_{cle} = \begin{Bmatrix} 0 & 0 \\ \beta & 0 \\ 0 & w \end{Bmatrix}_{cle} \tag{27}$$

The components of $\{ D_{S/\Sigma} \}_{nom}$ are small but of any value. The components of $\{ D_{S/\Sigma} \}_{cle}$ are bounded by the fact that the points of the linear generators $\mathcal{G}_u$ and $\mathcal{G}_l$ are confined by planes $\mathcal{P}_u$ and $\mathcal{P}_l$. For instance the contact between point U+ belonging to $\mathcal{G}_u$ (with coordinates (l, 0, r)) and plane $\mathcal{P}_u$ of unit outside normal $\underline{n}_\Sigma = -\underline{z}_{nom}$ ($\underline{n}_s = +\underline{z}_{nom}$) is expressed by:

$$\underline{U}_{S+/\Sigma} \cdot \underline{n}_S = \left(w - l\beta \right) \leq j \quad \Rightarrow \quad g_1(w, \beta) = -w + l\beta + j \geq 0 \tag{28}$$

All the possible cases have been recorded in the first two columns of Table 1.

Concerning the forces, the torsor has two non-zero components: F is the dual variable of w, and M is associated to β. Both components are null if functions g_l are strictly positive. Otherwise, F and M satisfy the Eqs. (7), depending on the values of w and β. For instance, if $g_1 = 0$ (unique contact in U+), one obtains for F and M:

$$\begin{cases} F = \lambda_1 \dfrac{dg_1}{dw} = -\lambda_1 \\ M = \lambda_1 \dfrac{dg_1}{d\beta} = +\lambda_1 l \end{cases} \quad \lambda_1 \geq 0 \quad i.e. \quad \begin{cases} F \leq 0 \\ M = -lF \geq 0 \end{cases} \tag{29}$$

All the other cases have been plotted on Table 1. The second and last columns define the constitutive equations of the joint under consideration.

Table 1. Forces and displacements relations for each contact case

Contact case	Additional constraints $g_l(w,\beta)$	Expression of F and M ($\lambda_i \geq 0$)	Relation between F and M
U+ / $\mathcal{P}_u$	$g_1 = -w + l\beta + j = 0$	$F = -\lambda_1$ $M = +\lambda_1 l$	$F \leq 0, M = -Fl$
U- / $\mathcal{P}_u$	$g_2 = -w - l\beta + j = 0$	$F = -\lambda_2$ $M = -\lambda_2 l$	$F \leq 0, M = +Fl$
L+ / $\mathcal{P}_l$	$g_3 = +w - l\beta + j = 0$	$F = +\lambda_3$ $M = -\lambda_3 l$	$F \geq 0, M = -Fl$
L- / $\mathcal{P}_l$	$g_4 = +w + l\beta + j = 0$	$F = +\lambda_4$ $M = +\lambda_4 l$	$F \geq 0, M = +Fl$
$\mathcal{G}_u$ / $\mathcal{P}_u$	$g_1 = 0, g_2 = 0$	$F = -\lambda_1 - \lambda_2$ $M = +\lambda_1 l - \lambda_2 l$	$F \leq 0, +Fl \leq M \leq -Fl$
$\mathcal{G}_l$ / $\mathcal{P}_l$	$g_3 = 0, g_4 = 0$	$F = +\lambda_3 + \lambda_4$ $M = -\lambda_3 l + \lambda_4 l$	$F \geq 0, -Fl \leq M \leq +Fl$
U+ / $\mathcal{P}_u$ and L- / $\mathcal{P}_l$	$g_1 = 0, g_4 = 0$	$F = -\lambda_1 + \lambda_4$ $M = +\lambda_1 l + \lambda_4 l$	$-\dfrac{M}{l} \leq F \leq +\dfrac{M}{l}, M \geq 0$
L+ / $\mathcal{P}_l$ and U- / $\mathcal{P}_u$	$g_2 = 0, g_3 = 0$	$F = -\lambda_2 + \lambda_3$ $M = -\lambda_2 l - \lambda_3 l$	$+\dfrac{M}{l} \leq F \leq -\dfrac{M}{l}, M \leq 0$

5. CONCLUSION

The general formalism of the constitutive equations of perfect joints has been applied to joints with clearance. From the small displacement assumption classically used in this case, we give the expression of the relevant virtual motions. Then we derive the relations existing between joint forces and displacements.

These constitutive equations complete the purely geometric approach of these joints. They can be helpful to find particular configurations associated with given loads. From a purely theoretical point of view, they represent the first step towards a complete theory of mechanisms including joints with clearance.

Further works will concern the extension of this model to finite displacements for the nominal degrees of freedom, and the definition of extended indicators (movability, redundancy, …) for multibody systems.

REFERENCES

1. M. Aublin, R. Boncompain, M. Boulaton, D. Caron, E. Jeay, B. Lacage, J.Rea, 1998, Systèmes Mécaniques - Théorie et Dimensionnement, Dunod, Paris
2. P. Clozel, 2003, 3D Tolerances analysis, from preliminary study, in: Geometric Product Specification and Verification : Integration of Functionality, P. Bourdet and L. Matthieu eds., Kluwer Academic publishers, Dordrecht, pp. 93-104
3. E. Ballot, P. Bourdet, F. Thiebaut, 2003, Determination of relative situations of parts for tolerance computation, in: Geometric Product Specification and Verification : Integration of Functionality, P. Bourdet and L. Matthieu eds., Kluwer Academic publishers, Dordrecht, pp. 63-72
4. A. Clement, A. Rivière, M. Temmerman, 1994, Cotation Tridimensionnelle des Systèmes Mécaniques, PYC Edition, Paris
5. M. Giordano, 1998, Modèle de détermination des tolérances, in : Conception de Produits Mécaniques, Méthodes, Modèles et Outils, M. Tollenaere ed., Hermes, Paris, pp. 305-321
6. S. Samper, M. Giordano, 2003, Simultaneous analysis method for tolerancing flexible mechanism, in: Geometric Product Specification and Verification : Integration of Functionality, P. Bourdet and L. Matthieu eds., Kluwer Academic publishers, Dordrecht, pp. 127-134
7. P. Germain, 1986, Mécanique, T. 1, Ellipses, Paris

Part 3

DESIGN TOOLS FOR PARTICULAR APPLICATIONS

AN ASSISTANCE TOOL FOR SPLINE COUPLING DESIGN

Adrien Barrot, Manuel Paredes and Marc Sartor
Laboratoire de Génie Mécanique de Toulouse, INSA, F-31077 Toulouse, France

Abstract: This research was motivated by the absence of a means to design a straight spline coupling in a reliable way. It is shown that the strain and pressure distribution in a spline coupling are not uniform. To predict these phenomena, an assistance tool would be of key interest. In this paper, a simplified finite element model has been made of a spline coupling to build a dimensioning tool for such couplings. This tool can give information about strain on the outside diameter of the sleeve and axial torque distribution. This study is based on an experimentally validated three-dimensional finite element model of the coupling under torsional load. Different sleeve sizes and loading methods are tested to highlight their effects on the axial torque distribution

Key words: Straight flank spline coupling; Finite element analysis; Strain distribution.

1. INTRODUCTION

In the industrial world, it is becoming more and more necessary to discover spline coupling behaviour with precision. There is little published information about the sizing of spline couplings. The studies presented here show that the axial load distribution on the teeth is far from uniform. The highlighted axial stress concentration contradicts the assumptions of standardization concerning the dimensioning of the spline coupling. This can have harmful consequences: mistakes in dimensioning can lead to a failure in spite of correct dimensioning in agreement with standards.

It is therefore essential to simulate such assemblies in order to evaluate the load and stress distributions, and thus to be able to adjust the standardized dimensioning according to the operating conditions. This work deals with creating and validating a tool that would allow the behaviour of a spline coupling to be quickly analyzed.

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 329-342.

Analytical approaches can be useful for developing such a tool. Volfson [1] performed an analytical study and highlighted the influence of shaft and sleeve inertia on the axial pressure distribution along the side of a tooth. Fig. 1 shows the different pressure distributions according to the inertia of the parts.

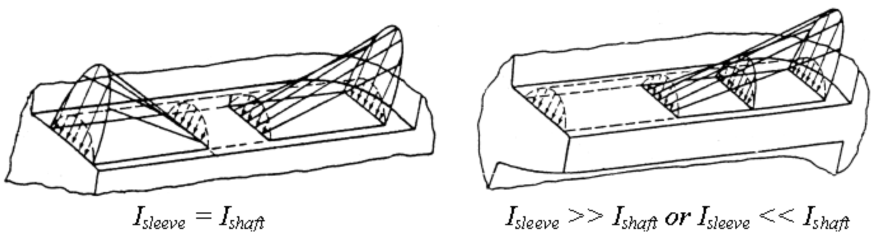


Figure 1. Diagram of pressure by Volfson (After [1]).

Volfson takes a parabolic pressure distribution as a hypothesis to resolve his problem. Moreover, a concentration coefficient of pressure distribution is imposed. Unlike Volfson, Tatur [2], tried to determine the axial distribution of the loading without assuming a coefficient of stress concentration. Finally, Orain [3] developed an analytical approach to axial pressure distribution that like Tatur also takes the effect of inertia into account. He then obtained a pressure distribution:

$$p(x) = A \,.\, e^{k\,x} + B\, e^{-k\,x}, \text{ where: } k = \sqrt{\frac{H.h.n.r}{G}.\left(\frac{1}{I_a} + \frac{1}{I_m}\right)}$$

A and B are constants depending on the boundary conditions, H corresponds to torsional rigidity of the spline contact, n is the number of teeth, G is the shear modulus, I_a and I_m correspond to the quadratic moment of shaft and sleeve, and h and r are described in Fig. 2.

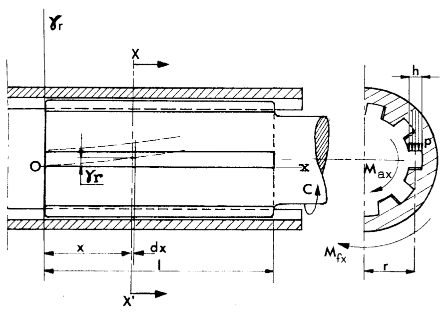


Figure 2. Representation of spline coupling with parameters (After [3]).

In this calculation, the radial pressure distribution is considered as constant (see Fig. 2). It is significant to notice that no calculation clarifying H (parameter describing the behaviour of the teeth) is described. Orain imposed H at the time of the application examples.

All these analytical studies can be used to model a basic spline coupling where the inertia of the shaft and the sleeve are constant in the axial direction. Moreover, only basic loading cases are to be considered (the torque has to be applied on one side of the coupling).

To improve the spline coupling model, a simplified finite element model can be built. Blanc [4] developed such a model to determine the axial load distribution of an involute spline coupling. To achieve this, the spline coupling is divided into twenty parts in the axial direction. Fig. 3 shows a description of a part of the sleeve or shaft.

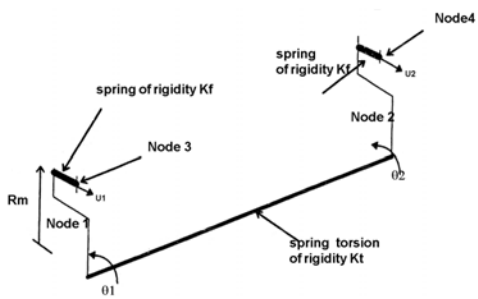


Figure 3. Description of degrees of freedom for a finite element model of a piece of shaft or sleeve (After [4])

Rm = sleeve radius
θ1, θ2 = extremities of rotation of the part
K_T = torsional rigidity of the shaft or sleeve for a length L
K_F = tangential rigidity in bending of a spline with a length of L / 2

The strain energy of a shaft or sleeve part is expressed as follows:

$$E=\frac{1}{2}K_T(\theta_2-\theta_1)^2+\frac{1}{2}K_F(R_m.\theta_1-u_1)^2+\frac{1}{2}K_F(R_m.\theta_2-u_2)^2+\frac{1}{2}K_C\left[(u_1-R_m.\theta_1)^2-(u_2-R_m.\theta_2)^2\right]$$

With $K_T=\frac{G.J}{L}$ and $K_C=\frac{G.S}{L}$, where J is the moment of inertia and S the surface area of the section.

The determination of K_F with an analytical method has not proved satisfactory. It has therefore been necessary to use a numerical method to find K_F. Blanc calculated a stiffness coefficient, and obtained a fixed value (K_F = 7,3 x 10^4N/mm) whatever the involute spline coupling.

With such a model, complex shaft and sleeve geometries can be tested and the external torque can be introduced at any axial position. It can quickly give the torque along the contact length at every node. Axial pressure distribution has been evaluated by Blanc [4]. The hypothesis is to take the radial pressure distribution as a constant.

For further results, exhaustive FE calculations can be performed. Leen [5] and Tjernberg [6] carried out simulations of an involute spline coupling using 3D finite elements to find the pressure distribution on the side of a tooth. The considered load is a torque applied to the side opposite to where the shaft is fixed. Fig. 4 [5] shows that the axial distribution of pressure is not uniform along the side of a tooth. It is significant to note the two peaks at the extremities of the spline.

Adey [7] used a boundary element method to study the axial load transfer mechanisms in involute spline couplings with a helix angle. The results were compared with strain gauge data and photoelastic torsion stresses.

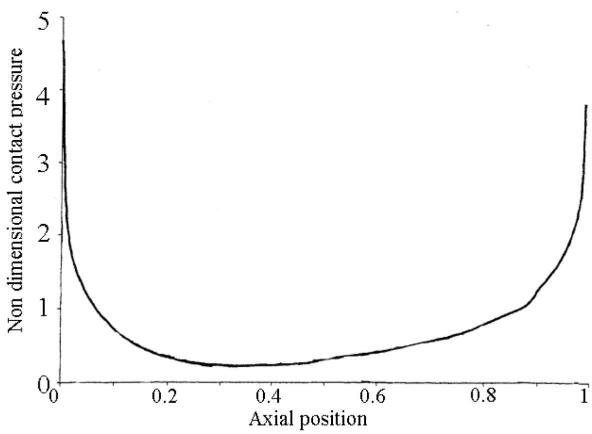


Figure 4. Axial pressure distribution along a tooth (After [7]).

These studies all show that the load is not uniformly distributed along the axial direction of a spline coupling.

To build an assistance tool, a simplified FE model based on Blanc's study [4] has been developed. This paper first presents the assistance tool using a simplified FE model, which can change the point of torque introduction and model load distribution in the axial direction of the spline coupling. Secondly, a three-dimensional finite element model with contact elements, built in order to validate the assistance tool, is described. An experimental study is then presented. Finally, a parametric study is performed and both numerical and experimental results are compared. In this paper a standard straight flank spline coupling is considered.

2. AN ASSISTANCE TOOL BASED ON A SIMPLIFIED FE MODEL

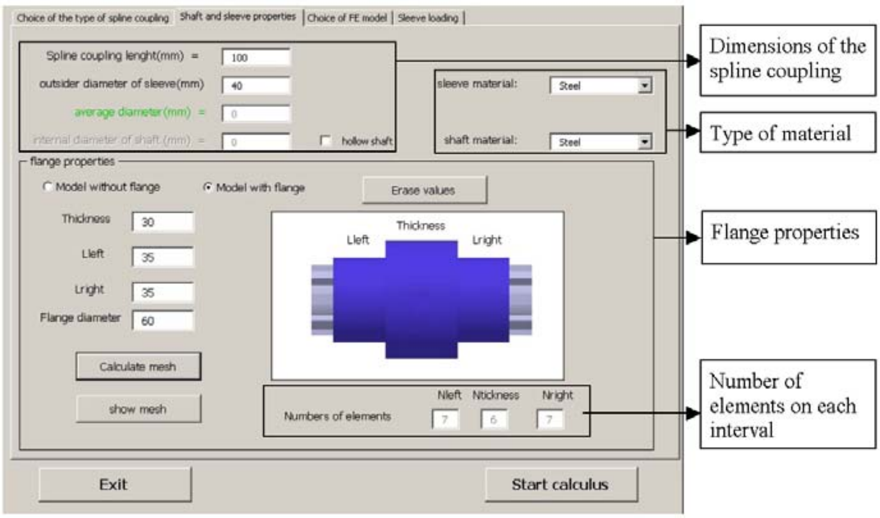


Figure 5. Main tool menu.

The assistance tool is based on the Blanc study. Visual Basic and Excel have been used to develop it. Fig. 5 shows the main menu where the sleeve and shaft characteristics can be defined. This tool takes into account different types of teeth, the sleeve and shaft characteristics (length, diameter, material). The flange geometry can be defined and placed at any axial position on the sleeve, enabling torque to be introduced at any axial position. The spline coupling is divided in the axial direction into twenty intervals. An additional option allows for the distribution of the axial elements on the three parts of the sleeve (see Fig. 5: L_{left}, Thickness, L_{right}). The buttons "calculate mesh" and "show mesh" are used to calculate the element distribution and to show the distribution of the elements.

To prepare our assistance tool, as in the Blanc study, FE analyses have been performed in order to define K_F. To find the associated coefficient for a given spline, a 3D model of part of the spline coupling (1mm width) has been studied with ABAQUS 6.3 (see Fig. 6). The coefficient is calculated for the shaft and the sleeve by using the average node displacements on a line in the thickness of each tooth. The nodes lines used are represented with an oversize in Fig. 6. Each spline coupling considered in the tool has different $K_{F\ Shaft}$ and $K_{F\ Sleeve}$.

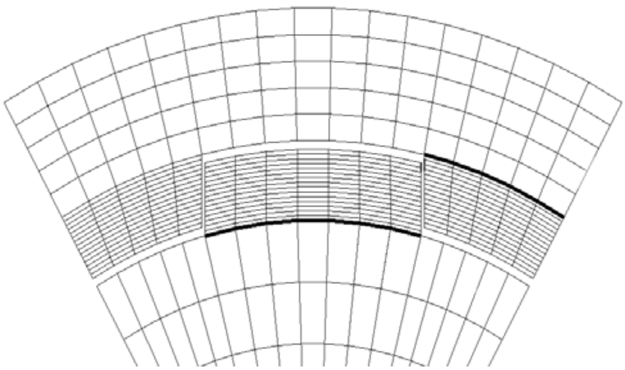


Figure 6. A part of 3D model to determinate the value of K_F.

To be able to validate this tool, results have to be compared with experimental data. Unfortunately, the axial pressure distribution that would be of key interest for designers can not easily be measured experimentally. For that reason, we have decided, as the results of the calculation, to draw the axial torque distribution (no hypothesis on the radial pressure distribution is made) and also the axial strain distribution on the outside sleeve diameter (this strain can be experimentally measured [7]). The sleeve strains are stored at every middle point of the finite elements and calculated as follows:

$$\varepsilon = \left(\frac{\theta_2 - \theta_1}{L}\right) R_o$$

θ_1, θ_2 = angle of twist at the node 1 and 2,
ε = strain at the middle point of the sleeve elements,
L = length of the element,
Ro = outside radius.

3. A 3D FINITE ELEMENT MODEL

Three dimensional finite element models using ABAQUS 6.3 have been made of different spline couplings considered in the parametric study (See Section 5).

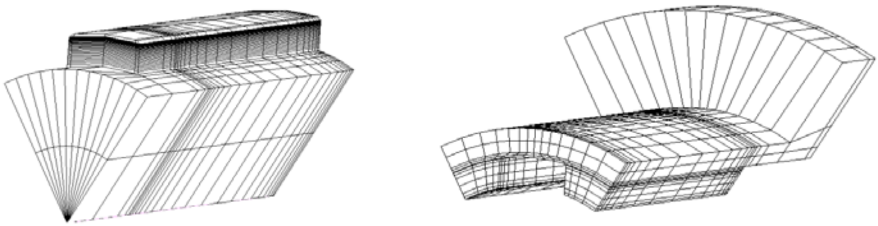


Figure 7. Part of the shaft and sleeve with the associated mesh.

A part of the shaft and sleeve are represented in Fig. 7. The model includes contact on the teeth flanks, and the gap between the contact surfaces is considered. The final model is the combination of copy and joining of those two parts. Fig. 8 gives an idea of the finite element model.

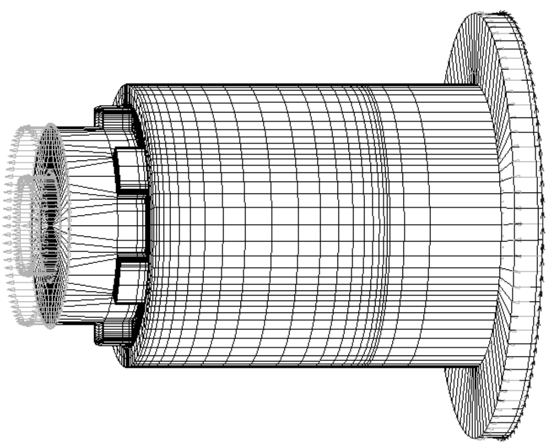


Figure 8. Finite Element Model of a shaft and a sleeve.

The boundary conditions are defined as follows: on the one hand nodes of a shaft's extremities are locked but the radial expansion is allowed, on the other hand nodes of a sleeve's extremities are forbidden to move in the axial translatory motion (see Fig. 8). The torque is applied to the outside edge of the flange.

4. AN EXPERIMENTAL STUDY

An experimental study has been conducted to validate the two presented models and to test two loading configurations.

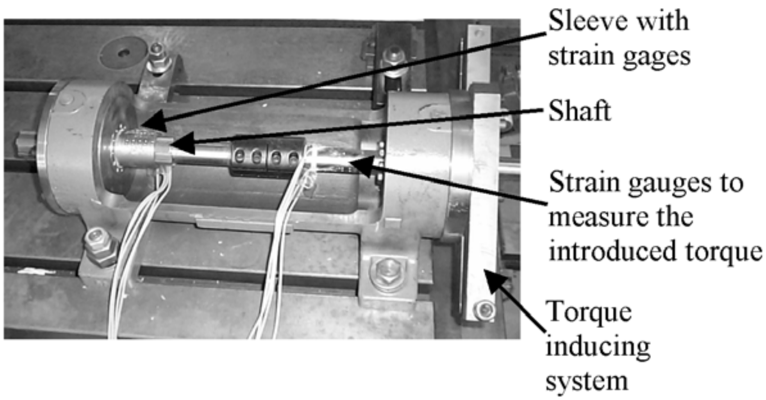


Figure 9. Experimental study.

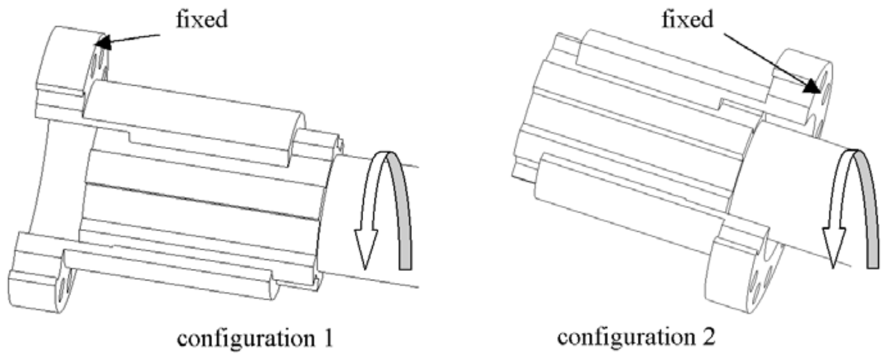


Figure 10. The two loading configurations.

Fig. 10 presents the two loading configurations. Configuration 1 is the loading case where power crosses the assembly. The sleeve is fixed to the side opposite to where the torque is applied. Configuration 2 is the opposite configuration, where the sleeve is fixed on the same side as the applied torque.

5. PARAMETRIC STUDY

5.1 Introduction

In our parametric study, different standard straight flank spline couplings have been considered. As an example, we give below the results related to a coupling that has been tested experimentally. The spline dimensions are described in table 1.

Table 1. Data for a straight flank spline coupling

	Shaft	Sleeve
Number of teeth	6	6
Root Diameter (mm)	27,92	34,15
Outside Diameter (mm)	33,78	
Tooth Thickness (mm)	6,92	7,12
Contact length (mm)	34	34

The tangential stiffness in bending (K_F) values of this spline, used in the assistance tool, are:

$K_{F\ sleeve} = 5{,}62 \times 10^4$ N/mm

$K_{F\ shaft} = 9{,}42 \times 10^4$ N/mm

In order to observe the influence of both inertia and axial loading position, the tests were carried out with two sleeves in the two loading configurations presented in Fig. 10. The sleeve geometry is represented in Fig. 11 with an outside diameter value of 40mm. The outside diameter value of the second sleeve is 50mm. The spline coupling is loaded with a torque of 200Nm. Five strain gauges are stick on the outside diameter of each sleeve, along a longitudinal line above the contact surface (See Fig. 9).

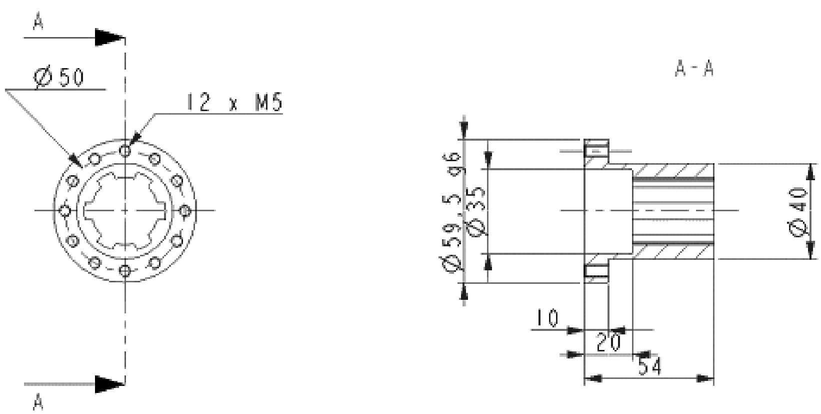


Figure 11. Sleeve geometry (values in mm).

The sleeve geometry (see Fig. 11) can be divided in three parts, a flange, a tubular section and a spline. The tubular section has been included in order to limit the influence of the flange to the spline.

Fig. 12 presents the obtained results. The Z axis represents the axial direction of the coupling. Value 0 represents the beginning of the contact tooth on the side where the torque is applied, and 34 the end of the contact (see Fig. 10).

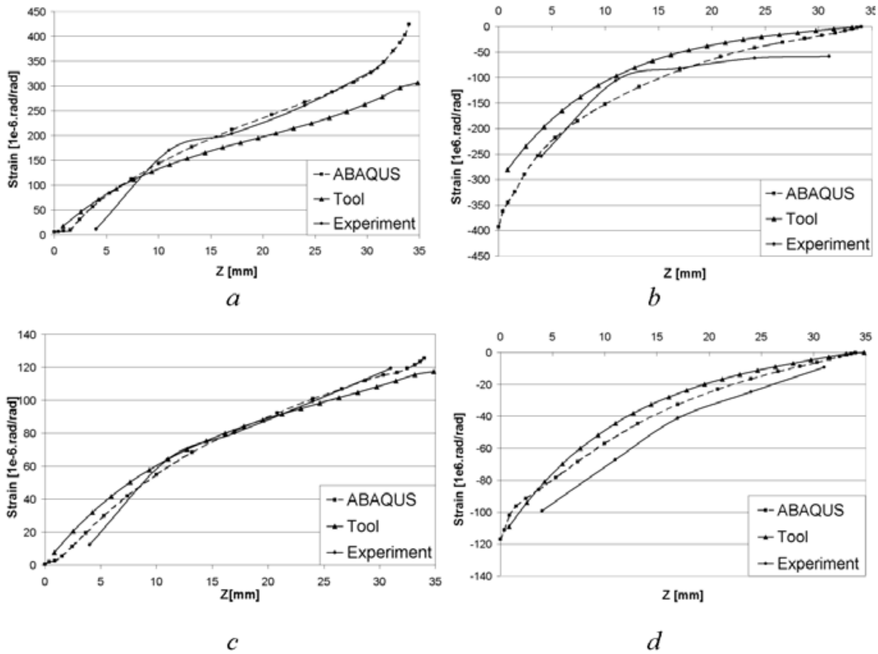


Figure 12. Strain of the sleeve for different outside diameters in two configurations: a = outside diameter 40 and configuration 1, b = outside diameter 40 and configuration 2, c = outside diameter 50 and configuration 1, d = outside diameter 50 and configuration 2.

