James A. Crowder · John N. Carbone
Shelli A. Friess

Artificial Cognition Architectures

Springer

Artificial Cognition Architectures

James A. Crowder • John N. Carbone
Shelli A. Friess

Artificial Cognition Architectures

Springer

James A. Crowder
Englewood, CO, USA

Shelli A. Friess
Relevant Counseling LLC
Englewood, CO, USA

John N. Carbone
Raytheon Intelligence and Information Systems
McKinney, TX, USA

ISBN 978-1-4614-8071-6 ISBN 978-1-4614-8072-3 (eBook)
DOI 10.1007/978-1-4614-8072-3
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013943011

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This book has been written to provide an avenue for discussion concerning the kinds of technologies, components, methodologies, architectures, etc., that will be necessary to create an actual fully autonomous artificial life form. An artificial life form is not just a collection of hardware and software/algorithms that magically becomes self-aware and begins to think, reason, learn, and make decisions like humans (as Hollywood would have you believe). It requires a cognitive ecosystem, similar to the human brain, central nervous system, etc., that all work and cooperate in unison to produce a complete "artificial brain." It is our opinion that you cannot design a truly artificial life form from the bottom up. It must be designed as a high-level cognitive entity, with all the components in place in the architecture, the information/knowledge models, communication mechanisms and methodologies, everything that is required in place in the high-level systems design first. Only then can you begin to decompose the system design into separate subsystems and begin to look at what is required for each lower-level entity within the ecosystem.

Along with creating synthetic models, designs, and architectures that represent neuroscience concepts adapted for artificial life forms, we must also take into account psychological concepts that explain interactions within the human brain and adapt those to their artificial life form counterparts. These topics have been a major focus of Dr. Crowder and Mr. Friess' research for the last 4 years. Dr. Carbone has spent many years deriving and architecting the information theoretics described in the book, in terms of knowledge formulation, retention, and retrieval within an artificial cognitive structure.

This book is a culmination of 18 years of research for all three authors. Each has concentrated on different aspects of Artificial Cognitive Architectures, bringing all the pieces together to form a complete picture and story of how an autonomous, thinking, learning, self-evolving life form could be designed and implemented. The authors have over 90 publications on various aspects of artificial intelligence, artificial psychology, information processing, and other concepts discussed here. These include journal publications, conference proceedings, books, and dissertations, many of which can be found online. One of the most important things to understand about the book is that it is not the final answer on Artificial Cognitive Architectures but represents the beginnings of the discussion on complete, fully autonomous artificial life forms.

We have strived to create a book that appeals to researchers in all fields but also to anyone who is interested in understanding artificial intelligence (AI) from a complete systems view.

Englewood, CO, USA — James A. Crowder
McKinney, TX, USA — John N. Carbone
Englewood, CO, USA — Shelli A. Friess

Contents

List of Figures

Chapter 1
Introduction

1.1 Striving for Artificial Intelligence

For more than 80 years, science and science fiction have been addressing the need and significant challenges of achieving truly autonomous machines that can act on their own. From the 1927 depiction of Maria in "The Eyes of the State," to Gort in "The Day the Earth Stood Still," to more semi-recent versions, HAL 9000 in "2001: A Space Odyssey," "Terminator," and Sonny from "I, Robot," the world has been fascinated, amazed, amused, and terrified at the thought of robots intertwined in our existence.[1] These creative depictions of artificially intelligent "robots" were supposedly capable of thinking, reasoning, learning, and making decisions. Unfortunately, the notion of a machine, run solely by software, no matter how sophisticated and creatively plausible, has continuously been met with significant theoretical, physical, and social challenges. Some would say, "troublesome" at its very core.

One of the reasons enabling machines with "human reasoning" is so difficult is that human learning is very dynamic in nature and hence, is somewhat fuzzy and random. There is no way to know when information is going to chaotically come our way, nor is it known how the information might apply to one of more simple or complex subjects, or topics in our memory pedigree [219]. The human condition is, and has always been, the true nature of a real-time system in a real-time environment. In order to create an Artificially Intelligent cognitive system that possesses the ability to ingest vast amounts of information content in real-time, then process and fuse it in order to learn, perceive, infer and ultimately evolve, a new synthetic humanistic and cognitively infused mathematical knowledge and relationship framework must be created. Recent disciplinary and trans-disciplinary advances in software, hardware, linguistics, semantic computing, cognitive computing, DNA computing, and neuroscience, suggest that computational devices can do as well as humans, especially across multiple information sources and information types.[2]

[1] http://www.filmsite.org/robotsinfilm1.html

[2] http://www.raytheon.com/technology_today/2011_i2/eyeontech_cognitive.html

J.A. Crowder et al., *Artificial Cognition Architectures*,
DOI 10.1007/978-1-4614-8072-3_1, © Springer Science+Business Media New York 2014

Human neuroscience research shows that generating new knowledge is accomplished via natural means: mental insights, scientific inquiry process, sensing, and experiencing, as well determining the context of this newly acquired knowledge, which characterizes the knowledge and gives it meaning [112]. True learning, therefore, can be a lengthy iterative process of knowledge discovery, experience, and refinement as new information is attained. This recursive refinement of knowledge and context occurs as a person's cognitive systems interact, over a period in time, with their environment; where the granularity of information content results are analyzed, followed by the formation of relationships and related dependencies. Ultimately, knowledge is attained from assimilating the information content until it reaches a threshold of decreased ambiguity and level of understanding, and is then categorized by the brain as knowledge, which acts as a catalyst for decision-making, subsequently followed by actionable activity or the realization that a given objective or inference has been attained [70]. Any real, functioning and evolving, autonomous, artificially intelligent system must have the cognitive system to perform these same activities.

Neuroscience studies also show that in order to understand the world we live in, humans synthesize models that enable us to reason about what we perceive. Fortunately for humans, we are able to deal with fuzziness [22]. We have the ability to perceive the world we see and form our own concepts to describe, prescribe, and make decisions. To accomplish this, language and communication is "fuzzily" applied, adapting and evolving our communications and processing to best align our needs, personal and conceptual views, and our strategic goals and vision for where we need to grow and evolve [191].

The purpose of this book is to describe concepts, theory, architecture, and practical designs of next generation Artificially Cognitive Systems (ACS). The ability to fully reason, learn, and self-evolve within an ACS connotes the need for a synthetic ability to infer about information, knowledge, observations, and experiences, and based upon these synthetic abilities, be able to affect changes within its synthetic, humanistic memories, and to allow the ACS to learn and perform tasks previously unknown, or to perform tasks already learned more efficiently [79]. The act of synthetically reasoning and inferring allows the ACS to construct or modify its internal representations of knowledge similarly to humans. Artificial human reasoning allows the ACS to flesh out skeletal or incomplete environmental and introspective information, called self-assessment, similar to the way a human brain functions when constructing a memory [63]. Throughout this book, we will present and describe processes, methodologies, architectures, and designs for the core components we believe are necessary to accomplish the creation of an Artificially Cognitive System. That being said, we have barely scratched the surface in the hunt for creating truly artificial life forms that mimic human reasoning and understanding. In some ways we have made great strides in understanding the human brain, its structure and functionality. Through trans-disciplinarily combining cognitive psychology, neuroscience, engineering, bioengineering, and computer science, we are making significant progress in bringing these human processes into artificially intelligent systems. However, our research community has barely scratched the surface in understanding everything that will be required to create a fully functional artificial evolving life form capable of human like thought, learning, reasoning, and self-evolving. After all, what we are striving for are artificial life forms that don't say things like: "No Dave, I really don't think I can do that." But instead when asked to jump, say, "How high? and How far"?

1.2 Historical Concepts of Intelligent Robots

1.2.1 Ancient Automatons

Since the dawn of humankind and the original inventions of mechanical devices there has been human fascination with creating machines that would act on their own, but on our behalf. The original mechanized "automatons" utilized hydraulics, pneumatics, and mechanics formed into devices which astounded audiences since most were originally created for entertainment or play. As far back as 3500 B.C. humans have surrounded themselves with animated creatures, either made of stone, clay, or precious metals, that looked and acted like humans. It makes one ponder the social motivations behind what would eventually become known as robots or androids that could do our bidding. According to Eric Wilson [211], the human fascination, or possibly obsession, with humanoid machines, has resulted from what is known as "the fall." Wilson's view is that the fascination derives from human dejection which he believes cannot be separated from human self-consciousness. Therefore, he explains there is a,

> …painful rift between mind and matter, knowing and being. To heal these splits, humans have created mechanistic doubles untroubled by awareness of self.

Is Wilson right? Is the fascination with humanoid beings built into our DNA, to identify with beings that are devoid of the problems of morality, self-awareness, or emotions? The next section briefly summarizes the last 5,500 years of fascination with artificial beings to allow you, the reader, to make the call.

3500–100 B.C.

Ancient Greece – The picture to the right is a statue of the Greek God Hephaestus of blacksmiths, craftsmen, artisans, and technology. Greek myths recount Hephaestus creating ancient robots to help him. In Roman myths, Hephaestus' Roman counterpart, Vulcan, is said to have created golden slave girls. In the Greek myth, the translation says:

> … and in support of their master moved his attendants. These are golden, and in appearance like living young women. There is intelligence in their hearts, and there is speech in them and strength, and from the immortal gods they have learned to do things. These stirred nimbly in support of their master…

~1000 B.C. – In ancient China it is reported that a mechanical engineer called Yan Shi presented King Mu of Zhou with a human figure he created. The story says the figure would walk and move its head up and down. When the king touched its chin, it began to sing. When he touched its hand, it began to gesture.

800–700 B.C. – Homer's Iliad includes the first mention of an automatae known as a *simulacra* (what we would later call a robot). Egyptians advanced this notion when, in the ancient Egyptian city of Napata, a statue was created of the great Amun, constructed to move its arm up and down and to speak to onlookers as well. Although the statue was not actually "intelligent," it is said to have had an impact on Egyptians of the time, portraying the perception of intelligence within their God.

384–322 B.C. – Aristotle mused about machines that did the work for humans. "… If every tool, when ordered, or even of its own accord, could do the work that befits it… then there would be no need either of apprentices for the master workers or of slaves for the lords.[3]" …

200 B.C. – The ancient Chinese created elaborate automations, including an entire mechanical orchestra.

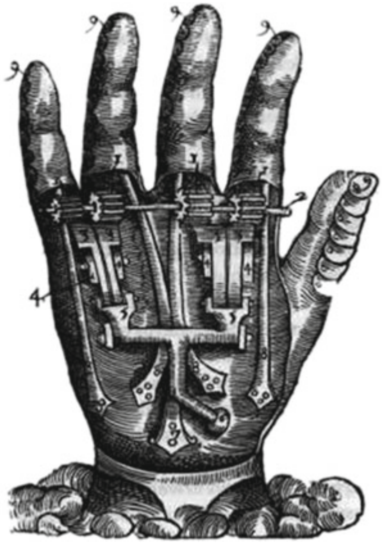
100 B.C. – The development of the Antikythera mechanism for calculating positions of astronomical objects.

100 A.D. – A hero of Alexandria wrote, in detail, about several automata that were used in the theater for religious purposes opening and closing gates, based upon hydraulic principles.

1495 A.D. – Leonardo da Vinci designs robots. Around this time da Vinci designed the first humanoid robot. The picture to the right is of a model, based on the drawings by Da Vinci. This is from the Mench-Erfinder-Genie exhibit, Berlin 2005.

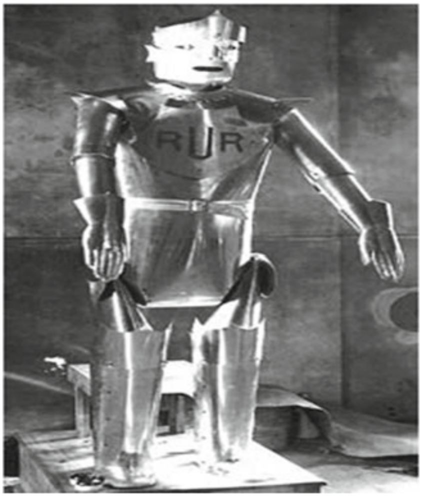
[3] http://it.toolbox.com/wiki/index.php/History_of_Artificial_Intelligence

1564 A.D. – In his work *Dix livres de chirurgie*, Pare Ambroise designs and published the design of a mechanical hand, which included mechanical muscles.



1801 A.D. – Joseph Jacquard builds an automaton loom controlled by "punch cards." Punch cards were used as the input for the twentieth century's earliest computers. This loom is on display at the Museum of Science and Industry in Manchester, England

1921 A.D. – Karel Capek coined the term 'robot' in the play R.U.R. The play was called "Rossum's Universal Robots." The term robot came from the word 'robota' which means tedious labor.



1950 A.D. – Isaac Asimov – I, Robot

1.2.1.1 Isaac Asimov's Laws of Robotics

The Three Laws of Robotics were introduced by Isaac Asimov:

1. A robot may not injure a human being, or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

The Three laws were first introduced in Asimov's book, "Runaround", which was written in 1942. Artificial entities circumventing these three laws has been the subject of many a science fiction story/film since their introduction. Example: a robot is rooted with the command to protect humans. The robot, seeing that humans argue and kill each other decides that in order to protect humans, they must be controlled. To not do so, would violate the first law. Its inaction, would allow humans to be harmed.

1.3 Hollywood's Views on Robots and Artificial Intelligence

Beginning with the movie "2001: A Space Odyssey" artificially intelligent entities were able to think, reason, and analyze the way humans do, generally with bad results. "Terminator" depicted intelligent machines overcoming their need for humanity and attempting to destroy them. Twentieth century robot infused movies spanned the spectrum of good (e.g. Lost in Space, I Robot, Star Wars C3PO, Space Camp's Jinx, and WALL-E) and evil (e.g. Battle Star Galactica's Cylons, Logan's Run, Blade Runner, etc.).

Here, we have attempted to explore some of the reasons why from ancient Greece, to current Hollywood screen plays, robots, androids, mechanical people, or whatever your favorite term is, have continuously been apart of human culture, literature, and myths. The next section will build on this foundational background and will define what artificial cognitive systems are, why we should care, and what they can do for us?

1.4 What Are Artificial Cognitive Systems and Why Do We Need Them?

On August 5th, 2012, Curiosity landed on Mars. It is estimated that 3.2 million people watched on line as the rover landed safely on the Martian surface. Curiosity is a marvel of technology, developed by NASA and the Jet Propulsion Laboratory (JPL). And while Curiosity can obey commands given from earth, and can utilize rules that are programmed into its system to follow those commands, Curiosity is not a Synthetic, Evolving Life Form (SELF). It cannot "think" for itself. It does not have the capability to dynamically adapt to things it has never encountered. Curiosity therefore relies upon aerospace engineers and software engineers to have written in as many of the possibilities of what might be encountered.

This is not because NASA and JPL do not desire these capabilities in systems like Curiosity, but we have not progressed in the ability to create and test such systems to a degree of certainty that we can trust them to be on their own (autonomous) in a place where we cannot effectively get to them if problems arise.[4] We cannot just give Curiosity goals like, look for water, or look for signs of life on Mars, and let it loose to determine how it thinks it can best accomplish the goals, make up its own tasks, and execute them as it seems best. However, what if we could? As we push the bounds of technology and develop more and more complicated and dangerous missions, the need to have systems that can think for themselves becomes crucial. Examples are deep space probes or deep undersea probes.

Our current and future space, air, and ground systems will continue to grow in complexity and capabilities, thus creating a serious challenge to monitor, maintain, and utilize them. The push toward autonomous cognitive systems makes this problem increasingly difficult. On-board systems must contain cognitive skills that can monitor, analyze, diagnose, and predict behaviors in real-time as the system encounters its environment. This requires creation of a SELF that can:

1. Act on its own behalf;
2. Perform autonomous reasoning, control, and analysis;

[4] We asked AAA and they definitely do not go out that far.

3. Find and fix problems within itself (self-assessment, self-regulation, and self-healing)
4. Predict future situations and determine its own internal recommended actions, and create or modify its own internal automated complex memories and processes.

Real cognitive intelligence has several manifestations, including the ability to adapt to unknown situations within dynamically changing environments. Without the ability to adapt to new situations, an intelligent system is left to rely on a previously-written set of rules to handle every possible contingency. Therefore, one of the main reasons for this book is to describe architectures and designs for a Synthetic, Evolving, Life Form (SELF) which can dynamically adapt to environments, with a simple set of a priori defined objectives and life rules. Artificial Cognition or Artificial Cognitive Perception (ACP) is intended to provide a SELF with the ability to mimic human reasoning during information processing and knowledge development. This differs from a classical definition of Artificial Intelligence systems which imitate and are measured by how similar they create humanistic results. For in humans, in order to understand our environment, without realizing it, we synthesize environmental models enabling us to reason about what we perceive. Information is absorbed from a variety of diverse sources and through each of our finely tuned senses. Much of this information is imprecise, or fuzzy, in nature and does not have a consistent basis. Information is also riddled with vagueness and ambiguity; is inexplicit, and contains content often initially unclear to us.

As humans, we observe, perceive, and reflect upon the world we see from context developed over time via actions and experiences, internalizing what we see, hear, feel, etc. To continuously form internal context we communicate fuzzily with language, adapting and evolving our grammar to best fit our needs. Our communication becomes dynamic and diverse; fuzziness becomes an essential feature used to communicate what we experience. Today, humans create virtual and cyber physical systems to emulate and interact with our physical environmental domain. If we truly desire systems which can more dynamically interact with the environment than systems today; to think, reason, act, and communicate in ways similar to humans, then we propose creating systems that mimic human cognitive processes. Today, systems closely interact with humans, indirectly and directly, peripherally and are even placed internally. Then why is it that the systems we build are devoid of "human" characteristics put into the system. What follows is a short description of the remainder of the book, structured to provide information, architectures, and designs we propose to realize such a system.

1.5 Layout of the Book

We have arranged the book by areas of human consciousness and cognitive skills for a SELF. Below is a short description of each chapter and its relevance to our objective. The book is arranged as follows:

Chapter 2 – The Information Continuum: describes the initial theory for Artificial Cognitive Architectures (ACAs). ACA is conceptually visualized by standing in the middle of the brain, at a given node, with a 360° view of information flowing in and out of each neural node. Subsequently, the conceptual view results in the derivation of the Information Continuum theory and description of the humanistic information flow emulated via an artificial node in a SELF.

Chapter 3 – The Psychology of Artificial Intelligence: describes evolution of an artificial cognitive system, and discusses constructs that mimic human processes, including psychological concepts of consciousness, sub-consciousness, and intuition, as well as concepts like the prefrontal and neo cortex. Additionally, this chapter discusses how such a system might be perceived and reacted to by others and vice versa.

Chapter 4 – Cognition Intelligence and the Brain: describes our approach to a SELF architectural framework that provides the abilities to organize information semantically into meaningful fuzzy concepts and information fragments (objects) similar to the human brain, which enables the creation of cognitive hypotheses as part of the SELF cognitive processing topology. This chapter also includes the description of the SELF Artificial Cognitive Neural Framework (ACNF) that provides the architectural constructs for artificial cognition and consciousness.

Chapter 5 – Artificial Memory Systems: describes the process of SELF decomposing, processing, encoding, storing, and retrieving information similar to human memory processes. At its heart, memories involve the acquisition, categorization, classification, storage, and retrieval (reconstruction) of information. This chapter will describe theory and architecture of artificial memory systems that will allow a SELF to capture information and information fragment objects, understand and categorize the context of the information, store it in short-term and long-term memories, and then, when required, construct memories from the information stored [175].

Chapter 6 – Artificial Consciousness: this chapter deals with the definition of consciousness, intuition, and the notion of artificial perception. To create a SELF, we will describe concepts to achieve the ability to perceive the environment. To take in information, make sense out of it, learn from it, and then act on it. Here we will discuss the architectures and methodologies to effect Artificial Consciousness, including the Intelligent information Software Agent (ISA) framework, as well Fuzzy, Contextual, Self-Organizing Topical Maps that are used to classify and categorize information.

Chapter 7 – Learning in an Artificially Cognitive System: this chapter describes the process of learning analogously with discussion on lower brain function and higher brain function learning. Also described is how a SELF can achieve abilities to learn from its experiences. Hence, discussed is a mathematical framework and methodologies required to provide a SELF with human-like learning.

Chapter 8 – Synthetic Reasoning: reasoning within a SELF implies the ability to infer about information, knowledge, observations, and experiences, as well as affect changes within the SELF cognitive framework (evolution). This chapter will describe the reasoning and inference architectures required to allow a SELF to effectuate and construct representations of its environment as it experiences and learns. Included are deductive, inductive, and abductive reasoning frameworks.

Chapter 9 – Artificial Cognitive System Architectures: Many have put forth architectures that facilitate cognition, learning, and information processing, but insufficient to create comprehensive autonomy. Discussed here is a cognitive learning and reasoning framework, knowledge and cognitive ontologies, as well as cognitive management structures, to facilitate autonomous, self-aware, self-assessing SELF [173]. This chapter describes a SELF processing and management framework known as the Polymorphic, Evolving, Neural Learning and Processing Environment (PENLPE) that provides the Cognitive Management architectures required to manage the SELF's cognitive processing algorithms in order to self-evolve.

Chapter 10 – Artificial Cognitive Software Architectures: Humans function via genetic material and biological neurons. A SELF runs on software. This chapter introduces the software architectures proposed to host and facilitate artificial consciousness and implement the architectures described in the rest of the book. This includes a software framework known as Intelligent information Software Agents for Artificial Consciousness (ISAAC), a dynamic neural software infrastructure which continually evolves neural system structures utilizing Neural Hyper-Threads.

Chapter 11 – SELF Physical Architectures: This chapter is dedicated to potential new hardware architectures to achieve a Synthetic, Evolving Life Form (SELF). Specifically, this chapter describes three-dimensional hardware structures theoretical concepts involving continuously recombinant hardware structures called Trans-Parallel Neural Hyper Strings.

Chapter 12 – Cyber Security within a Cognitive Architecture: This chapter describes a three-dimensional, real-time encryption scheme that provides encryption based upon a combination of information fragments, topics, need-to-know, and context. This chapter will also describe the three-dimensional Quantum Fractal Encryption scheme to provide security both from outside sources, and security within, to keep corrupted information from permeating the self-evolving cognitive framework.

Chapter 13 – Conclusions and Next Steps: This chapter describes challenges and benefits of achieving SELF. The chapter discusses the forces involved in pushing towards truly "autonomous" systems; robotic human accompaniment smart enough to do what we ask, when we ask, and to perform tasks without intervention or supervision. The chapter discusses awareness factors and some of the ramifications of what we might ask of SELF. Lastly, After all, we really don't want to hear our SELF reply to us: "I'm sorry Dave I really don't think I can do that."

Chapter 2
The Information Continuum

Research for the development of credible solutions within the Information Continuum has been a 17 year journey that began in the mid 1990s when the authors of this book were designing new ways to perform data capture, processing, analysis, and dissemination of high volume, high data rate, streams of information (what today would be called a “big data” problem). Hence, data analysis and lack of quality user interaction within that process are not a new problem. Users have continued to be challenged with keeping up with the vast volumes and multiple streams of data that have had to be analyzed. By the time a user had grabbed a time-slice of data, plotted it, and analyzed it, 100s of Gigabytes of data had passed through the system. In order to provide some semblance of rational attack against the onslaught of data we created what could be likened to a virtual window into the system that allowed the analysts to “walk into the middle” of the data and look at it as it flowed through the system. Analysts could reach out and grab a set of data, rotate it through its axes, and perform automated analysis on the data while remaining within the system data flow. This way analysts could intelligently and rapidly hop between portions of data within multiple data streams to gain pattern and association awareness.

The developed capability resulted in a realization that each point in time within the rapid data flow was an independent and discrete information continuum with specific and qualitative state. Subsequently, analogous thoughts began to emerge from research in artificial intelligence and artificially cognitive system theory. Envisioned was a virtual view within a portion of the human brain where one could view a given neural node, or a given neuron, and subsequently view data flow as data/information traveled in and out of the neuron. Once gathered, a hypothesis emerged that the analysis of brain locale, data, and study of brain processes through this type of virtual environment, could lead to important understanding of learning, inferring, storing, and retrieving (reconstruction) and/or all aspects of human neural processing. This led to the possibility of a theoretical Neural Information Continuum (NIC). This book builds upon the NIC concept as it applies to a SELF. Thus, the thoughts described in this book are described in terms of what components would be necessary to construct such a synthetic system, and what and how each artificial neural node in the system would be constructed from an information systems aspect.

J.A. Crowder et al., *Artificial Cognition Architectures*,

DOI 10.1007/978-1-4614-8072-3_2,

2.1 Information Flow Within a Synthetic Continuum

One of the first areas that must be investigated when considering a SELF is the flow of information. Humans take in ~200,000 pieces of sensory information each and every second of every day of our lives. Our senses (see, hear, smell, touch, etc.) are constantly receiving and processing information, correlating it, reasoning about it, assimilating it with what we already know, and finally leading to decision making based upon what was learned. For a system to become dynamic, self-evolving and ultimately autonomous, we propose to provide these same abilities; although the sensors and sensory perception systems may be synthetic and different, sensing a variety of information types that humans can't sense (e.g., infrared or RF information), the processes for autonomy, which correlate, learn, infer and make decisions, are the same. Besides receiving information from a variety of sources and types (e.g., auditory, visual, textual, etc.), another important aspect of information, is that the content is received at different times and at a variety of latencies (temporal differences between information). Additional characteristics include, a variety of associations between the information received and information the system may have already learned, or information about subjects never encountered. Therefore, these information characteristics and the challenging real-time processing required for proper humanistic assimilation, help us form the theory of the Autonomic Information Continuum (AIC). One of the first steps in developing our theory of synthetic autonomic hypotheses is observing/understanding the information continuum and the associated characteristics and operational relationships within the human brain. Hence, as we develop understanding of information flows into and out of neural nodes, types of information, processing mechanisms, distributions of information, enable us to establish foundational mathematical representations of these characteristics and relationships.