5.2 Comparison between the experiment and the 3D model

The 3D model provides a satisfactory estimation of the real situation.

When considering the contact pressure as constant, the strain curve should be a straight line. All the curves of Fig. 12 show that it is not the case. Indeed, according to Volfson [1], the inertia has a significant importance in the pressure distribution, and also in the strain distribution. When a gradient on the strain curve appears, at the same point, a gradient on the pressure distribution curve exists too. Thus, a gradient on a curve shows that the associated section is transferring torque.

The next values are the inertias of the shaft and the different sleeves:

$I_{shaft} = 2{,}96 \times 10^{-2}\ mm^4$

$I_{sleeve} = 9{,}03 \times 10^{-2}\ mm^4$ for an outside diameter of 40mm

$I_{sleeve} = 0{,}23\ mm^4$ for an outside diameter of 50mm

The study, which gives the curves of Fig. 12a, is made when power crosses along the spline (configuration 1). As the outside sleeve diameter is 40mm, the I_{shaft} value is almost the I_{sleeve} value. At the beginning and at the end of the contact, great gradients on the strain curves exist, as Volfson predicted. They are highlighted in Fig. 13 in which only the 3D model results are shown.

Fig. 12b represents the same spline coupling but when the power does not cross all the contacts (configuration 2). The curve has only one region of high gradient which is highlighted for the 3D model results in Fig. 13.

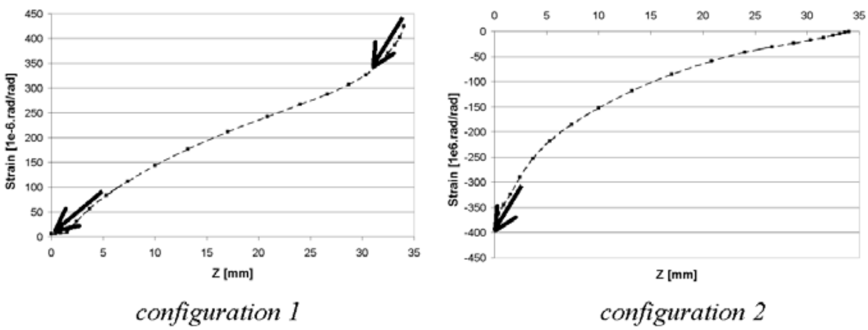


Figure 13. Gradients on the strain curves for the two configurations.

This shows that for the same spline coupling design, the torque is mostly transferred by the two axial extremities of the spline in configuration 1 and by only one extremity in configuration 2. The axial position of the loading is thus a very important factor in the distribution of the strain.

The curves in Fig. 12c and Fig. 12d were obtained for a spline coupling, which has a larger outside diameter for the sleeve (50mm). It induces $I_{shaft} < I_{sleeve}$ thus according to Volfson [1] only one region of high gradient should appear. This is in fact the case, whatever the configuration. Thus, the torque is mainly transferred by only one axial extremity of the coupling.

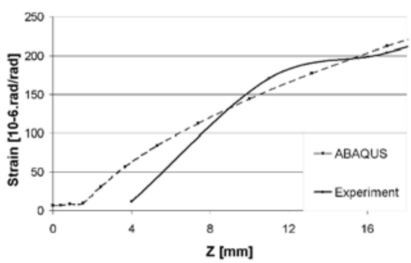


Figure 14. Detail of the sleeve strain for outside diameter 40 mm and configuration 1.

Moreover, Fig. 14, which is an enlarged detail of Fig. 12a shows that between 0 and 1.5mm, at the beginning of the contact, the finite element model curve is horizontal. It is a local phenomenon that also appears on the experimental curve. In fact, the strain value of the experiment at 4mm is almost null. We can explain this because of a local bending on the teeth.

The local bending at the beginning of the contact is shown in Fig. 15.

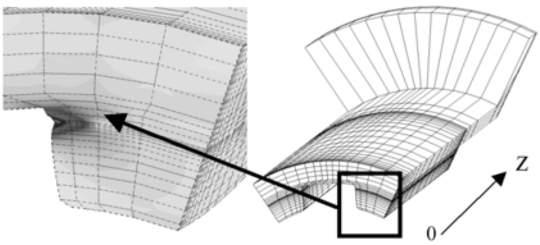


Figure 15. Local bending at the beginning of the contact.

The influence of this local bending on the external strain is shown in Fig. 16. Fig. 16c represents a detail of the beginning of the deformed sleeve (z = 0). It shows that at the beginning of the contact, the first elements of the external radius are similarly deformed and thus induce a local horizontal curve in Fig. 12a.

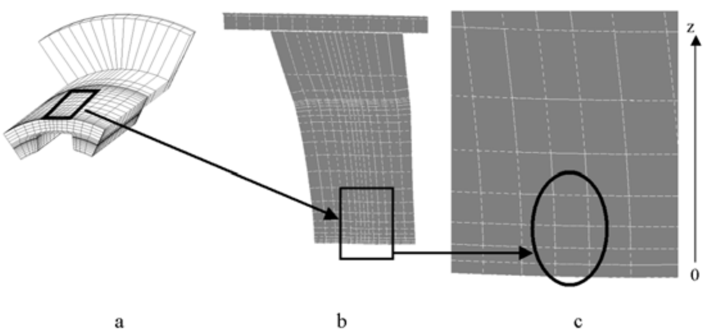


Figure 16. Visualization of the external strain.

All these results show that the 3D model can be considered as correct.

5.3 Comparison between the 3D model and the simplified tool

In Fig. 12a at the end of the contact and Fig. 12b at the beginning of the contact, there is a significant difference between the finite element model results and the tool results. This is mainly due to the influence of the tubular section of the sleeve. In fact, the theoretical strain of the tubular section is 409µrad/rad. The 3D model solution is close to this with a strain value of 425µrad/rad, unlike the tool, which gives a strain value of 300µrad/rad. The origin of this difference comes from the change in sleeve geometry between the tubular section and the spline section. The simplified tool considers each section individually and does not take into account local phenomena due to the change of geometry. This phenomenon does not appear in Fig. 12c and Fig. 12d. Indeed, the change of geometry is less important than in the previous case.

Let us now consider the case where the outside diameter value of the sleeve is 60mm.

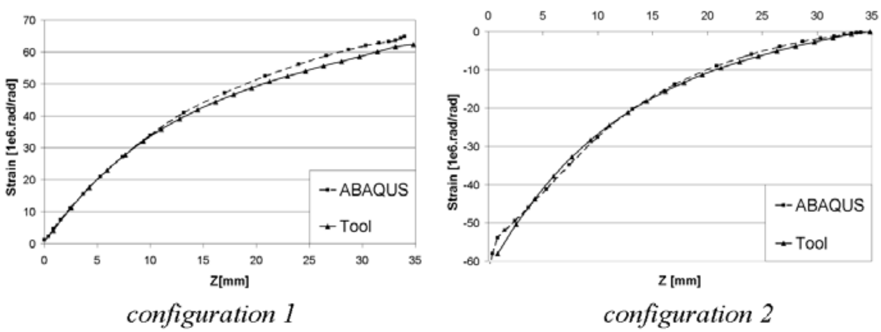


Figure 17. Sleeve strains for outside diameter 60mm for the two configurations.

Once again, the influence of the inertia and the loading introduction is visible: only one peak appears in the two loading cases (see Fig. 17). Thus the assistance tool provides an accurate estimation of the spline coupling behaviour.

6. CONCLUSIONS

To study spline coupling behaviour, an assistance tool based on a simplified FE model has been developed. It takes into account part inertia and the axial position of the applied torque. It quickly determines the axial torque transfer and the strain distribution. This tool requires the calculation of the tangential rigidity in bending of each different spline to be determined by a separated FE analysis.

The assistance tool has been validated by comparing its results with a 3D model and an experimental study. The three-dimensional finite element models, built with ABAQUS, allow different types of loading to be introduced.

The experiments have proved that both models give an accurate estimation of the real situation. The limits of the assistance tool are also highlighted.

The results obtained from the parametric study show that both the inertia of shaft and sleeve and the axial position of the applied torque play a key role in the axial torque transfer distribution. It has been shown that torque is mainly transferred by the two contact axial extremities. Moreover, in some cases, the torque can only be transferred by the contact extremity where the torque is applied on the shaft.

We suggest that future research might focus on the pressure distribution and the sliding at the contact in order to define a fretting corrosion criterion in order to modify and improve the spline coupling design.

REFERENCES

1. B.P. Volfson, Stress Sources and Critical Stress Combinations for Splined Shaft, Journal of Mechanical Design 104, 551-556 (1982).
2. G.K. Tatur, and A.G.Vygonnyi, Irregularity of load distribution along a spline coupling, Russian Engineering Journal XLIX(4), 23-27 (1969).
3. M. Orain, Contribution à l'étude de mécanismes homocinétiques, thèse de docteur ingénieur de l'université Paris III, 260-276 (1976).
4. H. Blanc, T. Kachler, and M.Pons, Un nouveau modèle pour l'étude et le dimensionnement des liaisons par cannelures axiales, 4th World Congress on Gearing and Power Transmission 2, 1217-1227 (1999).
5. S.B. Leen, T.R. Hyde, E. Williams, A.A. Becker, I.R. McColl, T.H. Hyde, and J.W. Taylor, Development of a representative test specimen for frictional contact in spline joint couplings, Journal of strain analysis, Imech 35(6), 521-544 (2000).
6. A. Tjernberg, Load distribution in the axial direction in a spline coupling, Volvo construction equipment components, Engineering Failure Analysis 8, 557-570 (2001).
7. R.A. Adey, J. Baynham, and J.W. Taylor, Development of analysis tools for spline couplings, Proceedings of the Institution of Mechanical Engineers, Imech 214, 347-357 (2000).

PART SURFACES 3D DIGITISING

An approach to accurate characteristic lines measurements

Charyar Mehdi-Souzani[1] and Claire Lartigue[1,2]

Laboratoire Universitaire de Recherche en Production Automatisée

Ecole Normale Supérieure de Cachan, 61 av. du Pt Wilson, 94235 Cachan cedex – France

IUT de Cachan, Université Paris-Sud 11

Abstract: The accurate measurement of characteristic lines using contact-less sensors is an important issue for it conditions the further point exploitation. For applications such as surface reconstruction the characteristic lines represent the 3D boundaries of point sets that can be fitted with surface models. The problem combines two specific issues: evaluating the accuracy of data obtained using non-contact sensors and identifying characteristic lines from discrete data. The first problem is solved through the use of quality indicators that are representative of data density, data completeness, and data accuracy. A particular attention is given to the digitising noise that greatly influences the precision of the acquired points. The second problem is solved through the 2D identification of points that are characteristic of 3D contours. The proposed approach relies on a voxel-space representation of the 3D digitised data that allows both the extraction of voxels belonging to the contour and the evaluation of voxels not satisfying the specified index precision. To improve data accuracy a method for rescanning non-accurate zones is detailed on the calculation of new sensor orientations so that the sensor is as normal as possible to the contour voxel and the digitising noise is limited.

Key words: Laser scanning; 3D data accuracy; Characteristic lines; Voxel representation.

1. INTRODUCTION

From now on, product functional or aesthetic requirements lead to the design of more complex free-form surfaces. The realisation of such surfaces is henceforth made possible with such new manufacturing processes as rapid prototyping. Once the free form is performed, the recurring problem is the measurement of the 3D object surfaces. Indeed, a throrough inspection of the part is recommended for free form surfaces [1].

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 343-356.

With the advances of 3D digitising systems using optical sensors, it is now possible to acquire the 3D object surface as a large set of 3D points, in a relatively short period of time. This set of 3D points, generally called a *large cloud of points*, is a discrete representation of the surface geometry. However, this large cloud of points is dense, inhomogeneous and noisy, which makes it difficult to use it for applications such as surface reconstruction, direct machining or part inspection [2]. Therefore, the evaluation of the point cloud accuracy represents an important issue. Nevertheless, as the expected quality level depends on the point exploitation, the quality of the delivered point cloud must be considered as regards its exploitation. Little work has been done in this direction. Most generally, accuracy is evaluated by the measurement of an artefact, including effects of both the depth of view and the view angle [3]. The influence of the scan planning is thus highlighted, and attention is generally focused on finding the best scan planning so that the surface inspection is complete [1, 4, 5] while optimising the accuracy of the acquired 3D points [6, 7].

Previous works provide keys to the user's evaluation of point cloud quality as regards its application (or exploitation) throughout the definition of indicators. These quality indicators are representative of the sensor accessibility relatively to the part surface, the digitising noise, and the accuracy [8]. They qualify the variable density of data (ρ-dense), the completeness of the point cloud as regards the initial object surface (κ-complete), the accuracy of the data in relation to the digitising noise and its position in the CCD space (δ-noisy and τ-accurate). The latter indicators are specific to the means used: a laser-plane sensor.

Most of the applications rely on the identification of characteristic lines, such as surface contours or styling lines [2]. Indeed, the extraction of surface contours gives the curves limiting subsets of points. This step is generally called point segmentation and simplifies surface reconstruction as well as direct manufacturing. In the first case, surfaces are fitted to point subsets, and in the second case, point subsets represent machining features. Styling lines or character lines are also essential in design, considering that they act as structural lines [9]. Therefore, extracting characteristic lines is an essential problem. Some automatic methods have been developed to extract characteristic lines from discrete data, but due to the inaccuracy of 3D digitising, they generally fail when applied to points issued from a real digitising [10, 11]. The method developed in previous works relies on the extraction of characteristic lines through the identification of characteristic points of 2D discrete profiles [2]. These profiles correspond to the *cutting* of the point cloud by planes. The use of the voxel-space representation of point clouds allows defining *contour voxels*: a contour voxel contains at least one characteristic point. The growing of contour voxels according to a proximity criterion defines an image of one characteristic line [2]. Basically, the characteristic lines are measured with less accuracy than the whole set of points. For instance, measuring an edge involves the sensor to be “as normal as possible to the real edge” in order to measure points really belonging to the edge. Edges play a pivotal part in surface reconstruction for data segmentation as well as for data meshing [12].

The paper deals with an approach to evaluate as accurately as possible, sets of points which are representative of characteristic lines. The 3D digitising system used is a CMM equipped with a plane-laser sensor supported by a Renishaw PH10 head. The approach consists in a first 3D digitising of the object surface with a given sensor orientation. The quality of the point cloud is evaluated through quality indicators [8]. Simultaneously, the identification of the 3D characteristic lines using the voxel-space representation from the discrete data is performed [2]. Using the voxel-space representation, we can compare the quality indicators associated to the characteristic lines to given thresholds. If necessary, an automatic calculation of new sensor orientations is carried out defining new digitising allowing the expected quality to be reached.

This approach imposes to deal with the whole process of 3D digitising. Therefore, the paper is organised as follows: Section 1 is dedicated to the presentation of our experimental measuring system, and details the step of sensor calibration for a given orientation [13]. Then, in section 2, the steps “analysis of data accuracy” and “identification of the characteristic lines are explained. Finally, section 3 summarises the step of generating “a new scanning process” leading to accuracy improvement. The paper ends with some conclusions.

2. 3D DIGITISING OF FREE FORM SURFACES

3D digitising of an object surface consists in defining a digital surface representation through sets of 3D points. The 3D digitising system used is a CMM equipped with a laser-plane sensor, mounted on a motorized indexing head PH10 from Renishaw (http://www.renishaw.com). Such a configuration, seldom used as it might be, enables the sensor to be oriented to repeatable positions increasing its accessibility space. The laser-plane sensor (KLS51 from Kréon Technologies; http://www.kreon3D.com) consists of both a transmitter and a receiver.

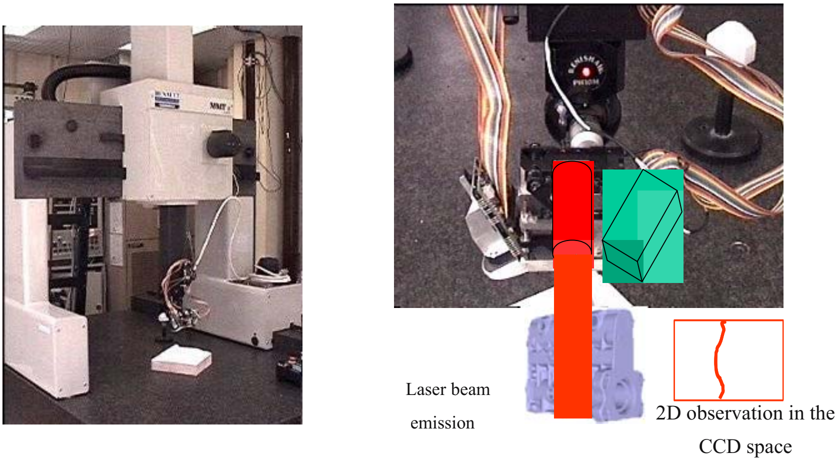


Figure 1. Measuring system

The transmitter sends a thin laser beam whose width allows to cover a large part of the surface object; the receiver, a CCD (Charged Couple Device) camera, visualizes the plane under incidence, with a given fixed triangulation angle. The observation is thus carried out in the 2D space of the CCD camera (Fig. 1). Therefore, the 3D digitising requires a calibration step in order to set up the relationship between the 2D data acquired and the 3D coordinates of points belonging to the object surface (Fig. 2). Basically, current calibration processes are split up into two stages, one for the geometrical parameters of the camera, and the other one for the sensor orientation. Moreover, orientations are generally limited to 4 indexed orientations.

Due to the specific measuring system, a calibration process has been developed which leads to define a global transfer function[13] which expresses the link between the 2D coordinates, N(R, C), (R: row, C: column), in the CCD space, towards the 3D ones M(X, Y, Z) :

$$\begin{cases} X = X_s + \dfrac{a_1 R + a_2 C + a_3}{a_{10} R + a_{11} C + 1} \\ Y = Y_s + \dfrac{a_4 R + a_5 C + a_6}{a_{10} R + a_{11} C + 1} \\ Z = Z_s + \dfrac{a_7 R + a_8 C + a_9}{a_{10} R + a_{11} C + 1} \end{cases} \qquad (1)$$

(X_s, Y_s, Z_s) are the coordinates picked up during measuring and correspond to the sensor location within the CMM space. The transfer function brings out 11 parameters that are identified by the measurement of a specific artefact, the *facet* sphere [13]. With such a method, both the sensor geometrical parameters calibration and the sensor orientation calibration are performed through the same identification process.

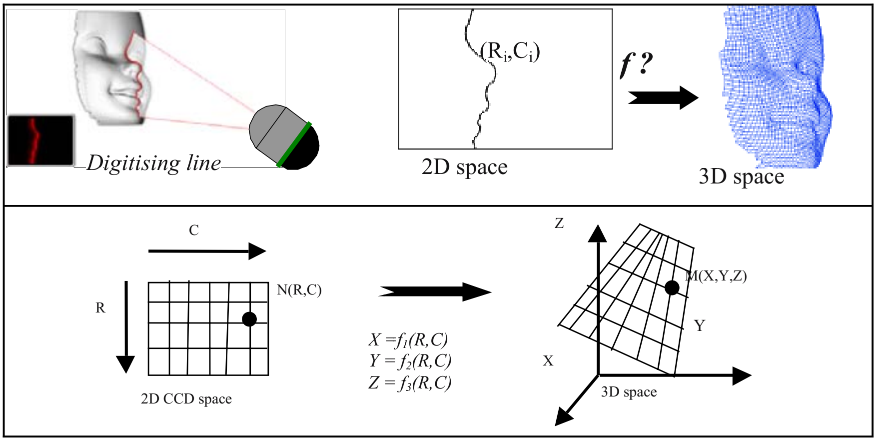


Figure 2. Calibration: expression of the 3D coordinates data in function of (R,C).

Nevertheless, due to optical phenomena, some of the measuring system parameters may affect measurements. In particular, the location in the 2D CCD space combined with digitising noise influences data accuracy. It is thus possible to estimate the variability of the identified parameters that impacts the global transfer function and leads to a dispersion on the location of 3D points [13]. However, this calibration process is robust and allows to calibrate the whole range of possible orientations of the laser-plane sensor, which is particularly convenient for new scanning steps.

3. ACCURATE MEASUREMENTS OF SPECIFIC SETS OF POINTS

The point cloud resulting from the 3D digitising using non-contact sensors is clearly dense, inhomogeneous, noisy, featuring gaps, and less accurate than when obtained using usual measuring techniques. Nevertheless, the 3D points obtained must be exploited for applications such as surface reconstruction, free form copying, or inspection, which means that the quality must be in line with the application. To solve this problem, quality indicators have been defined that allow the analysis of data accuracy [8]. These indicators have been detailed in §3.1. The use of a voxel-space representation allows a local representation of the indicators. In the same way, the identification of characteristic lines is essential for applications, since they generally represent the first step of point exploitation, whatever its nature. The method we have developed takes advantage of the voxel-space representation of the point cloud. In particular, characteristic lines are built from contour voxels. The latter are identified using the extraction of characteristic points in 2D. The method is explained in §3.2.

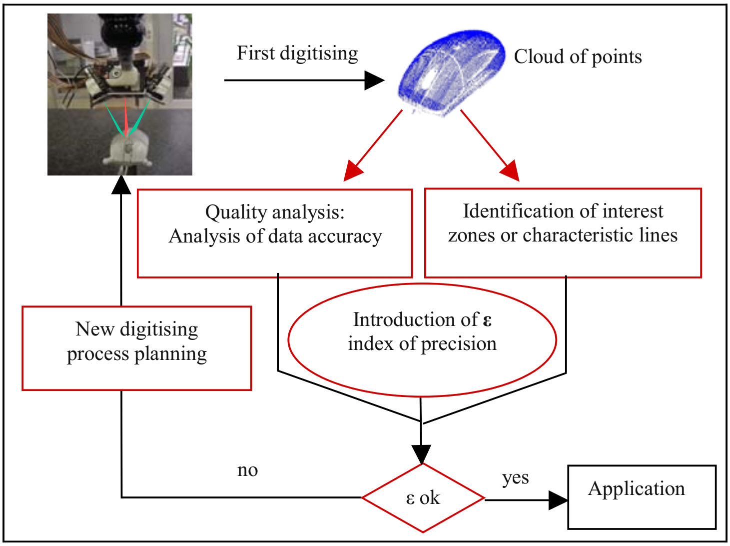


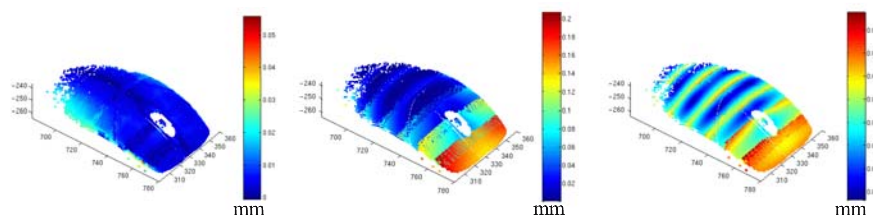
Figure 3. Description of the approach for accurate measurements

However, the evaluation of the indicators for acquired sets of points only gives an indication of the quality, but does not provide accurate measurements. In parallel, the identification of characteristic lines does not give information about the accuracy of such lines. To answer the problem of accurate measurements of characteristic lines, we propose to compare an index of precision ε, which corresponds to one of the quality indicators, to a threshold ε_{max}, only for measured points corresponding to characteristic lines. From the voxel representation, contour voxels corresponding to the characteristic lines are identified. On the other hand, from the whole set of voxel contours, those meeting the criterion $\varepsilon<\varepsilon_{max}$ are separated from those that do not. Following, the rescan of these particular zones can be carried out. Obviously, the new scanning process must incorporate the quality criterion to lead to more precise zones (Fig. 3).

3.1 Analysis of data accuracy

As previously explained, the conformity of the 3D digitised data is evaluated through quality indicators: δ-noisy, ρ-dense, κ-complete and τ-accurate, as suggested by Contri [8]. The first indicator is linked to data sampling errors. It is generally evaluated considering the deviations between the points and a geometric model fitted to the points. The density indicator is directly linked to the sampling density, and strongly depends on the scan planning and the digitising system used. The completeness figures out the importance of the gaps existing within the point cloud. The accuracy relates to the common notion of measurement uncertainty applied for a 3D point resulting of non-contact measuring. Methods have been developed to evaluate all these indicators [14].

An illustration is given in Fig. 4, which presents maps of indicators for a set of measured points obtained but with a unique sensor orientation. Fig. 4b and 4c are accuracy maps associated to the points. τ_1 corresponds to the dispersion of the 3D location linked to the variability of the parameters identified during the calibration step. Note that for some zones, the dispersion can reach 0.2 mm. τ_2 is the dispersion linked to the anisotropic behaviour of the CCD space. Indeed, a same 3D point can be observed in the 2D space through various location (R,C). The closer to the CCD centre the position of 2D picture (R,C), the better the 3D point accuracy.



δ-noise linked to the view angle — τ_1-accuracy linked to the calibration — τ_2-accuracy linked to the CCD

Figure 4. Quality indicators applied to measured points

Fig. 4a shows the evolution of the digitising noise δ in function of the view angle α. δ is evaluated as kσ, where σ is the standard deviation of the points relative to the geometrical model. To limit the values of δ, k is chosen equal to 1. As it is well known [1, 3, 6], the view angle α, influences the quality of acquired data. This influence has been experimented using a test part. The noise δ can be described as a polynomial function of the incidence angle. In particular, with the measuring system used, a “forbidden” zone is brought out: if $\alpha > 60°$, δ becomes greater than 0.02 mm (Fig. 5).

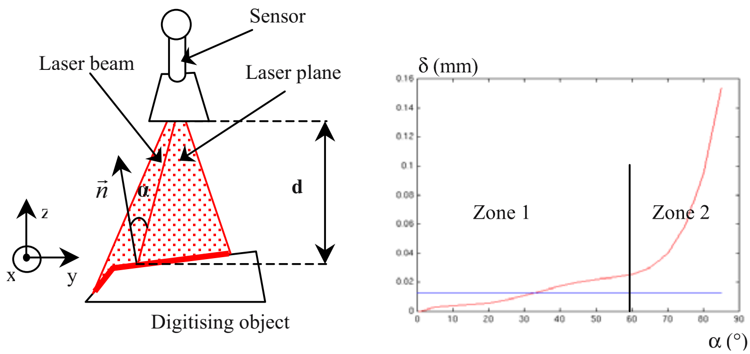
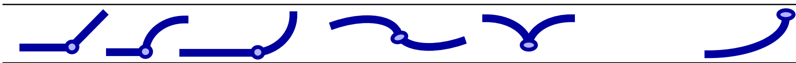


Figure 5. Digitising noise vs. view angle

It is thus recommended to choose a sensor orientation so that the view angle α is as small as possible to limit the digitising noise. To obtain the map presented in Fig. 4a the normal to the digitised surface must be computed at each point in order to calculate the incidence angle. The calculation is conducted considering a local Delaunay triangulation at the neighbourhood of each point. The use of a voxel map of the point cloud is also possible. For the orientation under consideration, the digitised noise is limited to 0.02 mm except for a few outliers. In parallel, the identification of the characteristic lines from the discrete data is carried out.

3.2 Identification of the characteristic lines

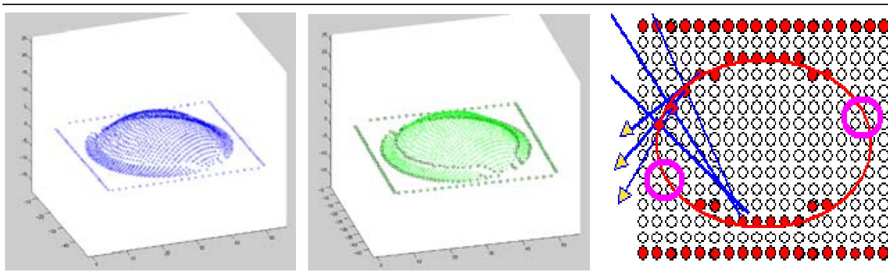
The method used is the method developed by Osty [2] that relies on the detection of characteristic points marking geometrical singularities. Basically, this approach originates from an observation of the *continuous* world. When a plane cuts an object, a point of the boundary can be identified by its singularities within the 2D profile: change in curvature, sharp angle, and inflection point. Therefore, a wise sweeping of the data by cutting planes allows to determine points that are characteristic of the 3D contours. This observation can be extended to the *discrete* world, that means to digitised data. The first step is the definition of the characteristic points. Six types of characteristic points are defined that permit to identify angular points, inflection points and extremities from the 2D profile. They represent limits between linear regions and regions featuring a same convexity (Fig. 6). The algorithm used to extract the points from discrete data is the directional coding [15]. It can be noticed that extremity points define characteristic points. Therefore, the boundary of a hole is a characteristic line.



Angular points — Inflection points — Extremity

Figure 6. Characteristic points

As far as the sweeping step is concerned, the simplest remains the sweeping of the data by parallel planes to a given direction. Nevertheless, points so determined only represent a basis for the extraction of characteristic lines. Indeed, experience shows that extracted characteristic points are more significant when the cutting plane is normal to the contour (characteristic line) [15]. Therefore, according to a proximity criterion, characteristic points are gathered to define the *initialisation line*: a polynomial cubic curve fits the points according to the least-square criterion. The tangent to the curve in construction (or cutting direction) defines the normal to the cutting plane in which the extraction of characteristic points can be performed (Fig. 7). The closure of the characteristic lines gives the 3D contour.



Digitised points — Characteristic points — Identification of the characteristic lines

Figure 7. Identification of the characteristic lines

However, the identification of the contour is only effective if the neighbourhood of a point can be defined. This leads us to base our method on a voxel-space representation of the 3D digitised data. A voxel map is defined on points. The size of a voxel is defined so that it contains at least 20 points. Then, we classify each voxel in function of the point set that is included within its space, a voxel that contains a portion of a characteristic line is called a *contour voxel*. To each contour voxel, we associate the attributes *normal* and *cutting direction*. The construction of the characteristic lines is performed by the union of contour voxels according to the proximity criterion and the concordance of attributes (Fig. 8).

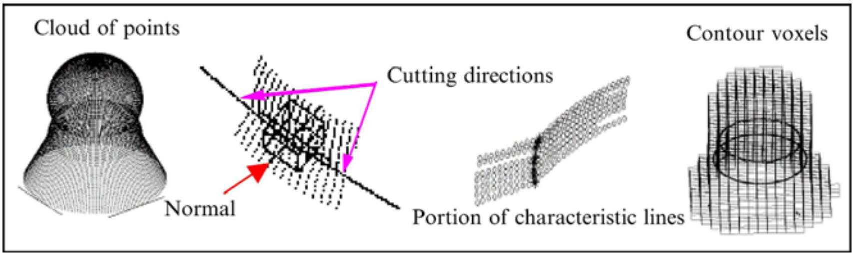


Figure 8. Illustration of the extraction of contour voxels

3.3 Synthesis

The following step is the analysis of the precision (according to the index ε) of the characteristic lines. This is simply carried out using the voxel map. Indeed, we check only for the contour voxels if the value of the index is less than the given threshold. This is illustrated in Fig. 9. Suppose that ε_{max} = 0.02 mm. On the left, the map is the voxel representation of the digitising noise for points obtained with 4 different views. The scale of colours shows that for some voxels, the noise may vary from 0.02 to 0.08 mm. For these points, the precision index ε exceeds the value authorised. On the right handside, pink voxels mark voxel contours. As the precision of the pink voxels is greater than the authorised one, a new scanning process must be defined so that ε is decreased.

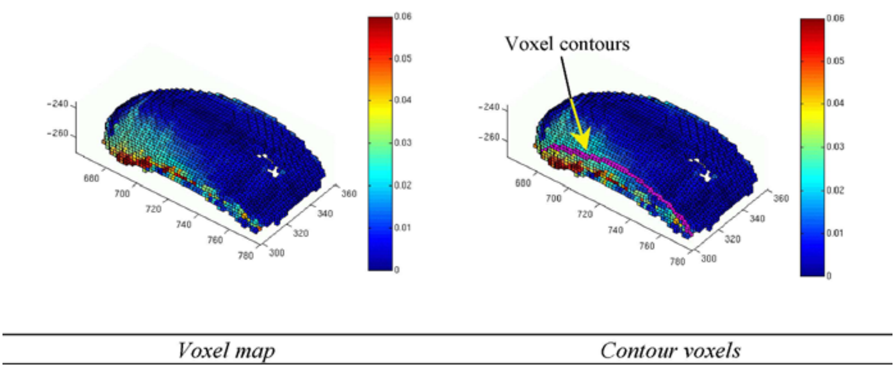


Figure 9. Evaluation of the digitising noise of contour voxels

Another illustration concerns hole boundaries. In this case, characteristic lines are representative of limits between both visible and hidden zones. The point cloud, which is studied for the characteristic point identification, results from a first digitising process for which there is no information whatsoever about the surface contours. Even if the sensor orientation is defined according to the normal of the part surfaces, the points belonging to the contours can be inaccurate as normals abruptly vary near the surface boundaries. As an example, in Fig. 10 it is not possible to precisely define the exact boundary between zone 1 and zone 2. Therefore, the characteristic line initially identified represents "the interest zone characterised by the presence of one characteristic line".

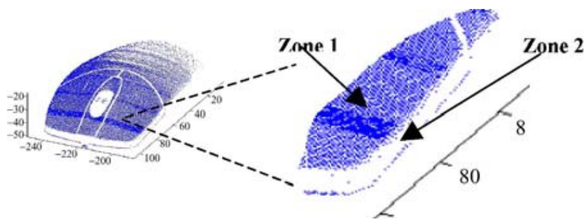


Figure 10. Point cloud after a first digitising

4. NEW SCANNING PROCESS

For the new scanning process various methods can be implemented. Some can be specific to the quality indicator. For instance, Contri [14] suggests to rescan portions for which τ_2 (dispersion linked to the anisotropy of the CCD matrix) is less than the given threshold calculating the new scan planning so that the barycentre of the digitised lines are centred in the CCD space. This is well adapted for the τ_2 accuracy that depends on the location within the CCD space of the point. observed Some authors consider that one of the most significant factors to improve the accuracy of 3D acquired data is the view angle [3, 6, 7]. The scanning process is calculated so that the number of views is minimal while optimising the orientation of the sensor relatively to the surface object. They generally base their algorithms on a CAD model of the object surface from which they calculate the local normals.

In our work, the first digitising plays the role of the CAD model. Indeed, a voxel map is defined from the point clouds, and as previously stated, it is thus possible to calculate local normals. Considering the specific case of the rescan of characteristic lines, the local normal is the normal to the voxel contour. Therefore, various possibilities are available to define the new scanning method. The simplest one is to use the polynomial cubic curve in construction. The parallel curve by a distance, which is in accordance with the maximum depth specified for the measuring means, can define the trajectory of the sensor. However, this imposes that the parallel curve can be calculated for all points, and does not present any self-intersection. This may give rise to a problem for concave curves.

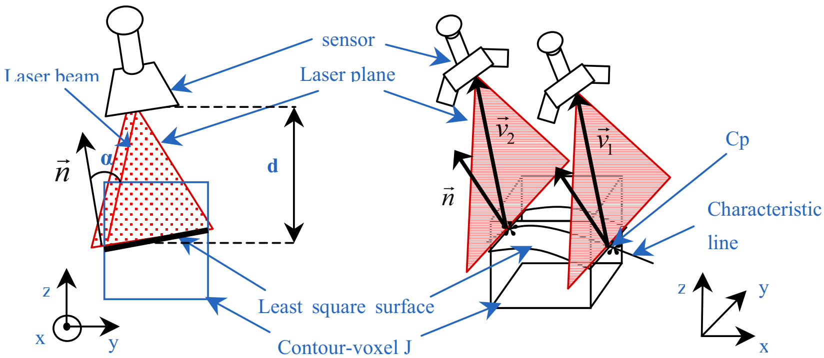


Figure 11. Orientation of the sensor according to the normal to the contour voxel

Another solution consists in using only the normal to the contour voxel. The sensor orientation is defined considering the position vectors $\vec{v}_1$ and $\vec{v}_2$ within the CCM space. These positions are calculated from the extremities of the characteristic lines, and incorporating the recommended depth of view, d (80mm for our measuring system). For each position we calculate:

$$Ip_j = \frac{\vec{v}_j \cdot \vec{n}}{\left|\vec{v}_j\right|} \quad (2)$$

Where $\vec{n}$ is the normal to the voxel. $\vec{n}$ is calculated considering the set of points belonging to the voxel space, and corresponds to the normal of the best-fit least-square plane. To maintain the view angle within the prescribed interval ($\alpha \in [-30°;30°]$), we impose that:

$$Ip_{vox} = \frac{Ip_1 + Ip_2}{2} > 0.85 \quad (3)$$

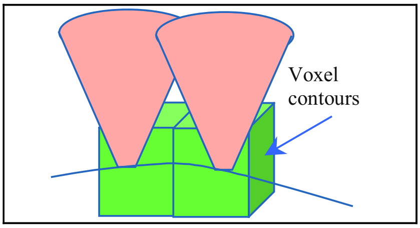


Figure 12. Space interval for the sensor's orientation

However, the initial digitising does not deliver accurate points. Moreover, for points close to 3D contours, the normal may vary abruptly. Therefore, it is necessary not to consider only the contour voxel for Ip_{vox} calculation, but to consider a mean value integrating the 18-neighbours of the voxel. Thus, a system of inequations, given by a set of Eq. (3) is set up, where the only undetermined variables are the sensor's position ($\vec{v}_j$). As it is shown in Fig. 12, the system solution is a cone shape space interval .

5. CONCLUSION

In this paper, the problem of accurately measuring sets of points that are representative of characteristic lines has been handled. This problem can be divided into two specific sub-problems. The first one concerns the quality analysis of data obtained using non-contact sensors. In order to cope with this issue, we analyse the quality of digitised points through indicators that are representative of some characteristics of the point cloud: noisy data, gaps in the digitising, variable densities, … The synthesis of the indicators is conducted using a voxel representation. The second problem relates to the identification of characteristic lines resulting from discrete data. It is solved by the extraction of contour voxels that contain 3D contour characteristic points.

The approach developed hereby relies on a first digitising that is the support of accurate measurements. An authorised precision is given in order to separate voxels meeting the precision requirements for those that do not. The evaluation of the precision of the contour voxels is performed using the indicators. A map of colours highlights voxels that do not meet the criterion. A new scan of these zones is planned taking advantage of the voxel representation, in particular for the calculation of the sensor orientation. The orientation vector of the sensor must be as close as possible to the normal at the contour voxel in order to limit digitising noise. Further work will concern the implementation and the assessment of the approach proposed on the experimental site.

REFERENCES

1. K. Lee, H-P. Park, Automated inspection planning of free form shape parts by laser scanning, Robotics and Computer Integrated Manufacturing, N°16, 2000, pp. 201-210.
2. G. Osty, Extraction de particularités sur données discrètes issues de numérisation 3D: Partitionnement de grands nuages de points, PhD Thesis, ENS de Cachan, January 2001.
3. H.Y. Feng, Y. Liu, F. Xi, Analysis of digitizing errors of a laser scanning system, Precision Engineering, N°25, 2001, pp. 185-191.
4. A. Bernard, M. Véron, Visibility theory applied to automatic control of 3D complex parts using plane laser sensors, Annals of the CIRP, N°49(1), 2000, pp. 113-118.
5. S. Rémy, A. Bernard, G. Ris, Numérisation automatique basée sur le concept de visibilité, IDMME'2002, Clermont-Ferrand(France), May 14-16 2002 (CD-Rom paper).
6. F. Prieto, T. Redarce, P. Boulanger, R. Lepage, Accuracy Improvement of contactless sensor for dimensional inspection of industrial parts, International Journal of CAD/CAM - Rapid Production, N°15, 2000, pp. 345-366.
7. N.A. Massios, R.B. Fischer, A best next view selection algorithm incorporating a quality criterion, 9th British Machine Vison Conference (BMVC'98), 1998, pp. 780-789
8. C. Lartigue, A. Contri, P. Bourdet, Digitised point quality in relation with point exploitation, Measurement, N°32, 2002, pp. 193-203.
9. J-P. Pernot, S. Guillet, J-C. Léon, F. Giannini, C-E. Catalano, B. Falcidieno, Design and styling features using surface deformation, IDMME 2002, Clermont-Ferrand (France), 14-16 may 2002.
10. A. Meyer, P. Martin, V. Guige, A method to divide digitized complex parts in elementary surfaces, IDMME 2002, Clermont-Ferrand (France), 14-16 may 2002.
11. S. Azernikov, A. Fischer, Surface Reconstruction of free form objects based on multiresolution volumetric method for distributed CAD, CIRP Design Seminar 2003, Grenoble(France), 12-14 may 2003 (CD-Rom paper).
12. H.W. Lin, C.L. Tai, G.J. Wang, A mesh reconstruction algorithm driven by an intrinsic property of a point cloud, Computer-Aided Design, Volume 36, N°1, 2004, pp 1-9.
13. C. Lartigue, P. Bourdet, L. Mathieu, C. Mehdi-Souzani, Algorithms for the calibration of laser plane sensors on CMMs, Applied Mathematical and Computational Tools in Metrology, Torino(Italy), 8-11 September 2003.
14. A. Contri, Qualité géométrique de la mesure de surfaces complexes par moyens optiques, PhD thesis, ENS de Cachan, November 2002.
15. G. Osty, C. Lartigue, Partition of 3D digitised data for CAD modelling, International Journal of CAD/CAM and computers graphics, N°13(4-5-6), 1998, pp. 263-272.

SIMULATION AND VALIDATION OF STATIC AND DYNAMIC MACHINE TOOL MEASUREMENTS

Mahbubur Rahman, Jouko Heikkala, and Kauko Lappalainen
Production Technology Laboratory, Department of Mechanical Engineering, University of Oulu, P.0. Box 4200, FIN-90014 Oulun yliopisto, Finland,

Abstract : Machine tools' accuracy level is an important factor for producing accurate dimensions in workpieces. To be sure about the suitable accuracy level, machine tool users use different kinds of measuring methods. There are two kind of measuring systems for machine tool measurements: static and dynamic measurement systems. There is a clear relationship between these measurements. In this work, a theoretical and practical relation has been established between static and dynamic measuring systems. These relations are important whenever machine tools are measured using different measuring devices to validate the measurement results. In this work the tool-tip trace obtained by one measuring system has been compared and simulated with the trace obtained by other methods. A number of systematic mathematical models have been developed and compared with the results obtained by alternative measuring methods.

Key words: Multi-axis machine tools, Accuracy, Error origins and measurements

1. INTRODUCTION

There are mainly two types of machine tool measurement systems: static and dynamic. Static measurement is such that a position deviation is measured along the linear axes using a comparator system, permitting conclusions exclusively on the machine's geometric accuracy. Machine tool conventional inspection and acceptance testing have been limited essentially

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 357-366.

to static measurement of the geometrical machine structure without load. Normally one axis is measured at a time.

Laser interferometers, linear comparators, inclinometers etc. are examples of devices for static measurements. Dynamic measurements, especially at high traverse speeds, provide information on contouring behavior that permits conclusions both on the conditions of the machine tool and on the parameter settings of the control loop consisting of CNC control, drives and position feedback systems.

Machining accuracy depends both on dynamic deviations from the nominal contour and on high acceleration in the machine tool. The Double Ball Bar (DBB) and Cross-grid Encoder (KGM) are examples of devices for dynamic measurements.

Measuring the NC machine tools motion error via existing measurement devices is difficult and needs long experience to understand the measurement results. NC machine tools experts use several methods [static or dynamic measuring systems] to measure the NC machine tools. All these methods are complex in their nature. There are currently no methods to compare the measurement results with each other. There should be a suitable method to find relationships among measurement systems to validate the measurement results. In this work, several mathematical algorithms have been developed for several measuring systems to compare and validate the machine tools motion error measurement results.

2. MACHINE TOOL ERROR MODELING

2.1 Ideal and real machine tool

Most machine tools are designed with the view that all of the joints will be either prismatic or rotary (for more than 3-axis machine tools) but it is physically impossible to construct a joint that will perfectly generate this type of motion. This kind of error always exists but due to servo tuning and calibration process these errors can be taken into consideration in CNC controllers.

In three-axis NC machine tools, if there are three linear axes, the tool tip position is expressed as a distance travelled from the reference point. For an ideal machine we get the tool tip point and work piece point via equations 1 and 2 [1].

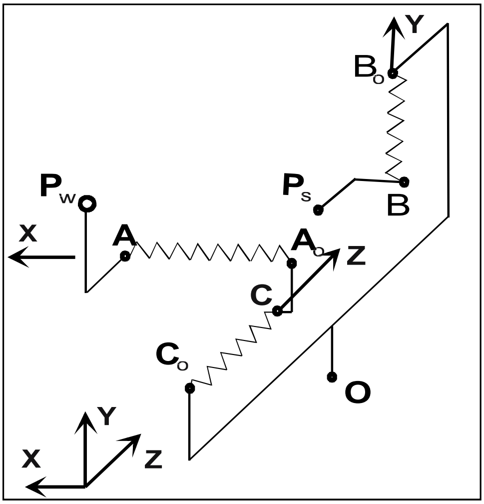


Figure 1. A three-axis machine tool.

$$OP_{wideal} = \begin{bmatrix} 1 & 0 & 0 & AP_{wx} + X + CA_{ox} + OC_{ox} \\ 0 & 1 & 0 & AP_{wy} + CA_{oy} + OC_{oy} \\ 0 & 0 & 1 & AP_{wz} + CA_{oz} + Z + OC_{oz} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

$$OP_{sideal} = \begin{bmatrix} 1 & 0 & 0 & BP_{sx} + OB_{ox} \\ 0 & 1 & 0 & BP_{sy} + Y + OB_{oy} \\ 0 & 0 & 1 & BP_{sz} + OB_{oz} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$$OP_{sideal} = OP_{wideal} \cdot T_{ideal} \quad (3)$$

$$T_{ideal} = \begin{bmatrix} 1 & 0 & 0 & -X \\ 0 & 1 & 0 & Y \\ 0 & 0 & 1 & -Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

Equation 4 indicates that tool tip can be calculated on the basis of the ideal X, Y and Z controls for an ideal machine [1]. If we consider the roll, pitch and yaw errors (RPY) of an axis we can write:

$$T_{Co}^{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -a_z & 0 \\ 0 & a_z & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & b_z & 0 \\ 0 & 1 & 0 & 0 \\ -b_z & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -c_z & 0 & 0 \\ c_z & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & \Delta X_{zx} \\ 0 & 1 & 0 & \Delta Y_{zy} \\ 0 & 0 & 1 & \Delta Z_{zz} + Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

Simplifying equation 5 we get:

$$\mathrm{T}_{Co}^{C} = \begin{bmatrix} 1 & -c_z & b_z & \Delta X_{zx} \\ c_z & 1 & -a_z & \Delta Y_{zy} \\ -b_z & a_z & 1 & \Delta Z_{zz} + Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

We can find T_{Ao}^{A} and T_{Bo}^{B} based on equation 6, which will represent the real coordinate for each axis.

The volumetric error can be defined between the differences of real tool tip position and the ideal tool tip position in the working space of the machine [2]. The actual tool tip position is given by:

$$\mathrm{OP}_{sreal} = \mathrm{OP}_{wreal} \cdot \mathrm{T}_{real} \quad (7)$$

The real tool tip can be found by integrating the transformation matrix into equation 7

Volumetric error can be expressed by equation 8.

$$\mathrm{E}_{vol} = \mathrm{T}_{real} - \mathrm{T}_{ideal} \quad (8)$$

By expanding equation 7, equations 9 and 10 are obtained:

$$\mathrm{OP}_{wreal} = \begin{bmatrix} AP_{wx} - AP_{wyx}c_z - AP_{wyz}c_z + AP_{wzx}b_z + AP_{wzz}b_z + \Delta X_{xx} + X + CA_{ox} - c_z CA_{oy} + b_z CA_{oz} + \Delta X_{zx} + OC_{ox} \\ AP_{wxz}c_z + AP_{wxx}c_z + AP_{wy} - AP_{wzx}a_z - AP_{wzz}a_z + c_z X + \Delta Y_{xy} + Xk_{xy} + c_z CA_{ox} + CA_{oy} - a_z CA_{oz} + \Delta Y_{zy} + OC_{oy} \\ -AP_{wxz}b_z - AP_{wxx}b_z + AP_{wyz}a_z + AP_{wyx}a_z + AP_{wz} - b_z X + \Delta Z_{xz} + Xk_{xz} - b_z CA_{ox} + a_z CA_{oy} + CA_{oz} + \Delta Z_{zz} + Z + OC_{oz} \end{bmatrix} \quad (9)$$

$$OP_{sreal} = \begin{bmatrix} BP_{sx} - c_y BP_{sy} + b_y BP_{sz} + \Delta X_{yx} + OB_{ox} \\ c_y BP_{sx} + BP_{sy} - a_y BP_{sz} + \Delta Y_{yy} + Y + OB_{oy} \\ - b_y BP_{sx} + a_y BP_{sy} + BP_{sz} + \Delta Z_{yz} + k_{yz} Y + OB_{oz} \end{bmatrix} \quad (10)$$

Volumetric error can be found by using equations 8 and neglecting those terms that do not directly affect the error.

$$E_{volX} = (-BP_{SZ} + Z) \cdot b_x + (Y - CA_{OY}) \cdot c_x + (-BP_{sz} + Z) \cdot b_Z + \Delta X_{yx} - \Delta X_{zx} + b_y BP_{sz} + c_z Y - \Delta X_{xx}$$
$$E_{volY} = (a_z - a_y + a_x) \cdot BP_{SZ} + (c_y - c_x - c_z) \cdot BP_{SX} + c_x (X + CA_{OX}) - Xk_{xy} - a_x Z + \Delta Y_{yy} - \Delta Y_{zy} - \Delta Y_{xy} - a_z Z$$
$$E_{volZ} = (b_x - b_y + b_z) BP_{sx} + (CA_{oy} - Y) a_x - \Delta Z_{zz} + \Delta Z_{yz} - \Delta Z_{xz} - b_x \cdot (X + CA_{ox}) + Yk_{yz} - Xk_{xz} - Ya_z$$

(11) (12) (13)

2.1.1 Error modeling for X-axis via laser interferometer measurement

Z and Y-axes are kept in fixed locations while X-axis moves from A_o to A. Positioning error for P_w is recorded with suitable intervals. Positioning error at point P_w includes the effects of RPY, any thermal expansion, elastic axis deformation, measuring system error ,etc. We define that the positioning error at P_w is the total positioning error and the positioning error at point A [a point on the ball screw center line] is defined as basic positioning error.

Total positioning error = Roll, Pitch and Yaw Effects + Basic positioning error (14)

When we measure the positioning error of P_w we actually measure the total positioning error of X-axis. The ΔX_{xx} in equation 18 is the basic positioning error.

The X component of equation 14 gives the total position error.

$$\Delta X_{totalxx} = -AP_{wy} c_x + AP_{wz} b_x + \Delta X_{xx} \quad (15)$$

where $\Delta X_{totalxx}$, ΔX_{xx}, c_x and b_x are total positioning error, basic positioning error, pitch and yaw error at a location A for X-axis. From equation 15 we can easily find the basic positioning error. When we measure the positioning error with laser interferometer we actually measure $\Delta X_{totalxx}$.

2.2 Error modeling for linear comparator measurement

In the Heidenhain VM182 linear comparator [3] the scanning head moves over the graduation scale without mechanical contact. When the scanning head moves along the rail, it reads the linear positioning error as well as the positioning error to the perpendicular direction. For example, for Y-axis movement it can measure the total positioning ΔY_{TotalY} error and the transverse error such as ΔY_{yx} and ΔY_{yz}. The linear comparator can measure but one transverse error at a time.

2.2.1 Linear comparator modeling for X-axis

In linear comparator measurement we do not have any information about RPY at any location (as we do with laser measurement). Instead we receive the total positioning error in three directions [one linear and two orthogonal]. We can express the real position for X-axis by the following equation:

$$T_{Ao}^{A} = \begin{bmatrix} 1 & 0 & 0 & \Delta X_{totalxx} + X \\ 0 & 1 & 0 & \Delta Y_{totalxy} \\ 0 & 0 & 1 & \Delta Z_{totalxz} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{16}$$

$\Delta X_{totalxx}$ is the total positioning error in X-direction when the slide moves in X-direction. $\Delta Y_{totalxy}$ is the total positioning error in Y-direction when the slide moves in X-direction. $\Delta Z_{totalxz}$ is the total positioning error in the Z-direction while the slide moves in the X-direction. Similarly we can model it for Y and Z-axes. Similarly we can find T_{Bo}^{B} and T_{Co}^{C}. We find the volumetric error for any point in the working space by the following equation for linear comparator measuring system [equation 8].

$$E_{volVM182} = \begin{bmatrix} \Delta X_{totalyx} - \Delta X_{totalxx} - \Delta X_{totalzx} \\ \Delta Y_{totalyy} - \Delta Y_{totalxy} - \Delta Y_{totalzy} \\ \Delta Z_{totalyz} - \Delta Z_{totalxz} - \Delta Z_{totalzz} \end{bmatrix} \tag{17}$$

2.2.2 Linear comparator and laser interferometer measurements

For a particular axis, whenever laser interferometer measurement technique is used, the total positioning error, roll, pitch and yaw are measured in separate cycles. When the slide or spindle moves along the axis, the machine also generates a transverse error but the laser interferometer cannot capture the transverse error on account of its working principle. On the other hand the linear comparator is capable to record linear errors (total positioning errors) as well as the transverse error sat the same time.