Processing, fusing, interpreting, and ultimately learning about and from received information requires taking into account a host of factors related to each piece or fragment of information. These include:

- Information Types
- Information Latencies
- Information Associations e.g.:
  - Time, State, Strength, Relationship Type, Source, Format etc.
- Information Value
- Information Context

Mathematically modeling the information continuum field surrounding a node within our synthetic AIC, is accomplished via inclusion of each discrete association for any node u, takes the form shown in Eq. 2.1:

$$C\frac{du(x,y,t)}{dt} = -\frac{1}{R}u(x,y,t) + \int_x\int_y w(x,y)z(x,y,t)\,dxdy + I(x,y,t) \tag{2.1}$$

where:

u *represents the unit node of the system,*
x *represents the preprocessed input to node u,*
y *represents the output from node u,*
w *represents the relative contextual information threads and association weight of u with its surrounding nodes, including a decay chat describes the relative contextual decay over time, where:*

$$w = \sum_{j=1}^{M} \frac{1}{r_j} T_j KD_j W_j \tag{2.2}$$

where:

T *represents the Contextual Information Thread j derived from Fuzzy, Self-Organizing Contextual Topical Maps*
KD *represents Knowledge Density j of Information Thread T*
W *represents Weighting for Contextual Thread j, and*

$$\sum_{j} W_j = 1 \tag{2.3}$$

I represents the processing activity for node u,
z represents the learning functionality for node u,
1/R represents the decay rate for node u,[1] and
C represents the capacity of node u.

This information field continuum equation (Eq. 2.1) allows us to analyze the equilibrium of nodal states within the AIC and to continuously assess the interactions and growth of independent information fragments within the system. Even in the most dense, most complex, cluttered information environments, each fragment of information and each action within the AIC is entropically captured explicitly and implicitly within Eq. 2.1. Equation 2.1 is the entropic engine which provides the ongoing analysis and virtual view into a synthetic AIC. Equation 2.1 enables us to assess the performance and quality of processing and to understand the capacities, information flows, associations, and interactions of knowledge and memories within the system, as well as, supporting analysis and inherent understanding of real-time system behavior. The variables in Eq. 2.1 can be interpreted as the average values in a heterogeneous assembly of information nodes, where Eq. 2.1 describes the behaviors of the interactions among n node assemblies within a synthetic AIC processing system. The objective is to have the ability to measure, monitor, and assess multi-level states and behaviors, and how and what kinds of associative patterns are generated relative to the external inputs received by an AIC system. Equation 2.1 provides the analysis needed to understand the SELF's ability for processing external content

[1] In this case, the decay represents the information's relative value over time.

within an AIC. Hence, real-time assessment and monitoring, and subsequent appropriate control, are expected to allow us to avoid developing a rogue AIC, much to the chagrin of Hollywood script writers.

2.2 Information Processing Models

Establishing a hierarchy of information flow within an AIC is a key objective for development of synthetic autonomic characteristics (e.g. cognition, thinking, reasoning, and learning). An AIC will need to be able to ingest and process a variety of inputs from many diverse information sources, dissect the information into its individual information fragments, fuse the information, and then turn this information into a formation which can be used to determine action- actionable intelligence. An AIC system must be able to assess situations previously not encountered, and then decide on a course of actions, based on its goals, missions, and prior foundational collected knowledge pedigree [183].

The underlying issues and challenges facing Artificially Intelligent systems today are not new. Information processing and dissemination within these types of systems have generally been expensive to create, operate and maintain. Other artificially intelligent system challenges involve information flow throughout the system. If flow is not designed carefully and purposefully, the flood of information via messages within these systems and between their software and hardware components can cause delays in information transfer, delaying or stunting of the learning process which can result in incorrect or catastrophic decisions.

Therefore, real-time decision making processes must be supported by sensory information and knowledge continuously derived from all cognitive processes within the system simultaneously, in a collectively uniform and cooperative model. Additionally, transformation from information to knowledge within an AIC system requires new, revolutionary changes to the way information is represented, fused, refined, presented, and disseminated. Like the human brain, the cognitive processes within an AIC must form a cognitive ecosystem that allows self-learning, self-assessment, self-healing and sharing of information across its cognitive sub-processes, such that information is robustly learned and rapidly reusable. This AIC ecosystem involves inductive, deductive, experimental, and abductive thinking in order to provide a complete Data-to-Information-to-Knowledge process explained in detail throughout the rest of the book. At a high-level, we are applying the theory of AIC and applying the constructs to the development of a humanistic analogous Synthetic Evolving Life Form (SELF). A SELF human brain analogy provides two-main layers of processing, a ***Deductive Process*** and an ***Investigative Process***. The ***Deductive Process*** is utilized for assembling information that has been previously learned and stored in memories (deductive and inductive logic), whereas the ***Investigative Process*** looks for patterns and associations that have not been seen before (abductive and experimental logic). Figure 2.1 illustrates the differences between deductive, inductive, abductive, and experimental inferences.

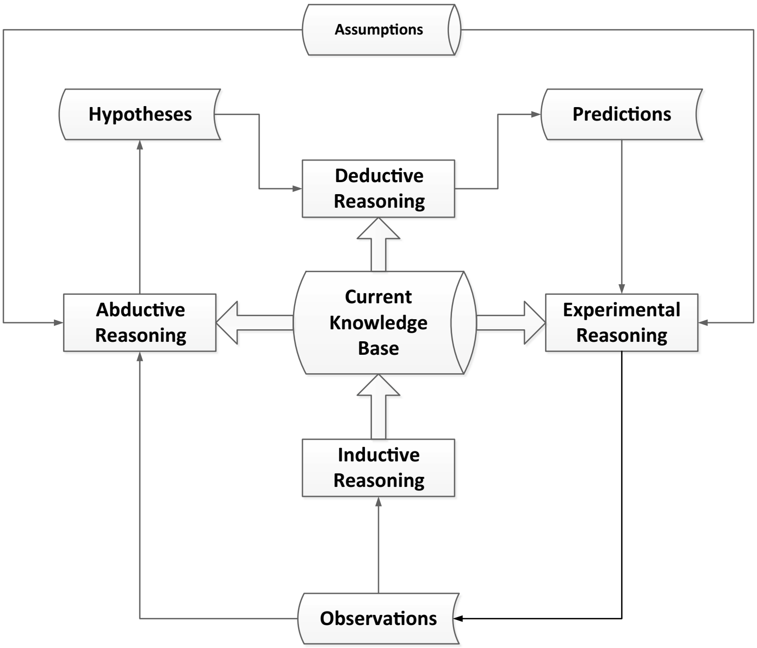


Fig. 2.1 Differences between logical inference systems

2.3 Discussion

If we desire to create an Artificial Cognitive Architecture that encompasses the AIC discussed above, in order to create a system that can truly think, reason, learn, utilizing the inferences shown in Fig. 2.1, we must consider the overall implications of such a system, including the psychological impacts and considerations both for humans and for the system itself. The next chapter will discuss the Psychology of Artificial Intelligence.Reference

Chapter 3
The Psychology of Artificial Intelligence

The preceding chapters focused upon introducing the characteristics within an Information Continuum and how they relate to a fully autonomous, learning, reasoning system (analogous to a synthetic brain), and how a SELF must possess constructs in its hardware and software to mimic humanistic processes and subsystems . This chapter will focus more upon designing and implementing these humanistic structures by understanding how they must interact and cooperate in order to form a comprehensive learning system. We employ concepts adapted from the domain of cognitive psychology as inputs into the formation of these interactive humanistic structures, sub-structures, and components. In short, Psychology helps us understand how these structures function within the human brain followed by translation efforts to design and implement these dynamic functions within an analogous synthetic brain. Hence, the foundational building blocks likened to "synthetic consciousness", comprised of cognition, intuition, and other capabilities that humans possess.

Creating a synthetic consciousness has significant technical challenges which are addressed throughout the book; however, this book will also explore adjacent cultural challenges which also need to be addressed. To create a complete artificial intelligent system, we need to understand how such a system would be received and perceived by people and how we expect any type of artificially intelligent system to react to and perceive people.

Therefore, here we explore the concept of "Artificial Psychology" where we will detail what it means to have a SELF resemble human intelligence and when and why the "Psyche" of the Artificial Intelligence system matters.

3.1 Artificial Psychology

Psychology is the study of mental processes and behavior of individuals. Artificial Psychology is then the study of the synthesized mental processes of the SELF similar to humans and the artificial cognitive processes required for an artificially

J.A. Crowder et al., *Artificial Cognition Architectures*,
DOI 10.1007/978-1-4614-8072-3_3,

intelligent entity to be intelligent, learning, autonomous and self-developing. In psychology there are several specialties or focused areas of study. Take for example cognitive psychology that studies how the brain thinks and works. This includes learning, memory, perception, language, logic. Developmental psychology considers the developmental stages in which an individual develops and what is appropriate to consider normal/standard for a human based upon these stages of development. Sports psychology considers mechanisms specifically to affect individual performance and how performance affects the individual. Hence, Artificial Psychology, in the context of this book, contains the artificial mental process considered necessary to create intelligent, autonomous, self-evolving, artificially cognitive systems.

Artificial Psychology is a theoretical discipline first proposed by Dan Curtis in 1963. This theory states that Artificial Intelligence approaches the complexity level of human intelligence when the artificially intelligent system meets three very important conditions:

- Condition 1: The artificially intelligent system makes all of its decisions autonomously (without supervision or human intervention) and is capable of making decisions based on information that is (1) New, (2) Abstract, and (3) Incomplete.
- Condition 2: The artificially intelligent system is capable of reprogramming itself (evolving), based on new information and is capable of resolving its own programming conflicts, even in the presence of incomplete information.[1]
- Condition 3: Conditions 1 and 2 are met in situations that were not part of the original operational system (part of the original programming).

Current engineering, bioinformatics, and computational science have evolved to a point where scalable processing power and real-time processing can, in parallel, perform operations to levels where we believe that not only the three conditions can be met, but that the possibility exists that an artificially intelligent system could have the ability to reach conclusions based upon real-time engineering which can aptly process newly acquired information, can infer upon it from learned and stored information in the form of synthetic memories. Therefore, we believe that enough criteria may exist, giving significant credence to the growing field of Artificial Psychology. This may require new theories and research to be explored in industry and at institutions of higher learning, specifically for addressing the rapidly expanding need for general human support systems to domains and environments where humans are still significantly challenged (e.g. space, deep sea exploration).

Artificial psychology, by definition, is required when the ability of the artificially intelligent system to reprogram, or self-evolve, through a process of self-analysis and decision, and based upon the comprehensive information available to the system, in real-time, does not and cannot provide the required mechanisms to process and resolve internal inconsistencies within the system.

The current theory of Artificial Psychology does not address the specifics of what those levels may be, but only that the level is sufficiently complex that the

[1] This means that the SELF autonomously makes value-based decisions, referring to values that the artificially intelligent system has created for itself.

intelligence cannot, at this time, simply be recorded by a software developer. Additionally, Artificial Psychology does not address the question of artificial consciousness.

3.2 Artificial Cognition: What Does It Mean to Be Cognitive?

Cognition is all about thinking. According to Ashcraft [7], cognition is the collection of mental processes and activities used in perceiving, remembering, thinking, and understanding, as well as the act of using those processes. Adding the term artificial identifies that the nonhuman synthetic system is a representation of the living intelligent system. Artificial Cognition is how the Artificial Intelligent (AI) machine (SELF) learns, integrates, recalls, ingests, processes, and uses the information that it receives. The challenges to create a SELF which is as complex as human thinking is driving towards Artificial Cognitive Science in order to achieve a better understanding of human processes and developing the truly intelligent machine [7, 84].

3.3 Artificial Intuition: What Does It Mean to Be Intuitive?

The word "intuitive" derives from contemplate in Middle English around 1400. Hence, the faculty or process by which humans consider previous information to make a judgment/decision about a given thought, action, or activity. In layman's terms, what does it mean to trust your gut? According to Anderson,[2] intuition is another way of problem solving that is not the same as logic [6].

> *Artificial intuition is not a high-level Logic model so there is no model to get confused by the illogical Bizarreness of the world. Systems with intuition then can operate without getting confused with things such as constantly changing conditions, paradoxes, ambiguity, and misinformation.*

Anderson also states that this does not mean that sufficient misinformation won't lead such a system to make incorrect predictions, but it means that the system does not require all information to be correct in order to operate. Intuition is fallible, and occasional misinformation makes failure slightly more likely. The system can keep multiple sets of information active in parallel (some more correct than others) and in the end, more often than not, the information that is most likely to be correct wins. This happens in humans, and will happen in Artificial Intuition based systems. The goal of the SELF is to provide the cognitive intuition required to deal with the world in a real-time, autonomous fashion. In subsequent chapters, a Dialectic Argument Structure is proposed, which is a methodology constructed for the SELF

[2] http://artificial-intuition.com/intuition.html

to deal with conflicting and ambiguous information and will allow the system to deal with our paradoxical and ever changing world. Other examples include, IntuView, an Israeli high-tech firm advertises the development of "artificial intuition" software that can scan large batches of documents in Arabic and other languages.[3] The espoused tool's capability "instantly assesses any Arabic-language document, determines whether it contains content of a terrorist nature or of intelligence value, provides a first-tier Intelligence Analysis Report of the main requirement-relevant elements in the document." This tool, like many others like it, provides semantic analysis and coded rules for the assessment. However, the human belief structure and memories comprise additional constructs of emotions linked to information which are included in the processes of intuition and human thought processes in general. Therefore, the subsequent chapter explores the effects of emotion and the development of synthetic emotion towards the end of achieving synthetic intuition.

3.4 Human Versus Machine Emotions

According to Minsky[4]: Human emotions are still about thinking:

> *The main theory is that emotions are nothing special. Each emotional state is a different style of thinking. So it's not a general theory of emotions, because the main idea is that each of the major emotions is quite different. They have different management organizations for how you are thinking you will proceed.*

Specifically, emotions can be thought of in terms of arousal states. Generally, when a person is calm and quiet they are more likely to be able absorb content and listen, learn, or problem solve. Contrary to being calm, an excited emotional state would make it less likely to be able to perform complex problem solving tasks [78]. With humans, that is why it is generally recommended to employ a predefined safety plan or practice evacuations. At the time of crisis, the brain doesn't have to perform as much problem solving, but instead follows a pre thought out plan. The instant a car accident occurs, the body is flushed with adrenaline, heart begins racing, and hands begin shaking. This would likely not be the right time to work out a calculus problem. Often, emotional states can influence our perception. A clinically depressed person would not likely perceive positive outcome of a given situation.

Similarly, for an artificially intelligent entity, emotions are states of being. If the system becomes overloaded, would it be likely to have the ability to determine what resources to allocate in order to return to the proper homeostatic state or a state of optimal performance? However, if enough indicators are sensed to arouse enough internal system urgency, could a system mediator or risk mitigation component keep operations performing optimally? Analogously, when an injury occurs or a virus is discovered within the human body, the immune system applies and/or moves resources

[3] http://www.wired.com/dangerroom/2008/10/tech-firms-face/

[4] http://www.aaai.org/aitopics/pmwiki/pmwiki.php/AITopics/Emotion

to fight the infection and/or heal a wound. The types of evolving system we are proposing would use similar constructs. An application of this type of system and how it might operate could be realized as part of monitoring data sources for imminent threats. Take terrorist threats for example. If an autonomous information processing system crosses a certain threshold of perception while monitoring a significant volume of information to conclude an attack might be imminent, the system could increase resources to determine the best plan of action. Just as the human level of arousal may contribute to what decisions we make. A minor chest pain from a strained muscle may result in taking an anti-inflammatory or a severe chest pain may cause us to employ rapid and decisive resources to call a paramedic [84].

3.4.1 Basic Emotions

In his book on Emotion and Intuition, Bolte concluded the following: [21]:

> *We investigated effects of emotional states on the ability to make intuitive judgments about the semantic coherence of word triads… We conclude that positive mood potentiates spread of activation to weak or remote associates in memory, thereby improving intuitive coherence judgments. By contrast, negative mood appears to restrict spread of activation to close associates and dominant word meanings, thus impairing intuitive coherence judgments.*

Bolte found a clear relationship between emotions and the ability to have or exhibit intuition. This drives us to a model of basic emotions that allow a system to channel resources and find solutions, based on emotional responses to its interaction with its environment. For the purposes of this book basic emotions are emotions that are in simplest forms of arousal states or states of being (e.g. calm, alerted, stress, terror or trauma).

The debate continues over the ability to artificially create human like emotions within systems. Consider Maslow's well-known hierarchy of basic human needs, which links human emotions to human needs and defines human characteristics when needs are met or not. An example of a Maslow basic human need is belonging or friendship. When humans meet this need they feel valued, loved and a sense of belonging. A general human perception might be that this would be unnecessary for a machine. However, if a system was given constraints would those constraints then effectively operate as needs? If the goal was to meet the constraint or satisfy the constraint; would the SELF begin to feel? Would the machine reach a level of arousal based on a need or constraint?

Given the studies sited, can we give a system a sense of intuition without emotion? If we can, could it then exceed human performance on tasks that emotions generally influence? How separable is intuition and emotion? The comprehensive question being asked is: can a system be developed which can perform humanistic intuitive predictions or problem solving without using states of arousal. We believe the emphatic answer to this question is no. Later chapters in this book describe the concepts and implementation strategies of an autonomic nervous system and related arousal states within to provide "emotion-like" features to interact with external environments.

3.5 Human Perception of Artificial Intelligence

According to Nass and Moon [176], humans mindlessly apply social rules and expectations to computers. In their work on human perception of Artificial Intelligence, they conducted three experiments to illustrate human tendencies to consider when discussing human perceptions of Artificial Life Forms. Their first experiment attempted to show that humans overuse social categories by applying gender stereotypes and ethnically identifying with computers. Their second experiment illustrated that people engaged in over-learned social behaviors such as politeness and reciprocity when interacting with computers. Thirdly they conducted an experiment to illustrate human's premature cognitive commitments by how humans respond to labeling. Thus human tendencies are important when considering human-Artificial Intelligence interaction.

Harmon [129] shows humans paired characteristics with a computer that may have been affected by gender and embodiment. Harmon describes significant correlation between gender, basic human characteristics, emotion, and computers. Specifically:

- Passive and Likeable for the male
- Understandable and Pleasant for both male and female
- Reliable and Likeable for male.

Harmon [129] found that both computer terminal and humanoid robot had significant correlation for understanding/pleasant and friendly/optimistic characteristics assigned by humans. Yet only the computer terminal showed significant correlation in regard to understandable/capable, pleasant/reliable, and helpful/reliable. Thus, concluding that humans were willing to assign human characteristics to computers.

Considering the research described in this chapter one can conclude that how Artificial Intelligence is presented to humans will affect how it is perceived. As an example, when any inanimate object becomes embedded with even a small amount of AI and then that system is given a name to embody it with human characteristics, or when a navigation system in a car is presented to a user with different types of voices the systems take on a whole different meaning. Clearly there are many variables influencing human perception of computers and continued research is required to understand interactive perceptions with AI systems to optimally benefit humans.

3.6 Human Acceptance of Artificial Intelligence

It seems that the non-intelligent robotics have had both a positive and negative reception from humans [129]. In varying degrees, humans already accept what is perceived to be artificial intelligence today: smart home software which secures, alerts, and automates the car that cools, heats, and can start itself. These are a few of many existing examples. On one hand technology helps humans function more efficiently: help to detect threats to national security, or virtually train our forces, and help solve other complex problems. On the other hand technology can take over human functions. Consider the effects of robotics embedded in the auto industry. Machines perform the work that humans used to do. Hence, because of a machine's ability to out-perform humans in certain conditions, we must consider the research required to set thresholds of human cultural acceptability. As with any new technology, there can be varying degrees of usage and learning curves. Human interaction with a SELF would be similar. The internet and cell phone technology has also clearly exposed generational differences in use and acceptance. Thus, it may take time for humans to accept a SELF's on a daily basis.

Historically, there has been continued concern with technologies and what they mean for human kind, whether the discussion involves cloning, public access to personally identifiable information, or embryonic stem cell use. Hence, in many scientific areas we have continuously evolved ethical guidelines for science to follow as technology has evolved. Therefore, ethical research will continue to be required as self-evolving concepts and cognitive technologies become more commonplace.

3.7 Artificial Intelligence Perception Design

It is generally accepted that humans are emotional beings and inanimate and animate computer systems are not, even artificially intelligent ones. Hence, let's consider human emotional intelligence. According to Mayer, Salovey, and Caruso [170], Emotional Intelligence (EI) entails the capacity of humans to reason about their emotions and emotions required to enhance thinking. They reasoned that Emotional Intelligence includes the abilities to:

- Perceive emotions,
- Access their emotions,
- Generate emotional knowledge,
- Regulate their emotions by reflecting on them,
- Use their emotions and emotional memories to promote emotional and intellectual growth.

In short, Emotional Intelligence allows humans to operate on and with emotional information[5] gathered from interactions with their environment and other people.

Therefore, the hypothesis of this book proposes that in order for artificially intelligent systems to comprehensively interface with humans in a human qualitative manner, the observations and perceptions of these systems must be driven by humanistic cognitive emotional growth architectures which can provide a foundation for qualitative interaction. Additionally, we propose that the architectures will be significantly influenced by the perception humans have of these systems as described

[5] Emotional information concerns the meaning of emotions, emotional patters and sequences, and the appraisals of relationships they reflect [170].

in the previous chapter. Hence, this allows us to extrapolate that a SELF should require parts of and architecture to address some levels of social intelligence. This will likely effect how humans perceive a SELF as well. Social intelligence as well as many other cognitive and psychological aspects of humanity will most logically have relevance in the modeling and development of cognitive architectures of one SELF (e.g. depression, the group context, peer pressure, sense of security)

Chai [31] describes a project in which the objective was:

> *…to build a software module for the analysis of cultural differences. The module is designed for incorporation into a decision-support environment in which real world actors with whom the user is interacting are "avatarized" into agents whose movements appear within a graphical user interface. The purpose of the module is to help members of multinational coalitions operate better.*

Sun-Ki Chai [31] goes on to say:

> *For the immediate future, I would argue that artificial intelligence needs social theory as much or more than social theory needs artificial intelligence.*

After giving thought to emotional intelligence, social intelligence, roles, and interfacing, can these lead to modeling and implementation of artificial personality? Can there be artificially designed traits, developed from a set of interoperability rules, which allow for internal preferences and behavior so SELFs can interoperate together as an ecosystem? We will continue to explore these concepts throughout the book as we describe the proposed mechanisms and integrated cognitive psychology required to build, test, and collaborate.

3.8 The Psychology of Human-Robot Collaboration

Historically, the purpose of robotics has been to perform some type of services on behalf of humans. Hence, to help define optimal human-robot-SELF interactions, we must look to the characteristics of human interactive behavior. Human collaboration, with other humans, fundamentally comprises trust and knowledge of another's abilities and limitations. In short, it's not possible to have an interaction between two human entities without there being some level of expectation of the interaction (discussed in more detail in Chap. 6) [200]. Let's consider a simpler example of human interaction with animals. Humans, for example, cannot completely predict an animal's behavior. However, it is still important to know how the animal will typically behave in order to predict and plan for the proper interactive response (e.g. give food, play, run to safety). Again, it comes down to human expectations. Understanding the animal's abilities and limitations will reduce frustrations of trying to meet a goal (e.g. taming a lion). Knowing the abilities of the animal changes our expectations. Bulldogs can't swim because of the shape of their nose, similar for dogs with large chest. Humans can accommodate for these limitations when they know about them. Understanding the expectations, abilities, and limitations of a SELF as well as the cognitively designed understanding of SELF expectations, abilities, and limitations of humans, is vital to efficient, and useful collaboration. Collaboration is much more than a mere working relationship. It is both a process and an outcome. The process is a coming together to work on a common problem while understanding that each other has influence on the other. The collaborative outcome is a solution where all parties can agree on the final solution [147]. Typically collaboration happens because an individual cannot accomplish the same goal alone. It is more than an association relationship it is more like a partnership.

So what is required for humans and robots, machines, to have a partnership? Likely, many of the same things as previously discussed; a sense of predictability, safety, reliability, trust, communication, knowledge, understanding, and accommodation just to name a few. We propose that everything collaborating with humans does not necessarily need to be human-like but as a minimum a need for some essential characteristics. Hence, it follows that some of the useful characteristics might be the ones that keep humans committed to the collaboration. Who will tolerate the constant attack of a lion, or the abusive coworker, or a laptop that continues to freeze in the middle of writing documents? Each will eventually be regarded as untrustworthy and would most likely be replaced.

Several research systems exist which are important to consider when thinking of the psychology of human-SELF collaboration. In their work on intelligent mechatronics, Harashima and Suzuki [128, 201] concluded that communicative artificial intelligence models must be equipped with mathematical models that touch on theory of mind, mind reading, and social common sense. This level of machine must also include eye contact robots and attempt to communicate intuitively and instantaneously. Such mechatronic systems have been able to perform as Ball Room dance partners and therapy Seals. There are many mechatronics designed to augment and/or enhance human skill. One example is a machine that assists as a scrub-nurse. Just the thought of a SELF assisting in any surgery implies a huge amount of trust particularly if ultimately allowed to perform surgery autonomously. Suzuki, Pan, Harashima, and Furuta [201] stated: "…knowledge and human psychology cannot be written sufficiently by computer algorithms; hence, the present intelligent mechatronics cannot understand the human perfectly". Later in the book we discuss the concepts and challenges with training a synthetically engineered evolving life form and vice versa using the process Human Mentored Software (HMS) [86], and Human Interaction Learning (HIL) [85].