2.2.3 Relation for X-axis

$\Delta X_{totalxx}$ (of equation 14) is measured with linear comparator when linear positioning for X-axis is measured. $\Delta Y_{totalxy}$ and $\Delta Z_{totalxz}$ are the total transverse errors in Y-direction and Z-direction while the slide moves to X-direction. All these three total positioning errors can be computed on the basis of the laser measurement result.

$$\Delta X_{totalxx} = -AP_{wy}c_x + AP_{wz}b_x + \Delta X_{xx} \tag{18}$$

where $\Delta X_{totalxx}$, ΔX_{xx}, c_x and b_x are total positioning errors (measured by VM182), basic positioning error, pitch [laser measured] and yaw [laser measured] error at a location A for X-axis. $\Delta Y_{totalxy}$ and $\Delta Z_{totalxz}$ can be found by using laser measurement results. Thus we have:

$$\Delta Y_{totalxy} = AP_{wx}c_x - AP_{wz}a_x + \Delta Y_{xy} + Xk_{xy} \tag{19}$$

$$\Delta Z_{totalxz} = -b_x AP_{wx} + a_x AP_{wy} + \Delta Z_{xz} + Xk_{xz} \tag{20}$$

Left hand side of equations (18), (19) and (20) are computed based on the right hand side components (laser measured), thus we could obtain or simulate trace for linear comparator measurement systems. Similarly, we can find relations between Y and Z-axes.

2.2.4 Converting the laser interferometer measurement results into DBB measurement

With laser interferometer the total positioning error is measured with suitable interval along the axis.

$$E_{des\_pt} = E_i + \frac{E_{i+1} - E_i}{\Delta} \times F \tag{21}$$

The error between two measurement points can be found by the equation (21) where E_{des_pt} is the error at a desired point, Δ is the measuring interval and F is mantissa.

$$\Delta R = \frac{(C_x \cdot X + C_y \cdot Y + C_z \cdot Z)}{R} \tag{22}$$

In DBB-measurement we make a circular interpolation on a plane while the radial changes are recorded with certain intervals (for example 1200 points around a circle) [4]. If we calculate the radius changes for DBB measurement by DBB formula we can simulate the DBB measurement.

3. MEASUREMENT AND SIMULATION

3.1 Simulating the DBB measurement by laser measurement results

By using error information obtained from the laser interferometer we can plot circular interpolation. The derived trace pattern [simulated DBB trace] obtained from laser interferometer measurement is shown in Fig. 2 and Fig. 3 shows the actual machine measured [by DBB] with the same condition. These trace patterns are quite similar so we can think that the measurement conversion is correct for major geometric errors.

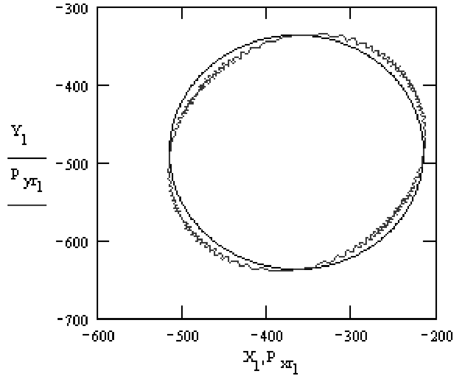


Figure 2. Simulated trace obtained from laser measurement.

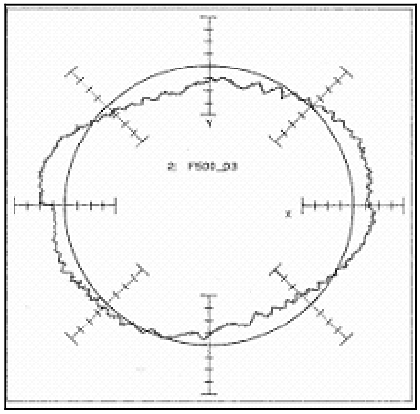


Figure 3. *Trace obtained from DBB.*

Fig. 2 is plotted based on the laser interferometer measurement results. With these results we have theoretically predicted the DBB trace. Fig. 3 is obtained by DBB measurement. In the simulated trace, backlash has not been considered, but there is clearly a backlash in negative X direction. On the other hand from the real DBB trace we see that DBB trace is more elliptical in the direction of X-axis. The reason could be a combination of the following reasons [4]:

- Local X-axis thermal expansion [Uniform expansion of X-axis]
- Scale error between X and Y-axes.
- Temperature effects have not been included in Fig. 2.
- When DBB detects the tool tip, it actually detects the resulting effects of all geometric errors of those axes and related Abbe effects. Resultant errors could be magnified or cancelled by individual error components. This is the main reason why there are big differences between circular and radial deviation.
- Controller compensates some of the errors.

Still there are some differences between Fig. 2 and 3. Most important differences are radial deviation and circularity.

Table 1. Comparison of simulated traces (from laser measurement) with DBB measurement results.

Method	Feed rate mm/minute	Scale (μm/division)	Circularity	Max. Radial deviation (μm)	Min. Radial deviation (μm)
Simulation from laser measurement to DBB measurement	N/A	20	43	22	-21
DBB measurement	500	5	15	8	-7

4. CONCLUSIONS

Three different kinds of measurement modeling have been studied in this research. We have plotted the actual measurement results obtained from the DBB measurement and predicted them from the parameters obtained from laser interferometer measurement. If we were absolutely correct in our modeling, then Fig. 2 and 3 should have been exactly the same. As we noticed that these two figures are not alike, this tells us that our modeling is not 100% reliable. The main reasons for these deviations are:

- Temperature effects are not considered in modeling.
- Two measurements were made at two different times, which means that they have different heating effects onto the machine tool.
- Cross coupling effects are not considered in modeling.
- Mass, velocity and acceleration effects have not been considered.

Converting the measurement results from one measuring system to another by software can increase the measurement confidence. This kind of conversion software can be developed based on the mathematical models developed within this research programme.

REFERENCES

1. M. Rahman, J. Heikkala, K. Lappalainen, Modeling, measurement and error compensation of multi-axis machine tools. Part I: theory, Int. Journal of Machine Tools and Manufacturing 40(10), 2000, pp. 1535-1546.
2. N. Jun, CNC machine accuracy enhancement through real-time error compensation, Journal of Manufacturing Science and Engineering, 119, 1997, pp. 717-725
3. A.G. Heidenhain, (1997) Measuring Systems for Machine Tool Inspection and Acceptance Testing, Product brochure, pp. 0-16.
4. Y. Kakino, Y. Ihara, A. Shinohara, Accuracy inspection of NC machine tools by Double Ball Bar method. Hanser publishers, Munich, Germany 1993 p. 191.

THE ROLE OF CMM AS A TOOL FOR DETECTING, MEASURING AND EVALUATING PERFORMANCE CHANGE IN THE MECHANICAL DESIGN PROCESS

Ian Egan[1], James M. Ritchie[2] and Paul D. Gardiner[3]

[1]UK Astronomy Technology Centre, Royal Observatory, Edinburgh, Scotland; [2]School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, Scotland; [3]School of Management and Languages, Heriot-Watt University, Edinburgh, Scotland.

Abstract: This paper reports on follow-up assessments carried out at industrial sites that had been previously benchmarked using a derivation of the Carnegie-Mellon/SEI Systems Engineering Capability Maturity Model® (SE-CMM®) called the Process Capability Model – Mechanical Design (PCM-MD). Two sites that had been previously assessed were revisited to carry out another assessment of the mechanical design process in use. The results were initially surprising in that it had been expected that key findings from the first batch of assessments would have been implemented as management had indicated when previous assessment results were released. It was found however that due to significant reorganizations at Group and Company levels that the impetus to improve at site level had not occurred to the degree expected. This showed up as stagnation and in some cases a retrograde movement in many of the capability levels measured. The continued development of the assessment method to include assessing people from key functions that interface closely with mechanical engineers has highlighted the communication issues that can arise across interfaces..

Key words: Process improvement; Process measurement; Mechanical engineering; CMM; PCM-MD.

1. INTRODUCTION

The measurement of the effectiveness of the mechanical design process can be seen as a significant challenge for companies that have a requirement for this design function. Inevitably, questions arise on what should be

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 367-376.

measured, particularly with the potential for making bad investment judgments should the wrong metrics be used; this is particularly the case when people strive to meet the designated metrics to avoid redress. The use of a well structured measurement tool such as the Capability Maturity Model® [1, 2] can help avoid such issues by providing a model for companies' processes to be measured against.

A new model called the Process Capability Model-Mechanical Design (PCM-MD) based on a derivation of the Systems Engineering Capability Maturity Model (SE-CMM), was created to specifically address the mechanical design process [3, 4]. The model retained the structure and engineering principles described in the SE-CMM but was substantially adapted to meet the working practices and nuances of the mechanical engineering design environment. A significant number of different engineering orientated disciplines (e.g. manufacturing, procurement, printed circuit board, production, stress, thermal engineers) can be involved in the production of a mechanically engineered product when compared say to a less involved software product; therefore it was important to find out when a site was reassessed whether this model could be used to capture the true state of the mechanical design process taking into account the many dependencies between such functions necessary for an effective mechanical design process.

The aim of this research is to determine if the PCM-MD can be used to evaluate the effectiveness of the process areas in the mechanical design process model. This paper follows on from previous work [4] to show how applying the model in a follow-up assessment can highlight differences from previous assessments and to determine if both negative and positive changes in the design process can be detected from people involved in the mechanical product creation chain.

2. BACKGROUND

Many techniques and standards have been created to help companies address the target of continued process improvement such as benchmarking, EFQM, Baldridge, ISO 9000 [5-9]. The availability of scarce resources and the time required for implementation means that models selected for such use must be carefully chosen in order to generate the best return and be repeatable.

The various CMM techniques developed by Carnegie Mellon University such as the Software CMM (SW-CMM), Integrated CMM (CMMI) and SE-CMM have been designed to provide a structured means of assessing process efficiency. The PCM-MD is based on the CMM approach, specifically the

SE-CMM, and provides a means to chart the existing engineering, project and organisation support capabilities of the mechanical design process thus helping to identify the areas where change can best be made.

Table 1. Capability level description

Capability Level	Description
Level 0; Not Performed	Base Practices not performed
Level 1; Performed Informally	Base Practices (Level 1 questions) are performed
Level 2; Planned and Tracked	The base practices are carried out and meet the following levels of performance - The Base Practice activities are: Planned Tracked Disciplined Verified
Level 3; Well Defined	Levels 1&2 are performed and a Standard Process is defined for all to follow and carry out
Level 4; Quantitatively Controlled	Levels 1-3 are performed and measurable quality goals (process and product) are established to allow objective management of performance
Level 5; Continuously Improving	Levels 1-4 are carried out with the addition of improving the organization capability and process effectiveness

Within this tool, each Process Area (PA) is applied at one of 6 levels as shown in Table 1 depending on the results of the analysis [4]. Eighteen process areas (PAs) are investigated into, covering 3 categories of interest, namely 'Engineering' (PAs 1-7), ‘Project’ (PAs 8-12) and 'Organization/Support' (PAs 13-18); these are plotted in Table 2.

After adapting the SE- CMM to meet the needs of the mechanical design environment on-site studies were carried out with Company A from 1999 to 2001. These initial studies had been deemed sufficiently successful by the company involved that in 2002 they requested that a second series of assessments be carried out to re-measure the mechanical design process in the light of the many changes that had taken place, including major high-level company activities such as physical relocations and mergers.

Table 2. Process Area numbering and description

PA#	Process Area Title
	ENGINEERING
PA 1	Understand Customers Needs and Expectations
PA 2	Derive Specification and Allocate Requirements
PA 3	Analyze Possible Solutions

PA#	Process Area Title
PA 4	Generate Design
PA 5	Integrate Product Modules
PA 6	Verify and Validate Product
PA 7	Integrate Disciplines
	PROJECT
PA 8	Ensure Quality
PA 9	Manage Risk
PA10	Plan Product Development Effort
PA11	Monitor and Control Product Development
PA12	Manage Configurations
	ORGANISATIONAL SUPPORT
PA13	Define Organization's Mechanical Design Process
PA14	Improve Organization's Mechanical Design Process
PA15	Manage Mechanical Design Support Environment
PA16	Provide Ongoing Skills and Knowledge
PA17	Manage Evolution of Product
PA18	Co-ordinate With Suppliers

3. METHODOLOGY

1. Select participants and run questionnaire sessions

▼

2. Analyze questionnaire data using Excel assessment tool

▼

3. Determine initial process area ratings

▼

4. Conduct individual interviews

▼

5. Prepare initial comments for feedback Discuss individual PAs and update ratings where necessary.

▼

6. Conduct group interviews

▼

7. Discuss individual PAs and update ratings where necessary.

▼

8. Transcribe interviews.

▼

9. Analyze transcripts and update ratings where necessary.

▼

10. Compile draft final report

▼

11. Issue and agree distribution of final report with company

▼

12. Present assessment findings

Figure 1. Assessment methodology used at Site 1 & 2

As a proof of concept the initial pilot study carried out at Site 1 concentrated on applying the model to the mechanical discipline (designers and drawing office staff) and a few closely associated disciplines such as stress and thermal analysis. Although this was successful it was believed that more would be gained from broadening the base from which participants were drawn to include people from the various stages of the mechanical product lifecycle - from conception to customer delivery - who were in a position to contribute towards the success of the mechanical design product created.

Consequently, the pilot study was supplemented by another assessment carried out in the Autumn of 1999. The results highlighted the value of this approach as several areas that were originally perceived to be reasonable from a purely mechanical engineering stance were found to be lacking when 'seen' from the perspective of people involved in the next stages of the product creation lifecycle process. This approach is analogous to a view recently taken within CMMI [10] which advocated 'an integrated approach across the enterprise for improving process.'

The methodology applied to the site reassessments followed closely the one used for the initial pilot study work and site assessments [4] including the use of a broad base of people and associated disciplines that had positive links with the mechanical design process. The method used is shown in Fig.1.

4. ANALYSIS AND RESULTS

4.1 Site 1 Results

Using the methodology outlined in the previous section, the sample of personnel involved in the re-assessment process for Site 1 in Sep 2002 comprised:

- 3 senior mechanical engineers;
- 2 mechanical engineers;
- 4 drawing office personnel;
- 13 members from other disciplines including quality, stress, thermal, production engineering, configuration management.

After filling out the questionnaires and analyzing the results six senior engineers were interviewed individually with the rest being interviewed in two groups of eight people. Process Areas for which the results were close to reaching either Level 1 or 2 or where there tended to be some disagreement amongst the questionnaire respondents, were then targeted for discussion. Many interview sessions also covered other PA issues due to the nature of the mechanical design function and the fact that problem cause-and-effect relationships became apparent during discussions. This provided fertile ground for building up a picture of the way that products evolved on both sites.

The results of the second survey at Site 1 in September 2002 can be seen at Fig. 2. Surprisingly, a general reduction in capability level is evident in many areas from when the model was applied during the previous assessment. As can be seen from the diagram, PAs 14 ('Improve Organization’s Mechanical Design Process’) and 15 ('Manage Mechanical Design Support Environment') in particular displayed dramatic reductions in level as well as illustrating several reductions in the other ‘Project’ and 'Organization' PAs (PA8-PA18). However, after discussions during the interviews, and through post-analysis talks with company management, the reasons for this apparent decline in process performance became evident.

Site 1 had become demerged from one internal unit and merged into another. Whilst this was going on a major move was taking place into a new building with the integration of other design groups from other divisions and sites into Site 1's mechanical design offices. So it would appear that both initiative fatigue and 'change' fatigue were having an adverse impact on mechanical design process performance. In this context, PCM-MD as a tool gives some measure of how employees respond to new ideas implemented and how they affect performance in a creative environment such as engineering design. It is often quoted that the only constant in business now is change but does industry really understand - or indeed know how to measure - the real effects of such an approach? Also, many people expected improvements identified after the first study to be implemented; this was not followed through.

Another key weakness from the results, and as has been found in all studies carried out in the mechanical engineering environment by the Heriot-Watt group, is that relating to PA7 ('Integrate Disciplines'). Due to the lack of formality in the mechanical design process and the fact that, due to 'cost cutting', many disciplines are not involved in the product development process at an early enough stage by project managers, there is a distinct feeling of poor communication between the functional groups. Mechanisms do exist for people to work in design teams - co-located or virtual - however, people on this site tend to resent the fact that the product development process itself does not trigger their involvement early enough in the life cycle management process. Also, the loss of quality in the mechanical design process was apparent by the score in the relevant process area (PA8).

After completing this detailed analysis a number of recommendations were made to the company with regard to improving its mechanical design process, some of which are now given:

- There was a need to formalize the product design (not just mechanical design) process and drive it through a system of milestones and gates using some form of product data management system (workflow). Staff members needed to be well educated in this process.
- This system should incorporate the company's product specification, project management and configuration management tools into the formalized and system driven process.
- The mechanical design workflow system should be tailorable, traceable and flexible enough to handle different kinds and sizes of project.
- Integrated Disciplines was particularly weak and real efforts have to be made to pull in various affected disciplines early on in the product development process, e.g. at the bid stage. It was recommended that the workflow system contain a series of positively enacted opt-outs for different disciplines to decide upon from the start of a bid or contract.
- Functional learning was inhibited by working in teams; this needs to be addressed.
- Engineers lower down the hierarchy should be involved in the bid stage and experience all aspects of the product development process.
- PCM-MDs carried out at all sites should enable a generic process to be developed for all company sites.
- CMM studies should be carried out for Systems Engineering and Project Management, understanding the problems in these areas will greatly enhance improvement in the overall mechanical design process.

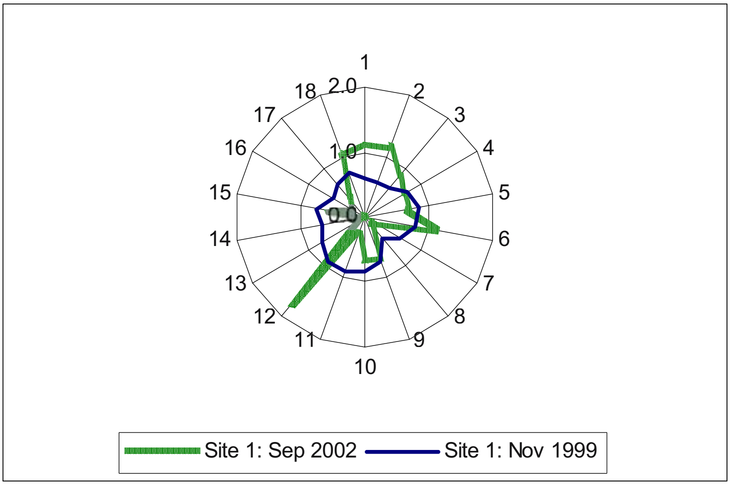


Figure 2. Site 1 results

4.2 Site 2 Results

Following the reassessment process at Site1, Site2 was reassessed. The sample of personnel involved in the re-assessment process comprised:

- 3 senior mechanical engineers;
- 4 mechanical engineers;
- 4 drawing office personnel;
- 12 members from other disciplines including quality, stress, thermal, production engineering, configuration management.

In keeping with previous studies, a positive reply rate of 90% was taken as the level required for each question in the questionnaire to be regarded as being carried out robustly on site. A key element that came out of the discussions during the interviews and post-assessment discussions with management was that, although the site had proactive mechanical engineering management, they had been blocked from releasing updated mechanical procedures shortly after the first assessment so that a group-wide approach could be taken. Interestingly, the biggest improvement was in the area of PA18 ('Co-ordinate with Suppliers') where issues raised by the last assessment were addressed (Fig. 3). However, PA7 ('Integrate Disciplines') was again very weak, for similar reasons to those outlined for Site 1.

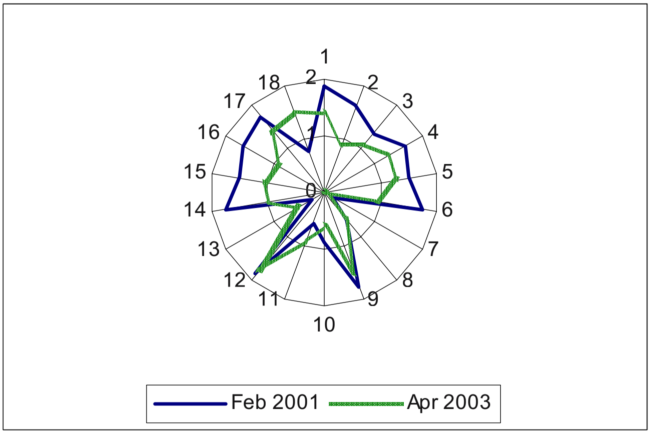


Figure 3. Site 2 results

The key findings from the study for Site 2 are very similar to those for Site 1. Some of the Site 2 findings are listed below:

- A group-wide formalized mechanical design process urgently needs rolled out amongst all the company's sites.

- Tailoring the new design process to individual projects is critical when deemed essential for business needs.
- There is a need to create mechanical design process metrics for measurement, analysis, reporting and continuous process improvement.
- There is a need to embed project management, product definition and configuration management tools into the workflow process.
- Co-location, physically or virtually, should be driven into the project culture so that mechanical engineers are not seen merely as a 'service' but as part and parcel of the solution within projects. This has to carefully balance the needs for multi-disciplinary team working and functional learning.
- There is a need to make more physical prototypes available for engineers.
- The level of interaction between different disciplines during design seems to vary according to the strength of personal interaction.
- Access and visibility to relevant project planning information and plan modifications should be given in a more timely manner to all concerned. Project planners should liaise more closely with mechanical leads, especially with regard to plans used for forward loading of effort.
- Ownership of estimates and budgets should be flowed further down to people who are closer to doing the work.
- A more structured and formal mechanism should exist for cross-project and cross-discipline learning.
- Senior mechanical engineers should be tasked with leading their more junior colleagues through the web based available information.

5. CONCLUSIONS

As can be seen from the site analyses carried out in this work the PCM-MD provided a rich source of data for quantitative and qualitative analysis when investigating the mechanical design processes within this company. The recommendations, some of which are given above, have led to the company taking two major steps in attempting to improve its process:

- As well as these reassessments the company also informed the Heriot-Watt team that they wished to roll out this process to at least their three other key mechanical engineering sites. Due to the time constraints in which they wished this completed, a team of senior mechanical engineers were trained to carry out this process internally with external auditing of the process being done by the University.
- A formalized, web-based mechanical design process workflow is being developed and introduced at all of these sites to give them better, consistent control of their overall mechanical design process.

Although these analyses did not show the expected overall improvements but more typically regression in the way in which these sites' processes operated, the company involved was able to measure the effect of some of its strategic business decisions where products have to be designed and data prepared for manufacturing. Improvements were measured where relevant and best practice can be observed from one site's results and investigated at another site if cross-site comparisons of results are made for specific PAs. Indeed, they are able to use these data as a strategic planning tool to help target their investment in process improvement.

To conclude, due to the way in which this company has adopted the PCM-MD and has been better able to understand its mechanical design process and function as well as the feedback obtained on the final results, this work has significantly validated the proposition that a modified CMM can be used to effectively evaluate the mechanical design process.

REFERENCES:

1. M.C. Paulk, B. Curtis, M.B. Chrissis, C.V. Weber, "Capability Maturity Model for Software V1.1" *CMU/SEI-93-TR-024*, Carnegie Mellon University. Pittsburgh, 1993.
2. D.A. Kuhn, et al. "A description of the Systems Engineering Capability Maturity Model Appraisal -Method V 1.1' CMU/SEI-96-HB-004, 3-21 Carnegie Mellon University. Pittsburgh (March 1996).
3. I. Egan, I. Black, J.M. Ritchie, P.D. Gardiner, "Process improvement in mechanical design using a derivation of the systems engineering capability maturity model." 33rd International MATADOR Conference, UMIST, Manchester, UK, July 2000.
4. I. Egan, J.M. Ritchie, P.D. Gardiner 'The Novel Use of a Capability Maturity Model to Evaluate the Mechanical Design Process' Recent Advances in Integrated Design and Manufacturing in Mechanical Engineering, Kluwer Academic Publisher, 2003, pp3-12.
5. B. Povey, "The development of a best practice business process improvement methodology." Benchmarking for Quality Management and Technology, Vol 5, No 1, 1998, pp27-44.
6. "European Foundation Quality Model", http://www.efqm.org.
7. "Baldridge Quality Award," http://www.quality.nist.gov.
8. A. Neely, M. Gregory, K. Platts, "Performance measurement system design a literature review and research agenda," International Journal of Operations and Production Management, Vol 15, No 4, 1995, pp80-116.
9. "ISO 9000-9004 Quality Management and Quality Assurance Standards, 1987" http://www.iso.ch/
10. Concept of Operations for the CMMI, 15 January 2001, www.sei.cmu.edu/cmmi/background/conops.html.

THE DESIGN OF FOOD PROCESSING SYSTEMS FOR IMPROVED RESPONSIVENESS AND LATE CUSTOMISATION

Cora Fisher, Glen Mullineux, Tony Medland and Ben Hicks
Department of Mechanical Engineering, University of Bath, Bath BA2 7AY, UK

Abstract: In designing food manufacturing systems, it is important to have design models to help increase understanding and to predict the behaviour of proposed processes. However, understanding of the properties of the relevant food materials may be limited and the underlying relationships are often highly non-linear. When coupled with models of the manufacturing system itself, the sensitivities of the various design parameters can be very different. Constraint modelling is a technique that has been used successfully in the design of machine systems. Here it is applied to create an initial model of a yoghurt processing line. The approach is to identify the constraints which bound the application and then to use optimisation techniques to resolve these.

Key words: Systems design; Complexity; Product customisation; Processing.

1. INTRODUCTION

Currently most products in the food industry are not developed using modern engineering techniques. Reliance is placed instead upon long established processes. This often leads to virtually no scope for small batch size or product customisation. For example, the ability to add flavour to foods as late in the production process as possible whilst still maintaining product quality and integrity is highly desirable. In the case of yoghurt, this would allow the base product to be made in bulk with the individual flavouring and fruit added to meet variation in customer demand. Such improved responsiveness requires the design of production machinery in order to allow late customisation whilst maintaining product quality.

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 377-388.

In order to design production machinery effectively, determine design trade-offs, and allow an overall optimal procedure to be established, the consideration of food technology and process engineering during the design and construction of production machinery is essential. The food technology issues impose constraints upon how the food material can be handled and processed without detriment to its taste, appearance and nutritional values, whilst the process engineering constrains certain operations such as heat transfer, mixing and product flow. The collective consideration of all these aspects is further frustrated by the fact that the modelling approaches and support tools used in the various disciplines are very different and are often incompatible. Food technology typically relies upon the empirical results arising from tasting sessions. Machine design is often based upon CAD data which is generally geometric and parametric, and food processing systems may be governed by fluid flow analysis that relies upon continuously varying parameters.

In order to address some of the above problems, the use of constraint-based methods is being investigated. These have been used successfully in mechanical design applications such as the design of mechanisms and machine subsystems. In such applications the design parameters usually relate to aspects of the system geometry, which are generally comparable in size and have roughly the same sensitivity. In wider applications such as food processing, the sensitivity of the governing design parameters may be widely different for different parameters. This complexity causes significant problems for the numerical methods used to analyse such systems.

This paper deals with a specific case involving the production of yoghurt. An additional difficulty here is that not only is the behaviour of the material non-linear in its response to deformation, but also the parameters involved in modelling that behaviour vary as more and more processing is applied. It is found that there is some recovery of properties when the material is allowed to rest.

2. CONSTRAINT MODELLING

The ideas of constraint modelling arose from investigations into the mechanical design process. In the early conceptual stages of a new design task, the precise rules by which the design is to operate are largely unknown. What are more apparent are the constraints which limit the performance and physical attributes of the design. For example, a part may need to be less than a particular size but still be sufficiently strong. As the design evolves, more and more constraints emerge, partly from a better understanding of the specific design problem. The aim of the designer is to find a design solution which meets all the imposed constraints. Often this is not possible since the constraints are in conflict and the skill of the designer is in deciding which constraints can be relaxed without compromising the overall design.

To support design work, a software environment has been created [1]. This incorporates a command language in which design parameters can be declared and constraints between them imposed. In order to resolve constraints, optimisation schemes are used to determine the best compromise solution even in cases when the constraints are conflicting. The user presents constraints as expressions between the design parameters. When each expression is true its value is zero, otherwise its values represents the error. During the resolution phase, the system forms the sum of the squares of these values and tries to minimise this by varying parameters specified by the user. Within the language of the system, constraints are specified within user-defined functions. The following is a simple example.

```
function solve
{
   var  x, y;
   rule( x + y - 4 );
   rule( 2*x - y - 5 );
}
```

Here `x` and `y` are have been declared elsewhere as global real variables. When this function is invoked, the system automatically adjusts the values of the two parameters and effectively solves a pair of linear equations.

In principle, any of the standard optimisation methods [2] for handling problems with continuous variables can be used in the constraint resolution process. As the system does not know beforehand what form of constraints are to be presented, robust general purpose methods are preferable. In particular, the derivatives of the constraint expressions are not available (but could be determined numerically). For these reasons, direct search methods have been implemented in the current version of the constraint modelling software. The two main ones are the method of Hooke and Jeeves and the extension to this, which is Powell's direct search method [2]. Both of these methods have been found to work well with applications arising in mechanical design.

As previously mentioned, the constraint modelling approach has been successfully used in the area of mechanism and machine design [3]. Here the constraints typically relate to the assembly of the parts of the machine and to its kinematic performance. Simple graphics allows the machine model to be viewed in wire-frame and the incorporation of the ACIS solid modeller allows solid objects to be handled and constraints applied to them.

The purpose of the work described here is to investigate the use of the constraint modelling approach in a complementary area, namely food processing systems. There are constraints here in terms of how the food product is processed and how it is perceived by the consumer. Modelling of the properties of food material often involves the use of non-linear algebraic equations and the solution of these seems to be a natural application of the resolution techniques within the constraint modeller.

As a simple example, consider the modelling of a Newtonian fluid through a simple pipe network. The network is shown in schematic form in Fig. 1 and comprises six branches which meet at four internal nodes. The model is described as a series of constraint rules. Some of these specify conservation of volume flow at the nodes. The typical nodal rule of this type has the form

```
rule( v01*a1 - v12*a2 - v13*a3 );
```

where the `v` terms represent velocities in the branches and the `a` terms are cross-sectional areas.

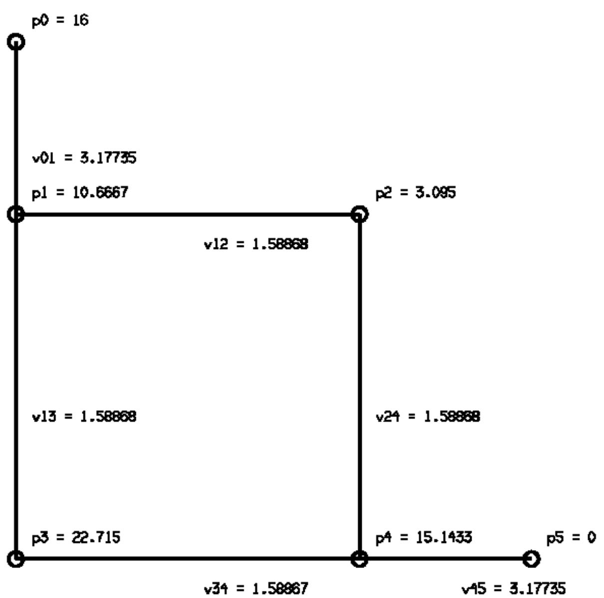


Figure 1. Flow of Newtonian fluid through pipe network.

The Bernoulli equation for a typical branch has the form

```
rule( (p0/(rho*g))+y0-h01-(p1/(rho*g))-y1 );
```

where `p` terms give the pressures at the nodes, the `y` terms relate to the height of a node above a given datum, and the `h` term represents a head loss along the branch. For the example shown here, there are four nodal constraints and six branch constraints. When the system is allowed to vary the nodal pressures and branch velocities a solution to the set of constraints is achieved as shown in the figure.

3. HERSCHEL-BULKLEY MODEL

Attention is now turned to the modelling of food production systems. Many food substances exhibit highly non-linear behaviour and a plethora of models have been proposed to deal with these [4]. Of these, one commonly used is the Herschel-Bulkley model, and this is usually assumed to hold for yoghurt. The model relates the shear stress τ to the strain rate $\dot{\gamma}$ by the formula

$$\tau = \tau_0 + K \dot{\gamma}^n \qquad (1)$$

where τ_0 is the yield stress of the material, K is the consistency factor, and n is the flow behaviour index. Fig. 2 shows the graph of τ against $\dot{\gamma}$ when τ_0 = 10 Pa, K = 10 Pa s^n and n = 0.3.

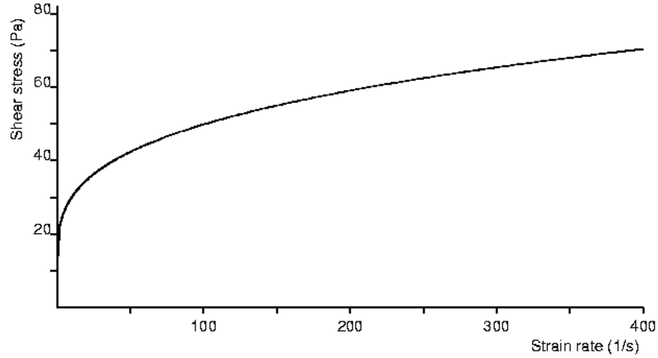


Figure 2. Typical Herschel-Bulkley curve with τ_0 = 10 Pa, K = 10 Pa s^n and n = 0.3.

The *apparent viscosity* η of a material is defined as the quotient of τ and $\dot{\gamma}$. So, from Eq. (1), the apparent viscosity is given by

$$\eta = \frac{\tau}{\dot{\gamma}} = \frac{\tau_0 + K\dot{\gamma}^n}{\dot{\gamma}} \tag{2}$$

In the case when τ_0 is zero, this expression reduces to

$$\eta = K\dot{\gamma}^{n-1} \tag{3}$$

It is necessary to note that the value of apparent viscosity is dependent upon the current value of $\dot{\gamma}$. In particular, the above equation says that, for τ_0 non-zero (or for $\tau_0 = 0$ and $n < 1$), the apparent viscosity is infinite when $\dot{\gamma}$ is zero. Fig. 3 shows a graph of viscosity against strain rate in the case when $\tau_0 = 10$ Pa, $K = 10$ Pa s^n and $n = 0.3$. In the example given in the next section, the apparent viscosity is evaluated at the point when the strain rate is 32 s^{-1}.

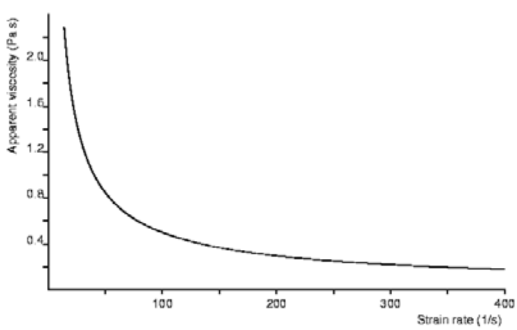


Figure 3. Typical graph of apparent viscosity obtained from the Herschel-Bulkley model with $\tau_0 = 10$ Pa, $K = 10$ Pa s^n and $n = 0.3$.

The Herschel-Bulkley model can be used to investigate the flow of material along a circular pipe of radius R. A relation between the volume flow rate Q and the imposed pressure gradient p' can be derived. This result has reported previously [4] but has an error in sign. The correct result is the following.

$$Q = \frac{\pi R^3}{K^{1/n}} \frac{(\tau_\omega - \tau_0)^{\frac{n+1}{n}}}{\tau_\omega^3} \left[\left(\frac{n}{3n+1} \right) \tau_\omega^2 + \frac{2n^2}{(2n+1)(3n+1)} \tau_0 \tau_\omega + \frac{2n^3}{(n+1)(2n+1)(3n+1)} \tau_0^3 \right] \tag{3}$$

Here τ_ω is the shear stress at the wall of the pipe. This is related to the pressure gradient p' by

$$\tau_\omega = \frac{1}{2} R p' \tag{4}$$

One of the difficulties in dealing with yoghurt is that its properties vary while it is being processed. Shear thinning of the material occurs. Initial experimentation has produced results corresponding to Herschel-Bulkley curves for repeated application of the test. As the repetition number increases, so thinning occurs and the curves tend to flatten and rise to lower heights (Fig. 4).

The following table gives the values of the parameters associated with some of the repeat curves in this process and it is clear that *K* decreases and *n* increases. The value of τ_0 has here been assumed to stay constant.

Table 1. Herschel-Bulkley parameters over various curves

Curve	τ_0	K (Pa s^n)	n
1	10	12.59	0.268
5	10	4.11	0.437
10	10	2.97	0.476
15	10	2.71	0.482
21	10	2.45	0.491

Let the curve number be *m*. The above data have also been interpolated with respect to *m*. The following formulae for the three parameters have been obtained.

$$\tau_0 = 10 \tag{5}$$

$$K = 1.944 + \frac{10.791}{m + 0.0136} \tag{6}$$

$$n = 0.5103 - \frac{0.4094}{m + 0.6891} \tag{7}$$

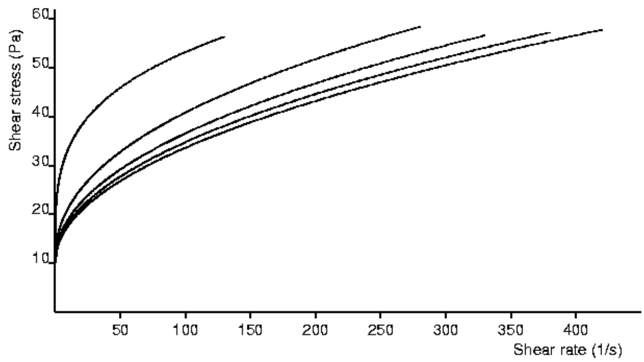


Figure 4. Graphs from the Herschel-Bulkley model using interpolated values of the parameters for m = 1, 5, 10, 15, 21.

Figure 4 shows the graphs from the Herschel-Bulkley model using parameters derived from the above interpolations. These compare well with the original experimental results. These formulae are used in the model of pipe flow and of mixing. To do this it is assumed that m related to the time of processing. Let f be a time factor related to the time of each repeat. Then the processing time t is related to m by $t = (m-1)f$. The $(m-1)$ term ensures that time t is zero at the start of curve number 1.

Other preliminary experimental results have revealed that viscosity decreases exponentially as the temperature rises. In practice, yoghurt is cooled after being processed and packed and so there is some recovery of its viscous properties as this happens. What has been found is that the amount of recovery is reduced by extensive working during the processing stages. Again a formula has been fitted to the experimental data. If the processing time is t (in minutes) and θ is the temperature (in °C), then the viscosity η is given by

$$\eta = \left(14.9 + 45.5e^{-0.178t}\right)e^{-0.282\vartheta} \tag{8}$$

Figure 5 shows graphs of η against θ obtained from the above function in the cases when t is 5, 10, 15 and 20 minutes.

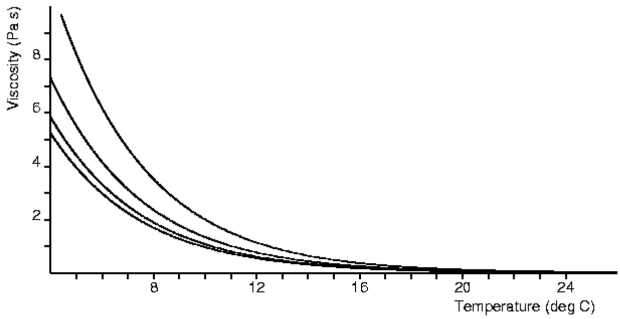


Figure 5. Graphs of interpolated formulae for cooling and recovery for processing times of 5, 10, 15 and 20 minutes.

4. THREE STAGE MODEL

Three stages of yoghurt processing have been modelled using the experimental results. The first stage is flow along a pipe. This has the effect of reducing the viscosity. The pipe diameter is 0.75mm and its length can be varied between 6m and 10m. The flow rate is between 6000 and 8000 litres per hour. The second stage is a mixing stage (during which fruit is added), which has the effect of reducing the viscosity considerably. The final stage is a cooling and recovery stage. Here the yoghurt is allowed to rest and cool. Let $T_0 = 0$ be the time at the start, and let T_1, T_2 and T_3 be the times at the ends of the three stages. Let η_0, η_1, η_2 and η_3 be the corresponding viscosities.

Eqs. (5)-(7) are assumed to hold in both stages 1 and 2. Since the processing is more intense in the mixing in stage 2, the time factor is assumed to be lower. The following equation is now used for determining the effective curve number *m*.

$$m = \begin{cases} 1 + \dfrac{t}{f_1} & \text{when } t \le T_1 \\ 1 + \dfrac{T_1}{f_1} + \dfrac{t - T_1}{f_2} & \text{when } t > T_1 \end{cases} \tag{9}$$

The values taken for the two time factors are: $f_1 = 60$ s and $f_2 = 5$s.

There are two parts to the first stage of the model. One is to find the pressure gradient p'. This is done by applying a single constraint which is derived from Eq. (3). The constraint specifies that the flow rate should reach the given desired value. To resolve the constraint, the modeller is allowed to vary the value of the pressure gradient p'. The other part is to find the reduction in the viscosity. This is done by using the imposed flow rate to determine an average flow velocity in the pipe and hence an average time. This time is converted to a curve number by Eq. (9) and then the viscosity is determined by Eq. (2) using the parameters obtained from Eqs. (6)-(7). The time is of the order of 20s and hence the values are those associated with the first curve (m = 1) of Fig. 4.

In the second stage of the model, the value of T_2 is effectively determined by the required final viscosity. If T_2 is increased then the final viscosity is reduced. The time is found by applying a constraint to ensure that the final viscosity reaches the required value. To resolve the constraint the processing time for the second stage is allowed to vary. Once the value of T_2 is known, the other parameters of the yoghurt can be determined. Figure 6 shows the basic output from a single run of the model. This is for the following values.

Table 2. Values for one run of the model

Parameter	Value
Initial viscosity	1.302 Pa s
Initial temperature	25°C
Pipe diameter	0.075 m
Pipe length	10 m
Flow rate	0.00222 $m^3 s^{-1}$
Final viscosity	12 Pa s
Final temperature	5°C
Cooling time	125 s

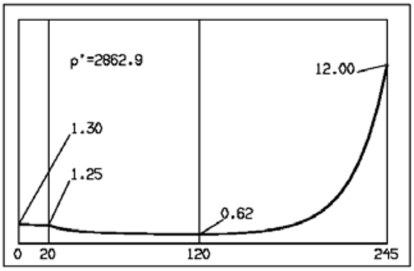


Figure 6. Graph of viscosity over the three stages of the model.

Figure 7 shows the effect of changing the final required viscosity η_3. The final values used are: 10, 12, and 14 Pa s. The lower the final viscosity value the more time is spent in the second stage, thus elongating the graph.

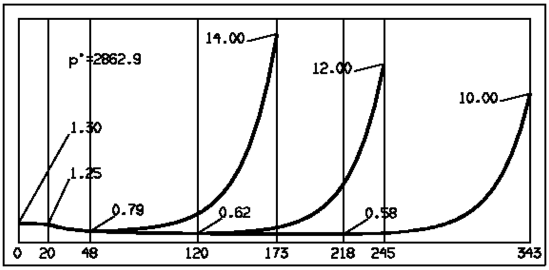


Figure 7. Graph of viscosity for three different prescribed final viscosities.

Figure 8 shows the effect of varying the pipe length. The values used are the same as for Fig. 6 except that the pipe length is reduced from 10m to 6m.

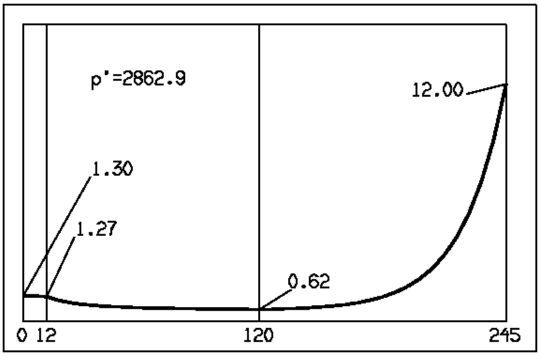


Figure 8. Graph of viscosity for reduced pipe length of 6m.

5. CONCLUSIONS

The design modelling and analysis of food processing systems is a difficult task for a number of reasons. Understanding of the properties of the food materials may be limited and what models do exist can involve highly non-linear relations between the parameters. These relations may change significantly while processing is being undertaken. Modelling of the food system itself brings in other parameters which may have different sensitivities to those of the food material itself.

This paper has looked at the application of constraint modelling techniques to an application related to yoghurt processing. As the approach is based upon optimisation and the constraints can deal with bounding conditions, it opens up the possibility of creating useful models with only incomplete understanding of the underlying processes. Furthermore it allows consideration of the food properties, the processing systems, and the machinery to be undertaken concurrently in the same modelling environment.

ACKNOWLEDGEMENTS

The work described in this paper has been carried out with the support of a DEFRA research grant given as part of the Food Processing Faraday. The authors gratefully acknowledge this support and that of other collaborators, in particular colleagues at the University of Birmingham who provided the experimental results.

REFERENCES

1. G. Mullineux, Constraint resolution using optimisation techniques, Computers &Graphics 25, 483-492 (2000).
2. G.R. Walsh, Methods of Optimization (Wiley, London, 1975).
3. L.P.J. Kenney, A.H. Rentoul, B.R. Twyman, D.R. Kerr, and G. Mullineux, A software environment for conceptual mechanism design, Proc. Instn Mech. Engrs - Part C: J. Mechanical Engineering Science 211, 617-625 (1997).
4. S.D. Holdsworth, Rheological models used for the prediction of the flow properties of food products: a literature review, Trans. IChemE. Part C. 71, 139-179 (1993).

EVOLUTIONARY MORPHOLOGY

A structured approach towards inventive engineering design

Grigore Gogu

Laboratory of Research and Applications in Advanced Mechanics, Institute Français de Mécanique Avancée et Université Blaise Pascal, Campus de Clermont-Ferrand / les Cézeaux, B.P. 265, 63175 Aubière, France

Abstract: The paper presents a new structured approach to inventive engineering design, called evolutionary morphology (EM), as proposed by the author of this article. In the first part of the paper we present the main paradigms and the formalization of this new method integrating morphological and evolutionary approaches. EM and evolutionary algorithms are in fact complementary methods. Evolutionary algorithms start from a given initial population to obtain an optimum solution with respect to a fitness function. EM creates this initial population to enhance the chance of obtaining a "more global optimum" by non-quantified diversification of the initial population. Evolutionary algorithms are optimized oriented methods; EM is a conceptual design oriented method. EM is formalized by a 6-tuple of final objectives, primary elements, morphological operators, evolution criteria, morphologies and a termination criterion. The main EM paradigms are inspired by the synthetic theory of evolution developed by modern biology. In the second part, an application to inventive design of orthogonal anthropomorphic robotic manipulators with only isolated singularities is presented.

Key words: Evolution, Morphology, Morphological research, Design, Robots

1. INTRODUCTION

In general, any abstract task can be thought of as solving a problem, which, in turn, can be perceived as a search through a space of potential solutions. The generation of this space is an inventive activity inherent to any innovation. Whether this is in product, manufacturing or service design, in structuring a company, a marketing network, innovation remains a

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 389-402.

challenge. Too much invention is guided by mystery. An inventive structured approach eliminates mystery, errors and time bombs, and reveals new solutions while remaining adaptable to future changes. In an inventive structured approach, the inventive effort becomes an inherent judgment process, traceable and reproducible. Inventive structured approach is ideally suited to engineering design and inventive problem solving [1-2]. Evolutionary morphology (EM) is such a structured approach to inventive design combining features of morphological research and evolutionary algorithms.

We use “morphosis” as the morphological terminology for change, replacing one system by the next one. This concept is based on the fundamental scientific method of alternating between analysis and synthesis. It can be trusted as a useful, non-quantified method for investigating complex problems which cannot be processed by direct mathematical formalization (equations), causal modelling and simulation [1]. We differentiate between static, dynamic and evolutionary morphosis. In static morphosis, a system is disassembled into very tiny components, and reassembled into a completely different system that has no obvious similarity with the previous one. In dynamic morphosis an existing system changes by enhancing, adding or reducing existing properties (parameters). Evolutionary morphosis has the characteristics of both static and dynamic morphosis.

2. MORPHOLOGICAL AND EVOLUTIONARY APPROACHES

The morphological approach is a methodology, first proposed by Zwicky [3] to enhance the systems’ engineering process. The morphological approach determines the requirements and conceptual design solutions of a system by considering the "totality" of all solutions included in a morphological box. This is an initial phase before the optimization process starts looking for “the best solution”. The evolutionary algorithms are stochastic techniques whose search methods for optimal solution model some natural phenomena: genetic inheritance and Darwinian striving for survival.

The term morphology comes from antique Greek (*morphe*) and means shape or form. The general definition of morphology is "the study of forms or patterns", i.e. the shape and arrangement of parts of an object, and how these "conform" to create a “whole” or *Gestalt*. The "objects" in question can be physical objects (e.g. an organism, an anatomy, a geographical or an ecological entity) or mental objects (e.g. linguistic forms, concepts or systems of ideas).

The term morphology, as an explicitly defined scientific method, was first mentioned by J. W. Goethe (1749-1832). In addition to poetry, dramas, and novels, Goethe continued his work in science, studying structures and shapes in mineralogy, geology, and osteology (the study of bones). In 1784, he had made the discovery, by methods which foreshadowed the science of comparative morphology, that the human jawbone contained traces of a structure similar to the intermaxillary bone found in other mammals. In 1790, he wrote *Versuch, die Metamorphose der Pflanzen zu erklären* (Essay on the Metamorphosis of Plants), which further developed his ideas on comparative morphology and to some extent foreshadowed Darwin's ideas on organic evolution. His "comparative morphology" of botany describing how animals and plants changed shape and evolved into new species was a breakthrough in thinking.

Today, morphology is associated with a number of scientific disciplines. In linguistics, morphology is the study of word formation and how the parts of words relate to each other. Morphology is used in disciplines where formal structure and structural interrelations, and not necessarily quantity, are the central issues, e.g. geology, anatomy, botany, zoology etc. Fritz Zwicky proposed a generalised form of morphological research [3]. He proposed to generalize and systematize the concept of morphological research and include not only the study of the shapes of geometrical, geological, biological, and generally material structures, but also to study the more abstract structural interrelations among phenomena, concepts, and ideas, whatever their character might be. Essentially, general morphological analysis is a method for identifying and investigating the total set of possible relationships or "configurations" contained in a given complex problem. In this sense, it is closely related to typology construction [4], although it is more generalised in form and conceptual range.

Although Zwicky coined the term morphological analysis, the technique predates him and can be traced back to the "universal method of thinking" presented by the Catalan philosopher Ramon Lull (1233-1315) in his work *Ars Magna* (The Great Art) published in 1275. Zwicky was the first to use the technique in modern-day applications. He used morphological research in new product and process ideation such as jet engines, propulsive power generation and astronomical investigation [3]. After his main work on the subject [3] originally published in 1966 by Droemersche Verlagsanstalt Th. Knaur Nachf., München/Zürich, under the title *Entdecken, Erfinden, Forschen im morphologischen Weltblid*, morphological research was extensively applied to engineering design, manufacturing and process design, systems engineering, technological forecasting and scheduling [5-6].

The evolutionary algorithms are stochastic techniques whose search methods for optimal solution model some natural phenomena: genetic inheritance and Darwinian striving for survival. There are several variants of evolutionary algorithms (genetic algorithms, evolution strategies, evolutionary programming, genetic programming) and there are also many hybrid systems which incorporate various features of these paradigms [7-12]. However, the structure of any evolutionary method is very much the same [7-8]. The beginning of the evolutionary algorithms can be traced back to the early 1950s when several biologists used computers for simulation of biological systems [9].

An evolutionary algorithm maintains a population of individuals for iteration t. Each individual represents a potential solution to the problem at hand, and is implemented as some data structure S. Each solution is evaluated to give some measure of its fitness. Then, a new population (iteration $t+1$) is formed by selecting the fittest individuals. Some members of the new population undergo transformations by means of genetic operators to form new solutions. There are unary transformations m_i (mutation type), which create new individuals by a small change in a single individual ($m_i : S \rightarrow S$), and higher order transformations c_i (crossover type), which create new individuals by combining parts from several (two or more) individuals ($c_j : S \times ... \times S \rightarrow S$). After a number of generations the algorithm converges to a "best individual". It is hoped that the best individual represents a near-optimum (reasonable) solution.

The genetic algorithms, as they are known today, are closely related to the work performed in the late 1960s and early 1970s at the University of Michigan under the direction of John Holland [9]. Genetic algorithms (GAs) were designed to model adaptation processes, mainly operated on binary strings and used a recombination operator with mutation as a background operator. Mutation flips a bit in a chromosome and crossover exchanges genetic material between two parents. The fitness of an individual is assigned proportionally to the value of the objective function for the individual. Individuals are selected for the next generation on the basis of their fitness. The combined effect of selection, crossover and mutation gives the so-called reproductive schema growth equation. The crossover operator enables structured yet random information exchange. The mutation operator introduced greater variability into the population.

To apply a GA to a particular problem, it is necessary to design a mapping between a space of potential solutions for the problem and a space of binary strings of some length. This is not a trivial task and, often, it has an important impact on the final solution [10-11]. EM contributes to solving this task. Conceptual aspects inspired by evolutionary theory are integrated in EM to generate non-quantified diversification of the initial population.

3. EVOLUTIONARY MORPHOLOGY

In this section we present the structure of EM and its specific differences with respect to morphological research and evolutionary algorithms. To facilitate understanding of the proposed formalization, the main paradigms of EM are presented in a closed relation with the biological background of the synthetic theory of evolution.

EM can be defined as 6-tuple

$EM_t = (\Phi, E, \Omega_t, \Psi_t, \Sigma_t, \iota)$

applicable to each generation *t* of the morphological evolutionary process, in which, there are :

$\Phi = (\phi_1, ..., \phi_n)$ - the set of final objectives,

$E = (\varepsilon_1, ..., \varepsilon_n)$ - the set of primary elements,

$\Omega_t = (\omega_1, ..., \omega_n)$ - the set of morphological operators applicable at each generation *t*,

$\Psi_t = (\psi_1, ..., \psi_n)$ - the set of evolution criteria from generation *t* to generation *t+1*

$\Sigma_t = (\sigma_1, ..., \sigma_n)$ - the set of solutions at each generation *t*, called morphological product or morphologies at generation *t*,

$\iota : \varphi_i \rightarrow \{true, false\}$ - a termination criterion for EM.

The set of final objectives represents the final solution's required characteristics, which could be expressed by any kind of formal language: mathematical (equalities, inequalities), semantic (textual description), graphical (parametric design), etc. We can see that the final objective is not necessarily expressed by an objective function as in genetic algorithms. EM is not an optimization technique; it is an approach to generating the total set of possible solutions matching the complete set of final objectives. EM operates preferably with non-quantified characteristics or with quantified characteristics limited to a low number of discrete values.