Current human-robot interaction technology and design has developed from master–slave type interactions toward more collaborative. Karami, Jeanpierre, and Mouaddib, [147] described a model where the robot is able to consider human intentions and operate without communication. Karami, et al., also discussed how robots can build beliefs about human intentions by observing, collecting, and perceiving human behavior. Although the experiment shown was a seemingly simple task of moving objects, the results showed further promise for human-robot collaboration more advanced than in the previous master–slave paradigm.

Research shows that humans adapt to how they respond to robots over time [122]. Initially, humans tend to use simplistic communications with robots until they learn how the robots adapt to higher order types of communication. In later work, they investigated human robot interaction, illustrating how language and gestures help humans and robots collaborate during spatial maneuvering [121]. They concluded that over time humans used more complex language and gestures as they learned that the robot could successfully respond to them. Giving credence to the hypothesis that as humans and robots interact, increased understanding of constraint and limitation characteristics grows and directly affects qualitative collaboration.

Trends in human-robot interaction [127] show that several characteristics increase human trust in robots, among which reliability is a major factor. Also influencing trust is type, size, proximity, and behavior of the robot. Later research indicates that human characteristics such as ability and personality, and environmental characteristics such as task and team, along with robot performance characteristics/attributes effect training and design implications, thus, affecting human-robot collaborative team trust [19].

Since existing bodies of research indicate clearly that trust and clear expectations are important in human robot collaboration, significant challenges lay ahead for human adaptation to recent increases in capabilities of more highly autonomous cognitive systems. Similar to human-human or human creature relationships, little collaboration or cooperation will occur until understanding, expectations, and/or predictability become well defined in context of environment, enhanced trust, and collaboration.

3.9 Discussion

We have discussed the human desire for a SELF and the psychology involved in dealing with such a life form as part of our everyday lives. The rest of the book is dedicated to describing how to create an Artificial Cognitive Architecture that has the capabilities to learn, think, reason, infer, remember, and make decisions like humans.

Chapter 4
Cognitive Intelligence and the Brain: Synthesizing Human Brain Functions

In order for a SELF to function autonomously, and have the abilities to learn, reason, infer, evolve, perform self-assessment and self-actuation, we propose a cognitive framework similar to the human brain. What we describe in this chapter is an ***Artificial Cognitive Neural Framework*** (ACNF) that provides the ability to organize information semantically into meaningful fuzzy concepts and information fragments that create cognitive hypotheses as part of a SELF's topology [129], similar to human processing. This approach addresses the problems of autonomous information processing by accepting that the system must purposefully communicate concepts fuzzily within its processing system, often inconsistently, in order to adapt to a changing real-world, real-time environment. Additionally, we describe a processing framework that allows a SELF to deal with real-time information environments, including heterogeneous types of fuzzy, noisy, and obfuscated data from a variety of sources with the objective of improving actionable decisions using Recombinant kNowledge Assimilation (RNA) processing [70, 71] integrated within an ACNF to recombine and assimilate knowledge based upon human cognitive processes. The cognitive processes are formulated and embedded in a neural network of genetic algorithms and stochastic decision making with the goal of recombinantly minimizing ambiguity and maximizing clarity while simultaneously achieving a desired result [58, 95].

4.1 The Artificial Cognitive Neural Framework (ACNF) Architecture

The ***Artificial Cognitive Neural Framework*** (*ACNF*) processing infrastructure is a hybrid computing architecture that utilizes genetic, neural-network, fuzzy, and complex system components, that allow integration of diverse information sources, associated events, and iterative learning combined with artificial human-like memory systems to make observations, process information, make inferences, and ultimately, decisions. Within the *ACNF*, Continuously Recombinant Neural Fiber

J.A. Crowder et al., *Artificial Cognition Architectures*,
DOI 10.1007/978-1-4614-8072-3_4,

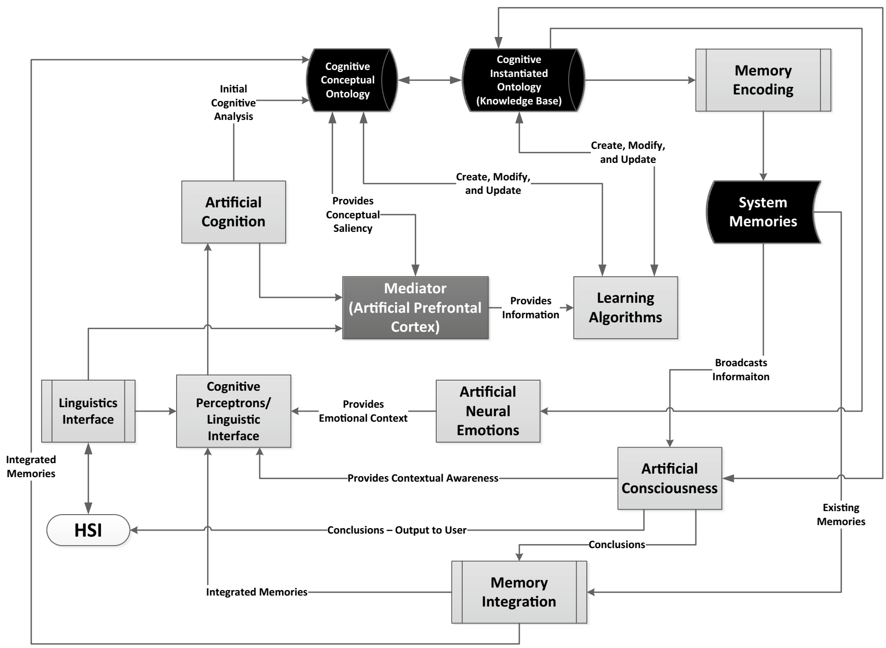


Fig. 4.1 The ACNF high-level architecture

Networks are utilized to map complex memory and learning patterns as the system learns and adapts [20]. The entire system "lives" and communicates via software agents called ***Cognitrons***, which are cognitive intelligent information software perceiving and processing agents that have the ability to mimic human perception and reasoning by understanding how to create and develop hypotheses [57, 214, 215]. The Cognitron comprises the knowledge and context of numerous perceptrons which intelligently collect and carry individual grains of perception within the ACNF. The architecture provides a collection of constraints, building blocks, design elements and rules for composing the cognitive aspect of a SELF. Figure 4.1 illustrates the ACNF architecture.

The three main domains of the ACNF are:

1. **The Cognitive System Components**: this consists of the Artificial Cognition, Learning Algorithms, Artificial Neural Emotions, Artificial Consciousness, and Cognitrons that make up the consciousness structures. These are responsible for the cognitive functionality of perception, consciousness, emotions, information processing, and other cognitive functions within the SELF ACNF.
2. **The Mediator (the Artificial Prefrontal Cortex)**: The Mediator takes information from the Cognitrons, processed through the Artificial Cognition processes, and forms coalitions of perceptrons that are used to update the short-term, long-term, and emotional memories.

3. **The Memory System**: the Memory System consists of the different memories (sensory, short-term, long-term, and emotional), and Memory Integration functionality. Here, memories are integrated together and information that is available within the ACNF memories (what the system has learned and "knows") and continually broadcasts it to the conscious perceptrons that form the cognitive center of the SELF. It also integrates information into current short-term memory to provide Integrated Knowledge (world data) to the Cognitrons to analyze incoming sensory information.

4.1.1 Cognitrons

The Cognitrons provide the ACNF with the ability to mimic human reasoning in processing information and developing knowledge. This intelligence takes the form of answering questions and explaining situations that the ACNF encounters [70, 77, 142]. These are persistent software components that perceive, reason, act, and communicate. Cognitrons are software structures that provide the following abilities to the ACNF:

- Allows the ACNF to act on its own behalf
- Allows autonomous reasoning, control, and analysis
- Allows the ACNF to filter information and to communicate and collaborate with other Cognitrons.
- Allows autonomous control to find and fix problems within the ACNF
- Allows pattern recognition and classifications
- Allows the ACNF to predict situations and recommend actions, providing automated complex procedures

Figure 4.2 illustrates another view or slice through the ACNF consciousness framework. In Fig. 4.2 we provide the ACNF Cognitron Ontology [191, 203]. This Ontology illustrates how the Cognitrons are intended to function within the ACNF.

4.2 The Artificial Prefrontal Cortex (The Mediator)

As described above, specialized Cognitrons are autonomous "Conscious" software agents that range in functionality and are situated in the processing environment. They sense the environment via fine-grained perceiving units, known as perceptrons, and act on them over time, in pursuit of an agenda, based on their evolving constraints. As they evolve it is possible for them to change what they sense at a later time. These "conscious" agents are also "cognitive" agents, in that they are equipped with constructs for concept formation, consciousness, basic emotions, and short & long-term memories [108]. The long-term memories provide identification, recognition and categorization functions, as well as identification of feelings [218].

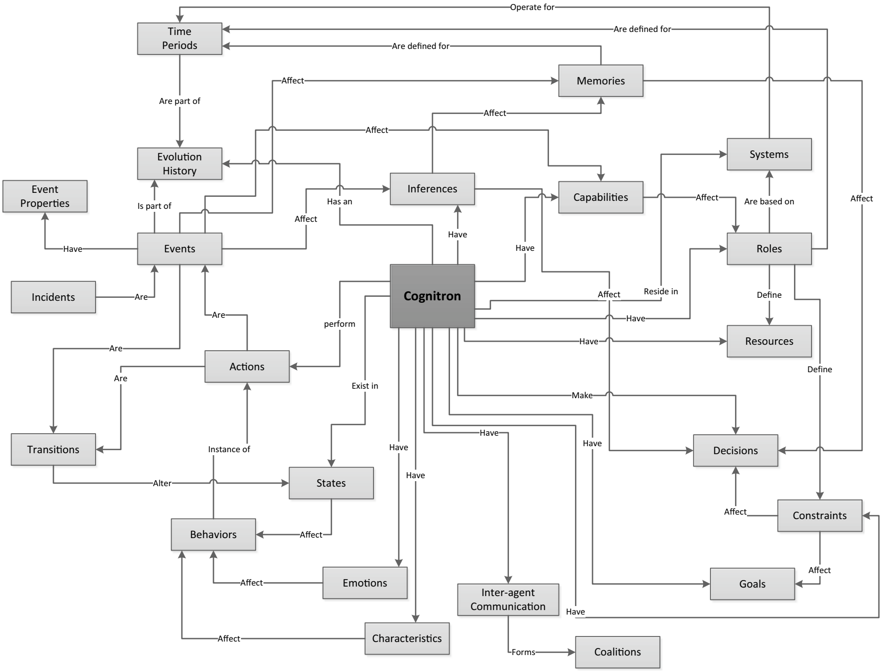


Fig. 4.2 The Cognitron conceptual ontology

The short-term memories provide preconscious buffers as a workspace for internal activities. A transient episodic memory is also provided as a content-addressable associative memory with a moderately fast decay rate.

This provides the architectural framework for an Artificial Prefrontal Cortex (APC), which provides cognitive intelligence for AI processing systems and allows for rapid analysis, reasoning, and reporting capabilities. The APC facilitates information, intelligence, and memory integration and allows faster accommodation and delivery of knowledge and knowledge characteristics across the system [156].

4.2.1 Artificial Prefrontal Cortex and Cognitive Control

The prefrontal cortex has long been suspected to play an important role in cognitive control, in the ability to orchestrate thought and action in accordance with internal goals. Cognitive control stems from the active maintenance of patterns of activity in the prefrontal cortex that represent goals and the means to achieve them [157, 158]. They provide bias signals to other cognitive structures whose net effect is to guide the flow of activity along neural pathways that establish the proper mappings between inputs, internal states, and outputs needed to perform a given task. The Prefrontal Cortex is integral to planning complex cognitive behaviors, personality expression,

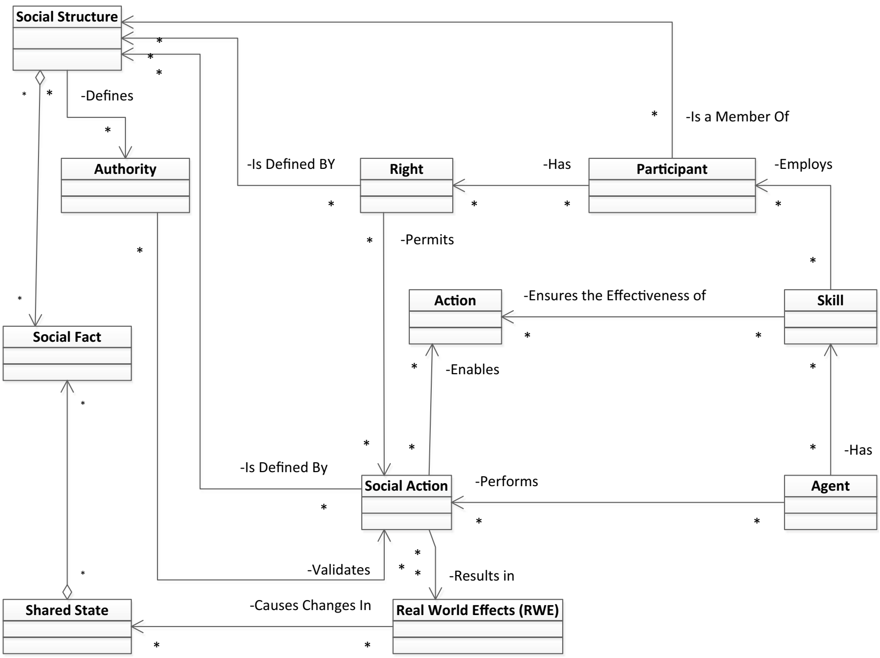


Fig. 4.3 Structure and control of social behavior

decision making and moderating correct social behavior [132, 218] (See Fig. 4.3). The basic activity of this brain region is considered to be orchestration of thoughts and actions in accordance with internal goals [172].

Here we describe and artificial structure within the architecture to provide an APC humanistic functionality and identify the structure, context, artificial feelings, emotions, and their roles within the SELF for performing real-world tasks. These SELF Cognitrons would be actively involved in every instance of action selection and in each learning event [177]. The pervasive, central role that feelings and emotions play in our proposed control structure of these conscious software agents mimics the roles they play in human cognition, and over time, give rise to clarifying hypotheses about human decision-making and several forms of human learning [93, 94].

4.2.2 Artificial Prefrontal Cortex Framework

Executive functions carried out by an artificial prefrontal cortex region are represented as core management functions related to overarching abilities that can manage and differentiate among conflicting thoughts, determine good and bad behavior, better and best, same and different, future consequences of current activities, working toward a defined goal, prediction of outcomes, expectation based on actions, and social "control" [99]. The prefrontal cortex is of significant importance when top-down processing is needed. Top-down processing by definition, is when a specific given or requested behavior is guided by internal states or intentions otherwise known as the cognitive concept of "mindfulness [168]:"

- Mindfulness: an awareness that lets us see things as they truly are without distortion or judgment, giving the most insightful explanation of how mindfulness can change not only our lives, but the very structure of our brains.

In order for our SELF to be autonomous, we propose to give it "executive functions" abilities. One of the cognitive concepts that employed for an autonomous system is the ability to perform top-down processing. To develop an understanding of a given mission or task at hand, and from this define an internal perception of needed goals/gaps along with prediction of possible outcomes, and subsequently utilize this knowledge to define the system processing behaviors needed to ultimately meet that mission or task goal. Executive management of autonomous system processes involves planning, monitoring, evaluating and revising the system's own cognitive processes and discrete outcomes. Strategic knowledge involves knowing what tasks or operations to perform (factual or declarative knowledge), knowing when and why to perform the tasks or operations (conditional or contextual knowledge) and knowing how to perform them (procedural or methodological knowledge). Both executive management and strategic knowledge capabilities are required for the system to autonomously self-regulate its own thinking and learning [46].

Hence, we propose a model for an Artificial Prefrontal Cortex sub-framework as part of our overall Artificial Cognitive Neural Framework (ACNF) and discuss the utilization of the Hidden Markov Model and related fuzzy possibilistic logic to drive the system between cognitive states.

4.2.3 *Artificial Prefrontal Cortex Architecture*

Architectural components within the Artificial Prefrontal Cortex provide the governance capabilities that enable definition and enforcement of cognitive policies governing the content and usage of cognitive maps and topical maps. Together these maps define the knowledge and context relationships processed by the Cognitron framework within a SELF. The logical architecture flow for the Artificial Prefrontal Cortex (APC) is shown in Fig. 4.4.

To understand the cognitive interactions that occur within an Artificial Prefrontal Cortex, a model was built to drive the Cognitron framework that provides linkage between the major cognitive states within the cortex [78, 79]. Figure 4.5 illustrates this cognitive processing model, rooted in foundations based on upon Artificial Intelligence interpretations of Dr. Peter Levine's Autonomic Nervous System States [163].

Detecting cognitive process information within an ACNF begins with sensors and sensory perception techniques that capture internal and external information about the system's physical or cognitive state or behavior. The information is gathered and interpreted by Cognitrons similar to how humans utilize cues to perceive

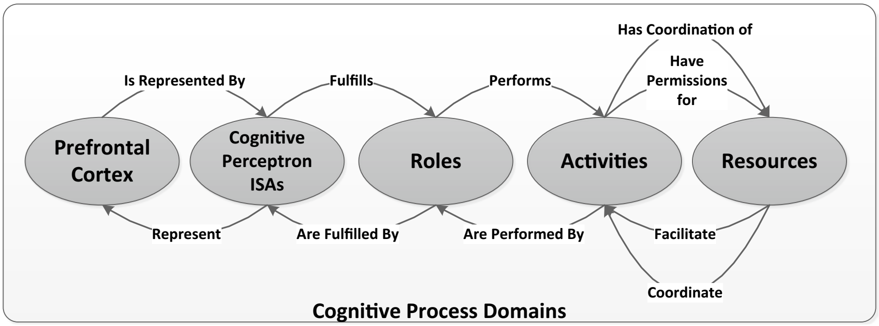


Fig. 4.4 The APC inference flow

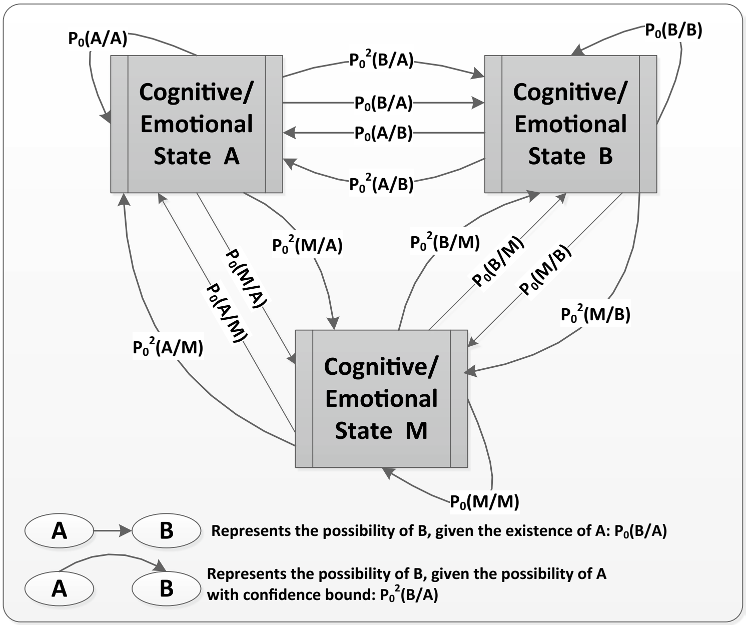


Fig. 4.5 The APC affected state model

cognitive states or emotions in others. The Artificial Prefrontal Cortex (APC) provides possibilistic inferences for a system to transfer between cognitive states. For simplicity, the APC shown in Fig. 4.5 illustrates only three cognitive states. Extending the model to include additional states is simply a function of possibilistic state transitions. The objective operation of an APC is to rapidly transition between cognitive states at any instant, and transition between states based upon possibilistics. These possibilistic parameters evolve over time, driven by learning algorithms expressed in later chapters and continuously reevaluated and affected both by normal and emotional memories [171].

Cognitive state transition related conditional possibilistics provide the APC with abilities to make executive-level plans and processing to move between cognitive states, each of which has its own set of priorities, goals, and motivations. An APC helps meet the objective to create an internal environment of self-evolving autonomy that could be used in a variety of applications. Humanity continues to desire systems that can explore regions of our earth and beyond where environments are dangerous and prohibitive. Many of these austere locations are unknown quantities, for which no astrophysicist or computer scientist could ever develop enough a priori planning or source code. Therefore, these types of systems must have enough a priori knowledge as we can give them, along with the ability to discern and infer how to land on a distant planet autonomously. Many existing system, like Unmanned Aerial Vehicles, intelligence information processing systems, cyber monitoring and security systems, all continue to have the "human-in-the-loop" making ultimate decisions, but are making strides toward autonomous operations every day. However, these systems are all developed with the goal of thinking all the possible causalities processed by the infamous IF, Then statements that the best software engineers can devise to prepare each of these systems for what it might someday encounter. Therefore, in order to evolve beyond this paradigm, we propose a system employ an APC comprising the following capabilities, process, and execution behaviors similar to a human prefrontal cortex:

Cue Familiarity: cue familiarity is the ability of the system to evaluate its ability to answer a question *before* trying to answer it [218]. In cue familiarity, the question (cue) and not the actual memory (target) become crucial for making cognitive judgments. This implies that judgments regarding cognitive processing and decisions would be based on the system's level of familiarity with the information provided in the cue. This executive-level, top-down cognitive judgment requires APC abilities that allow a SELF to judge whether the answer to a question is known, or whether the system is already familiar with the topic or mission, allowing the system to judge unfamiliar terms or conditions.

Cognitive Accessibility: suggests that a system's memory will be more accurate and more rapidly available for use when the ease of cognitive processing (accessibility) is correlated with emotional memories. For an APC, we propose that the quality of information retrieval depends on the system's density of knowledge on the topic or subject or individual elements of informational content about a topic. Individual elements of topical information can differ in strength while the speed of access is tied to both density of knowledge and level of emotional memory when a system responds to the information cues.

Cognitive Competition: comprises three principles:

- An AI cognitive processing system (the brain) is activated by a variety of inputs (sensors), perceiving text, audio, and visual pictures and video. Hence, different types of information are sensed simultaneously.
- Competition develops over time as simultaneous data is processed within the multiple cognitive processing subsystems and is adjudicated by the intelligent software agents; Cognitrons.
- Competition is assessed utilizing top-down neural priming.

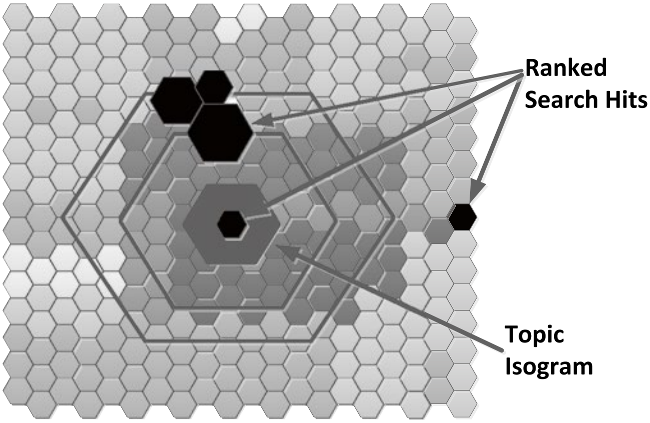


Fig. 4.6 The fuzzy, self-evolving semantic topical map

Cognitive Interaction: Combines cue familiarity and cognitive accessibility. In cognitive interaction, once cue familiarity fails to provide enough information to make cognitive inferences, cognitive accessibility accesses extended memories and may employ stored emotional memory cues to access additional information to attempt to make the required cognitive inferences. This may result in slower response time than with cue familiarity alone. Even in humans reaction times can be slower when the situation requires additional learning [227].

4.2.4 Artificial Prefrontal Cortex Processing

In order for an APC to process the challenges of cognitive competition as described above, processing constructs must be in place to allow cognitive inferences to be made, inferences and decisions learned, and simultaneously comply with an overall sense of priorities, goals, and needs. The following constructs are proposed to allow a viable APC to be constructed. The first construct employed is a topic map, specifically, a Fuzzy, Self-Evolving, Contextual topical map (FUSE-CTX). A topic map is an organized hierarchy of information, which crosses a threshold of similarity with other information to form a name comprising the general similarity.

A FUSE-CTX topical map is a general cognitive method for visualizing underlying analysis, and providing context for inferencing complex, multi-dimensional sensory information (e.g. textual, auditory, and visual). The FUSE-CTX is actually built by a two-step process employing a Fuzzy, Self-Evolving Semantic topical maps (FUSE-SEM) and then superimposing FUSE-SEM topical map onto a FUSE-CTX topical map. First, FUSE-SEM, topical maps organize information semantically into categories, or topics, based on derived eigenspaces of features discovered within the information. Figure 4.6 illustrates an FUSE-SEM topical map with information and topical "closeness" density for a series of responses received to a search query. The larger hexagons denote topical sources that best fit the search criterion. The isograms denote how close returns are to a particular cognitive information topic.

The FUSE-SEM information and topical closeness map have several important attributes:

- Searches employing FUSE-SEM topical maps use contextual information to discover links in relevant memories and stored information
- Image processing algorithms can be utilized to automatically analyze the visual output of the FUSE-SEM.
- The FUSE-SEM topical map is self-maintained and automatically locates input from relevant Cognitrons.
- FUSE-SEM topical maps operates unsupervised.