The set of primary elements represents non-quantified "protoelements" (from the Greek "*proteios*" or primary and element), used to build the solutions along the morphological evolutionary process.

The primary elements belong to homogeneous groups, called primary groups, which could be: a) material parts (subsystems) of the final system, b) geometrical, physical or chemical characteristics, c) geometrical, physical, chemical or organizational principles, d) natural phenomena, e) human behaviour, f) interrelations among certain parts, characteristics, principles, natural phenomena and human behaviour.

This list of primary groups is non exhaustive. It could be completed with problem-specific groups. The protoelements are non-quantified elements. For instance, the distance, as a geometrical characteristic, could be a non-quantified primary element when it is important as a qualitative dimension and not as a specific value with a specific unit of measurement. The distance, as a primary element, is equivalent to a length in parametric design. In conceptual design what really matters is to consider or to ignore the existence of this distance as a design parameter. The absolute value and the measurement unit are not important at this design stage.

The primary elements represent the evolving unit in EM. They are equivalent to a cell structure's nucleic acids and proteins. EM is equivalent to chromosome generation, a common gene pool included in the genotype of the individuals. The genetic algorithms manipulate the genotype of the initial population to find the "best" individual. EM generates the initial population genotype.

The identification and study of genes are of great interest to biologists. All genes of an organism, distributed over several chromosomes, are summarized under the term genome. The identification and study of primary elements in EM is equivalent to the generation of the "genome of the technical system" in engineering conceptual design. This generation could help both in understanding and avoiding functional flaws as well as in diversifying and improving technical systems designed by the engineer.

A primary element has a monary or a binary structure. The monary structure indicates that the primary element (ε_i) maintains the same feature. The binary structure indicates that the primary element has two different features. In conceptual design we often use ($\varepsilon_i = 0$ or $\varepsilon_i \neq 0$) which correspond to the lack ($\varepsilon_i = 0$) or the presence ($\varepsilon_i \neq 0$) of a feature in a solution. Any other two distinct features could also be used; for instance the parallel and perpendicular geometric position ($\varepsilon_i \Rightarrow ||$ or $\varepsilon_i \Rightarrow \perp$). The solutions exclusively built up with primary elements featuring monary structure are equivalent to haploid organisms in evolutionary theory. The solutions exclusively built up with primary elements featuring binary structure are equivalent to diploid organisms. Haploid organisms are modelled by one set of genetic information per individual which is unicellular from a biological point of view. In the diploid organisms (e.g., in the case of humans) each body cell includes two sets of chromosomes, such that for each chromosome two homologous forms exist. The structure of both sets of chromosomes is identical, but their genetic information contents may be different. For diploid organisms, corresponding genes in homologous chromosomes are called alleles and they may be identical (homozygotic) or different (heterozygotic). The primary element with a monary structure in EM is homozygotic for a particular gene in evolutionary theory. A primary element with a binary structure is heterozygotic for the particular gene. Both alleles are carried by the primary element, as in genetic material of the individual, but if one is dominant, only that one will be expressed. The other allele is called recessive.

The morphological operators are: (re)combination, mutation, migration and selection. These operators are deterministic and can be applied at each generation t of EM.

Morphological combination is applied directly to the primary elements to obtain the complete set of solutions at the first generation. Morphological recombination is used to combine solutions of generation t with solutions of the previous generations, including the primary elements, to obtain the complete set of solutions at generation $t+1$. We use the morphological box to generate the total set of possible solutions at each generation. In EM, (re)combination is a deterministic operator and not a probabilistic one as in genetic algorithms. The mechanism of morphological (re)combination is combinatorial. The evolution is made by adding up parts/characteristics to the previous morphology and not by interchanging (crossing-over) characteristics as in intrachromosomous recombination (crossover) specific to genetic algorithms. The (re)combination in EM is equivalent to chromosome generation in ontogenesis.

We know that chromosomes contain the genetic blueprints for a specific organism. Several thousand genes are arranged in a single line on a chromosome, a threadlike structure of nucleic acids and protein. The variation present in individuals is a reflection of the genetic recombination of these sets of chromosomes from generation to generation in phylogeny.

Morphological mutation modifies variability in the process of solution generation, as in genetic algorithms; the difference is that morphological mutation is a deterministic operator (all mutations are predefined) and not a probabilistic one. Morphological mutations act in two ways:

a) increasing variability by interchanging a dominant allele by the recessive one,

b) diminishing variability by changing a binary primary element into a monary one (by cutting up the dominant or the recessive allele).

The morphological migration acts by interchanging a dominant primary element (ε_i) by a new dominant primary element (ε_j). The primary elements must belong to the same group of homogeneous elements. The morphological migration, such as the mutation in genetic algorithms, introduces more variability in the solution generation process.

The morphological selection acts to eliminate the non-evolving intermediary solutions and the incompatible final solutions. The non-evolving intermediary solutions are the solutions in a generation t that did not evolve with respect to the previous generation. The incompatible final solutions are the solutions which do not match the set of final objectives. The set of evolution criteria from generation t to $t+1$ allows us to eliminate the non-evolving solutions at generation t and to stop spreading that to the following generations.

procedure EM
begin
  initialize $\Phi = (\varphi_1,...,\varphi_n)$
  $t \leftarrow 0$
  initialize $E = (\varepsilon_1,...,\varepsilon_n)$
  while (not termination condition) do
  begin
    $t \leftarrow t+1$
    combine $E = (\varepsilon_1,...,\varepsilon_n)$ with $\Sigma_{t-1} = (\sigma_1,...,\sigma_n)$ to obtain
      $\Sigma_t = (\sigma_1,...,\sigma_n)$
    evaluate $\Psi_t = (\psi_1,...,\psi_n)$
    modify $\Sigma_t = (\sigma_1,...,\sigma_n)$ by mutation
    modify $\Sigma_t = (\sigma_1,...,\sigma_n)$ by migration
    eliminate non-evolving solutions from $\Sigma_t = (\sigma_1,...,\sigma_n)$
  end
  procedure mutation
  procedure migration
end

Figure 1. Structure of evolutionary morphology

The set of solutions is also called morphology at generation *t*. The set of solutions in the final generation represents the result of EM called morphological product. It corresponds to the first generation that satisfies all the final objectives. When the morphological product is obtained, the termination criterion stops the evolution algorithm.

The general structure of EM is shown in Fig. 1.

EM enables us to make discoveries and inventions in a systematic way, and it is a guide to the exploration of the basic evolutionary interrelations among conceivable families of objects and physical, chemical and biological phenomena that govern their interactions. We can see that EM could be considered as genetic manipulation of the "genome of the technical system".

4. SERIAL 6R ROBOTS INVENTIVE DESIGN WITH ONLY ISOLATED SINGULARITIES

We present the application of EM on a concrete conceptual design example. We have chosen the structural design of a new class of serial robots matching the following set of final objectives $\Phi = (\phi_1,...,\phi_6)$:

ϕ_1 - robotic manipulator with a serial kinematic chain; only one elementary open kinematic chain in which the extreme links are linked by only one joint and the intermediary links are linked by two joints,

ϕ_2 - anthropomorphic robotic manipulator; only revolute joints in robot structure - *R*,

ϕ_3 - orthogonal robotic manipulator; with perpendicular ($\perp$) or parallel (//) relative position of axes of two revolute pairs adjacent to the same link,

ϕ_4 - six degrees of freedom ($M=6$),

ϕ_5 - no structural redundancies; the degree of spatiality/connectivity must have the same value as the degree of mobility, i.e $M = S = rank(J)$, where J is the jacobian matrix mapping the joint vector space and the operational vector space,

ϕ_6 - only isolated singularities; the determinant of the Jacobian matrix can be factorised in functions of one single joint variable, that is $det(J) = \prod_{i=1}^{4} f_i(q_i)$, where $f_i(q_i)$ is a function of single joint variable q_i.

The final objectives are either non-quantified or quantified by the mathematical conditions above mentioned.

Serial orthogonal anthropomorphic robotic manipulators with six degrees of freedom are the most commonly used industrial robots. The standard structural synthesis approach to serial six degrees of mobility robotic manipulators is based on the concatenation of a positioning kinematic chain (called the arm) and an orientation kinematic chain (called the wrist). The use of these two distinct kinematic chains considerably reduces the number of structural solutions currently used in industrial robots. To overcome this inconvenience, we have used EM successfully. It gave the author the opportunity to design a new class of anthropomorphic robots featuring but isolated singularities. This capacity was recently presented for the first time in literature [15]. By using EM, the author achieved the structural synthesis of 126 serial orthogonal anthropomorphic robotic manipulators with only isolated singularities. These new robots have very few singular positions, which is an important advantage in task planning and motion control. This advantage is corroborated with very simple structural solutions. No serial anthropomorphic robot with six degrees of mobility produced today by the robot manufacturing industry benefits from this important characteristic of having only isolated singularities.

We apply EM step by step to generate all inventive solutions of serial orthogonal anthropomorphic robotic manipulators with six degrees of mobility and only isolated singularities. In Fig. 2 we present the structural graph of a serial anthropomorphic robotic manipulator with six degrees of mobility. We can see that such a robot contains one fixed link (*0*) and six kinematic links (*1,2,…,6*) connected by six revolute joints (*R*). The relative positions p_i (*1,2,…,5*) of the axes of two revolute joints adjacent to a kinematic link *i* are orthogonal ($\perp$) or parallel (//) in the case of orthogonal robotic manipulators. Each kinematic element *i (i=1,2,..,6)* of a *6R* orthogonal robotic manipulator can have two dimensions [13]: the axial offset a_i, measured on the axis of the joint *(i-1,i)*, and the link length b_i defined by the length of the normal to the joint axes adjacent to element *i*. The dimensions a_i and b_i can be zero or non-zero.

By applying a simple combinatorial approach we obtain 2^{17}=131 072 potential solutions by combining the two positions (orthogonal or parallel) of p_i and the two values (zero or non-zero) of the dimensions a_i and b_i. Each potential solution must be checked with respect to quantified final objectives ϕ_4-ϕ_6 although few solutions match them. We can see that this combinatorial approach (morphological box type) combined with symbolic calculation of conditions defining ϕ_5 and ϕ_6 became deceptively complex. The main disadvantage of the morphological research by using the morphological box method is real although few parameters exist in this design problem - 17 parameters (p_i, a_i and b_i) with two generic values each. To bypass this disadvantage of morphological research we apply EM.

The set of primary elements, which represents the non-quantified "protoelements" used to build up the solutions by morphological evolutionary process, is composed by : ε_{1i} - revolute joint (R_i) between the links *i -1* and *i*, ε_{2i} - dimension a_i of the kinematic link *i*, ε_{3i} - dimension b_i of the kinematic link *i*, ε_{4i} - relative position p_i of the axes of two revolute joints adjacent to a kinematic link *i*.

The revolute joint belongs to a primary group representing material subsystems of the robotic manipulator (the joints). Dimensions a_i and b_i belong to a primary group representing geometrical characteristics of the links and the relative position p_i belongs to a primary group representing interrelations among the axes of the revolute joints.

We note that the protoelements are non-quantified elements. Here the distance is important as a qualitative dimension and not as a specific value with a specific unit of measurement.

We consider the following primary elements with monary structure: $\varepsilon_{1i}=1$, $(i=1,\ldots,6)$, $\varepsilon_{21}=0$, $\varepsilon_{26}=0$, $\varepsilon_{36}=0$. These monary features indicate that only revolute joints exist in an anthropomorphic robotic manipulator as well as the fact we always consider $a_1=0$, $a_6=0$ and $b_6=0$. This is possible due to a proper choice of reference and final frames used in kinematic modelling and Jacobian matrix calculation. The primary elements with monary structure are not affected by the morphological mutation operators. They contribute to the reduction in the number of non-evolving solutions.

The following primary elements have binary structure:

$\varepsilon_{2i}=0$ if $a_i=0$ and $\varepsilon_{2i}=1$ if $a_i \neq 0$, ($i=2,\ldots,5$),

$\varepsilon_{3i}=0$ if $b_i=0$ and $\varepsilon_{3i}=1$ if $b_i \neq 0$, ($i=2,\ldots,5$),

$\varepsilon_{4i}=0$ if $p_i \Rightarrow ||$ and $\varepsilon_{4i}=1$ if $p_i \Rightarrow \perp$.

We consider $\varepsilon_{ji}=1$ as dominant and $\varepsilon_{ji}=0$ as recessive alleles

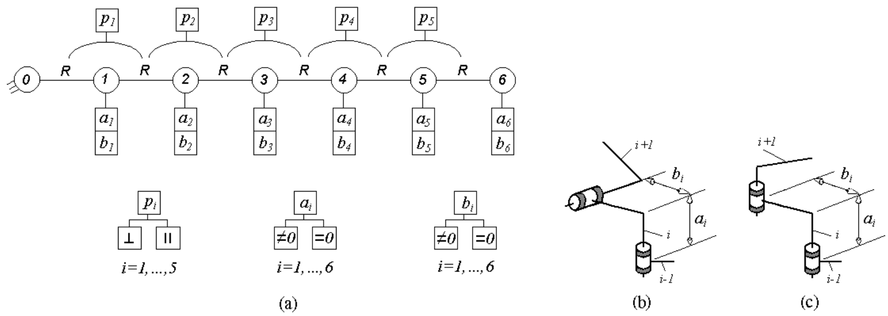


Figure 2. Structural graph of a serial anthropomorphic robotic manipulator with six degrees of mobility (a) and the primary elements of the link *i* (b) and (c)

($j=2,3,4$ and $i=2,...,5$). Fig. 2 (b) and (c) illustrates the geometrical meaning of the primary elements defined for the link *i* by $(\varepsilon_{1i}=1,\varepsilon_{2i}=1,\varepsilon_{3i}=1,\varepsilon_{4i}=1)$ and $(\varepsilon_{1i}=1,\varepsilon_{2i}=1,\varepsilon_{3i}=1,\varepsilon_{4i}=0)$.

The evolutionary process from generation *t* to generation *t+1* ($M_t<6$) is analyzed by the set of evolution criteria $\Psi_t=(\psi_1,\psi_2)$, where ψ_1 is a mobility criterion defined by $M_t<M_{t+1}$ and ψ_2 is a spatiality criterion defined by $S_t<S_{t+1}$. The solutions of the generation *t+1* that do not match those two criteria are eliminated from the evolutionary process. This allows us to avoid the proliferation of non-evolving solutions in the next generations. These evolution criteria are applicable to the generations with $M_t<6$.

The evolutionary process from generation *t* to generation *t+1* with $M_t=6$ is analyzed by the spatiality condition $S_t=S_{t+1}=6$, that can be considered a non-involution criterion. In fact, this criterion indicates that the process is stationary.

The set of morphological operators is composed by three operators: ω_1 - (re)combination, ω_2 - selection and ω_3 - mutation.

Morphological combination is applied directly to the primary elements to obtain the complete set of solutions at the first generation. At this stage, we consider only the dominant alleles of the primary elements. Evolution is achieved by adding parts (revolute joints - *R*) and geometric characteristics (a_i, b_i and p_i) to the previous morphology. At least three generations are necessary to evolve from the primary element with $M=S=1$ (Fig. 2,b) to a serial anthropomorphic robotic manipulator with $M=S=6$ (Fig. 3,a). At each generation ($M_t<6$) the solution matches the evolution criteria ψ_1 and ψ_2.

We first apply the mutation operator to the solutions with $M_{t=3}=S_{t=3}=6$ obtained at the end of the recombination process. This operator interchanges the dominant allele $\varepsilon_4=1$ with the recessive one $\varepsilon_4=0$. By applying this operator, we obtain a fourth generation with $2^5=32$ solutions having $M_{t=4}=6$. Only 23 solutions match the spatiality criterion ψ_2. The 9 incompatible solutions with $S_{t=4}$<6 are eliminated from the evolutionary process. The remaining 23 solutions of the 4th generation comply with the evolution criteria and implicitly the final objectives -yet they do not match the final objective ϕ_6.

The solutions with two and three consecutive parallel joint axes of the 4th generation are subjected to mutation operators which change some predefined binary primary elements into monary ones. The 5th generation still contains 23 solutions with the following extra monary structure primary elements: $\varepsilon_{2i}=0$ ($a_i=0$) if $\varepsilon_{4i}=0$ ($p_i \Rightarrow ||$), $i=2,...,5$ and $\varepsilon_{3i}=1$ ($b_i \neq 0$) if $\varepsilon_{4i}=0$ ($p_i \Rightarrow ||$), $i=1,...,5$. New mutation operators are then applied to the 23 solutions with $M_{t=5}=S_{t=5}=6$ to interchange the dominant alleles $\varepsilon_{2i}=1$ and $\varepsilon_{3i}=1$ with the recessive ones ($\varepsilon_{2i}=0, \varepsilon_{3i}=0$) in the remaining primary elements with binary structure. By applying these two operators, we obtain a final generation with 1528 solutions complying with to the final objective, 1450 solutions matching to the final objectives ϕ_1-ϕ_6 and 126 solutions in compliance with all final objectives. The selection operator checking the compliance with the final objective has ruled out no fewer than 1402 solutions. The set of 126 solutions represents the final morphologies of new architectures of serial orthogonal anthropomorphic robotic manipulators with six degrees of mobility and only isolated singularities. These new architectures belong to 5 groups of solutions: 40 solutions with three parallel joint axes, 34 solutions with two parallel joint axes, 30 solutions with no parallel joint axes, 18 solutions with three parallel joint axes plus a pair of two parallel axes, 4 solutions with two pairs of parallel axes. Fig. 3,b presents an example belonging to the second group [16]. This solution has $b_1 \neq 0$, $b_2 \neq 0$ and $b_5 \neq 0$. The solutions of the fourth group have also been published recently [15]. The other solutions will be published in subsequent papers.

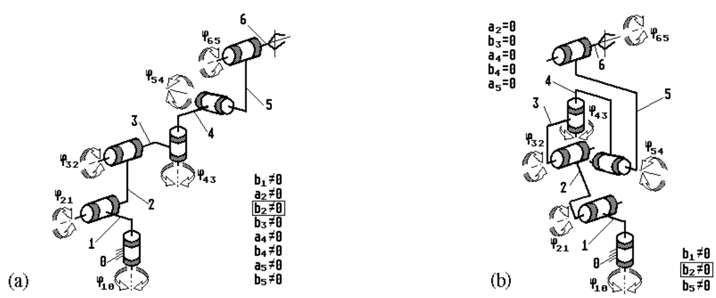


Figure 3. General architecture of a serial 6R orthogonal anthropomorphic robotic manipulator (a) and an architecture with only isolated singularities (b)

5. CONCLUSIONS

EM is based on the fundamental scientific method of alternating between analysis and synthesis. For this reason, it can be trusted as a useful, non-quantified method for investigating inventive problems, which cannot be solved by formal mathematical methods, causal modelling and simulation. Furthermore, EM gives a fairly clear “audit trail”, which makes the judgment processes relatively traceable, reproducible and inventive. EM “solution space” consists of the complete subset of configurations matching some criteria (final objectives). Of course, as is the case with everything else, the output of an EM depends on the quality of its input. The final solutions depend a lot on the choice of primary elements. This choice is a problem-dependent task. The main advantage of this approach consists in finding the complete set of possible solutions in an evolutionary sequence based on a set of protoelements. EM may help us to discover and to invent new relationships or configurations, which may not be that obvious, or which we might have been overlooked by other – less structured – methods. By using EM, the author has “discovered” the possibility to design serial orthogonal anthropomorphic robotic manipulators with only isolated singularities. Most importantly, it encourages the identification and investigation of boundary conditions, i.e. the limits and extremes of different contexts and factors. The evolutionary sequence limits the generation and the proliferation of incompatible solutions.

The main EM paradigms are inspired from the synthetic theory of evolution developed by modern biology. EM integrates some features of morphological research and evolutionary algorithms but it also represents a radically different approach. Morphological approaches are based especially on static morphosis, the evolutionary algorithms on dynamic morphosis and EM on evolutionary morphosis.

The morphological box represents a static morphosis based on a simply combinatorial approach. However, Zwicky’s highly systematic approach to this field should not be underestimated. Used properly – and on the right types of problems – the method is both deceptively complex and rich.

Evolutionary algorithms start from a given initial population to obtain an optimum solution with respect to a fitness function. EM creates this initial population to enhance the chance of obtaining an optimum “more global” by non-quantified diversification of the initial population. EM is a new approach acting prior to the application of any evolutionary algorithm. Genetic information is quantified in evolutionary algorithms and non-quantified in EM. Evolutionary algorithms are optimized oriented methods; EM is a conceptual design oriented method. EM and evolutionary algorithms are complementary methods.

EM is formalized by a 6-tuple of final objectives, primary elements, morphological operators, evolution criteria, morphologies and a termination criterion. EM was gradually developed by the author of this paper in the last 20 years and it was applied especially to structural synthesis of new mechanisms used in machine-tools, robots, car suspensions, published by the author in more than 100 articles and 30 patents protected in about 80 countries. In spite of this long period of development and application [14], the author has generalized and formalized EM very recently. This is the first paper in which the author presents the main paradigms and the formalization of EM as a structured approach to inventive engineering design.

REFERENCES

1. T. Ritchey, General morphological analysis: A general method for non-quantified modelling, 1-th EURO Conference on Operational Analysis, Brussels, 1998.
2. G. Gogu, Méthodologie d'innovation: la résolution des problèmes créatifs, Revue française de Gestion industrielle, 19(3), 35-62 (2000).
3. F. Zwicky, Discovery, Invention, Research (Macmillan, New York, 1969).
4. K. Bailey, Typologies and Taxonomies - An Introduction to Classification Techniques (Sage Publications, Thousand Oaks, 1994).
5. K.W. Norris, The Morphological Approach to Engineering Design, in: Conference on Design Methods, edited by J. C. Jones and D. G. Thornley (Pergamon Press, Inc., New York, 1963), pp. 115-140.
6. L. Gerardin, Morphological Analysis: A Method for Creativity: A Guided to Practical Technological Forcasting (Prentice-Hall, New Jersey, 1973).
7. D.E. Goldberg, Genetic algorithms in Search, Optimisation and Machine Learning (Addison-Wesley, Reading, MA, 1989).
8. D. Dasgupta, Z. Michalewicz (Eds.), Evolutionary Algorithms in Engineering Applications (Springer-Verlag, Berlin Heidelbeerg, 1997).
9. J.H. Holland, Adaptation in Natural and Artificial Systems (University of Michigan Press, Ann Arbor, 1975).
10. Z. Michalewicz, Genetic Algorithms + Data Structures=Evolution Programs (Springer-Verlag, 3rd edition, 1996).
11. T. Bäck, Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms (Oxford University Press, New York, Oxford, 1996).
12. H.-P. Schwefel, Evolution and Optimum Seeking (John Wiley, Chichester, UK, 1995).
13. G. Gogu, P. Coiffet, A. Barraco, Représentation des déplacements des robots (Hermès, Paris, 1997).
14. F. Dudiță, D.V. Diaconescu, G. Gogu, Mecanisme articulate : inventica si cinematica in abordare filogenetica (Ed. Tehnica, Bucuresti, 1989).
15. G. Gogu, Families of 6R orthogonal robotic manipulators with only isolated and pseudo-isolated singularities, Mechanism and Machine Theory, 37(11), 1347-1375(2002).
16. G. Gogu, P. Coiffet, Structural synthesis of anthropomorphic robotic manipulators with only isolated singularities, 35th International Symposium of Robotics, Paris, 23-26 April 2004.

COMPUTER ASSISTED DESIGN OF ACTUATORS FOR HIGH PRECISION ADJUSTMENT IN MICRO TECHNOLOGY

Hinnerk Hagenah, Thomas Wurm and Manfred Geiger
Chair of Manufacturing Technology, University of Erlangen-Nuremberg, Egerlandstraße 11, 91058 Erlangen, Germany

Abstract: Micro systems technology is expected to play a pivotal part in future product development both in scientific and industrial fields. Decreasing size and tolerances make mounting and joining operations quite a challenging task. Laser adjustment is a suitable technology to achieve the necessary accuracy by processing specifically designed actuators on which the functional elements have been mounted. Nowadays, the specifically tailored design of these actuators is performed by experts through an iterative process. In this paper we will present the concept and parts of the design and development of a computer based assistance system helping the designer define new actuator geometries. We will introduce a data model to describe already existing actuator geometries and summarise them in a construction catalogue. The adjustment task at stake is described by means of kinematic chains. The design of actuator geometry therefore boils down to the design of a kinematic chain. A suitable kinematic chain is sought for and computed using the Denavit-Hartenberg method which has been extended to handle closed kinematic chains as well.

Key words: Actuator design, Manufacturing processes, Laser adjustment, Micro systems technology

1. INTRODUCTION

Both scientific and industrial experts agree that micro systems technology is one of the future key technologies. It will be the pre-requisite for the development of new product generations. Micro systems technology aims at integrating electronic, optical and mechanical features into micro components and workpieces. Besides the ever increasing pressure on prices

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 403-416.

explains the importance of this technology in terms of the rising demands on the part of customers who keep demanding small high functionality low price products.

Therefore a necessary condition to make the most of the potentialities of micro system technology is the development of comparably inexpensive but reliable concepts and technologies for hybrid micro systems manufacture and assembly. For this purpose an optimisation and development of all the technologies involved has to be carried out. Special attention has to be paid to the adjustment processes. Due to miniaturisation it is becoming very difficult and expensive to assemble workpieces within tolerances. Often an adjustment step is needed to ensure product quality. The adjustment process should be flexible to be applicable to a large variety of products. In addition, it is desirable to use a process that might be automated. Metallic material laser beam forming is a process that satisfies these demands. The integration of laser beam forming into the process chain of adjusting preassembled systems leads to a new technology called laser adjustment that allows the highly accurate and inexpensive production of high precision parts [1].

Basically, laser adjustment replaces the highly accurate assembly process with a less accurate and thus faster and less expensive assembly step and a subsequent highly accurate laser adjustment process. One or more specifically designed actuators are a prerequisite for the laser adjustment step. Through planned irradiation of these actuators local plastic deformations can be achieved which lead to the functional element position adjustment. A major advantage of this procedure is its ability to be controlled by functional signals in both static and moved or closed systems [2].

The design of assemblies with integrated actuators is very challenging for both the design of the geometry and the process. Complex interdependencies of process, geometry and material have to be taken into account. Especially the definition of a suitable actuator providing the degrees of freedom needed for the adjustment process while fitting into the space available and having the required stability is quite a demanding task. So far human experts have to design the actuators depending solely on their intuition [3] in an iterative process that is time consuming and thus cost intensive. So far there has been no computer assistance specifically developed and suited for this task available. The missing system should take into account the boundary conditions resulting from the space available as well as the demanded stability. It should then guide and assist the designer in finding an actuator geometry that can fulfil the required adjustment task. The remainder of this paper will describe the developed methods and concepts in some detail.

2. LASER BEAM FORMING

Laser beam forming is based on the base of the introduction of thermal stresses into sheet metal workpieces through local heating. Through the stresses and the heating the flow limit is reached and plastic deformation occurs. Since the ambient material prevents lateral expansion local compressive stresses are introduced. When cooling, this area shrinks and the corresponding stresses deform the workpiece. A detailed description and analytical investigation of laser beam forming can be found in [2], [4]. Depending on geometry and irradiation strategy this basic mechanism leads to different forming mechanisms. Today four different forming mechanisms are distinguished [5]. The fundamental mechanisms applied in laser adjustment are the temperature gradient mechanism (TGM) and the upsetting mechanism (UM). Both are depicted in Fig. 1.