As topics develop during the data mining/sensing process, topical spaces are compared, within the APC, to stored emotions to determine derived “eigenmoods” within the emotional memory, as each topic is analyzed. The resulting eigenspaces determine topics that can be compared to the FUSE-CTX topical map to look for “closeness” of topics determined by cognitive processing algorithms to find the cognitive state to be used to make inferences about a question or task being posed. The eigenspaces are estimated under a variety of emotional memory conditions, dependencies, external inputs, and cognitive factors. Eigenvector trajectories are then characterized, capturing the dynamic aspects of relationship intersections between topical closeness and the information and memories available.

Once the FUSE-SEM is created, the resultant topical eigenspaces are mapped to the larger FUSE-CTX to show cognitive influences and ties to larger cognitive processes and other memory information, as depicted in Fig. 4.7. The value of superimposing the FUSE-SEM onto the FUSE-CTX is that the process defines the cognitive information domain’s hierarchical ontology in real-time, and hence enables the use of a real-time Topic Map Query Language (TMQL) to rapidly search more accurately, enabling sophisticated dialectic searches of only the information that has been deemed most important.

The need to mimic human intelligence demands a polymorphic architecture that is capable of both hard and soft computing. The APC FUSE-CTX topical map soft computing structure, utilizes the ACNF framework to evolve and grow context as it learns about its environment. The act of learning about a completely unknown environment; sensing, observing, processing, inferring, with no a priori information can be challenging. This requires processing obscure and diverse streams of terse information, thus, providing terse vectors for FUSE-SEM topical maps and cognitive mapping [153]. However, to a SELF, embedded with an APC, the amount of information content obscurity or terseness is only seen as another state, set of topics, and level of ambiguity to be resolved [17, 18].

The FUSE-SEM topical map processing resolves these ambiguities by performing a critical role, collapsing multiple dimensional relationships between pieces of information onto a two-dimensional space; a form that may be more easily computed and

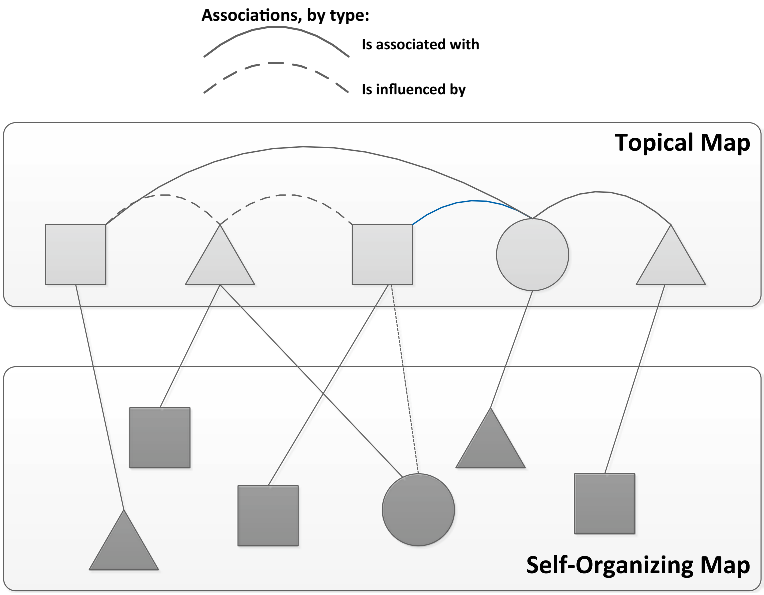


Fig. 4.7 Superimposing the FUSE-SEM onto FUSE-CTX

understood by an emotional memory enhanced APC. As more information is continuously acquired, it is iteratively mapped into previously understood knowledge and context structures within the ACNF [67, 79].

4.3 Self-Evolving, Cognitrons: The Heart of the SELF

As described in previous sections, the cognitive framework within a SELF is facilitated by Cognitrons that are used to mimic human reasoning and processing within the ACNF cognitive framework. As we push toward a completely autonomous SELF we require the SELF's on-board system contain cognitive skills that can monitor, analyze, diagnose, and predict behaviors in real-time as the SELF encounters its environment. Described here is a cognitive system of autonomous, learning, self-evolving software agents that provide a SELF with the ability to mimic human reasoning in the way it processes information and develops knowledge [79, 80]. As explained previously, Cognitrons are persistent software components which perceive, reason, act, and communicate. Cognitrons provide a SELF the following abilities [16]:

1. Act on its own behalf,
2. Autonomous reasoning, control, and analysis,

3. Allows the Cognitrons to filter information, communicate, and collaborate with other Cognitrons,
4. Autonomous control to find and fix problems within the SELF,
5. Situational predictability and offerings of recommended actions.

4.3.1 Self-Adapting Cognitrons

Intelligence reveals itself in a variety of ways, including the ability to adapt to unknown situations or changing environments. Without the ability to adapt to new situations, an intelligent system is left to rely on a previously written set of rules. If we desire to design and implement an autonomous SELF, it cannot depend on precisely defined sets of rules for every possible contingency. The questions then become [198]:

- *How does an autonomous AI system construct good representations for tasks and knowledge as it is in the process of learning the task or acquiring knowledge?*
- *What are the characteristics of a good representation of a new task or a new piece of knowledge?*
- *How do these characteristics and the need to adapt to entirely new situations and knowledge affect the learning process?*

As explained above, Cognitrons mimic human reasoning to process information and develop intelligence. The ACNF Cognitron architecture is comprised of a Java framework for constructing systems of Cognitrons, each with a specialized purpose, or talent. The architecture includes a Cognitron API that includes intelligent software processing functions for building multi-Cognitron intelligent autonomic systems. The Cognitron API also includes the framework for providing business rules and policies for run-time systems, including the autonomic computing core within a multi-Cognitron infrastructure. Figure 4.8 illustrates an overview of the Cognitron architecture framework operational process flow for a SELF.

The upper portion of Fig. 4.8 is broken into three separate combined internal/external system interface inputs comprising: commands, solutions, data, and problems. From left to right, interface 1 includes Data/Information and Command inputs, interface 2 describes the process flow of inputs arriving from a human user via an external portal, and Interface 3 describes receiving solutions, commands, and approaches from within the system. Each data input spawns a search process, “Se”, to detect if we already know the command and already have a solution and/or if we don’t understand the command. If we don’t fully understand what has been given as input into the system Cognitrons begin to spawn processes to develop hypotheses to determine either a new solution, or attempt to refine the requirement within the Evolution Domain with what we already know from our box of “Memories”. When an internal or external directive has been satisfied, the output response includes submission of the satisfactory solution to the memory repository and/or is submitted for external dissemination as specified by the solution.

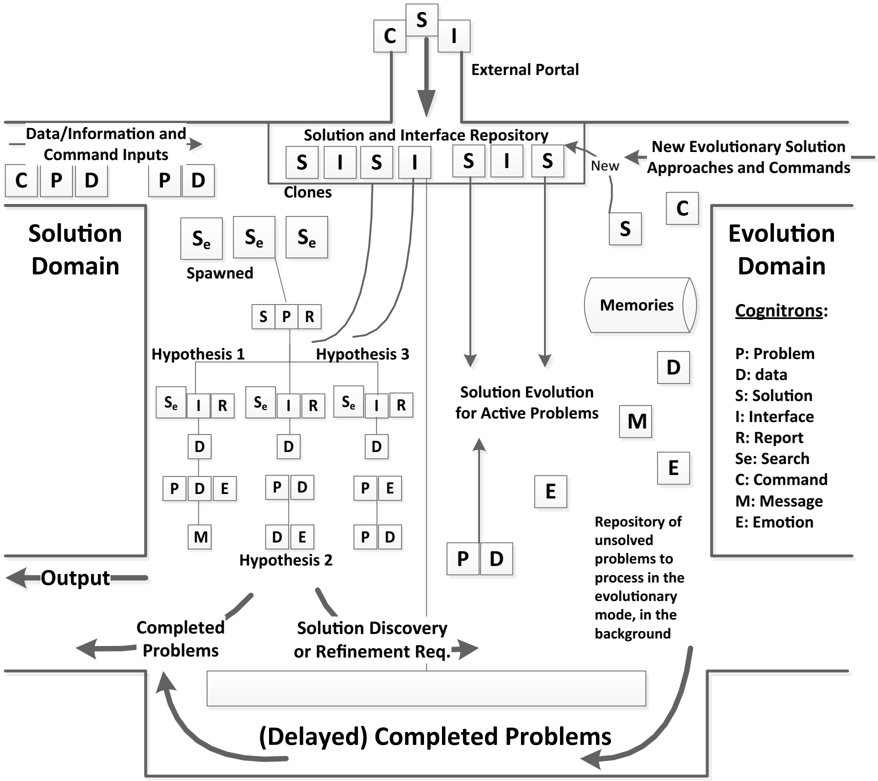


Fig. 4.8 The Cognitron high-level flow diagram

4.3.2 Cognitron Tasking

Cognitrons have the ability to learn from experience and can be used to actually predict future states (prognostics). They are able to analyze sensor data using classification and clustering techniques to detect complex states and diagnose problems (anomaly detection and resolution). Cognitrons interface with other Cognitrons and components and have the ability to reason using domain-specific application objects and are given autonomous (proactive) behavior and goals. Lastly, they have the ability to correlate events to situations, reasons, and take action.

The Cognitron computing architecture uses genetic, neural-network and fuzzy logic to integrate diverse sources of information, associate events in the data and make observations. When APC processes are combined with a dialectic search [57], information processing accuracy and speed show significant promise. The dialectic search seeks answers to questions that require interplay between doubt and belief, where our knowledge is understood to be fallible. This 'playfulness' is key to hunting within information and is explained in more detail in the section that address the Dialectic Search Argument (DSA).

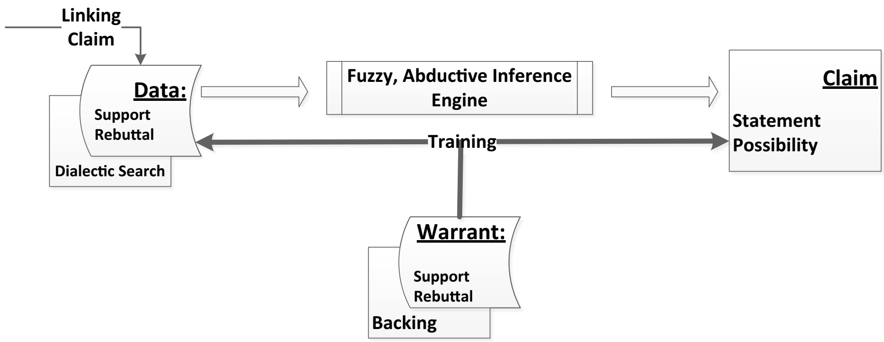


Fig. 4.9 The SELF DSA structure

4.3.3 The Cognitron Dialectic Search Argument (DSA)

The Dialectic Search uses the Toulmin Argument Structure to find and relate information that develops a larger argument, or intelligence lead. The Dialectic Search Argument (DSA), illustrated in Fig. 4.9, has four components:

1. Data: in support of the argument and rebutting the argument.
2. Warrant and Backing: explaining and validating the argument.
3. Claim: defining the argument itself.
4. Fuzzy Inference: relating the data to the claim.

The argument serves two distinct purposes. First, it provides an effective basis for mimicking human reasoning. Second, it provides a means to glean relevant information from the FUSE-SEMs [91] and transforms it into actionable intelligence (practical knowledge.) These two purposes work together to provide an intelligent system that captures the capability of a human Intelligence Operative to sort through diverse information and find clues.

This approach is considered dialectic in that it does not depend on deductive or inductive logic, though these may be included as part of the warrant. Instead, the DSA depends on non-analytic inferences to find new possibilities based upon warrant examples (abductive logic). The DSA is dialectic because its reasoning is based upon what is plausible; the DSA is a hypothesis fabricated from bits of information.

Once the examples have been used to train the DSA, data that fits the support and rebuttal requirements is used to instantiate a new claim. This claim is then used to invoke one or more new DSAs that perform their searches. The developing lattice forms the reasoning that renders the intelligence lead plausible and enables measurement of the possibility.

As the lattice develops, the aggregate possibility is computed using the fuzzy membership values of the support and rebuttal information. Eventually, a DSA lattice is formed that relates information with its computed possibility.

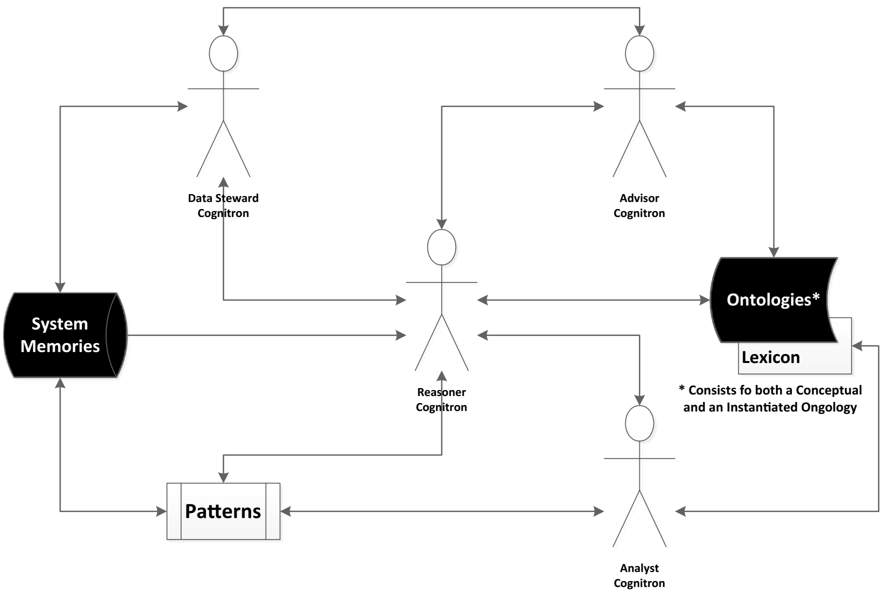


Fig. 4.10 The Cognitron cognitive ecosystem structure

The computation, based on Renyi's entropy theory, uses joint information memberships to generate a robust measure of Possibility, a process that is not achievable using Bayesian methods.

4.3.4 The Cognitron Software Architecture

Within the ACNF, each Cognitron provides different cognitive capabilities (called cognitive archetypes) that form a cognitive ecosystem within the SELF cognitive framework, allowing inter-Cognitron communication, collaboration, and cooperation. Figure 4.10 illustrates this ecosystem. Each Cognitron archetype, while having separate capabilities, has a defined cognitive structure, or ontology [191, 203] that was shown in Fig. 4.2.

Each Cognitron is a self-contained software unit (software agent) comprised of one or more services, shown in Fig. 4.10 [91]. The combination of services defines a Cognitron's capabilities. There are five currently defined Cognitron types within the ACNF processing infrastructure:

1. **Data Steward**: this Cognitron acquires raw data from a variety of sources, including sensors, and prepares incoming data for use by other Cognitrons. The Data Steward Cognitron generates and maintains metadata required to find and extract data/information from heterogeneous sources.
2. **Advisor Agent**: this Cognitron disseminates the right information to the right place at the right time; it provides capabilities that allow collaborative question

asking and information sharing by agents and end-users. Advisor Cognitrons generate and maintain topical maps required to find relative information fragments, memories, and "expert" Cognitrons.

3. **Reasoner Agents**: The Reasoner Cognitron interacts with the Data Steward and Advisor Cognitrons and utilizes the ontologies and lexicons to automate the development of domain-specific encyclopedias; it provides a mixed source of information and question answering that is used to develop an understanding of questions, answers, and their domains. Reasoner Cognitrons analyze questions and relevant source information to provide answers and to develop cognitive ontology rules for the SELF reasoning framework (explained in Chap. 8).
4. **Analyst Agents**: The Analyst Cognitrons are fed by Reasoner Agents and utilize the developed ontologies and lexicons to expand upon questions and answers learned from collected information.
5. **Interface Agent**: The Interface Cognitron assesses the correctness of major decisions and adjusts the decision processes of the Advisor Cognitrons. Interface Cognitrons also accommodate human-in-the-loop structures.

The ACNF Cognitron architecture provides the SELF with the following high-level features:

1. An Intelligence Network: this includes mechanisms for gathering information, learning, inferences, and providing decision support and situational analysis to the SELF APC.
2. Answer Extraction: these are mechanisms for posing hypotheses about situations and providing answers.
3. Situational Analysis: mechanisms for finding situations that require active investigation and provide actionable intelligence to the ACNF APC.

Figure 4.11 illustrates the capabilities of the various Cognitrons.

4.3.5 Teaching Cognitrons to Learn and Reason

As explained above, Cognitrons have the ability to predict future states (prognostics). They are able to analyze sensor data using classification and clustering techniques to detect complex states and diagnose problems (anomaly detection and resolution). Cognitrons can interface with other autonomic Cognitrons and components, and have the ability to reason using domain-specific application objects and have autonomous (proactive) behavior and goals. They have the ability to correlate events to situations, reasons, and take actions.

Creating Cognitrons which are capable of learning and reasoning about information provides a robust, adaptive information processing system capable of handling new situations [131]. When we use the term reason, we refer primarily to abductive logic, sometimes called critical thinking, to discriminate it from the formal logic methods of deduction and induction. For example, data mining uses induction to

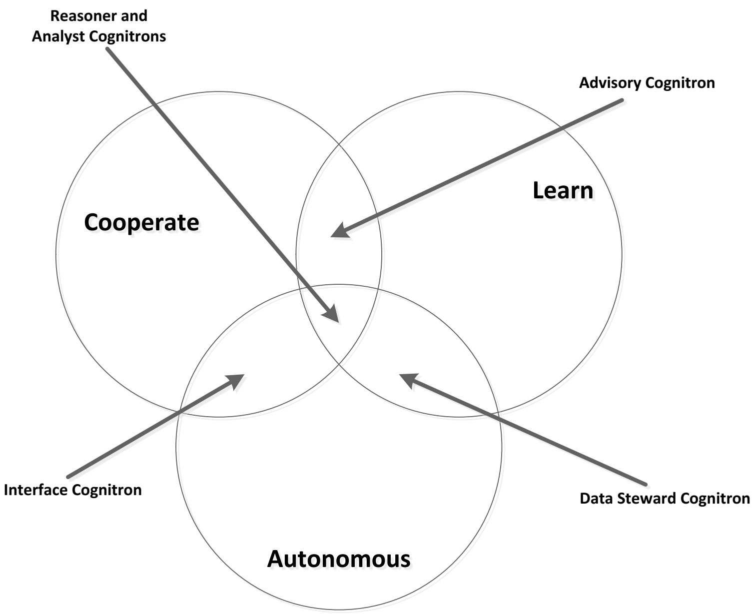


Fig. 4.11 Cognitron Venn diagram

develop assertions that are probably true. The dialectic search uses abductive logic to develop propositions that are possibly true. As explained earlier in the book, Bayesian methods cannot be used to measure possibility; in its place we use a method that is based upon Renyi's entropy theory.

As explained above, Cognitrons mimic human reasoning, and continually searches for relevant information, formulates inferences and provide leads. A key value of the Cognitron is that it provides its ability to learn from users and from data. Using this learning, the Cognitron has the potential to provide users and analysts more rapid, accurate, and actionable knowledge extracted from various diverse sources of information. As a software agent it can perform this function 24*7, and can be cloned/scaled to support as many operators as required and as system resources allow.

The approach to analyzing intelligence information utilizing Cognitrons is threefold. First the FUSE-SEM topic map processing is investigated to semantically organize the diverse information collected. Second, the topic map produced by the FUSE-SEM is used to enhance the user's comprehension about the situations under analysis. Third, as the user traverses the map to find related and relevant events, the results are used to train a Fuzzy, Active Resonance Theory Neural Network (FuNN) to replicate the approach.

This approach mimics human intelligence, learning from human agents utilizing a Conceptual Ontology to define particular domains, having experts (Cognitrons) to cartographically label the FUSE-SEM topical maps to capture the meaning of the integrated information thus capturing the knowledge of each Intelligent Information

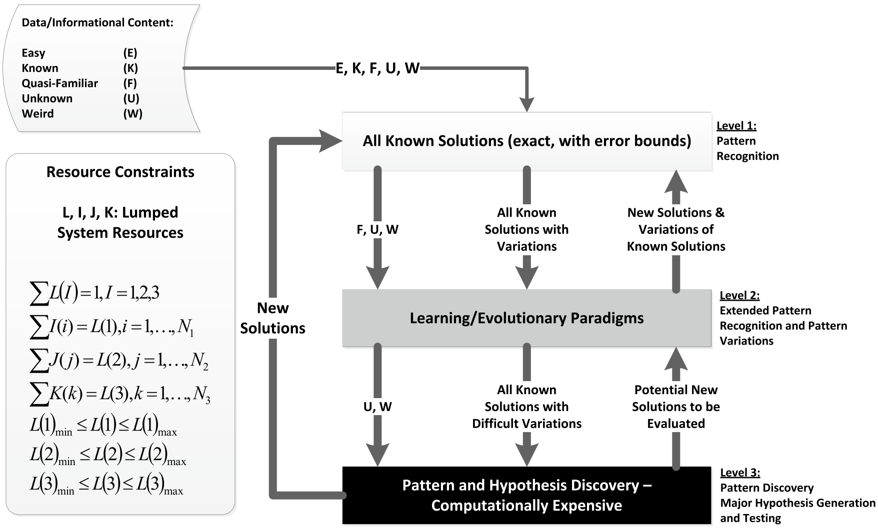


Fig. 4.12 Cognitron processing levels

in the FuNN [203]. The Cognitron processing environment has three processing levels, illustrated in Fig. 4.12. The first will identify patterns of behavior that have been seen (or behavior similar in a "fuzzy" relational way) before. The second is an expanded pattern recognition that involves pattern discovery algorithms that augment patterns that are similar to known patterns but need additional information to describe the pattern divergences. The third is a full up pattern discovery paradigm to make sense of information that has not been previously described (how do I find things I didn't know I was looking for).

These processing levels are necessary because information domains are too diverse and extensive for humans to comprehend in total, which is why we divide labor into to classifications of expertise. Similarly, we divide information using ontologies, or ontological views, where each view provides a certain perspective. By careful combination of such views we propose building a set of FUSE-SEMs that provide alternative, specialized maps of the information. These maps are designed to suit the different types of use, but they can be used in combination much like the dimensional views used for OLAP (Online Analytical Processing). Figure 4.13 illustrates three possible Cognitrons that could be used to implement the DAS: the Coordinator, the DAS, and the Search work together; each having its own learning objectives.

The Coordinator is taught to watch the FUSE-SEM topic maps, responding to new hits (input) that conform to patterns of known interest. When an interesting hit occurs, the Coordinator selects one or more candidate DAS Cognitrons and then spawns Search Cognitrons to find information relevant to each DAS. As time proceeds, the Coordinator learns which hit patterns are most likely to yield a promising lead, adapting to any changes in the FUSE-SEM topic map structure and sharing what it learns with other active Coordinators. The Search Cognitron takes the DAS

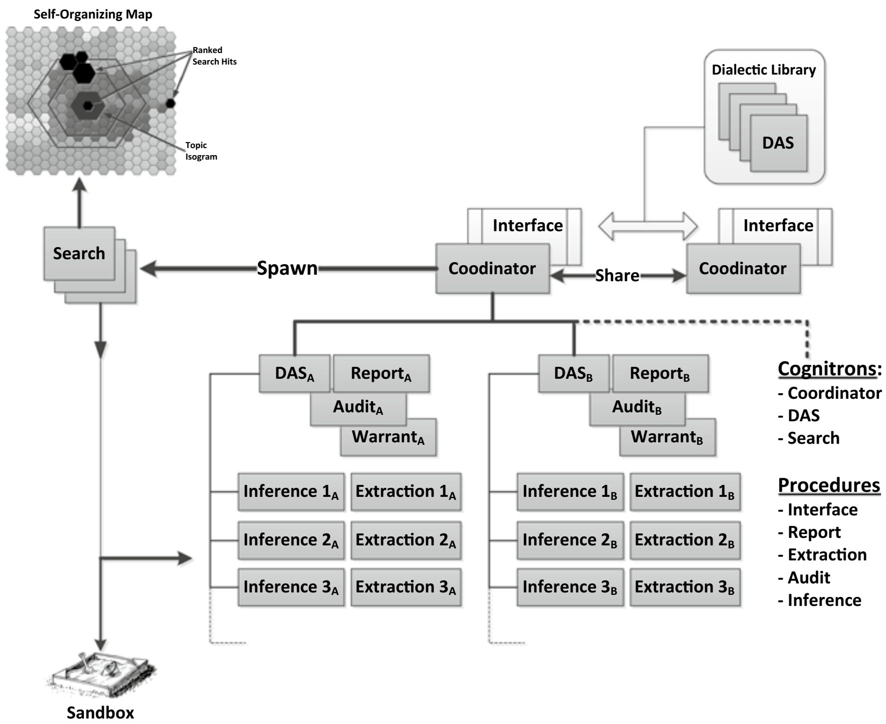


Fig. 4.13 Cognitron software agency

prototype search vectors and, through the FUSE-SEM topic map, finds information that is relevant and related. The Search Cognitron learns to adapt to different and changing source formats and could include parsing procedures required to extract detailed information.

The final Cognitron, the DAS, learns fuzzy patterns and uses this to evaluate information found by the Search Cognitron. Any information that does not quite fit is directed to a sandbox where peer Cognitrons can exercise a more rigorous adaptation routine to search for alternative hypotheses. The principal requirements addressed by the use of agents are:

1. Learn to adapt to changes in the surrounding environment.
2. Capture the knowledge as its cognitive framework processes information.
3. Sharing of information and learning between Cognitrons.
4. Hypothesize through the use of an Abductive Neural Network (discussed later).
5. Remember and capture relevance in contextual threads so as to avoid old mistakes and false leads.