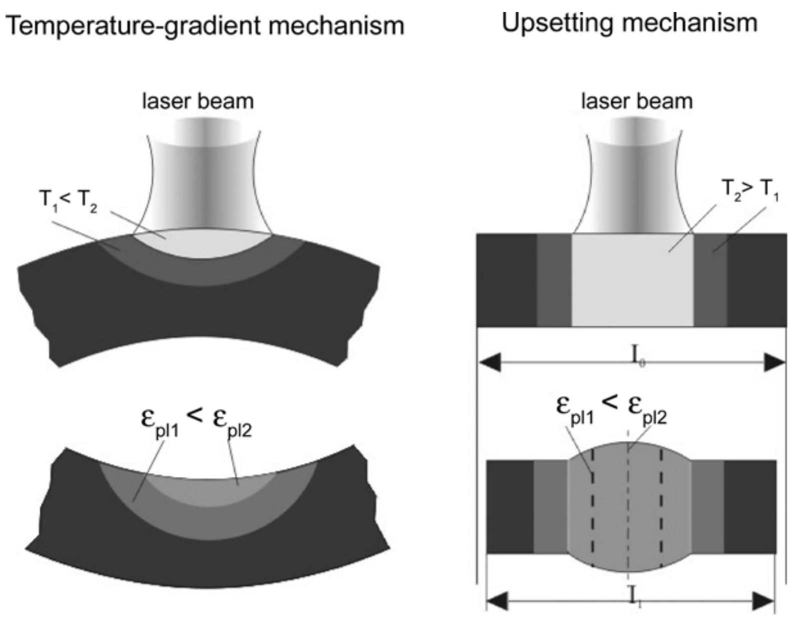


Figure 1. Two basic mechanisms of laser beam forming

The TGM as depicted on the left half of Fig. 1 is the most used of the existing mechanisms for laser beam forming. It is effective whenever a significant temperature gradient is introduced between the upper side of the workpiece facing the laser source and the lower side. To achieve this temperature gradient, taking into account the usually satisfactory metal heat conductivity, short irradiation times are needed. The workpiece usually is irradiated along a straight line. The width of the upset area on the upper side depends on the applied temperature distribution. It is usually of the same magnitude as the sheet thickness. Details on the effects leading to the bending towards the laser are described in [4]. The potential of the TGM for adjustment tasks is already in use in micro optics [6].

The UM as shown on the right hand side of Fig. 1 is effective whenever a homogeneous temperature field is given over the sheet thickness and the width of the plastic area is similar to the sheet thickness. So far the UM has been mainly used in tube and extrusion forming. Lately it has also been introduced to complement TGM into adjustment as a second mechanism [7].

The geometry resulting from laser beam forming is influenced by many parameters. They can fall into three different groups: material/laser/ part geometry parameters. Heat conductivity, density, heat capacity, thermal expansion coefficient, surface quality, inherent stress condition, state structure as well as yield stress are material parameters. Laser parameters such as laser power, path feed rate, beam diameter and laser operational mode are of special importance. Not only are they are operator dependent but they also determine the forming mechanism, since they cause the temperature field development within the interactive area. Sheet thickness, irradiated path length, bend area geometry and actuator overall are noted as geometry parameters. These are the parameters likely to be influenced by the designer. They should be optimised to the adjustment task at stake.

3. ASSISTANCE SYSTEM CONCEPT

As mentioned before, the construction of such an optimised actuator that fulfils the geometric constraints of the surrounding workpiece while providing the necessary degrees of freedom for the desired adjustment step places high demands on the designer’s technical skills and intuition. A computer-aided system that assists the designer has to allow for the constraints and demands applicable to the actuator geometry being looked for.

The challenge of defining the actuator geometry can fall into two parts. One part is to find a sequence of axes that can be activated by laser irradiation and provide the needed ability to move the functional element. The other part is to transform this sequence into a geometry that can then be manufactured. The first part can be modelled as the search for a suitable kinematic chain as known from robotics. A kinematic chain is a sequence of rotational and transitional axes with some bars in between. Modelling actuators this way allows the computation of the movement and thus the adjustment that can be carried out with them. To design an actuator that consists only of the necessary axes to carry out the required transformations is done by combining basic actuator geometries. These basic actuator geometries have to be defined. They are then described using the same abstractions that will be used to model the resulting actuator. The basic actuators are archived in a construction catalogue. The first task in designing an actuator specifically tailored to solve a problem can thus be expressed as looking for a combination of basic actuator geometries that provide the demanded adjustment capability while fulfilling constraints imposed by the workpiece properties. The computer system will then use the geometric descriptions of the combined basic actuator to provide a first sketch of the resulting geometry. The information on the locations to be irradiated and the irradiation needed for a requested forming will also be provided by the system based on the archived information on the basic actuator types. The fine adjustment of this geometry that includes all the mandatory axes and bars must then be carried out by the designer. However the difficult task of combining the basic actuators is assisted by the computer system. The parts and their interoperation are depicted in Fig. 2. Due to their major importance for the overall function of the system, some very basic actuator types as well as the modelling and calculation of an actuator's abilities within the construction catalogue will be described in more detail.

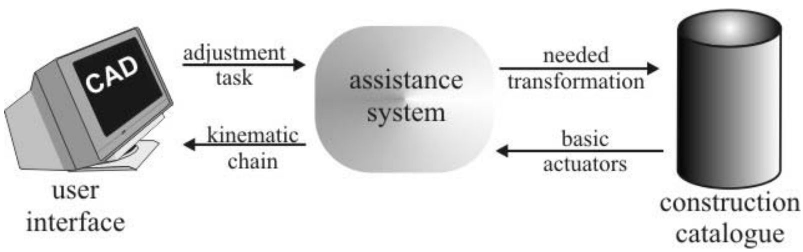


Figure 2. Models of the assistance system and their interaction

4. BASIC ACTUATOR GEOMETRIES AND THEIR REPRESENTATION

4.1 Classification of basic actuators

For the design of actuator systems an exact description of the basic actuators is necessary. A basic actuator is defined as the minimum unit, which leads to a controlled displacement or bend by means of laser irradiation. Fig. 3 shows the four basic types: the bridge and angle actuator with possible displacement in positive x-direction, the double-bridge and angle actuator with embossment, which allow a displacement both in positive and negative x-directions.

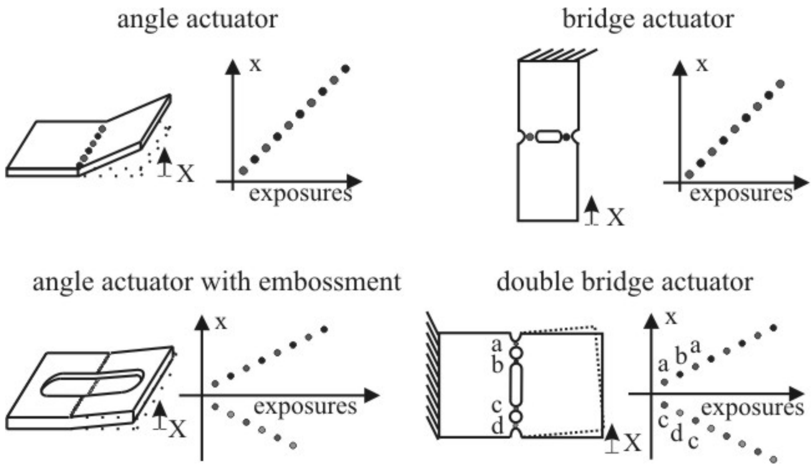


Figure 3. Basic actuator types and possible movement

The angle actuator is simply represented by a stripe which can be irradiated by single pulses, scans or lines. Bending occurs depending on the bending moment. Fig. 3 illustrates the bending angle as a function of the pulse number. If the irradiation side is not changed the angle actuator is limited to the adjustment in one direction. Therefore the angle actuator with embossment also shown in Fig. 3 has been developed. The embossment shifts the position of the neutral bending line of the sheet. Therefore it is possible to adjust towards or away from the laser position depending on whether the laser pulse is located inside or outside the embossment. The bridge actuator presents excellent linear behaviour at multiple irradiations. Applying the first pulse on one strip of the bridge generates compressive stresses in the irradiated area resulting into tensile stresses at the adjacent strip. During the cooling phase tensile stresses arise in the irradiated strip and therefore generate compressive stresses in the unprocessed strip. Therefore the pulse on the second stripe starts with a highly effective compressive stress field, since the yield point is reached quite fast. The procedure described can be repeated many times. A limitation is given by the initial strip length. The double bridge actuator is based on the same mechanism as the bridge actuator. The most important difference between both is that the double bridge actuator is not limited to the adjustment in one direction.

Further actuator geometries that have been developed and used for industrial applications can be generated by combining these basic actuators. To assist the designer in defining an actuator geometry for the needed adjustment, the effects of the actuator have to be mathematically formalised. This enables the combination of several actuator geometries already archived in the construction catalogue with a new actuator. The possible adjustment with this actuator can then be calculated.

4.2 Mathematical description of the actuator behaviour

As mentioned before, an actuator system can be described as a kinematic chain of single geometric elements. The basic actuators can be regarded either as joints in the out-of-plane (TGM) case or as guides for the in-plane case, depending on the effect they use. Thus it is possible to describe an actuator by means of a bar model as a multi-body system with *n* partial bodies. The coordinate system describes a single partial body initial position coordinates.

Each single actuator can be regarded as an open kinematic chain or as a tree structure. For the recursive description of the actuator's orientation the well known Denavit-Hartenberg method [8], [9] can be applied. With this method, a base B_i is assigned to each bar. The relative position of base B_i to base B_{i-1} is described by the Euler angles ψ_i and θ_i and the line segments u_i and v_i. Angle ψ_i is the rotation around the z-axis of base B_{i-1}. Angle θ_i results from the rotation of base B_{i-1} around the new x-axis of base B_{i-1} until the z-axes of the bases *i* and *i-1* coincide. The segment lines u_i and v_i are perpendicular to each other. v_i is defined as the length of the linking straight line of the generally curved joint axes *i-1* and *i*. The distance between the intersection of this linking straight line and base B_{i-1} is u_i. Using these definitions each angle in the actuator, either an edge, which was produced by the manufacturing process of the actuator, or an irradiation position for laser beam application in the out-of-plane basic form, can be described by a rotational matrix T_i as given in Eq. (1).

$$T_{i-1,i} = \begin{bmatrix} \cos\psi_i & -\sin\psi_i & 0 \\ \sin\psi_i & \cos\psi_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\vartheta_i & -\sin\vartheta_i \\ 0 & \sin\vartheta_i & \cos\vartheta_i \end{bmatrix} \quad (1)$$

With rigid forming edges angles ψ_i and θ_i are fixed. The angles resulting from laser beam application depend on the initial angle present before any irradiation, the number of irradiation n_i at any given spot *i* and a transfer coefficient C_i which depends on process parameters and geometric details. The present angle at any given state of the actuator is then given as $\psi_i = \psi_{i,init} + C_i\, n_i$ and $\theta_i = \theta_{i,init} + C_i\, n_i$. C_i. This approach uses the linear behaviour as shown in Fig. 3, which can be assumed for the number of irradiations that are to be applied to a given actuator. Rigid bars, e.g. the transformation from B_0 into B_1 in Fig. 4, which only connect two active elements, are described by a translation vector t_i defined in Eq. (2).

$$t_{i-1,i} = \begin{bmatrix} v_i \cos\psi_i \\ v_i \sin\psi_i \\ u_i \end{bmatrix} \tag{2}$$

For active bars like B_1 in Fig. 4 two cases have to be distinguished. Either the active axis coincides with the rotational axis *i-1*. In this case u_i has to be replaced by $u_i = u_{i,init} - C_i \cdot n_i$. Or, if this is not the case, v_i has to be expressed by $v_i = v_{i,init} - C_i \cdot n_i$. Through the subtraction of $C_i \cdot n_i$ the shortening effect of the UM is described.

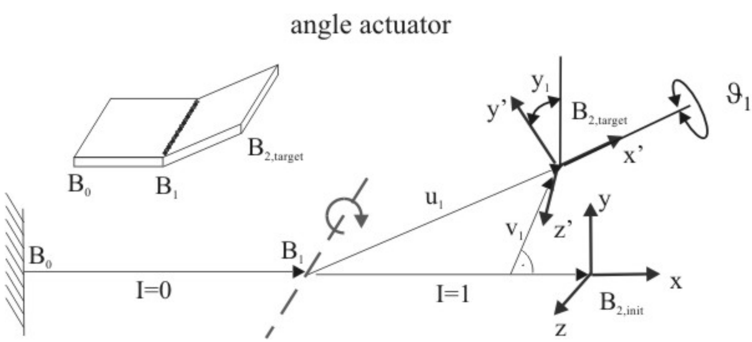


Figure 4. Kinematic chain of the angle actuator

Fig. 4 shows a description of the angle actuator by means of kinematic chains. This actuator provides a bending which transforms the initial coordinate system of the functional element (x,y,z) to the target frame coordinate system (x',y',z'). For this, a translation in terms of u_i and v_i of the frame B_0 is indicated. Furthermore, a rotation around the z-axis needs to be modelled as well in order to approach the target frame $B_{2,target}$. Consequently, the value of ψ_i must be greater than zero. In contrast θ_i is equal to zero keeping in mind that no rotation around x-axis can be achieved by the irradiation of an angle actuator (see Fig. 3).

In order to describe the behaviour of an actuator geometry in a general way, the rotational matrices T_i and the translation vector t_i are combined in the matrix D_i. Considering the formulae presented in [8] the total matrix for the actuator can be drawn up by Eq. (3).

$$D_{i-1,i} := \begin{bmatrix} T_{i-1,i} & t_{i-1,i} \\ 0 & 1 \end{bmatrix} \tag{3}$$

While the Denavit-Hartenberg method works for open kinematic chains, the assumption of actuators being connected to the surrounding workpiece at one point like robots does not hold for real world applications. Thus closed kinematic chains have to be taken into account. Therefore the Denavit-Hartenberg method has been extended to closed kinematic chains as depicted in Fig. 5.

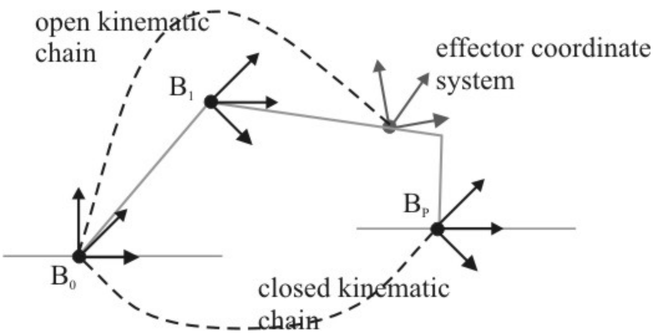


Figure 5. Extension of method of Denavit-Hartenberg for closed kinematic chains

To handle the closed kinematic chain a virtual basis B_P has been introduced into the system. Therefore we get an open kinematic chain connecting the basis of B_0 and the basis of the effector's coordinate system and a closed kinematic chain connecting B_0 and the virtual basis. For the calculations it is important to have B_P unchanged after finishing. This is achieved by thinking of B_P as a clamping device that is attached to some fixed point and thus cannot be changed throughout the simulations. Thus the calculation of a closed kinematic chain can basically be reduced to the calculation of an open kinematic chain.

5. REPRESENTATION OF THE DESIGN PROBLEM

The basic function an actuator has to provide is the movement of the object that has to be adjusted from an initial position to a target position. The frame for all initial positions is provided by the manufacturing and handling accuracy. The frame of the target positions is given by the tolerances of the workpiece (Fig. 6).

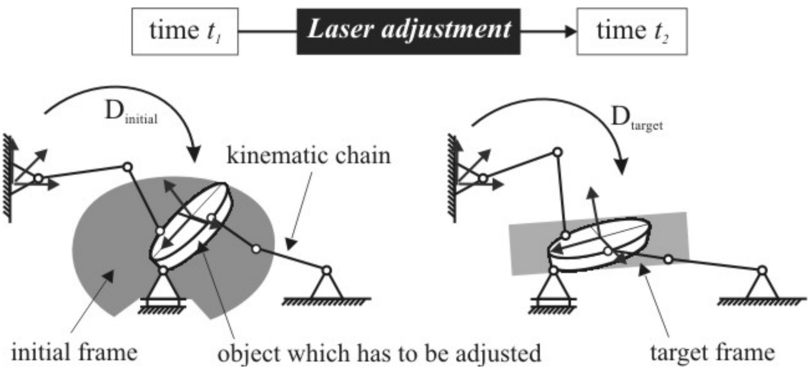


Figure 6. Transformation from the initial into the target frame

The construction engineer developing the actuator geometry has to ensure that it can move the object to be adjusted from any point within the initial frame to a position within the target frame. Describing actuators as kinematic chains and using Denavit-Hartenberg's matrices as shown above, the two positions can be described with Eqs. (4) and (5). T and t describe rotation and translation before ($T_{initial}$, $t_{initial}$) and after (T_{target}, t_{target}) irradiation of the actuator as defined in Eqs. (1) and (2). For the initial state of the matrices the number of irradiations in each position $n_i = 0$. The construction of a complex actuator geometry thus has been reduced to finding a combination of basic actuator geometries presented as $D_{i,i+1}$ that fulfil both Eqs. (4) and (5) with a certain combination of n_i.

$$D_{initial} = D_{0,1} \cdot \ldots \cdot D_{n-1,n} = \begin{bmatrix} T_{initial} & t_{initial} \\ 0 & 1 \end{bmatrix} \quad (4)$$

$$D_{target} = D_{0,1} \cdot \ldots \cdot D_{n-1,n} = \begin{bmatrix} T_{target} & t_{target} \\ 0 & 1 \end{bmatrix} \quad (5)$$

The resulting linear equation system can be solved using the numerical Newton procedure. The constructing engineer is thus given the basic actuator geometries from which a suitable kinematic chain can be built up. The linear equation system can be enlarged to incorporate further limitations e.g. available space or rigidity requirements.

6. APPLICATION OF THE SYSTEM

Fig. 7 shows an actuator with three bars, an active angle and a rigid angle. This actuator is used for the adjustment of a lens, its three dimensional shape being also depicted in Fig. 7.

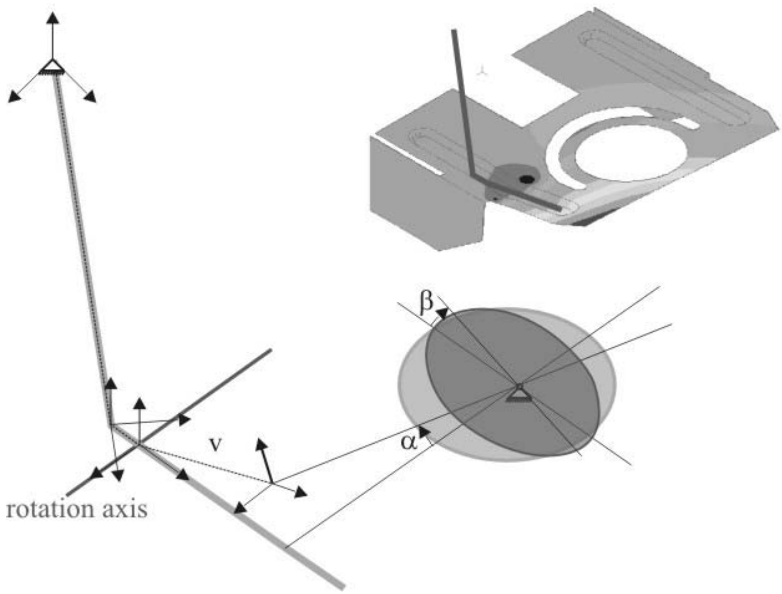


Figure 7. Kinematic chain for an actuator

The part of the actuator represented by the kinematic chain is also marked in the volume model. The kinematic chain has one rotational axis that will be taken into account for this example. The dimension in question is the length of bar *v*. All the other rotational axes and bars provide the ambient structure in order to fulfil the constructive constraints. Some points are to lie on a straight line thus limiting the design possibilities. The major constraint concerns the adjustment area, which has to cover the range from -0.75° to +0.75° in relation to α and β. As a minimum step size and thus a final 0.10°

precision is required [10]. All other dimensions are considered as fixed. Both angles α and β can be influenced by irradiating the indicated rotational axis. The magnitude of the influence depends on the length of *v* if all other parameters are kept constant. The actual actuator has more than one rotational axis and can vary in more than one bar length, but for the demonstration of the general function of the presented approach this degree of freedom is satisfying. Formulae for α and β can be derived from the geometric relationships formalised as Denavit-Hartenberg matrices for the target and the initial frame, shown in Fig. 7 by the grey shaded areas. By inserting different values for the length of *v* into these formulae the abilities of the actuator geometry can be examined through simulation.

The results obtained from simulating the actuator geometry behaviour have been plotted in Fig. 8. Three different values for the length of *v* have been investigated. The resulting adjustment possibilities in α and β are shown in Fig. 8. The requirement was a minimum step width of 0.10° and a working range from -0.75° to +0.75°. By calculation of the actuators reaction resulting from different irradiation pulses, a point field has to be set. To meet the requirements this point field should have a minimum point distance <= 0.10° and size of target space >= 0.75°.

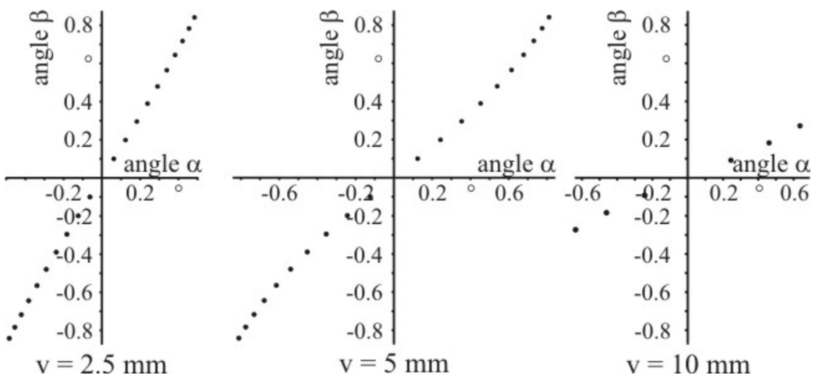


Figure 8. Transfer behaviour of an actuator dependent on the bar length, *v*

It can be seen that with a bar length of *v*=5mm an almost linear behaviour can be achieved with simultaneous equivalence of the size of space in α and β. The demanded range is therefore covered. Reducing the length *v* to just 2.5mm improves linearity but also has the effect that the size of space in the α-direction does not meet the requirements any longer. If the bar length is further increased to *v*=10mm the minimum step size, which can be achieved by laser adjustment with a pulse in α-direction, is far above the demanded value of 0.10°. Thus from the calculation it can be seen that 5mm is a satisfying solution for the length of bar *v*. Furthermore, the irradiation pulses and position are also known from the calculations in order to carry out the adjustment later on.

7. SUMMARY AND OUTLOOK

Motivated by the complexity of the design task in defining actuator geometries for laser adjustment in micro technology, a concept and methodology to assist the construction engineer in this challenging task has been developed. In this paper a method to mathematically describe the behaviour of actuator geometries has been presented. These are understood as kinematic chains and are described using Denavit-Hartenberg matrices. Thus the demanded adjustment abilities of a new complex actuator can be guaranteed in the construction process by solving the linear equation system that results from describing both initial and target frame using the respective matrices. The abilities of the thus defined actuator geometry can be verified in simulation runs. A first sketch of the actuator's geometry can be obtained by combining the basic actuator geometries archived in the construction catalogue.

Fine-tuning of this geometry to rigidity and other constraints and limitations has to be carried out by the construction engineer in the current state of the system. Future research should lead to assistance in this step of the design process as well. Topologic simulation and optimisation are quite promising tools for this task. Furthermore the present system does assist the construction engineer in the definition of a suitable kinematic chain. However, this kinematic chain has not been optimised in any way and thus usually does not really provide an optimum solution. Therefore optimisation algorithms should be integrated into the system. Preliminary research in this direction suggests genetic algorithms may be an appropriate tool for this task. While the modelling and calculation of actuator geometries from basic actuator geometries provide a first step towards assistance and automation of this vital step in laser adjustment for micro technologies, additional research is needed to complete the assistance system.

8. REFERENCES

1. W. Hoving. Laser applications in micro technology, Laser Assisted Net Shape Engineering 2, Proceedings of the LANE '97, M. Geiger, F. Vollertsen, Bamberg: Meisenbach, 1997, pp. 69-80
2. F. Vollertsen. Laserstrahlumformen - Lasergestützte Formgebung: Verfahren, Mechanismen, Modellierung, Bamberg: Meisenbach, ISBN 3-87525-071-0, 1996

3. M. Geiger, B. Müller. Coupling of analytical and numerical models for laser adjustment, Lasers in Engineering, Nr. 7 (1998), pp. 275-286
4. F. Vollertsen. An analytical model for laserbending, Laser in Engineering, Vol. 2, 1993, pp. 261-276
5. A. Huber. Justieren vormontierter Systeme mit dem Nd:YAG-Laser unter Einsatz von Aktoren, Bamberg: Meisenbach, PhD-Thesis, ISBN 3-87525-153-9, 2001
6. C. Hamann, H.G. Rosen. Relaisfederjustierung mittels gepulster Nd: YAG-Laser, Laser /Opto-elektronik in der Technik, W. Waidelich, Berlin: Springer, 1990, pp. 661-665
7. T. Hennige. Laser forming of spatially curved parts, Laser Assisted Net Shape Engineering 2, Proceedings of the LANE '97, M. Geiger, F. Vollertsen, Bamberg: Meisenbach, 1997, pp. 409-420
8. F. Pfeiffer, E. Reithmeier. Roboterdynamik, Stuttgart: Teubner, 1987
9. J. Denavit, R.S. Hartenberg. A Kinetic Notation for Lower-Pair Mechanisms Based on Matrices, ASME Journal of Applied Mechanics, pp. 215-221, 1955
10. A. Huber, B. Müller. Verkürzte Prozeßzeiten beim Laserstrahljustieren durch optimierte Regelkonzepte., Laser in der Elektronikproduktion & Feinwerktechnik, LEF'99, M. Geiger, A. Otto, Bamberg: Meisenbach, 1999, pp. 157 - 168

CONCEPTION OF RECONFIGURABLE MACHINING SYSTEMS

Design of Components and controllers

Karl-Heinz Wurst, Christian Kircher, and Michael Seyfarth

Institute for Control Engineering of Machine Tools and Manufacturing Units (ISW), The University of Stuttgart, Seidenstrasse 36, 70174 Stuttgart, Germany, Phone: +49 (0)711 / 121-2410, Fax: +49 (0)711 / 121-2413

Abstract: This article gives a definition of reconfiguration and evaluates former approaches towards reconfigurable machining systems. Derived from the lack of these approaches a concept for the design of a reconfigurable machining system, including the control system is presented. The basic principle is autarkic, mechatronic components with well defined mechanical, electrical and information interfaces. The machine's modularisation in mechatronic components and the demand for a fast and reliable reconfiguration process leads to a self-adaptable, platform-based control system. To maintain the machines production ability after reconfiguration processes an autarkic diagnostic system is introduced.

Key words: Reconfiguration, Open-control, Mechatronic components

1. INTRODUCTION

The demands on production companies increasingly vary due to unpredictable changes. The reasons are influences of the market (e.g. fluctuating demand or technical developments) but sometimes they are not pushed by the market (e.g. legal changes). This results in batch production of varying products and shorter life cycles of products and leads to difficulties in controlling the production sequences. In order to maintain optimal production processes the machining systems, including the controllers, must be adapted, when changes get far too drastic.