A similar diagram can be drawn for the FUSE-SEM topic map where the Search Cognitron draws information out of heterogeneous sources, the DAS is replaced by a Topic Map, and the Coordinator is a hyper-map with its own specific Ontology.

4.4 Continuously Recombinant Neural Fiber Threads

The underlying issues and challenges posed by the introduction of Artificial Intelligence into system designs are not new. Systems for information processing and dissemination are an expensive infrastructure to operate and more-often-than-not these systems fail to provide tangible and useful situational information, typically overwhelming the system's infrastructure with system messages and other low-level data. A real-time SELF that incorporates human decision making processes must be supported by information derived from an extensive fusion and inference process and must operate in a uniform and cooperative model, fusing data into information and knowledge, so the system's cognitive engine can make informed decisions [35].

Here we discuss a proposed modular architecture for our SELF, based on a mixture of neural structures that add flexibility and diversity to the overall system capabilities. We discuss the object architecture for a flexible, continually adaptable neural processing system capable of dynamically adding and pruning basic building blocks of the neural system as the real-time requirements of the system change. This modular architecture is based on a "mixture of experts" methodology [35]. The difference here is that in our architecture, an expert is defined as a particular fuzzy, genetic perceptron object, which has been created for a particular algorithm, and thus is an expert at processing a particular type of data in a particular manner [44, 45]. The algorithm for which the perceptron is generated may be predetermined or may have been evolved by the neural system itself, providing a continuously evolving neural architecture, based on genetic learning within the recombinant neural structure [1, 18].

One major piece to the puzzle of how to create a continuously evolving architecture is the design of the information flows through the system. As discussed, the SELF information processing system requires the fusion of data and information from a myriad of heterogeneous sensors and sources (e.g., visual, auditory, radar, textual, etc.) to effectively create situational awareness and other products, which satisfy enduring information correlation challenges. The application of data fusion in multi-data type systems requires mathematical and heuristic techniques from fields such as statistics, artificial intelligence, operations research, digital signal processing, pattern recognition, cognitive psychology, information theory, and decision theory.

This translation of data-into information-into knowledge requires revolutionary changes in the way data/information is represented, fused, refined and disseminated. One such new approach is a continuously recombinant genetic neural fiber network. We believe this new system representation can be used to capture and evaluate system codes for events (both simple and complex) and will provide the mechanisms for determining the metric resolution required for facilitating complex manipulation of heterogeneous data types/sensor types, even in cluttered information environments and is the mathematical basis for the cognitive architectures discussed later.

4.4.1 *Self-Adaptive Cognitive Neural Fibers*

Theory into human consciousness postulates that human cognition is implemented by a multitude of relatively small, special purpose processes, almost always unconscious [177, 205]. These processes are autonomous and narrowly focused. They are efficient, high speed, and make very few errors because their purpose is narrowly focused. Each of these human processes can act in parallel with others. In the SELF, this is accomplished with fuzzy-neural perceptrons. Each perceptron is accomplished by codelets, small pieces of code that each performs one specialized simple task. Codelets often play the role of waiting for a particular type of situation to occur and then acting as per their specialization. These perceptron codelets are themselves miniature fuzzy-neural structures with specific purposes accomplished through tight constraints; but have the ability to learn and evolve. These are called Cognitrons and contain both short-term and long-term memories, providing the ability to communicate with other Cognitrons as needed. In human cognitive theory, the Cognitrons can be thought of as cell assemblies or neuronal groups [178, 205].

In the ACNF, first the unconscious Cognitrons, each working toward a common goal, form a coalition. These coalitions vie for access to the information or problem to be solved. The system "consciousness" provides a mediator for coalitions of processes to communicate with other coalitions. Information, or problems, that enter the system broadcast information to all unconscious Cognitrons. This allows the conscious coalition to recruit other Cognitrons that can contribute to the coalition's goals. The coalitions that understand the broadcast (i.e., their Cognitrons perform processes which are applicable to the broadcast) and need to take action on the problem.

The ACNF architecture that was shown in Fig. 4.1 is designed to allow for system-wide action selection. The APC gathers information and facilitates communication between Cognitrons. The APC takes information from Cognitrons through the artificial cognition software and form coalitions of Cognitrons and updates the short-term, long-term and episodic memories through the learning algorithms. The information available in memory is continually broadcast to the Cognitrons that form the artificial consciousness of the system (i.e., they are responsible for the cognitive functionality of perception, consciousness, emotions, processing, etc.) [81].

The active Cognitrons are constantly broadcasting information to the system unconsciousness as the problem is solved and the system evolves, to see if any of the "unconscious" Cognitrons can help solve the current problem. One or more of them may decide they need to act and join the active coalition. In this case one or more of the currently active Cognitrons may rejoin the unconscious collective. As the system evolves, links are formed between Cognitrons, based on their joint applicability to more problems. Links are created and strengthened by the amount time Cognitrons spend in the system's active consciousness and by the system's overall motivation at that time.

Each of the Cognitrons carries behaviors and drives. If the system is operating on intelligence information, the Cognitrons might have the behavior to look for information that they perceive is from a particular source. If the system is operating as an autonomous health management system, say for an airliner, the Cognitrons might have the behavior to look for data that indicates a problem within a particular subsystem. Artificial Cognition, or overall perception, is an overseer and monitors the internal conditions of the various Cognitrons. If necessary, it can influence behaviors through the mediator. For instance, the Cognition subsystem can make one Cognitron more goal-oriented and increase the chances that a coalition of Cognitrons will make it to the active consciousness. Learning works with cognition and influences the ability to learn new behaviors [51].

Utilizing the concepts described above a Continuously Recombinant Neural Fiber Network was created to investigate the concept of utilizing Neural Fiber Networks to create agile neural structures. This network utilizes the FUSE-SEMs discussed earlier, Genetic Learning Algorithms, along with Stochasto-Chaotic constraints on the neural fiber connections to determine constraint optimization [18].

This *Recombinant Neural Fiber* is different from standard Neural Networks, in that the internal nodes are interconnected and learn from each other. These "interneurons" utilize *Stochasto-Chaotic* constraints that allow continuous adjustments in inter-neural perceptions (how they relate to each other) and adjust their perceptional processing accordingly. These recombinant neural fibers represent the continuously recombinant nature and learning nature of this Neural Fiber Network evolution. Layer n + 1, during its generational evolution develop neural fiber connections between layer nodes to aid in the learning process of the Neural Fiber evolution. These intra-neural layer connections allow the network to more efficiently evolve when intra-layer nodes communicate and learn from each other.

During genetic synthesis and recombinant neural fiber generations, connections (uni and bi-directional) are created and assessed. During successive generations of genetic neural structures, nodes may skip a neural generation, depending on the stochasto-chaotic constraints imposed on generational fiber evolution. The internal neural structure conformed to:

$$\tau_i \dot{y}_i = -y_i + \sum_{j=1}^{N} w_{ji}\sigma\left(g_j\left(y_j + \theta_j\right)\right) + I_i, i = 1,\ldots,N \tag{4.1}$$

- Where:
  - y is the state of each neuron
  - t is its time constant
  - $w$ji is the connection from the jth to the ith neuron
  - g is a gain
  - θ is a stochasto-chaotic bias term,
  - $\sigma(x) = \dfrac{1}{\left(1+e^{-x}\right)}$ is the standard logistic activation function,
  - And I represents an external sensor input (depending on the neuron)
- States are initialized utilizing a forward Stratanovich function (with a nominal integration step size of 0.1)

4.4.2 Stochasto-Chaotic Differential Constraints

In order to derive the Stochasto-Chaotic constraints required for the Fuzzy, Continuously Recombinant Neural Fiber Network, we look to Chaotic Calculus [111]. In particular, we produce Chaos expansions for Markov chains via orthogonal functionals that are analogous to multiple stochastic integrals [154]. By looking at environments that converge, orthogonally, to stochastic differentials and chaotic differentials we can capture the environment and determine the existence and connectivity of pulses that form intelligent sequences (in a stochastic and chaotic sense). We look for solutions in chaotic calculus (martingales), whose multiple stochastic and chaotic integrals can be expressed as polynomial solutions (utilizing Meixner and Krawtchouk polynomials), and therefore whose solutions can be constructed utilizing Renyi's mutual information theory. In this way, we can compute these stochastic and chaotic functionals as discrete iterated integrals with respect to a compensated binomial process [41, 42].

We start with deriving the Kreatchouk polynomial differential solutions by generating the Koekoek and Swarttouw function which is a stochastic process and allows us to construct orthogonal functionals of Markov chains. This construction is related to the chaos expansion:

$$\tilde{f}_n\left(k_1,\ldots,k_n\right)=\frac{1}{n!}\sum_{\sigma\in\sum_n} f_n\left(k_{\sigma(1)},\ldots,k_{\sigma(n)}\right),k_1\ldots,k_n\geq 1 \tag{4.2}$$

assuming finite Markov chains in continuous time (finite neural structures). The notion of orthogonal tensor Markov chains where one is stochastic and one is chaotic allows us to solve for the two main conditions of information and data evolution, both stochastic and chaotic evolutions. The pseudo-randomness of the data evolutions due to unknown but deterministic functions provides a standard Markov solution, while the stochastic input through non-linear conditions provides a Chaotic Markov solution that is orthogonal to the stochastic Markov solution. For the non-changing information in the SELF's environment that look like actual random processes, the solutions will be orthogonal. For the pseudorandom-looking processes they will be simultaneously solve a Stochastic and Chaotic equation and should converge in the solution space. The non-pseudorandom "noise" in the environment should solve distinctly orthogonal Stochastic/Chaotic pairs of equations and show up in the solution space as orthonometric pairs of solutions. The isometric Stochastic/Chaotic chain looks like:

$$J^n\left(1_{[1,N]}^{\circ n}\right)=\sum_{d=1}^{d=n}\sum_{\substack{1\leq i_1<\ldots i_d\\ n_1+\cdots+n_d=n}}\frac{n!}{n_1!\ldots n_d!}\prod_{k=1}^{k=d}\phi^{n_k}\left(S_{i_k}\mid S_{i_k-1}\right) \tag{4.3}$$

and from here the Stochastic Markov is constructed as:

$$\overset{\circ}{\varepsilon}_N(z)=\sum_{n=0}^{n=N} z^n J_n\left(1_{[1,N]}^{\circ n}\right)=\sum_{n=0}^{n=N} z^n\sum_{d=0}^{d=N}\sum_{\substack{1\leq i_1<\ldots i_d\\ n_1+\cdots+n_d=n}}\frac{n!}{n_1!\ldots n_d!}\prod_{k=1}^{k=d}\phi^{n_k}\left(S_{i_k}\mid S_{i_k-1}\right),z\in R \tag{4.4}$$

and the Chaotic Markov is constructed as:

$$J_n\left(f_n\right)=\sum_{1\leq i_1<\cdots<i_n} f_n\left(i_1,\ldots,i_n\right)\Phi^1\left(X_{i_n}\right) \tag{4.5}$$

With:

$$f_n=\sum_{1\leq i_1<\cdots<i_n} f_n\left(i_1,\ldots,i_n\right)e_{i_1}\circ\cdots\circ e_{i_n} \tag{4.6}$$

Knowing whether sensory information is random or pseudorandom allows the SELF to determine natural versus man-made information. Nature doesn't utilize pseudorandom sequences. Solutions to the orthogonometric equations becomes the constraints for the Genetic-Neural Fuzzy populations of Neural Fiber Threads, eventually forming a Neural Fiber Network that provides an internal neural structure that can process complex stochastic and non-linear data and information patterns that are encountered within the SELF's ever changing environment.

4.4.3 Continuously Recombinant Neural Fiber Topology

The SELF's internal neural fiber performance is highly dependent on its structure. The interaction allowed between the various Fiber Nodes of the network is specified using the structure only. A Neural Fiber Network structure is not unique for a given problem, and there may exist different ways to define a structure corresponding to the problem. Hence, deciding on the size of the Neural Fiber Network (number of nodes, number of interconnections, number of Fuzzy, Self-Evolving Topical Maps, etc.) is also an important issue. Too small a Neural Fiber Network will prohibit it from adequately characterizing and learning complex information/knowledge patterns; creating a Neural Fiber Network that is too large will be too complex to be of practical use and will consume too much of the SELF's always limited resources.

Determining the SELF's optimal Neural Fiber topology is a complex problem. It is even impossible to prove that a given structure is optimal, given that there may be many Neural Fiber structures that may be appropriate. Different combinations of nodes and connections are tried out so that it gives maximum level of response within the given Stochasto-Chaotic constraints. Such methods rely on overall performance of the Neural Fiber Network, so parts of the network that contributed well are difficult to identify. In human terms, every human's neural pathways are different. It is not possible to determine which structure is optimal, given that every human is different. The use of Evolutionary Programming within the ACNF provides the mechanism for defining its internal Neural Fiber Network topology, with their natural makeup of exchanging information. The search space here is also too big, similar architectures may have quite difference performance; different architectures may result in similar performance. This makes Evolutionary Programming a better choice as opposed to algorithms which start with a maximal (minimal) network and then deletes (adds) layers, nodes or connections when necessary [43].

The genotype representation of the SELF's Neural Fiber Network architecture is critical to the functionality of its Continuously Recombinant Network. Considerations have to be taken so that the optimal Neural Fiber structures are representable and meaningless structures are excluded. The Evolutionary Programming (EP) genetic operators yield valid offspring, and the representation do not grow in proportion to the network. Ideally, the representation should be able to span all potentially useful structures and omit unviable network genotypes. The encoding scheme also constrains the decoding process. A Neural Fiber Network requiring a Continuously Recurrent structure should have a representation expressive enough to describe recurrent networks. Also the decoding mechanism should be able to read this representation and transform it into an appropriate recurrent network.

The low-level or *direct encoding* techniques specify the Neural Fiber connections only. *Indirect encodings* are more like grammatical rules; these rules suggest a context free graph grammar according to which the Neural Fiber Network topology can be generated. Direct encoded genotypes increase too fast in length with a growing network. Thus, the maximum topological space has to be limited. This may exclude the fittest structure in the lot, or may result in networks with special connectivity patterns.

One of the major challenges with evolving the Neural Fiber Network was to find a meaningful way to crossover disparate Neural Fiber topologies. Usual genetic operators will fail to preserve the structural innovations occurring as part of the evolutionary process. Some kind of a speciation is required so that individuals compete primarily within their own niches, and not with the population at large. This is why EP was utilized to guarantee that the new parental population was not too far deviated from previous generations. The EP algorithms used utilize methods for historical markings, speciation, and incremental growth from minimal structure for efficient evolution of the Neural Fiber Network topology [35].

The EP algorithms divide the population into different species on the basis of a compatibility distance measure, utilizing the FUSE-SEMs. This measure is generally derived from the number of disjoint and excess genes between two individuals. If an individual's distance measure from a randomly selected one is less than a fuzzy membership value, then both individuals are placed into the same species. Once the classification is done, the original membership values are adjusted by dividing by the number of individuals in the species. A species grows if this average adjusted fitness is more than the population average, otherwise it shrinks in size. By doing so, the EP algorithms not allow any particular structure dominate over the whole population, but at the same time allows for the growth of the better performing ones, providing both local and global optimization of the Neural Fiber Network (L^2 vs. $L\infty$).

It should be noted that the same input–output mapping can be implemented by different Neural Fiber Network architectures. For a given data environment, the topology for Recombinant Neural Fiber Network is not unique, in that the genotype representation of two structurally different Neural Fiber Networks will be different even though the functional mapping they define may be same. EP algorithms are not able to detect these symmetries and hence a crossover in such a case would very often result in an unviable offspring. Moreover, in Neural Fiber Networks where more than one signal needs to be learned, there are chances of incompatible roles getting combined leading problems with Neural Fiber Network convergence. A simple solution to these problems is to restrict the selection operator to small populations, and to introduce intuitive biased measures in crossover and mutation.

The ACNF's Continuously Recombinant Neural Fiber Network is capable of learning very high-order possibilistic correlations that are present in a continuously changing data environment. The learning algorithms provide a powerful mechanism for generalizing behavior to new environments. For these Neural Fiber Networks, endogenous goals play an important role in determining behavior and EP methodologies are the appropriate mechanism for developing goals and purposeful behavior.

The EP algorithms are computationally expensive, but are necessary when little or no prior information about the data environment are available. More effectively, they are good algorithms to start with the design and once some knowledge is gained, other purposive algorithms may be designed to come up with the solution faster. Parallel implementations of these algorithms will also become more and more purposeful as the need for designing real world applications arises [92].

4.5 Discussion

Chapter 4 has laid out what it means to be cognitive and has described the concepts and architectures for an ACNF capable of providing synthetic, human-like, cognitive capabilities. One of the major components of the processing framework for the SELF that allows it to process, store, and retrieve information effectively is its overall memory system. The next chapter discusses Artificial Memory Systems required for the SELF to ingest, encode, store, and recall (construct) information and knowledge within its cognitive infrastructure.

Chapter 5
Artificial Memory Systems

At their very heart, memories involve the acquisition, categorization, classification and storage of information. The purpose of memory is to provide the ability to recall information and knowledge as well as events that have happened to us in the past. We base our current understanding of the world around us on what we have learned and stored in the past and we react to that same environment relying on the memories of what has happened before, and what has been learned in the past. Without our memories, day-to-day living is not manageable. It would require continuous abstract thought and continuous reiteration of the most basic functions, analogous to the symptoms of an Alzheimer patient. Without memories, we wouldn't be able to drive a car, brush our teeth, or perform any of the things we do "without thinking about them." Through our abilities of conceptual recollection of past memories we are able to reflect, infer, and even communicate with other people.

Thus, implementation of an autonomous SELF requires these same abilities. Memories are typically classified into three different types: Sensory, Short-Term, and Long-Term. Each memory type is designed to support different types of time based system context processing functions. We will explore each type of memory system and the implications to our SELF.

We begin our discussion of memory types with a look at the relationships between the three main types of memories. Figure 5.1 illustrates our SELF Memory Upper Ontology which describes these relationships.

5.1 Artificial Context in Memory Systems

In order for the SELF to mimic human reasoning and processing, it must be provided with real-time and recursive cognition-based information discovery, decomposition, reduction, normalization, encoding, and memory recall (knowledge construction) [229]. Thus, a SELF must be able to recombinantly assimilate information content into knowledge [229]. To accomplish this, the ability to develop context specific relationships is required. These relationships are developed using

J.A. Crowder et al., *Artificial Cognition Architectures*,

DOI 10.1007/978-1-4614-8072-3_5,

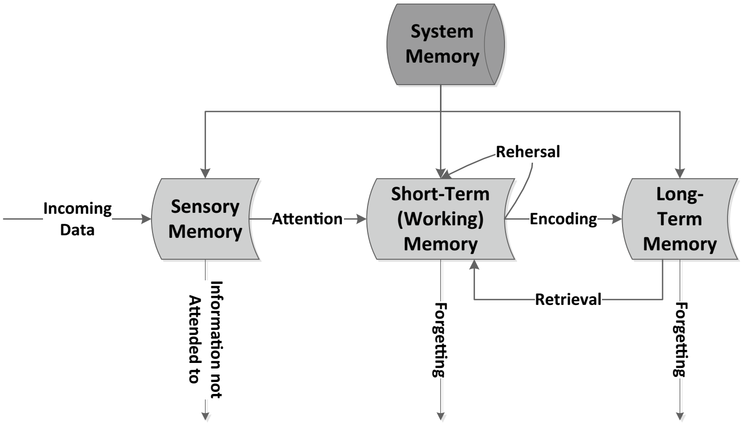


Fig. 5.1 SELF artificial memory upper ontology

knowledge relationship threads sewn together during SELF development of each system's cognitive objective [229]. Concept management of these objectives requires a layer of ongoing processing known as a Cognitive Conceptual Ontology (CCO) [90] in order to be able to "think" about, correlate and integrate information, over time, into its overall memories. A CCO's objective is to manage concepts using the internal semantic language of a SELF just as humans do when they count and/or conceive thoughts in their mind using their native language. Analogously, multilingual individuals are sometimes asked, "What language do you count, think, or dream in?" When describing how science integrates with information theory, Brillouin defined knowledge succinctly as resulting from a certain amount of thinking and distinct from information which had no value, was the "result of choice," and was the raw material consisting of a mere collection of data [26, 27]. Additionally, Brillouin concluded that a 100 random sentences from a newspaper, or a line of Shakespeare, or even a theorem of Einstein have exactly the same information value. Therefore, information content has "no value" until it has been thought about and thus turned into knowledge [27]. Subsequently, knowledge generated is ultimately used continuously for making decisions usually resulting in various levels of inference and specific activity. Each activity usually comprises some form of action and reaction, which leads rapidly to a discussion of appropriate inference and appropriate action by a SELF.

Decision-making is a great concern due to the requirement for handling ambiguity and the ramifications of erroneous inferences. Often there can be serious consequences when actions are taken based upon incorrect recommendations and can influence decision-making before the inaccurate inferences can be detected and/or even corrected. Underlying the data fusion domain is the challenge of creating actionable knowledge from information content harnessed from an environment of vast, exponentially growing structured and unstructured sources of rich complex interrelated cross-domain data. As expected, dealing with ambiguity is a major challenge for humans, as well as, the autonomous SELF. This will be discussed at length throughout the entirety of this book.

Dourish [104, 105] expressed that the scientific community has debated definitions of context and it's uses for many years. He discussed two notions of context, technical, for conceptualizing human action relationship between the action and the system, and social science, and reported, "ideas need to be understood in the intellectual frames that give them meaning." Hence, he described features of the environment where activity takes place [103]. Alternatively, Torralba [207] derived context based object recognition from real-world from scenes, described that one form of performing the task was to define the 'context' of an object in a scene was in terms of other previously recognized objects and concluded, that there exists a strong relationship between the environment and the objects found within, and that increased evidence exists of early human perception of contextual information. Dey [98] presented a Context Toolkit architecture that supported the building of more optimal context-aware applications. He argued, that context was a poorly used resource of information in computing environments. To him, context was information, which must be used to characterize the collection of states or as he called it the "situation abstraction" of a person, place or object relevant to the interaction between a user and the application. Similarly, when describing a conceptual framework for context–aware systems, Coutaz et al. [39] concluded that context informs recognition and mapping by providing a structured, unified view of the world in which a system operates. The authors provided a framework with an ontological foundation, an architectural foundation, and an approach to adaptation, which they professed, "…all scale alongside the richness of the environment." They concluded that context was critical in the understanding and development of information systems. Winograd [212] noted that intention could only be determined through inferences based on context. Hong and Landay [133] described context as knowing the answers to the "W" questions (e.g. Where are the movie theaters?). Similarly, Howard and Qusibaty [135, 136] described context for decision making using the interrogatory 5WH model (who, what, when, where, why and how). Lastly, Ejigu et al. [109] presented a collaborative context aware service platform, based upon a developed hybrid context management model. The goal was to sense context during execution along with internal state and user interactions using context as a function of collecting, organizing, storing, presenting and representing hierarchies, relations, axioms and metadata.

This chapter describes the importance of context in memory systems and the importance of applying a cognitive processing framework and memory encoding in conjunction with a storage methodology for capturing contextual knowledge, as well as, the importance of appropriate inferencing and having decision making constructs within the system which effectively use a knowledge repository that can efficiently manage the cognitive concept development process using internal language for each instance of specific context.

5.2 Sensory Memories

The Sensory Memory within the SELF's memory system are memory registers where raw, unprocessed data/information are ingested via a SELF's environmental sensors and placed into preconscious buffers to begin initial processing. An ACNF's sensory memory system should be designed to accommodate large capacity for large quantities of possibly disparate and diverse information from a variety of sources. Additionally, large sensory observations also have the characteristic of requiring processing within a short duration of time; hence, high volume, high processing rates. Analogously, a human ingests and processes hundreds of thousands of sensory perception events per second. Information buffered in sensory memory must be sorted, categorized, turned into information fragments, metadata, contextual threads, and attributes (including emotional attributes) and then sent on to the working memory (Short-Term Memory) for initial cognitive processing. Based on the information gathered in this initial Sensory Memory processing, Cognitron threads are generated, creating discrete initial "thoughts" about the data and aggregated into context specific CCO hypotheses. The thought process information, along with the sensory information is passed to working memory on the backs of Cognitrons while the Artificial Cognition processes within the ACNF are alerted. Figure 5.2 illustrates the data/information content primitives which make up a SELF's Sensory Memory Lower Ontology.

Sensory Memory, depicted in Fig. 5.2, as a minimum, is decomposed into a few of the senses which could be generally required for autonomous SELFs (e.g. audio, visual, textual, etc.) depending on the need/gap required. Touch and olfactory sensing are examples of others. Each sense requires a domain specific subsystem that can process the essential elements or data primitives. Each sub-system is responsible for providing context using the appropriate memory interfaces, existing knowledge, to the CCO for managing knowledge relativity threads discussed later in this chapter.

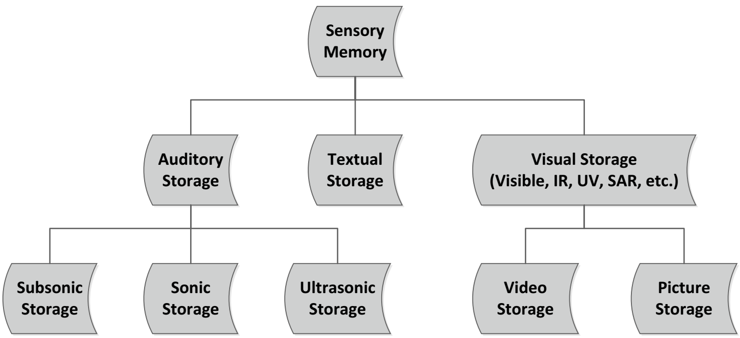


Fig. 5.2 SELF artificial sensory memory lower ontology

5.3 Short-Term Artificial Memories

Short-Term or "Working" memory within the SELF's ACNF is where new information is temporarily stored while it is being processed. Just as in humans, Short-Term Memory (STM) is also where most of the reasoning within a SELF occurs. STM is divided into two major constructs, known as, "rehearsals": Elaborate and Maintenance rehearsals. The ACNF continually refreshes, or rehearses, memories while they are being processed and reasoned upon, so memories do not degrade until they can be placed into Long-Term Memory. Figure 5.3 illustrates the Short-Term Memory Lower Ontology for the SELF.

Figure 5.3 illustrates the Short-Term Memory Lower Ontology for the SELF. Elaboration and Maintenance rehearsals are processes in which humans attempt to comprehend their current environment in context with the newly ingested information. The contexts are created by the senses and the inferences are generated by internalizations of time, space or current location, and what the possible transmitted/communicated response might be. Figure 5.3 also shows how each of the rehearsal types are tied to an Episodic buffer which is the event buffer of short term information content being processed.

5.3.1 *Short-Term Memory Attention Processing*

As explained above, in the human brain, Short Term Memory (STM) corresponds to that area of memory associated with active consciousness, and is where most

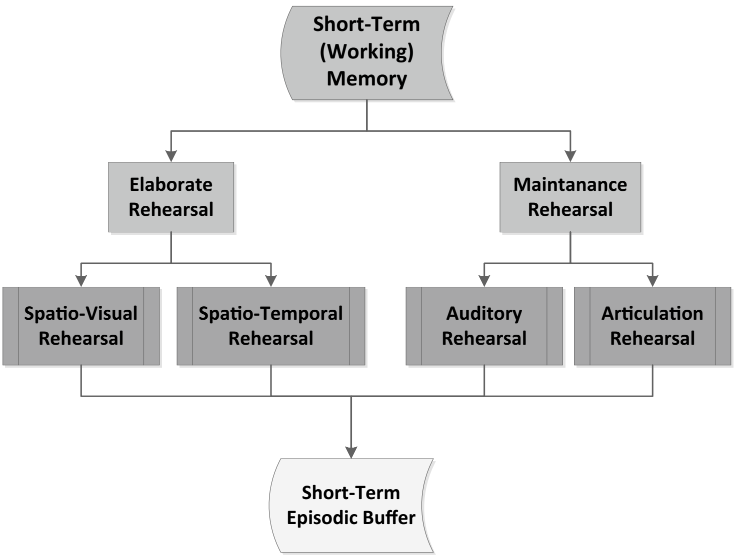


Fig. 5.3 SELF artificial short-term memory lower ontology

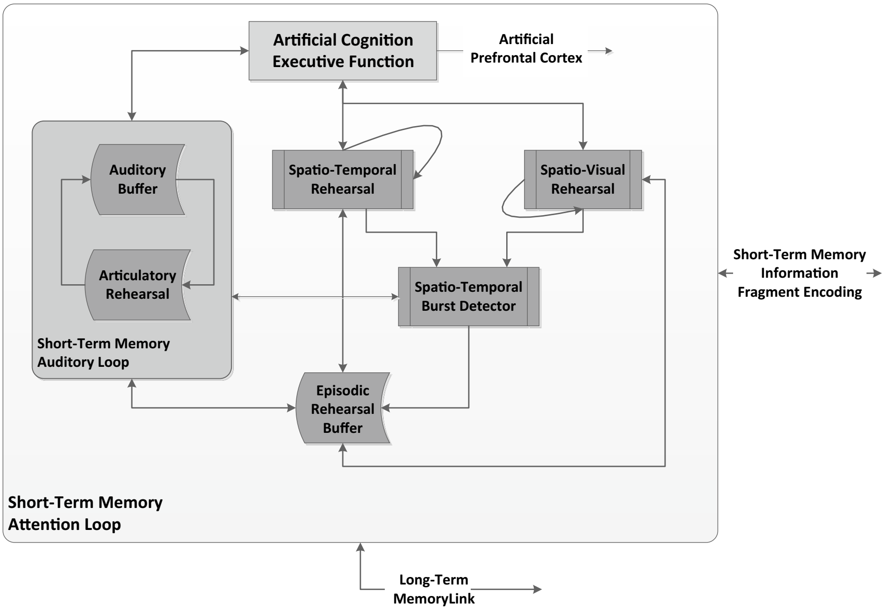


Fig. 5.4 SELF artificial short-term memory attention loop

cognitive processing takes place. The temporary storage requires rehearsal to develop and maintain context currency until placed into the Long-Term Memory (LTM) processing sub-system. A SELF's, memory system does not decay over time, however, the notion of "memory refresh" or context currency through rehearsal is still necessary to maintain real-time qualitative interaction with a SELFs environment. In the ACNF, the rehearsal comprises keeping track of many simultaneous "versions" of independent cognitive concepts within STM as they are processed and continuously evaluated by the Artificial Cognition algorithms. This is illustrated in Fig. 5.4, as the SELF STM Attention Loop.

One of the major functionalities within the STM Attention Loop is the Spatio-Temporal Burst Detector [192, 209]. Within these processes, Information Fragments are ordered in terms of their spatial and temporal characteristics. Spatial[1] and Temporal transitions states are measured in terms of mean, mode, median, velocity, and acceleration and are correlated between their spatial and temporal characteristics and measurements [5]. Rather than just looking at frequencies of occurrence within information, we also look for rapid increases in temporal or spatial characteristics that may trigger an inference or emotional response from the cognitive processes. State transitions bursts are ranked according to their weighting (velocity and acceleration), together with the associated temporal and/or spatial characteristics,

[1] Spatial in this reference can be geographically (either 2-D or 3-D), cyber-locations, or other characteristics that may be considered "spatial" references or characteristics.

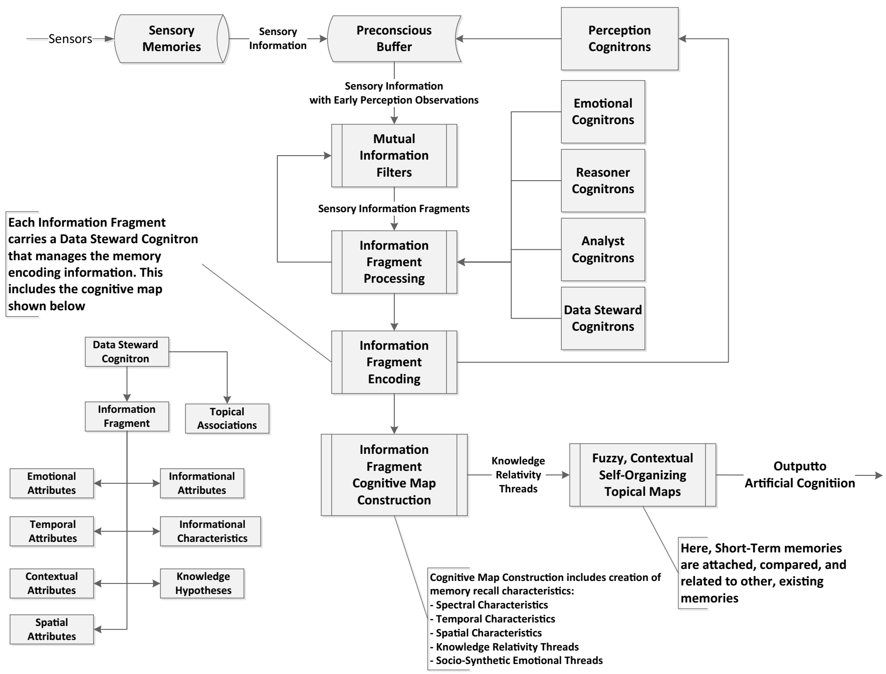


Fig. 5.5 SELF information fragment creation and encoding

and any emotional triggers that might have resulted from this burst processing. This Burst Detection and processing may help to identify relevant topics, concepts, or inferences that may need further processing by the Artificial Prefrontal Cortex and/or Cognitive Consciousness processes.

There are three distinct processes within STM that use initial Recombinant kNowledge Assimilation (RNA) [239] instructions to decompose, reduce, compare, contrast, and discover content to determine the flow of information transfer after cognitive processing. This process is shown in Fig. 5.5. These are:

- **Information Fragment Selection**: involves filtering incoming information from Preconscious Buffers decomposing them into separable information fragments and then determining which information fragments are relevant to be further processed through reduction, comparison, contrast, and then stored, and acted on by the ACNF. Once information fragments are created from the incoming sensory information, they are analyzed and encoded with initial topical information, tagged as metadata attributes that follow information content through the flow of cognitive processes, continuously organizing and integrating incoming information fragments into the SELF's overall LTM system. The information fragment encoding creates an initial, cognitive map used by organization and integration functions.
- **Information Fragment Organization**: processes within the ACNF create the internal structural representation of knowledge and context comprising attributes

within the information fragment cognitive map that allow it to be organized for integration into the overall SELF LTM sub-system. These attributes define how information is represented in LTM and determine how memory fragments are used to construct, or recall memories. These constructs are known as Knowledge Relativity Threads [80, 229] that capture knowledge context of information fragments.

- **Information Fragment Integration**: Once information fragments within the STM have been encoded, they are compared, related, and attached to larger, topical cognitive maps that represent relevant subject or topics within the SELF LTM system [199]. This reasoning process will be described in detail in Chap. 8. Once information fragment cognitive maps have been integrated, processed, and reasoned about, including analysis of emotional triggers and defined as emotional memory information, they are queued up for processing by both the LTM and Artificial Prefrontal Cortex sub-systems to determine required actions. At this point, the STM sub-system should have completed all memory encoding, mappings to topical associations, and their contexts captured. If the representations created are deemed relevant to "remember" they are stored in one of the Long Term Memory systems.

5.4 Long-Term Artificial Memories

Long-Term Memory (LTM), in the simplest sense, is the permanent place where we store our memories. If information we take in through our senses doesn't make it to LTM, then we do not "remember" it. The amount of knowledge and context stored should be made configurable, as well as, the value proposition that determines the threshold of importance. Information that is processed in the STM makes it to LTM through the process of rehearsal, processing, encoding, and then association with other memories. In the brain, memories are not stored in files, or in a database. Memories, in fact, or not stored as whole memories at all, but are stored as information fragments. The process of recall, or remembering, is a process of rapid reconstruction of memories from information fragments that are stored in various regions of the brain, depending on the type of information. Thus, we propose developing a SELF in a similar manner, mimicking human reasoning. The SELF LTM Lower Ontology is illustrated in Fig. 5.6, describing three main types of LTM:

- Explicit, or Declarative Memories
- Implicit Memories [34]
- Emotional Memories

5.4.1 Explicit or Declarative Long-Term Memories

Explicit or Declarative Memory is utilized for storage of "conscious" memories or "conscious thoughts." Explicit memory carries those information fragments utilized

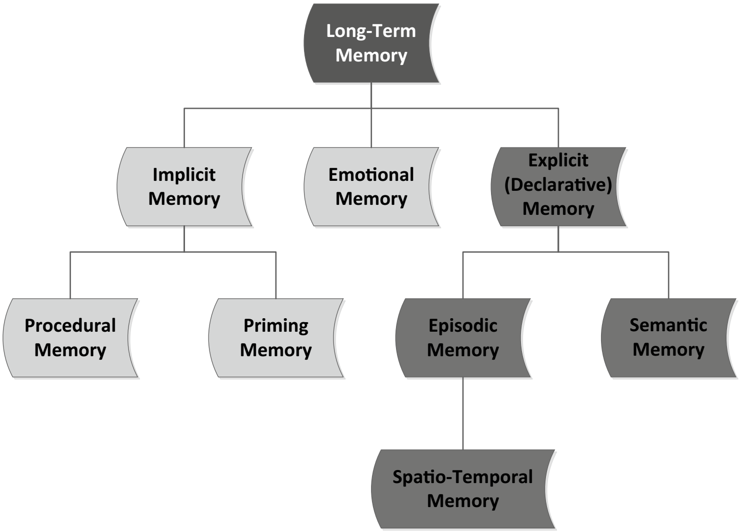


Fig. 5.6 SELF long-term memory lower ontology

to create what most people would "think of" when they envision a memory. Explicit memory stores, objects, events, and things that are experienced in the person's environment. Information fragments stored in Explicit Memory are normally stored with relationship attributes describing or pointing to other information fragments that relate in some fashion. The more meaningful the association, the stronger the memory and the easier the memory is to construct/recall when chosen. In the SELF, Explicit Memory is divided into different regions, depending on the type or source of information. This is because different types of information fragments within the SELF's memories are encoded and represented differently, each with its own characteristics that make it easier to construct/recall memories later. In the SELF LTM, we utilize FUSE-SEMs as described in Chap. 4, to associate currently processed Information Fragments from the STM with memories stored in the LTM. LTM information fragments are known as Binary Information Fragments (BIF) and are not stored in databases or as files, but encoded and stored as a triple helix of binary information (see Fig. 5.7) of continuously recombinant neural fiber threads that represent:

- Binary Information Fragment (BIF) object along with the BIF Binary Attribute Objects (BAOs),
- BIF Recombinant Knowledge Assimilation (RNA) Binary Relativity Objects,
- Binary Security Encryption Threads.

Each fragment of context is represented by an object which carries with it a set of attributes and characteristics used throughout the processing within a ACNF system to describe the essence of a given information fragment. Knowledge Relativity Threads (KRT), described later in this chapter, form relationships between developing objects of context. These objects are known as Binary Relativity Objects (BROs), a name given to them to simply represent a point in time in object development

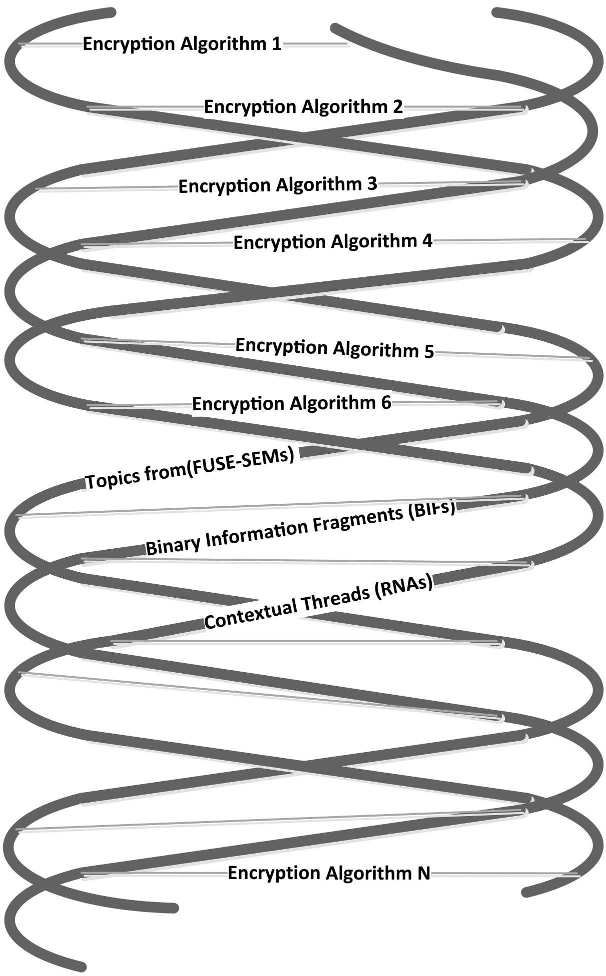


Fig. 5.7 SELF LTM triple Helix encoding

when an object's context is being related to another. Built into the RNA Binary Relativity Objects are another time based type known as Binary Memory Reconstruction Objects (BMRO), which are designated as such when based on the type and source of BIF, the memories are to prepositioned for recall purposes or reconstruction. There are several types of Binary Memory Reconstruction Objects, discussed later in this chapter:

- Spectral Eigenvectors that represent a memory used in the reconstruction process and characterized by implicit and biographical LTM BIFs
- Polynomial Eigenvectors that characterize memory reconstruction using Episodic LTM BIFs

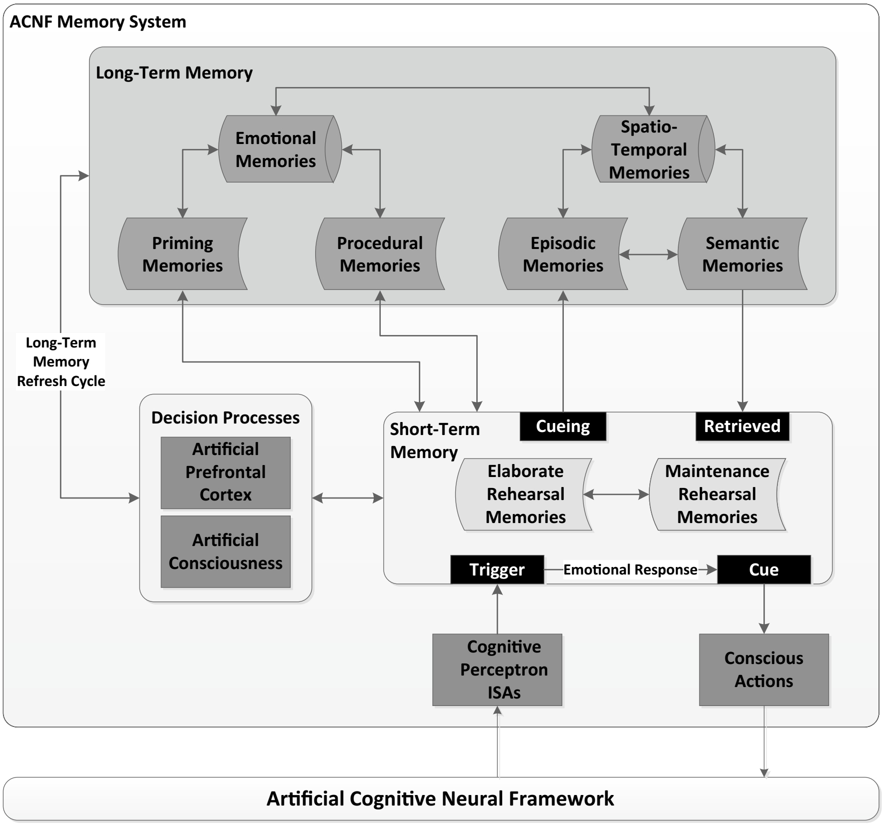


Fig. 5.8 SELF high-level memory architecture

- Socio-Synthetic Arousal State Vectors that characterize memory reconstruction using Emotional LTM BIFs
- Temporal Confluence and Spatial Resonance coefficients that characterize memory reconstruction using Spatio-Temporal Episodic LTM BIFs
- Knowledge Relativity Contextual Gravitation coefficients that characterize memory reconstruction using mathematical and Semantic LTM BIFs

5.4.2 Long-Term Spatio-temporal Memories

In Fig. 5.8, the SELF Spatio-Temporal memory is depicted as a special type of Episodic Memory specifically designed to store complex temporal (time-based) and/or spatial (geographically-based) memories [179]. These can be thought of as

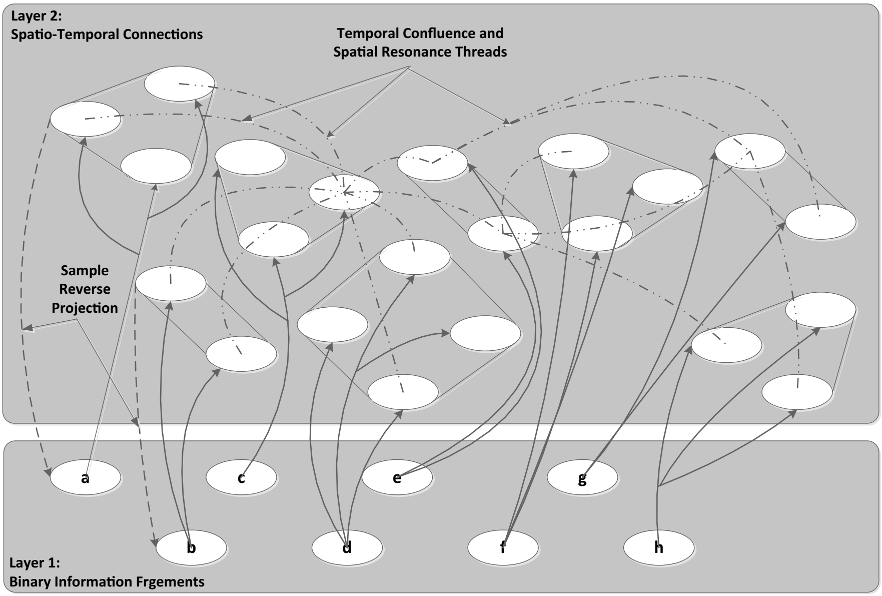


Fig. 5.9 SELF LTM spatio-temporal episodic memory

State-Sequence memories that track memory characteristics as they change over time and or space.

Figure 5.9 depicts temporal memory consisting of two layers: Layer 1, the lower layer, corresponds to Information Fragment Objects that represent topical object entities (e.g. car, person, lamp, etc.) and also SELF environment encountered events. Layer 2 represents each LTM Information Fragment Objects relationships to other LTM Information Fragment Objects relative to time and geolocation utilizing weighting coefficients [49, 50]. This provides complex time and spatial relationships to be learned, maintained, encoded, stored, and recalled within the SELF. These relationships are stored as weighted Temporal Confluence and Spatial Resonance coefficients, based on their relative strength [71, 74].

5.4.3 Long-Term Semantic Memories

Semantic Memory is where meanings, understandings, and concepts/concept-based knowledge are stored. Semantic memory is normally unrelated to experiences, context, or relevance. Semantic Memory is, in general, a collection of factual knowledge that has been learned about the world. The Information Fragments stored in Semantic Memory does not need to involve a specific event or emotion, but the memory is an object or relationship that is just fact. An example would be the relationship, “A screwdriver is a tool.” This memory may be stored without any relationship as to how this fact was learned. As explained above, these relationships are stored as Knowledge Relativity and Contextual Gravitation coefficients and are used during memory construction/recollection.

5.4.4 Long-Term Implicit Memories

Implicit memories carry information used to perform tasks without realizing that recall has happened. In humans, for example, once we learn to drive a car, we perform the actions while driving without specifically trying to recall how to perform them. We unconsciously recall how to perform the actions and act on them without "conscious" thought. This type of information is carried in Implicit Memory. There are two distinct types of Implicit Memories within the SELF, Procedural and Priming Memory.

5.4.4.1 Priming Implicit Memory

Priming Memories carry triggers that allow us to recall, or reconstruct memories faster. They act as a stimulus, or activation mechanism for memories. Exposure to a stimulus during learning results in an implicit trigger to be stored that may be triggered by a similar stimulus while reconstructing a memory later. This can occur during the storage of any type of *LTM*. In the SELF, Priming Memories occur during creation of the Knowledge Relativity Threads. Knowledge Relativity Meta-tags that are relevant to the correlation of sensory inputs, e.g., when both Visual and Auditory contextual correlation is present in memory creation, are captured and stored as Priming Memory cues that aid in memory recall/reconstruction under similar circumstances. This does not have to invoke an emotional response, just a correlation response to previously encountered situations. It, in essence, is a facilitated perception of sensory stimuli that is precipitated by earlier exposure to such stimuli. These triggers, or contextually correlated memory meta-tags allow the SELF to more rapidly recall/reconstruct memories that are particularly relative to sensory cues.

5.4.4.2 Procedural Implicit Memory

Willingly or unwillingly, consciously or unconsciously, humans live scripted or learned lives. This is not to say that one can look ahead into the next chapter and find out what one will have for dinner a month later. Life brings surprises and the scripts we discuss are not the "scripts of life," but rather much smaller groupings of events that represent familiar routines and are stored as procedural memory in our brains. Scripts as large structured chunks of information, typically sequences of events describing standard routines, permeate human life, society, culture. Humans are well aware of them and have the ability of thinking of and manipulating the whole scripts at any level or detailization, or grain size. Thus, when you buy a new iPhone, you must program or set it up. This is a script that your manual describes by chunking it up into setting up your calendar, email, GPS, etc., each of each is also a script. Within the script of date and time, a few clicks will set you the date and a few others the time of the day. Within the latter, there is a tiny subscript of setting up the hour, and another to set up the minutes. Other, less well-defined scripts seem to be capable of almost infinite grain size refinement.

Procedural Implicit Memories allow previously learned tasks to be performed without specific “conscious” memory recall/reconstruction of how to perform the task. Procedural memories tend to be inflexible, in that they are tied to the task being performed. For example, when we decide to ride a bike, we don’t unconsciously recall/reconstruct memories of how to drive a car, we recall/reconstruct unconscious Procedural Memories of how to ride a bike. In a SELF, tasks that are learned and are deemed ‘important’ to capture for future use will have Procedural Memories stored as steps, or “procedures” that are required to perform the same task in the future. The specific Procedural Memories would be tied to the particular domain for which a SELF is designed.