The adaptation process is called reconfiguration. The reconfigurability of a system is derived from its configurability. Whereas the configuration

A. Bramley, D. Brissaud, D. Coutellier and C. McMahon (eds.), Advances in Integrated Design and Manufacturing in Mechanical Engineering, 417-429.

process includes the design, the selection und the composition of modules of a modular construction system meeting the requirements and the user's demands. Reconfiguration is the modification of the structure, the capacity and technology at a later date. This is achieved by replacing, complementing or removing self-contained functional modules.

The idea of reconfiguration is not new. There have been several approaches to reconfigurable machining systems. But these approaches were not very successful and partly disappeared from the market. The following chapters evaluate the current approaches to reconfigurable machining systems and deduce requirements which are considered in the conception and design of mechatronic components and controllers for reconfigurable machining systems.

2. EVALUATION OF CURRENT APPROACHES TO RECONFIGURABLE MACHINING SYSTEMS

In many areas of daily life and in industrial manufacturing configurable and reconfigurable machining systems are used. Studying their current designs demands can be derived, which have to be met, so that a reconfigurable, metal-cutting machining system will be accepted in the market.

One of the most famous reconfigurable machines is the kitchen machine: existing of one "active" base-module (autonomous drive-unit) it can easily be reconfigured in a few seconds due to a mechanical adapter and different passive process modules (cutting, kneading, stiring, pressing, …). The kitchen machine becomes a process specific machine.

A similar principle underlies the first generation of do-it-yourself machines: an autonomous, active base module can be reconfigured within a few seconds to a specific machine by passive process modules (milling, turning, planing, drilling) due to a mechanical adapter. In contrast to the kitchen machine this system disappeared from the market, because it was not handy and the specialized machines were produced more cost-effectively.

In the area of production engineering modular robot systems [1, 2] and the development of CO_2-Laser machines [3] show the advantages and disadvantages of reconfigurability.

Reconfigurable robot systems are built up by active modules (kinematically oriented one- or more axes joint drives) which form the robot base, and are complemented with passive modules (robot arms) which mainly form the structure of the whole robot (*Figure 1*). The advantages of this concept are well defined mechanical interfaces allowing exchanges of active and passive modules within seconds and machine internal bus systems for communication which could be routed into the modules via an appropriate adaptor. But disadvantages outweighed, so that this reconfiguration concept was not successful. One problem was the machine-internal distribution of auxiliary energy. A fundamental reason for the reluctance of the end-user was the weak adaptability of the control system and the absence of appropriate diagnosis systems to check and document the static and dynamic behaviour, i.e. the usability of the reconfigured machine. Considering the non-machining-system-relevant aspects, there was a lack in construction- and simulation tools with adequate flexibility.

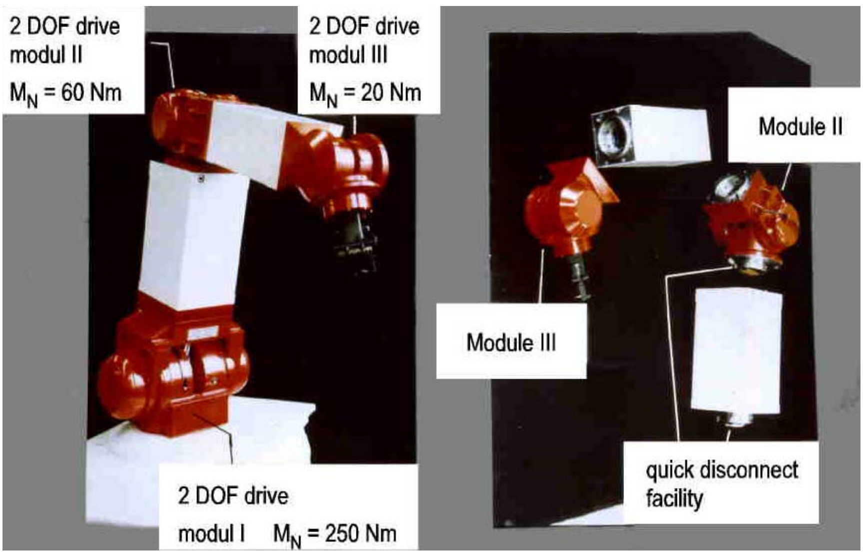


Figure 1. Six-degree of freedom robot with drive modules and link modules.

In the area of CO_2-Laser machines there was a similar approach of achieving (re-)configurability of the kinematic structure: the modularisation of the machine and the design of reconfigurable processing heads (*Figure 2*) [4]. The reasons for the low acceptance of modularised CO_2-Laser machines were just the same as why modular robot systems did not succeed. Additionally, there is a strong coupling of the tool “laser” and the mechanical structure of the CO_2-Laser machine by the components for guiding the laser beam. The user cannot handle the reconfiguration process and until now no really satisfactory software have been provided.

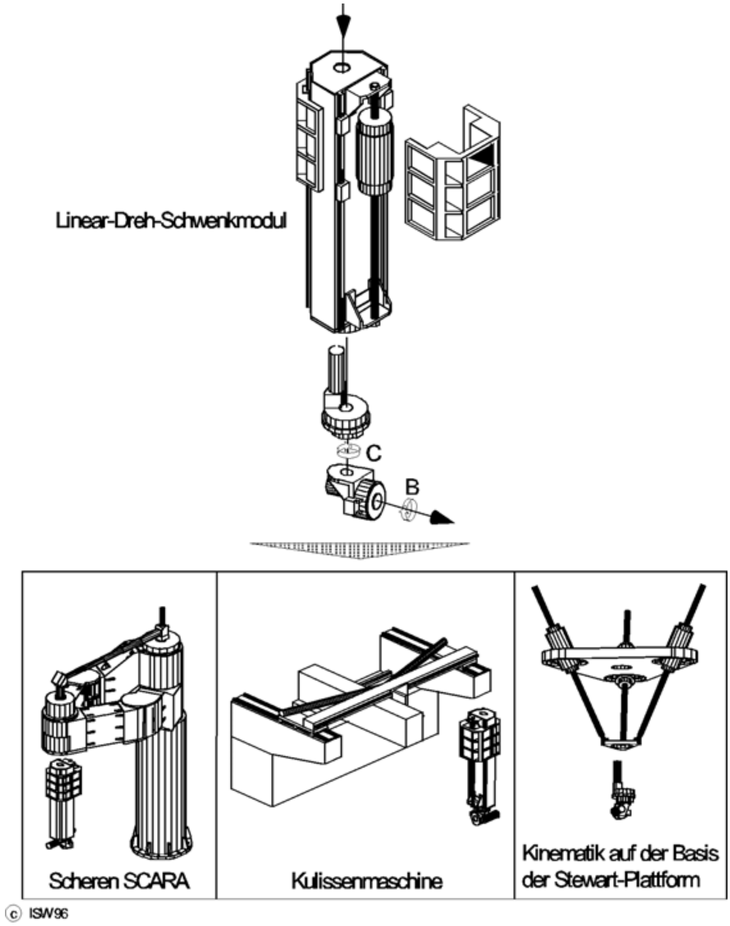


Figure 2. Application examples for CO_2 Laser processing head.

Recapitulating can be deduced that reconfiguration will only be established for such machining systems in which the manufacturing process (tool) is for the most part decoupled from the kinematic and structural design of the machine. Furthermore, it can be stated that the reconfiguration of machine tools is accepted and will gain practical success only if:

- The reconfiguration process can be carried out in a short time (from a few minutes up to a couple of hours).
- The reconfiguration process is manageable not only by experts.
- The cost for reconfigurable machines and the reconfiguration process are lower than purchasing specialized machines.
- The producibility after a reconfiguration process can be checked and documented smoothly and automatically.

Formulating requirements for reconfigurable machining systems these results must be kept in mind.

3. REQUIREMENTS FOR RECONFIGURABLE MACHINING SYSTEMS

The evaluation has shown that the reconfiguration process must primarily be manageable in a short time by non-experts. This means that:

- The mechanical interfaces and adaptors must be handy and easy-to-use.
- The interfaces for closing the supply systems for auxiliary energy, power, communication and information handling must be designed "plug-and-producible".
- The modules of the reconfigurable machining systems must be reactionless in order to simplify and accelerate the reconfiguration process performed by non-experts.

Mechanical reconfiguration of the machining system results in the need for changes in the control system, including control kernel and human machine interface. There are complex relations between the different application software modules and the mechanical assembly of the machining system, which can only be managed by experts. Accordingly it is not sufficient to look only at the mechanical requirements because the requirements for controllers must also be considered. The main requirements for controllers are [5]:

- Modularised application software: the controller software must be modularised and encapsulated into independent components.
- Scalability: the increase or decrease in functionality and capacity of the controller can be achieved, corresponding to reconfiguration of the mechanical part of the machining system
- Inter-operability: application software modules can be substituted by one another due to reconfiguration of the mechanical part of the machining system.

Additional requirements for controllers arise from the fact that the reconfiguration process must be manageable and carried out within a short time:

- Change of controller configuration: the number and type of the application software modules must dynamically be changeable, corresponding to changes in the mechanical assembly of the machining system.
- Automatic adaptation of the controller after mechanical reconfiguration: a controller integrated mechanism automatically adapts the controller to

the new machine structure and functionalities in order to achieve fast, fault-safe and high-quality reconfigurations.

4. CONCEPTION OF A RECONFIGURABLE MACHINING SYSTEM

A basic concept for designing a reconfigurable machining system is shown in *Figure 3*. The machine-structure is formed by mechanical modules divided into passive and active modules. Active modules are for example modules that generate motion by drives or other adaptive elements, whereas passive modules feature static and supporting functionality.

Each module aggregates a main- or auxiliary function like motion generation or workpiece exchange. The linkage of the modules can either be movable or fixed. The mechanical interfaces (i.e. attachment and load transmission), power interfaces (i.e. electricity, hydraulics, pneumatics) and information interfaces (i.e. bus systems like PROFIBUS, SERCOS Interface, Ethernet, IEEE 1394) of the modules are well defined, in order to guarantee their free exchange.

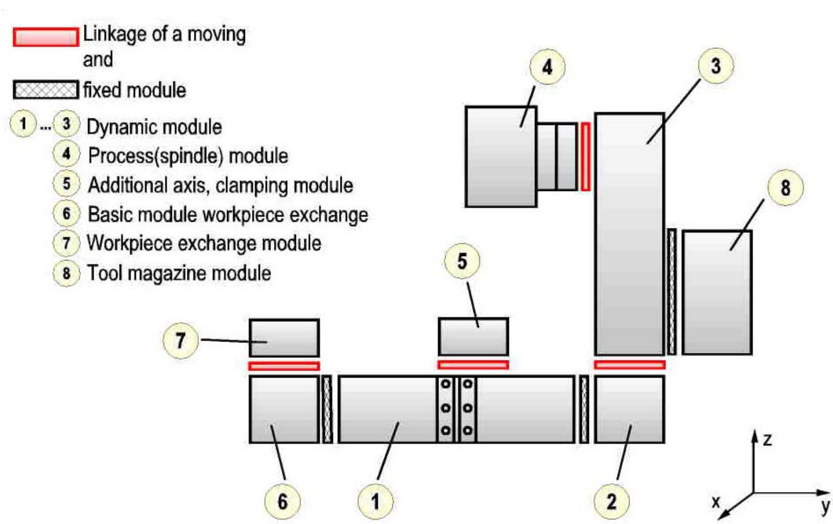


Figure 3. Arrangement of a machine tool with a modular design.

Furthermore, the reconfigurable machining system does not only consist of the mechanical modules but also of decentralised, linked controllers integrated into the modules. Thereby a module is no longer a mechanical module but a full mechatronic component.

In order to meet the requirements worked out in chapter 3, the conception presented is not sufficient. Particularly the manageability and the operability of the reconfiguration process are still very complex because of the lack of methods checking the machine's ability after a reconfiguration process. Tests have to be carried out whether the modules are suitably connected in and whether the reconfigured machine meets the accuracy requirements. Therefore, measurement methods and more particularly measurement devices have to be developed. In addition, tools for the (re-)configuration planning and execution are needed.

The following chapters describe the mechatronic components and controllers for reconfigurable machining system in details. Afterwards a special component for the automatic measurement after a reconfiguration process is presented.

4.1 Conception of mechatronic components and controllers

The partitioning of the machine into modules inflates the number of interfaces because each module has to be supplied with power and information. One example for the enlargement of the number of interfaces are process specific interfaces which have to be complied with by being passed through. This is contradictory to the requirements formulated in chapter 3 stating that the number of interfaces has to be small in order to maintain manageable "plug-and-produce"-modules. Therefore, module specific functionalities have to be integrated so that . the module becomes an autarkic module.

Autarkic modules integrate all elements used for the realisation of service functions or process specific functionalities. In order to minimize the number of interfaces parts of the control functionality are also integrated in the module. Therefore each active autarkic module has a controller (e.g. embedded controller, micro-PLC) that can locally process I/O and sensor signals without the help of a central control system. That way communication to a central control system is avoided.

Figure 4 shows an example for a conventional dynamic module (*Figure 4*a) and a (re-)configurable autarkic dynamic module (*Figure 4b*). Contrary to the conventional module the autarkic module integrates communication-, power-supply- and control components (including PLC); additionally closed autarkic circuits for lubrication and compressed air. This leads to the necessity of miniature pumps and compressors. But such components have already been developed, for example, in medical technology (high-power pumps as artificial hearts) or in automotive technology (miniature compressors). Also miniature power-supply components and miniature control systems are available.

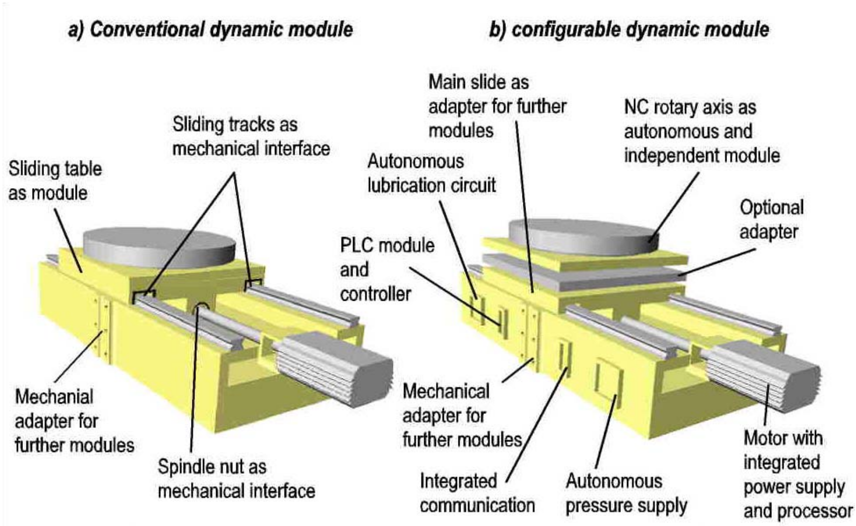


Figure 4. (Re-)Configurable design of machine modules.

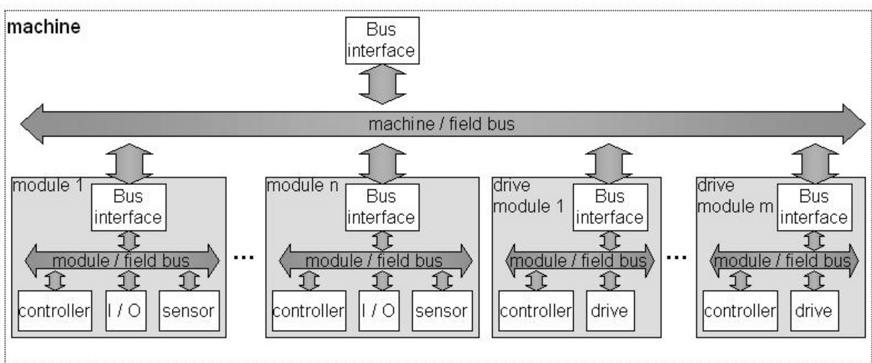


Figure 5. Structure of the communication system.

Since each module has a controller for decentralised functions, like sensor data processing, control of actuators or drives, a module-internal cross linking is needed. Therefore an autarkic module has a module bus which can be proprietary as long as it is encapsulated. In order to enable machine-internal communication between modules each module has the same bus interface connected to the machine bus system. *Figure 5* shows the conception of a communication structure with decentralised control functionality. As there is no difference between modules and drive modules the number of interfaces is reduced and standardised. The machine's communication structure looks like a module's communication structure with an internal bus system (machine bus) and a bus interface.

The illustrated communication structure (*Figure 5)* has no central control system. This is only possible with a decentralised control system based on a modular system platform which is described in the following chapter.

4.2 Conception of a platform based controller for reconfigurable machining systems

The modularisation of the machining components into autarkic and mechatronic modules leads to the decentralisation of the control functionality and demands new approaches from control systems as parts of the control software are integrated in the module and run on the module's integrated controller. Therefore a conception of a platform based controller for reconfigurable machining systems is presented which enables the distribution of control functionality.

The platform consists of specific hardware, a specific operating system and system software, which encapsulates the layers beneath. The system software provides the functionality needed to support the openness for application software, like hardware- and operating-system-independent communication (realtime-capable) and mechanisms for instantiating encapsulated application software modules. This so-called configuration runtime system allows combining of modules into a running system at the start-up of the control [6-8].

Figure 6 shows the modular system platform with different hardware and operating systems, an integrated communication system and control- and application software modules. The control functionality is implemented in the control- and application software modules. These modules have defined application program interfaces (API) and they are connected with a real time communication system. The communication system hides the underlying operating system and the hardware. From the module's point of view the hardware on which the module is processed is not visible. The hardware can even be distributed so that the use of a central control system is no longer needed. Consequently the control functionality is distributed among the modules. Nevertheless some modules with central functions are still needed, for example modules for coordinated motion or integrity test modules.

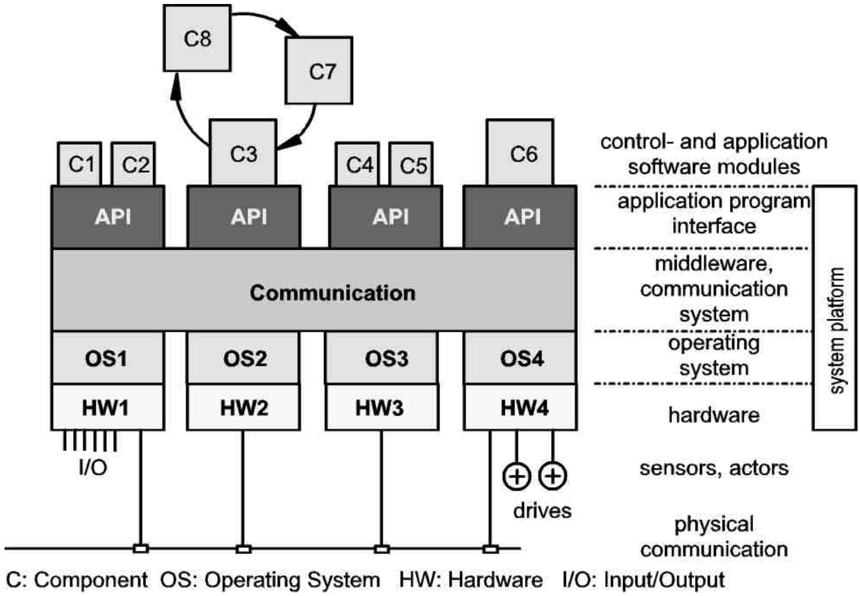


Figure 6. Structure of the system platform for an open controller architecture.

The presented platform enables the distribution of control functionality integrated in autarkic modules. But so far the conception is common. In order to achieve a control system which can react on mechanical reconfiguration and automatically adapt itself, additional modules and adaptation mechanisms are needed. These reconfiguration-specific modules are described in [5]. The main extension modules for a self-adaptation enabled control system are:

- The bus administrator, which detects new attached modules or removed modules.
- The identification module, which identifies detected modules and the whole mechanical configuration.
- The plausibility test module, which tests the configured control systems in respect of completeness, plausibility and contradictions.
- The configuration runtime system, which boots the control system, automatically loads the software application modules from a library and removes the software application modules unused.

During the reconfiguration process, the bus administrator detects module replacement, attachment or removal. Afterwards the new mechanical configuration is automatically detected by the identification module due to the modules' information crosslinking features and a reconfiguration command for the control system is generated. For that purpose an integrating mechatronic model is needed which associates the mechanical modules with the necessary control software application modules. The model also describes the inter-dependencies between mechanical and software modules with formalised rules allowing the automatic checking of the whole configuration. *Figure 7* shows the necessary modules for a self-adapting control system for reconfigurable machining systems.

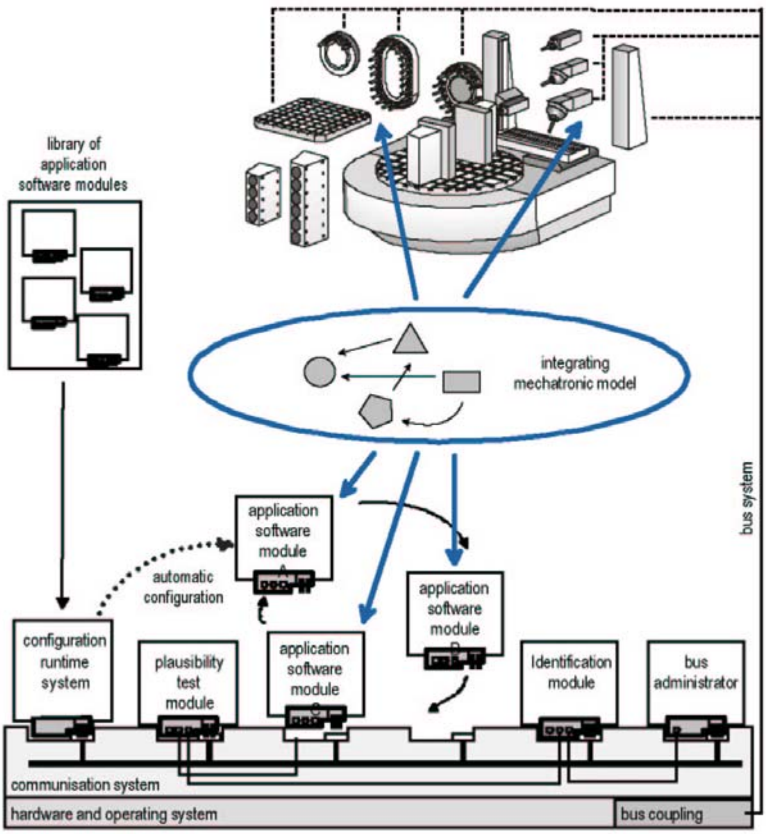


Figure 7. Self-adapting control system.

4.3 Conception of an autarkic module: A closed diagnostic system

Changes in the machining systems' structure after a reconfiguration process affect the quality and accuracy of the machine. Thus, it is necessary to check whether the dynamic and static characteristics of the machine meet the requirements on demand. In order to keep the reconfiguration process manageable for non-experts a diagnostic system which can analyse the accuracy between tool and workpiece is presented below.

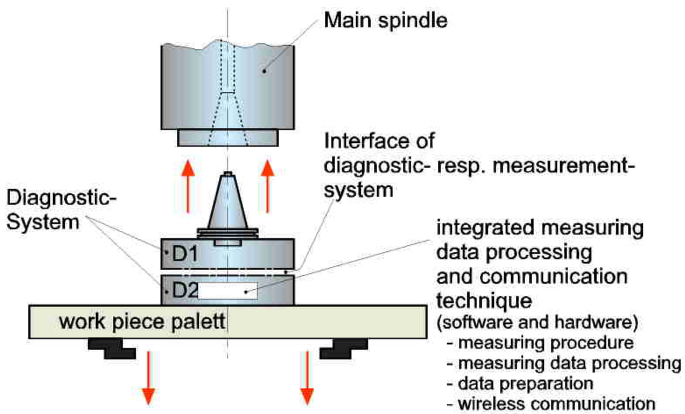


Figure 8. Autarkic diagnostics- and measurement system for reconfigurable machining systems.

Figure 8 shows the closed diagnostic system consisting of two components: Component D1 with a tool adapter and the complementary D2 component fixed on a pallet. The interfaces of the diagnostic system are intentionally designed like a tool and a workpiece pallet, so that the whole diagnostic system can be handled without any special apparatus. Based on the concept of autarkic modules the diagnostic system integrates a controller for local data processing and a communication unit which transmits the measurement's results to the control system. Radio communication or internet technology can be used for data transmission.

5. SUMMARY

Shorter product life cycles and the batch production of varying products lead to the demand for reconfigurable machining systems. Though former approaches to adaptable machining systems did not fulfil the requirements for reconfigurable machines, they provide an indication of how reconfigurable machining systems have to be designed. The evaluation of the approaches results in the main requirements of autarkic, manageable "plug-and-produce"-components integrating the mechanics, the electric devices and the necessary information technology. This leads to mechatronic components featuring well defined interfaces. To achieve a fully integrated approach the control system including self-adaptation mechanisms is decentralized and based on a open platform. To automatically test the machining ability after a reconfiguration process the concept is completed by an autarkic diagnostic- and measurement system. The consideration of the concept for reconfigurable machining systems could increase the acceptance of these systems on the market.

ACKNOWLEDGMENTS

We thank the German Research Foundation (DFG) for funding the project “General Conception of reconfigurable machining systems in consideration of control systems and mechanical construction” and supporting the collaborative research centre “Transformable Business Structures for Multiple-Variant Series Production”.

Furthermore, we would like to thank the European Union for funding the project “OCEAN: Open Controller Enabled by an Advanced Real-Time Network”.

REFERENCES

1. K.-H. Wurst, Flexible Robotersysteme – Konzeption und Realisierung modularer Roboterkomponenten, Springer-Verlag, 1991.
2. G. Pritschow, U. Kosiedowski, W. Schmid, Schnittstellen–dezentrales Steuerungskonzept für Roboter reduziert Verkabelungsaufwand, Maschinenmarkt, 51/52, 1993, pp. 52-55.
3. K.-H. Wurst, F. Mertin, Die „agile“ Laserbearbeitungsmaschine, Technica, 15/16, 1997, pp. 10-17.
4. K.-H. Wurst, U. Peting, PKM Concept for reconfigurable machine tools, The 3rd Chemnitz Parallel Kinemtics Seminar, Verlag Wissenschaftliche Scripten, 2002, pp.683-695.
5. G. Pritschow, G. Pritschow, K.-H. Wurst, M. Seyfarth, T. Bürger, Requirements for Controllers in Reconfigurable Machining Systems, CIRP 2nd international conference on reconfigurable manufacturing, Ann Arbor, USA, 2003.
6. G. Pritschow, Y. Altintas, F. Jovane, Y. Koren, M. Mitsuishi, S. Takata, H. Van Brussel, M. Weck, K. Yamazaki, Open Controller Architecture – Past, Present and Future, Annals of the CIRP, 50/1, 2001, pp. 463-470.
7. Y. Koren, U. Heisel, F. Jovane, T. Moriwaki, G. Pritschow, G. Ulsoy, H. van Brussel, , Reconfigurable Manufacturing Systems, Annals of the CIRP, 48/2, 1999, pp. 527-540.
8. G. Pritschow, G. Rogers, G. Bauer, M. Kremer, Open Controller Enabled by an Advanced Real-Time Network, CIRP 2nd international conference on reconfigurable manufacturing, Ann Arbor, USA, 2003.

Authors index