In his work on Procedural Memory and contextual Representation, Kahana showed that retrieval of implicit procedural memories is a cue-dependent process that contains both semantic and temporal components [144]. Creation of Procedural Memories is tied not only to task repetition but also to the richness of the semantic association structure [220]. Earlier work by Crowder, built on Landauer’s Procedural Memory computational models and Griffith’s topical models [221], theorized about the creation of artificial cognitive procedural memory models based on Knowledge Relativity Threads to create the semantic associations [73] and work in Fuzzy, Self-Organizing, Semantic Topical Maps [59, 60] counted on the topical model needed to create long-term procedural memories. These Knowledge Relativity models and Topical Maps are based on early work by Zadeh. Zadeh [219], described tacit knowledge as world knowledge that humans retain from experiences and education, and concluded that current search engines, with their remarkable capabilities, did not have the capability of deduction, that is the capability to synthesize answers from bodies of information which reside in various parts of a knowledge base. More specifically, Zadeh describes fuzzy logic as a formalization of human capabilities: the capability to converse, reason and make rational decisions in an environment of imprecision, uncertainty, and incompleteness of information. In their work in cognition frameworks, Crowder and Carbone [72, 75] expand on the work not only by Zadeh but also by Tanik [112] in describing artificial procedural memories as procedural knowledge gained through cognitive insights based on fuzzy correlations made through a labeled form of an FUSE-SEM (discussed in Chap. 4) that provides the following attributes:

1. Contextual algorithms explore the map visually for informational connection located by meaning.
2. Procedural searches utilize semantic contextual information to find links to relevant procedural information.

3. The informational maps autonomously locate temporal and semantic associations that provide procedural connections to a topic.
4. The FSSOM represents a normalized representation of any physical information content used in the development of the procedural knowledge and content.

5.4.5 Procedural Memory Description

In artificial intelligence, procedural information is one type of knowledge that can be learned and carried by a Cognitron [148]. From the initial research in the 1998 and 1999 [88, 89], work has continued on the development of artificial memory systems that mimic human processing, storage, and retrieval. It is believed that providing a cognitive framework that mimics human processing and reasoning also requires creating a constructive memory system similar to human memory storage and processing [87, 122]. The initial work in artificial memory systems involved the use of Cognitrons to create the overall artificial cognitive framework [44, 45]. This work led to investigation into Linguistic Ontologies used to facilitate conceptual learning in the creation of artificial neural memories [44, 45, 47].

5.4.5.1 Creation of Artificial Procedural Memory Scripts

Continued investigation, utilizing the work of Kahana [144] in associative episodic memories [43], led to the development of a Cognitron framework for creation, storage, and retrieval artificial implicit memories [78, 79, 81] (see Fig. 5.7). Based on this work, a systems and software architecture specification was developed for an artificial cognitive framework utilizing Cognitrons [91].

The main hypothesis here is that the procedural memory scripts can be detected and acquired with the combination of rule-based computational semantic techniques enhancing the SELF's understanding of repeatable and useful procedures. The objectives of artificial procedural memories are:

1. To identify the procedural memory script acquisition using a combination of meaning-rule-based techniques from the Ontological Semantic Technology with meaning- and cognitively-enhanced machine-learning techniques from Cognitive Artificial Intelligence.
2. To develop the principles of comparison of the comprehension of natural language by the SELF (see Fig. 5.10).

Crowder, in conjunction with Carbone and Friess, in researching artificial neural memory frameworks that mimic human memories, are creating computer architectures that can take advantage of Raskin and Taylor's Ontological Semantic Technology [191, 204] and create an artificial procedural memory system that has

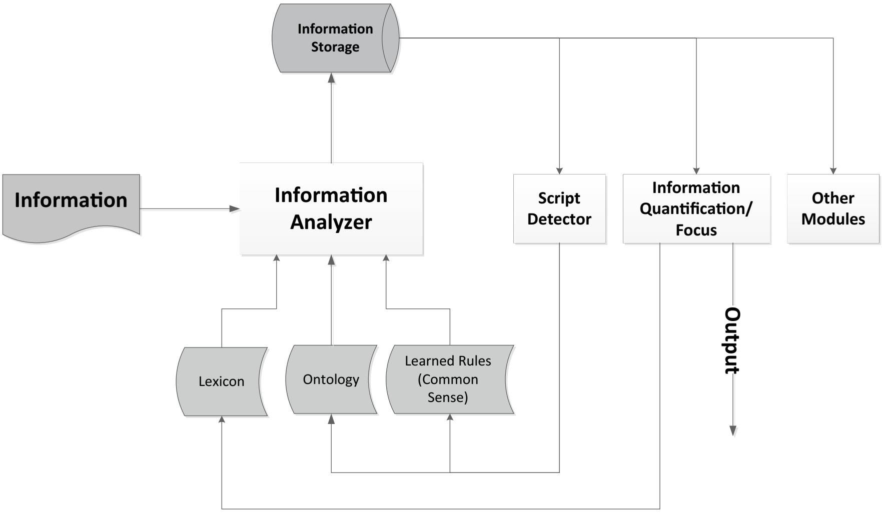


Fig. 5.10 SELF artificial procedural memory generator

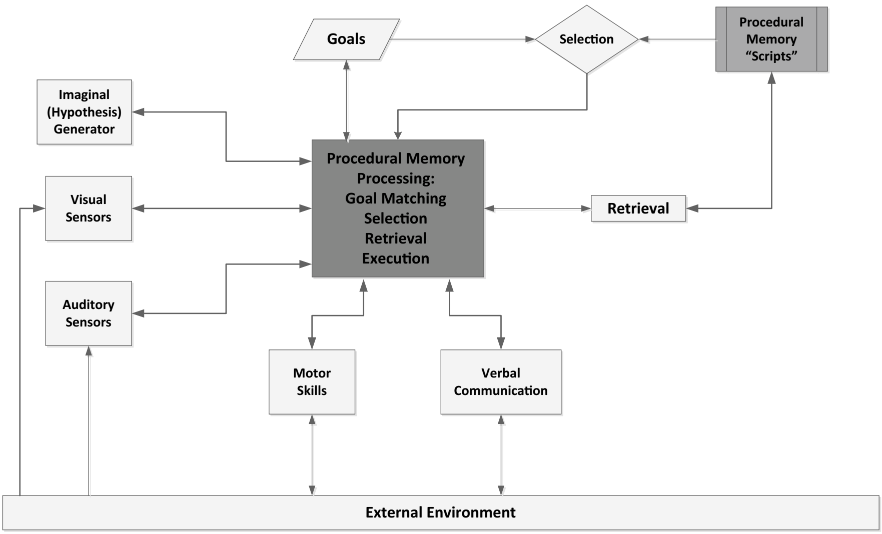


Fig. 5.11 SELF artificial procedural memory retrieval

human reasoning capabilities and mimics the fuzzy and uncertain nature of human cognitive processes. This new focus for Crowder [90] is to create processes necessary for the creation, storage, retrieval, and modification of artificial procedural memories (see Fig. 5.11).

5.5 Group Consciousness and Memory Sharing

In order to facilitate self-evolution within the SELF, each cognitive subsystem, each Cognitron, each and every part of the system must be able to learn from every other part of the system. In essence, the combination of all of the Cognitrons within the SELF form a collective, or group consciousness that drives how the system learns, reasons, and behaves. In order to facilitate this collective group consciousness, there must be memory sharing across the entire system. And while the system has a collective set of LTMs, these memories, their implications, their contexts, and their emotions must be broadcast, or transmitted, to each part of the system so that each Cognitron can evaluate how they are affected by learning and 'remembering' that goes on in other Cognitrons. SELF System-Level goals must be evaluated in terms of this collective group consciousness. The first step is to develop a Goal-Oriented Knowledge Ontology that can be used by the Cognitrons to evaluate their own goals and objectives in light of the collective goals and objectives of the entire SELF. Figure 5.12 illustrates the SELF Goal-Oriented Knowledge Lower Ontology.

In the ACNF, one of the functions of the Mediator, or Artificial Prefrontal Cortex is to manage this Group Consciousness by correlating LTM Information Fragments with the ongoing real-time Cognitive Consciousness comprised of behaviors, cognitive processes, current goal and objectives, emotions, contextual knowledge, etc. The Metacognitive and Metamemory processes correlate all of this information and send information to the reasoning processes and broadcast relevant information to the Cognitrons currently operating in the system. Figure 5.13 illustrates this process.

In order for the collective group consciousness to understand and make use of cognitive information from the collection of Cognitrons operating within the ACNF, capturing and relating the context of cognitive information is crucial, for it forms

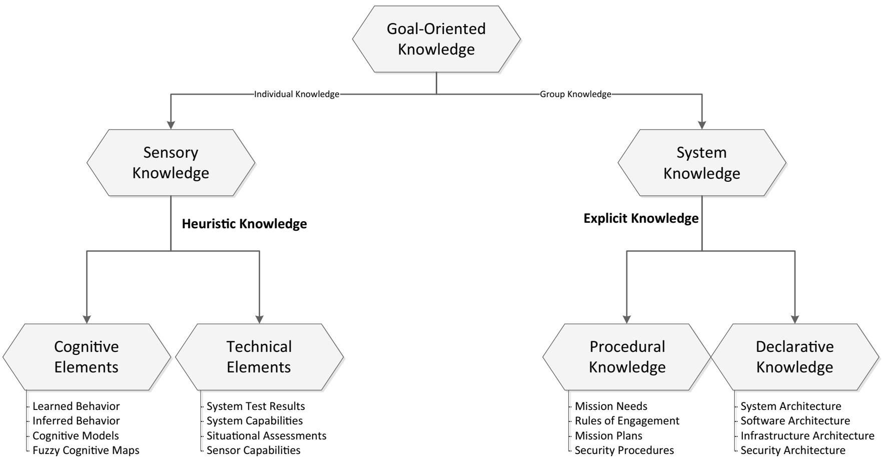


Fig. 5.12 SELF goal-oriented, collective knowledge

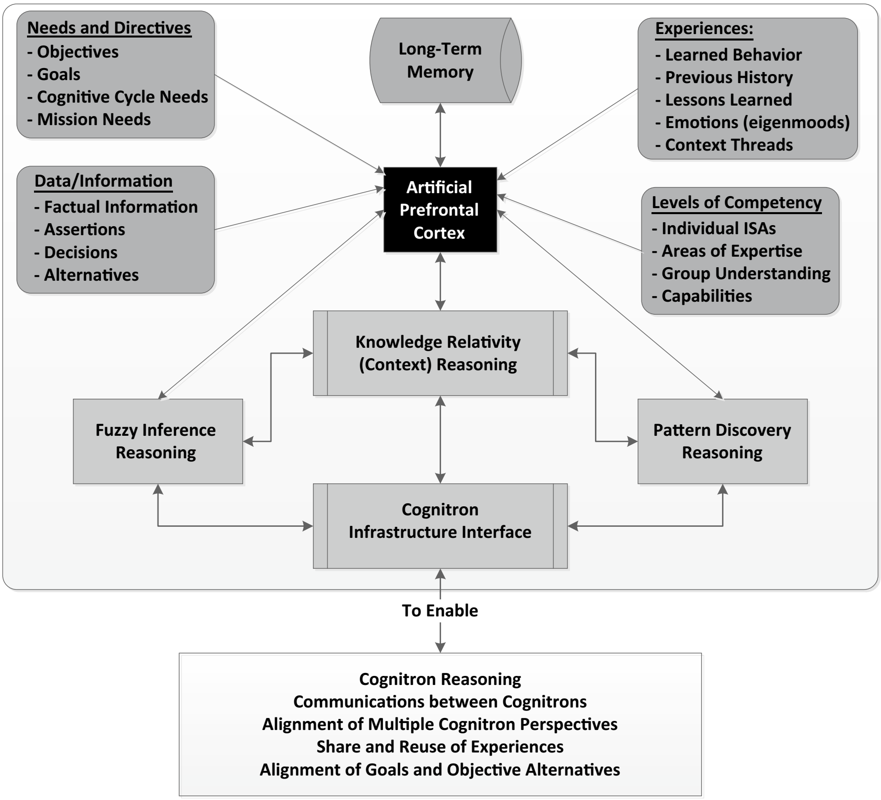


Fig. 5.13 SELF collective group consciousness architecture

that causal connectedness between the Information Fragments each Cognitron is processing and their relevance to other Cognitrons, and to past experiences captured in LTM.

5.6 Emotional Memory

As was discussed earlier, memories about emotional situations are often stored in both Explicit and Implicit *LTM* systems. Figures 5.14 and 5.15 illustrate the basic structure for the AIS *Emotional Memories*. Figure 5.16 shows the structure for storage of emotional memories and Fig. 5.17 illustrates construction/retrieval of emotional memories.

Figure 5.16 illustrates a high-level comparison of the *AIS Artificial Central Nervous System* with the human central nervous system. The Emotional Memories are based on Dr. Peter Levine's Autonomic Nervous System Trauma States [163, 167]. This provides the framework to map Levine's Nervous System States into artificial system states that can be tied to *Artificial Neural Emotions* and *Emotional Memories*.

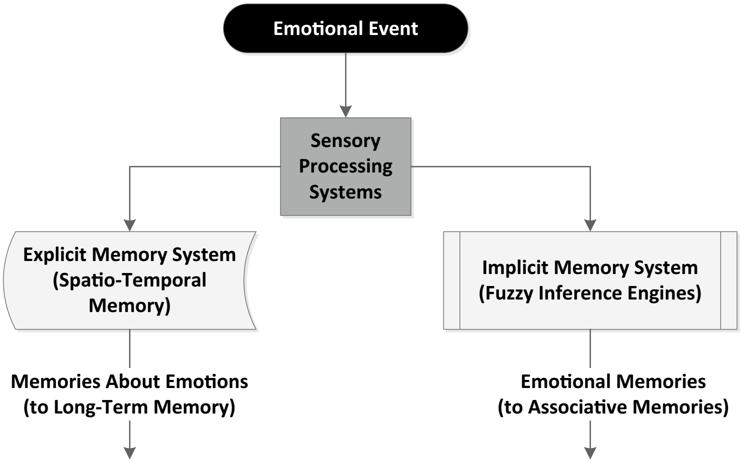


Fig. 5.14 Formulation of SELF emotional memories

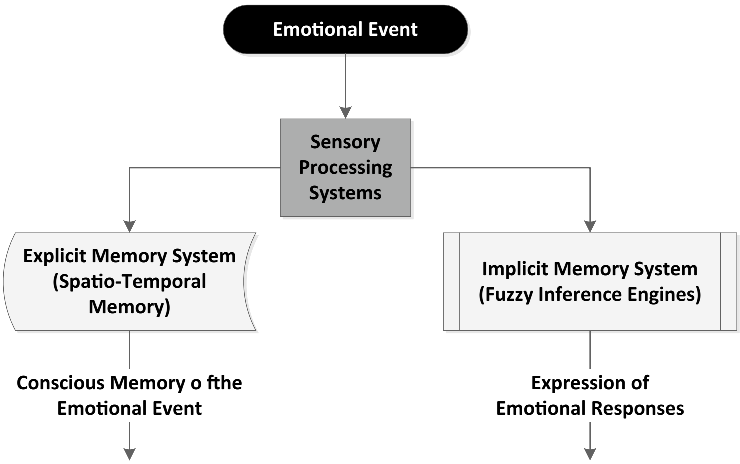


Fig. 5.15 Retrieval of SELF emotional memories

This, combined with Emotional Cognitron*s*, provide the constructs for artificial emotional control, as illustrated in Fig. 5.17. Combined, these provide the Cognitive Processing and Cognitron environment to allow artificial neural emotions and emotional learning within the SELF.

5.6.1 SELF Artificial Autonomic Nervous System States and Emotional Memories

Just as the amygdale and hippocampus are involved in implicit and explicit emotional memories, respectively, the ACNF and the Cognitron Coalitions become

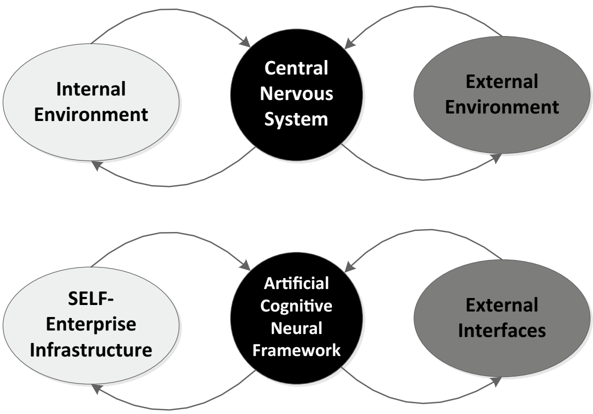


Fig. 5.16 Comparison of SELF connections versus human central nervous system connection

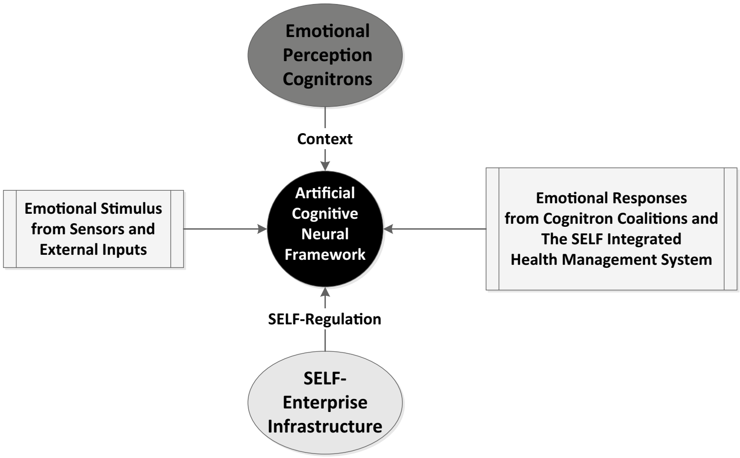


Fig. 5.17 SELF ACNF artificial emotional control structure

emotionally aroused when they form semantic and episodic memories about situations that cause "stress" within an artificial neural system. Stress situations may involve a loss of resources, new data environments that are unfamiliar or new interfaces that are introduced into the environment. These cognitive representations of emotional situations better referred to as memories about emotions rather than emotional memories.

In human emotions, emotional arousal often leads to stronger memories. This is a statement about explicit memories involving emotional situations (memories about emotions). The effects of emotional arousal on explicit memory are due to processes that are secondary to the activation of emotional processing systems in the ACNF. For example, in a situation of danger (say in an artificial neural system controlling a weapon system), processing of threatening environment stimuli would lead to the activation of the active cognitive emotion Cognitrons within the ACNF, which, in

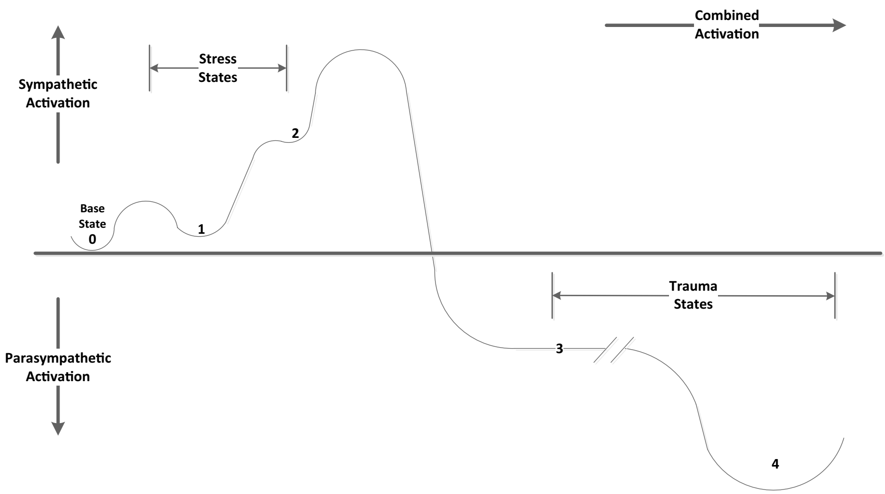


Fig. 5.18 Dr. Peter Levine's autonomic nervous system states

turn, would transmit information to neural structures within the infrastructure of the system and system network. Activity in these areas would be detected by the cognitive coalitions and would lead to increases in system emotional arousal (due to activation of modulation) within the neural structure that leads to the release of cognitive problem, solution, search, and emotion Cognitrons. The transmittal of informational content as well as emotional context allows information construction/retrieval performance to be greatly enhanced, allowing for "cognitive economy" within the artificial neural systems.

5.6.2 SELF Artificial Autonomic Nervous System States

Figure 5.18 illustrates the Autonomic Nervous System States described by Dr. Peter Levine [163]. However, the descriptions provided have been revised to fit within the context of the overall SELF and the ACNF. These descriptions are "fuzzily" encoded within the Information Fragments and within the Cognitive Consciousness.

These provide the artificial neural emotional states that correspond to system states and are stored, along with data and information to allow rapid retrieval and transmittal within the overall cognitive framework when similar situations present themselves for analysis and problem solving.

0: Base State: System is calm, current cognitive Cognitrons can easily respond to input (external interfaces). The artificial neural system is in a state of Pendulation (the SELF ACNF is in a natural rhythm supporting the basic process of contraction and expansion of system resources, corollary is the movement between tension and relaxation or inhalation and exhalation in human autonomic systems).

1: Mild Stress: Active, heightened state of Cognitive Awareness. The SELF ACNF will allocate an increased number of Cognitrons in order to solve the current situation. Actual evolution takes place in this state as the Cognitive Consciousness collects information and makes inferences. Inferences about the emotions connected with the situation are categorized and stored in Emotional Memory, while the informational content is stored in Spatio-Temporal Memories. Short-term emotional responses are stored in the STM for processing by Analyst Cognitrons for possible immediate response.

2: High Stress: A hyper-alert, panicky state that in humans provokes fight or flight responses. In the ACNF, it invokes a massive creation of Reasoner Cognitrons, as well as a massive increase in messaging Cognitrons to broadcast the emotional situation and information to as large a population of Cognitron Coalitions as possible. This promotes rapid thoughts and evolution of Cognitrons, and causes rapidly changing and extreme artificial neural emotions. This happens in an extreme situation when the system is in jeopardy of mission failure of shutting down completely. In this state it will consume large amounts of system resources. Emotional Memory will include and predict the need for system resources required for problem solving should the situation arise again.

3: Mild Trauma: The heightened feeling of panic and hysteria (in neural system terms) is still present, however it is now an underlying emotion and the system appears to be in a dormant state, not able to find a solution to the problem at hand. It human terms, this state is appropriate for a situation that might need to be passive emotions, i.e., after a trauma when it is important to rest and gather one's energy for a sudden outburst. In the ACNF, this is facilitated through an increased burst of genetic algorithms [36] that search every possible solution space (hypotheses generation) in order to provide an evolved solution that was previously unavailable and then allows a sudden burst of activity to provide mission solutions.

4: Severe Trauma: The artificial neural system is perceived dormant or shut down. There is a lack of cognitive activity and Emotion Cognitrons are suppressed in this state. There are eruptions of activity like those in State 3 and flashes of extreme evolution similar to State 2. This state is appropriate when the perceived threat to the system (either internally or externally) is overwhelming. This may occur in the ACNF when all external interfaces are unavailable and the system is devoid of input and no solution is imaginable within the current emotional and information states within the system memories. This causes a disconnection of the Cognitron from its current Emotional Memory and a flurry of evolutionary activity is required to allow solution spaces to be evolved without emotional influence that could interfere with the evolution of a possible solution space. When solutions are available, neural connectivity to the rest of the system is reestablished and a new set of neural emotions and Emotional Memories are established and new neural pathways are established and "remembered."

The use of Emotional Memories within the ACNF provides the constructs and mechanisms for rapid retrieval of memories as well as rapid broadcast of contextual information unavailable with non-emotion based cognitive systems.

5.7 Memory Recall in the SELF: Memory Reconstruction

As discussed above, the SELF must be able to store and recall (reconstruct) complex sequential patterns. We have discussed the episodic and semantic memories and the need for the SELF to demonstrate these memory properties within the SELF's cognitive domain for a large number of spatio-temporal memory patterns (episodes), given only a simple example or representation of such a pattern (a partial memory). The SELF's cognitive system must be able to construct memories from information fragments without significant interference, as well as exhibit similarity-based category generalization described as semantic memory properties in humans [146].

5.7.1 Constructivist Memory Theory

In order to design, develop, and implement the SELF to be truly autonomous, it must be provided with dynamic memory abilities [57]. Memories are typically classified into three different types: Sensory, Short-Term, and Long-Term. Each memory type has several instantiations, dealing with different types of information [15]. The SELF's cognitive processes are based on Constructivist Learning (which will be discussed in Chap. 7), in which the ACNF cognitive learning processes are a building (or construction) process in which the SELF's cognitive system builds internal illustrations of its learned knowledge-base, based on its experiences and personal interpretation (fuzzy inferences and conceptual ontology [191, 204]) of its experiences. The SELF's Knowledge Representation and Knowledge Relativity Threads [71], within the SELF's cognitive system memories are continually open to modification, and the structures and linkages formed within SELF's short-term, long-term, and emotional memories [77], along with its Knowledge Relativity Threads (KRT) [72], form the bases for which knowledge structures would be created and attached to the ACNF memories, which are stored as Binary Information Fragments [85] as discussed in previous chapters.

5.7.2 Artificial Memory Reconstruction

Given that the SELF's memories are not stored as database files, but as Binary Information Fragments with Knowledge Relativity Threads that provide contextual and meta-information, memory recall, similar to humans, is a process of reconstructing the memory, based on topical maps that map the topic or subject to be "remembered" into the Cognitive Conceptual Ontology (CCO) and Information Base to find those information fragments that are relative to the topics(s) to be pulled, associated, integrated, and presented as a memory to the SELF's conscious processes. The process outlined below illustrates the memory reconstruction

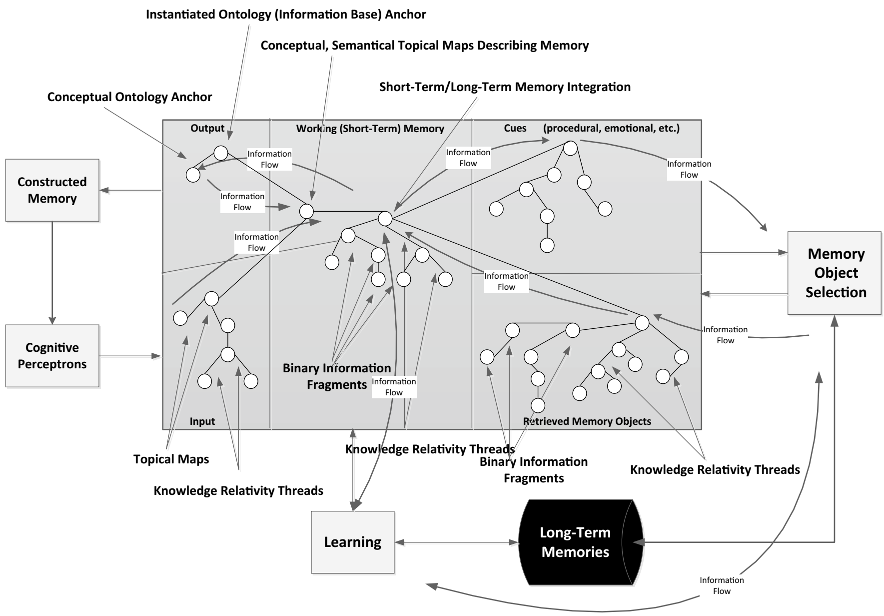


Fig. 5.19 The SELF memory reconstruction architecture

process (see Fig. 5.19) [68]. Figure 5.20 illustrates an example of the process of building a Short Term Memory construction.

1. For each memory reconstruction node from LTM, compute the total contextual relativity weights from the current set of Binary Information Fragments associated with the memory reconstruction node. These contain Knowledge Relativity Threads related to the concepts involved it the memory reconstruction:

$$\psi_{i,t} = \Sigma x_{j\in\Gamma_t} W_{ji} \tag{5.1}$$

2. Normalize the values from each Topical Map that relate to a concept within the Conceptual Ontology, related to LTM Binary Information Fragments. i.e., we find the maximum value above and divide by all the individual values by the greater of the maximum and the F-matrix threshold ${}^F\Theta_t$. This is done to ensure that the memory feed-forward signals are not amplified in subsequent normalization steps to avoid catastrophic interference in the memory reconstruction process:

$$\psi_{i,t} = \frac{\psi_{i,t}}{max\left(max(\psi_{j,t}), {}^L\Theta_t\right)} \tag{5.2}$$

3. Analogous to step 1 – only for Short Term Memories:

$$\phi_{i,t} = \Sigma_{j=\Delta_{i-1}}, t > 0 \tag{5.3}$$

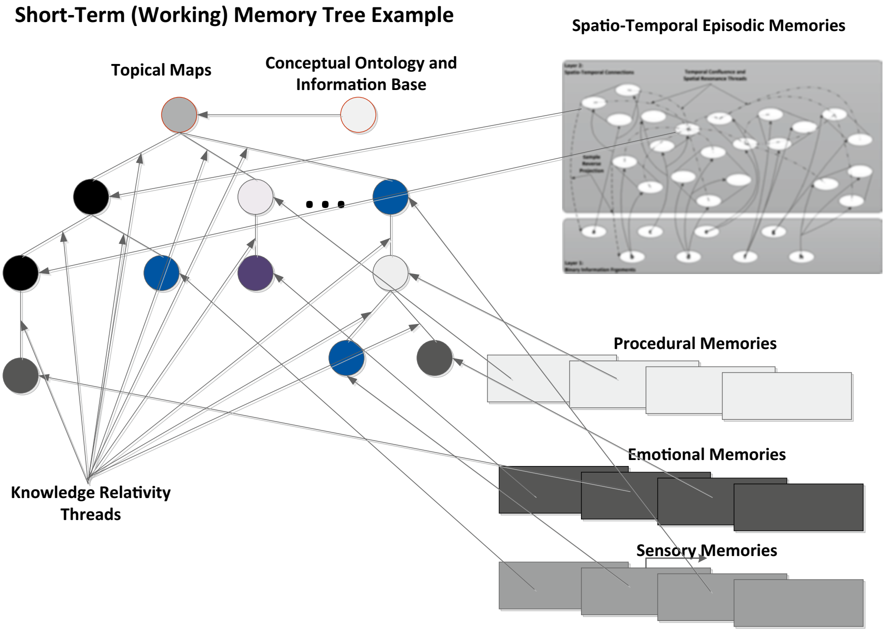


Fig. 5.20 The SELF short-term memory tree construction

4. Analogous to step 2 – only for Short Term Memories:

$$\Phi_{i,t} = \frac{\phi_{i,t}}{max\left(max_{j\in CO}(\phi_{j,t}),\ {}^{S}\Theta_{t}\right)} \tag{5.4}$$

5. The memory reconstruction process is an iterative process (assessing whether the reconstruction process meets the conscious memory's requirements and constraints). Here, the Knowledge Relativity Threads from LTM and STM are put through filters to produce a generalization gradient for their relativity to the memory to be constructed, based on the Semantic, Self-Organizing Topical Maps, utilizing adapted gravitational theory [71]. This gives them membership values between 0 and 1:

$$\chi_{i,t} = \Psi^{\mathrm{u}}_{\mathrm{i,t}}\Phi^{\mathrm{v}}_{\mathrm{j,t}} \tag{5.5}$$

6. In step 6, we look for those memory fragments that are most relevant within the context of the Conceptual Ontology and Information Base:

$$\Phi_{i,t} = \frac{\phi_{i,t}}{max\left(max_{j\in CO}(\phi_{j,t}),\ {}^{S}\Theta_{t}\right)} \tag{5.6}$$

7. Once we have all of the Binary Information Fragments from LTM and STM that are relevant, given the Topical Maps that may be relevant, we look for those

Topical Maps that make the most sense, given the strength of their membership functions for each set of combined ST and LT memories. Again, this is based on the SELF's current CCO:

$$\pi_{k,t} = max_{j \in CO} X_{j,t}, \quad 1 \leq \mathrm{k} \leq \mathrm{Q} \tag{5.7}$$

8. This is in the process in case of an "unexpected" result, based on cues from procedural, emotional and spatio-temporal memory creation. This triggers if the actual combined memory is unexpected, given the current temporal context of the SELF's past experience and clues. This may be an inference for a new concept not currently in the SELF's CCO. In this case, we get a compressive nonlinearity that maps the required information into a new concept within the CCO:

$$G_t = \Sigma_{k=1}^{Q} \frac{\pi_{k,t}}{Q} \tag{5.8}$$

9. Finally, we simply choose the winners of the nodes and provide the constructed memory out to the SELF's Cognitrons for use by the SELF's conscious processes and Cognitrons:

$$p_{i,t} = \frac{f\left(X_{i,t}, G_t\right)}{\Sigma_{j \in CO} f\left(X_{i,t}, G_t\right)} \tag{5.9}$$

5.8 Discussion

We have provided the architecture and methodologies for artificial memory encoding, storage, and reconstruction. This provides the SELF with the abilities to process, encode, and retrieve information (memories) similar to human memory processing. These architectures and processes are only necessary if the SELF has the capabilities of human-like learning and reasoning, along with an artificial consciousness that allows integration, control, and management of all of the SELF's cognitive capabilities. The next chapter will discusses and describe an architecture and framework for Artificial Consciousness within the SELF.

Chapter 6
Artificial Consciousness

To develop "Artificial Consciousness" for a SELF requires investigation and understanding of what it means to be conscious. The textbook definition of consciousness is:

> Knowing and perceiving; to have awareness of surroundings and sensations and thoughts; showing realization or recognition of oneself and ones surroundings.

Another view of consciousness is to be in total control of one's mind; in control of life and circumstances. Part of the notion of being conscious is to be able to perceive the world around you. Consciousness is ability to understand and react to your reality; your environment. Our conscious mind has an internal map of our experiences and our emotions, and affects our reactions to external stimulus or information. Our experiences assist in determining our decisions, our strategies, and in turn affect the way we perceive and sort incoming information and the way we store and remember.

A prerequisite for a SELF consciousness includes methodologies for perceiving its environment, take in available information, make sense out of it, filter it, add to internal consciousness, learn from it, and then act on it. Here, we describe Artificial Neural Perceptrons as the mechanisms to provide "perception" characteristics to a SELF. Information travels through the SELF on the backs of perceptrons which coagulate to form various granular cognitive thoughts, perceptions, and concepts known as Cognitrons. Cognitrons communicate with each other to form an artificial "collective consciousness" within a SELF. Information is gathered and sorted (filtered) via Knowledge Relativity Threads (KRT) which build system context and knowledge corpus while FUSE-SEM topical maps provide a SELF the internal formation of knowledge density that help a SELF decide what information is relative to things the SELF already knows, or has learned about and remembered (stored as memories).

J.A. Crowder et al., *Artificial Cognition Architectures*,
DOI 10.1007/978-1-4614-8072-3_6,

6.1 Artificial Neural Cognitrons

Theory into human consciousness postulates that human cognition is implemented by a multitude of relatively small, special purpose processes, almost always unconscious [87]. These processes are autonomous and narrowly focused. They are efficient, high speed, and make very few errors because their purpose is narrowly focused. Each of these human processes can act in parallel with others. In a SELF ACNF, this is accomplished with fuzzy-neural perceptrons (Cognitrons) [88]. Each Cognitron is accomplished through the use of codelets; small pieces of code that each performs one specialized, simple task. Codelets play the role of waiting for a particular type of situation to occur and then acting upon it per its own specialization. These Cognitron codelets are miniature fuzzy-neural structures each with a specific purpose and facilitated by well-defined constraints, and are designed to adapt or analogously learn and evolve. As mentioned earlier, the ACNF provides perceptrons access to short and long-term memories and have the ability to communicate with other codelets as needed. Hence, in human cognitive theory, codelets can be thought of as cell assemblies or neuronal groups [58].

In order to initiate artificial consciousness within a hybrid, synthetic cognitive neural structure, it is necessary, like in the human brain, to create an architecture that provides a constructive neural environment where network learning algorithms extend the neural architecture as needed [148]. This facilitates a massively parallel, highly interconnected network of loosely coupled, relatively simple processing Cognitrons, called "experts," in a hybrid, fuzzy, genetic neural system of "M expert" architecture [87]. The purpose for this constructive cognitive processing environment is to provide a hybrid, artificial neural environment that is adaptable to a variety of classes of operations:

- Language Processing
- Signal Detection and Processing
- Sensor and Information Fusion
- Inductive, Deductive, and Abductive Reasoning

The ACNF model, described in Chap. 4, suggests a number of possible specialized roles within the ACNF for artificial emotions and artificial cognition necessary to produce motivations and goals, and to facilitate learning within the SELF via ACNF Cognitron development. Each of the various components of the ACNF, perception, attention, behaviors, expectations, and interactions, are described below:

Perception: for perception, sensory stimuli (both external and internal inputs) is received and interpreted by the perception processes, providing meaning and context for the sensory inputs. There are several perception processes within the Artificial Cognition subsystem [69, 80].

(a) **Early Cognitive Perception**: Here input arrives via the sensors. Multiple specialized perceptrons attach to the sensory inputs and extract those features relevant to their specialty. If features are extracted, each perceptron will broadcast observations, analyses, and thought processes onto the Artificial Cognition processes.

(b) **Cognitron Coalition**: it is possible for multiple perceptrons to be activated and utilized for a given set of input from the sensors. The Attention Manager within the Artificial Prefrontal Cortex (the Mediator) forms coalitions of perceptrons known as Cognitrons, and facilitates convergence from the different sensory information, along with its context. In this process, relevant emotions and possible emotional memories are recognized and identified along with the objects and contextual information from the various memory systems within the ACNF. This emotional reaction to external inputs may entail a simple emotional response based on a single Cognitive Emotional Cognitron (CEC), or it may involve a very complex emotional memory or response that requires additional convergence of several CECs.

(c) **Cognitive Preconscious Buffers**: The perceptions gained from processing of external sensory inputs, along with its meaning and context, are stored in working memory, called preconscious buffers, before the information is sent on to the Artificial Cognition subsystem. Depending on the type of sensory information, these buffers could involve spatio-temporal, spatio-visual, auditory, or other types of information. Emotions and emotional memories may be part of this preconscious perception, depending on what features and triggers are extracted and discovered. These emotions, memories, and contexts are part of the preconscious perceptions stored in the preconscious buffer memories that are transferred to the Artificial Cognition processes during each cognitive development cycle.

(d) **Cognitron Associations**: The Artificial Cognition processes utilize the incoming Cognitrons, along with the preconscious buffer memories/information as cues for creating cognitive hypotheses to provide reasoning and inferences about incoming sensory information. This includes emotional and contextual information. The Artificial Cognitive processes form local Cognitron associations from information retrieved from the transient episodic memories and the long-term associative memories [13]. Emotional memories and emotional cues are utilized in order to add emotional context that aids in creating the local cognitive associations. These local associations contain recorded logs of Cognitrons past emotions and emotional memories that are contained in the associated situations close to or coincident with the sensory input hypotheses.

(e) **Cognitron Competition**: "Attention" Cognitrons bring relevant and/or urgent events to the Artificial Consciousness processes. These Attention Cognitrons search the long-term memories, based on input from the Metamemory processes that information may exist in long-term memories about the subject, topic, or hypothesis currently being processed. As information and context are retrieved, it is possible for Cognitron coalitions to be formed, some of which may compete for access to the consciousness processes. This competition may include a number of Attention Cognitron coalitions, including coalitions formed during previous consciousness cycles. It is possible the priorities assigned to coalitions are influenced by the emotional responses, trauma states, and/or other emotional memories. A strong affective emotional response will strengthen a coalition's priority and therefore increase the chances of getting access to the Artificial Consciousness processes.

(f) **Broadcasting Perceptron Consciousness**: A Cognitron coalition, once it has gained access to the Artificial Consciousness processes, has its contents broadcast within the ACNF. This coalition may include an Attention Cognitron, along with its coalition of relation Informational Cognitrons, which carry information content, along with informational context. This ‘consciousness broadcast’ will contain the complete contents of this consciousness object, including the affective (emotional) information. This consciousness broadcast updates the perceptual memory, including the emotional content, which may lead to new emotional memories. The stronger the affective information, the stronger the emotional memories and triggers that are encoded into memory. The Transient Episodic Memories are also updated with the contents of the current consciousness object. During long-term memory cycles, the contents of the episodic memory are consolidated and stored as long-term declarative memory. Procedural Memory may be updated, modified, or reinterpreted, depending on the strength of the affective portion of the consciousness object.
(g) **Cognitron Resource Management**: Behavior Cognitrons respond to the Consciousness Broadcasts. Behaviors are controlled by information from the Consciousness Broadcast that drives the creation of Attention Cognitrons. One type of Attention Cognitron is an Expectation Cognitron that may be created due to an unexpected hypothesis or result from a previous Consciousness Broadcast. In this case, a Cognitron Coalition may be created in order to handle the unexpected situation. This coalition may consist of many types of Cognitrons, allowing the Artificial Consciousness to handle resource management by recruiting resources through the creation of Cognitron Coalitions. The emotional, or affective, content of the Cognitron Coalitions will affect the attraction of relevant resources, including processor utilization, memory availability, and creation of other coalitions, in order to handle the current perceived situation.
(h) **Cognitron Action Selection**: Based on the reactions, analyses, hypotheses, and other information provided by the Artificial Consciousness processes, the Behavior Processes select a behavior, or action, driven by both conscious and unconscious goals carried within the ACNF. This may be the result of a current situation, or a previous situation that has gained higher priority within the Attention Manager. Again, the action selection can be heavily influenced by the emotional content of the coalitions. The relationships between previous, current, and possible future behaviors affect activation of actions, as does the residual activation levels (priorities) from the various choices of actions.
(i) **Cognitron Action Activation**: Based on the selection of action(s), the Behavior Cognitrons set into motion a chain of actionable events that may drive the SELF to perform both internal and external tasks in order to meet its current goals. This will also include a set of Expectation Cognitrons whose task it is to monitor the actions performed in order to provide success/failure information to the Artificial Consciousness processes, based on the expected

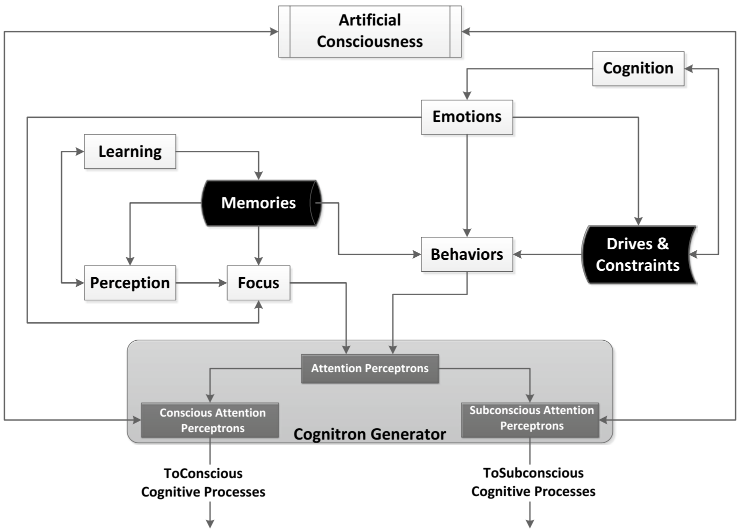


Fig. 6.1 Conscious, Cognitron connectivity

results. The success or failure information may create new hypotheses and Cognitron coalitions in order to deal with this new information from the Expectation Cognitrons.

The Cognitrons provide the ACNF with the ability to mimic human reasoning in processing information and developing knowledge. Figure 6.1 illustrates the ACNF Cognitron Artificial Cognition Infrastructure that drives the coalitions and provides the infrastructure for the hybrid neural processing environment. These Cognitron solution coalitions, in the end, will create additional memories within the ACNF, including emotional memories, based on the overall response of the system to the current situation. The Solution Domain is the front end processor of data and information in which solutions are matched to incoming data. Known solutions (answers to questions or understanding incoming situational information) may require minor adjustments to parametric values and memories, based on subtle changes to known solutions. The latency requirements in this domain are very short. Unsolved or inadequately solved problems or situations are moved to the Evolution Domain for further processing.

6.2 The SELF Mixture of Experts Architecture

The SELF ***Mixture-of-Experts*** architecture allows dynamic allocation of perceptron objects through a divide-and-conquer principle. The neural infrastructure employs genetic programs which possibilistically "***softly***" divide the input space into overlapping regions on which the perceptron "experts" operate [58].

Assuming at time *t*, there are ***M*** local perceptron experts at any given subsystem level of the SELF. Each of the ***M*** experts looks at a common input vector *x* to form an output, $\hat{y}_j(x)g_j(x), j = 1\ldots M$. $g_j(x)$ is a gating function which weights outputs of the perceptron experts to form an overall Cognitron output:

$$\hat{y} = \sum\nolimits_j y_j g_j(x) \tag{6.1}$$

The localized neural gating model is based on a Rammerti model:

$$g_j\left(x,v\right) = \frac{\alpha_j P\left(x|v_j\right)}{\sum_i \alpha_i P\left(x|v_i\right)} \tag{6.2}$$

Where:

$$P\left(x/v_j\right) = \left(2\pi\right)^{-n/2}\left|\sum\nolimits_j\right|^{-1/2} e^{\left\{-0.5\left(x-m_j\right)^T \sum_j^{-1}\left(x-m_j\right)\right\}} \tag{6.3}$$

Thus, the *i*th Cognitron's influence is localized to a region around m_j, with its sphere of influence determined by Σ_j. The formulation for estimating the Cognitron's "expert" network parameters is given by:

$$\begin{aligned}
h_j^{(k)}\left(y^{(t)} \mid x^{(t)}\right) &= \frac{g_j^{(k)}\left(x^t, v\right) e^{\left\{-0.5\left(y-y_j\right)^T\left(y-y_j\right)\right\}}}{\sum_i g_j^{(k)}\left(x^t, v\right) e^{\left\{-0.5\left(y-y_i\right)^T\left(y-y_i\right)\right\}}} \\
\alpha_j^{(k+1)} &= \tfrac{1}{N}\sum\nolimits_t h_j^{(k)}\left(y^{(t)} \mid x^{(t)}\right) \\
m_j^{(k+1)} &= \frac{1}{\sum_t h_j^{(k)}\left(y^{(t)}|x^{(t)}\right)} \sum\nolimits_t h_j^{(k)}\left(y^{(t)} \mid x^{(t)}\right) x^{(t)} \\
\sum\nolimits_j^{(k+1)} &= \frac{1}{\sum_t h_j^{(k)}\left(y^{(t)}|x^{(t)}\right)} \Sigma_t h_j^{(k)}\left(y^{(t)} \mid x^{(t)}\right)\left[x^{(t)} - m_j^k\right]\left[x^{(t)} - m_j^k\right]^T \\
\theta_j^{k+1} &= \arg \overset{\min}{\theta_j} \sum\nolimits_t h_j^{(k)}\left(y^{(t)} \mid x^{(t)}\right)\left\|y - y_j\right\|^2
\end{aligned} \tag{6.4}$$

6.2.1 Dynamic Cognitron Growing and Pruning

The purpose behind dynamic allocation of Cognitron experts within the SELF is to provide complete flexibility in the system as new data/information classes are encountered within the ACNF's Conceptual Ontology. In the SELF dynamic system, it is expected that growing and pruning changes are slow with respect to time, since the decision to add complexity or remove capability should be based on information that has been learned over many iterations of the system [87, 88, 178]. The dynamic error estimate for Cognitron expert is:

$$E_{j,t+1} = E_{j,t} + \lambda_{t+1}\left[\left(y^t - y^t\right)^2 - E_{j,t}\right] \tag{6.5}$$

If the error estimate corresponding to a particular Cognitron expert increases beyond a given threshold, the need for an additional Cognitron expert is detected. The dynamic procedure for adding a Cognitron expert is:

- Initialize the mean vector corresponding to the new Cognitron expert to be equal to the mean vector of the corresponding expert to be split.
- Add a small random perturbation to the two means.
- For a window of length T of input classes identified, make parametric updates to all the Cognitron experts except the expert being split and the new Cognitron expert.
- For a window of length T of input classes, if the posterior h_j, corresponding to one of the new Cognitron experts is the highest among the posteriors of all the experts for a given signal class, make parameter updates for this Cognitron expert also.

The window length T is chosen in such a way to separate the 2-means space. After this window length of T data samples, the two Cognitron experts become part of the normal "Mixture-of-Experts" system. Pruning is performed by monitoring the parameter α_j. When α_j becomes small, the corresponding Cognitron expert is pruned from the system. Figure 6.1 illustrates the connectivity between components of the overall SELF. The individual cognitive processes shown in Fig. 6.1 (e.g., behavior, perceptron, cognition, etc.) are driven by small, single task Cognitrons [54–56].

6.3 Artificial Metcognition: Cognitive Regulation

Recent advances in cognitive computing suggest that computers can do as well as humans, even across multiple information sources and information types [50]. This work by Crowder and Friess established the constructs within an ACNF to provide Metacognitive and Metamemory processing concepts similar to how the human brain operates to process, categorize and link information. Metacognition provides the SELF with a sense of Self-Analysis, or Introspection, allowing the system to "think about what it thinks." Metamemory is a concept of an Artificial Intelligence system's memory capabilities and strategies that can aid in memory representation, retention, mining, retrieval, as well as the processes involved in memory Self-Monitoring. Metamemory constructs for an Artificially Intelligent system has important implications about how the system learns and uses memory. For example, the system can make a judgment on whether it has enough information to complete a mission, known as "judgments of learning." Presented here will be the derivation and application of Metacognition and Metamemory concepts to real-time system utilizing Artificial Intelligence and Cognitrons, including capture and utilization of Emergent Behavior, or concept derivation, allowing a system to preserve its own attention, avoiding interruption and distraction during the analytical process [41, 42].

6.3.1 Artificial Cognition with Metacognition

Metacognition, or ***Knowledge of Cognition***, refers to what a system knows about its own cognition or about cognition in general. In short, it describes the system's ability to think about how and what it thinks. It includes three different kinds of metacognitive awareness: declarative, procedural, and conditional knowledge [93, 178]. They are described as:

- **Declarative Knowledge**: refers to knowing "about" things,
- **Procedural Knowledge**: refers to knowing "how" to do things, and
- **Conditional Knowledge**: refers to knowing the "why" and "when" aspects of cognition.

We can classify Knowledge of Cognition into three components [47, 48]:

- **Metacognitive Knowledge**: (also called Metacognitive awareness) is what the system knows about itself as a cognitive processor [108].
- **Metacognitive Regulation**: is the regulation of cognition and learning experiences through a set of activities that help the system control its learning [156]. This may be based on its understanding of its own "knowledge gaps."
- **Metacognitive Experiences**: are those experiences that have something to do with the current, on-going cognitive endeavors (current mission).

As in humans, we assume that direct inferences are necessary for cognitive processing and Metacognition [94]. For a normal, healthy system, if a cognitive function, $\boldsymbol{F}$, activates a functional cognitive area $\boldsymbol{A}$, it always activates $\boldsymbol{A}$. We utilize this phenomenon to create an initial Conceptual Ontological Framework (Fig. 6.2) of cognitive instances for the SELF.

Each cognitive instance contains expectation derived from experiences the system has had [158, 159]. We can deconstruct the cognitive instance expectations, based on the concepts and skills that drive the expectations, as depicted in Fig. 6.3.

The nature of Metacognition is to understand the structure of one's own cognitive processes, i.e., the ability to think about what we think. To an AI system, the beginnings of Metacognition are founded in understanding that each action the system might take is broken into sub-actions, each of which may have a separate and different contexts (see Fig. 6.4).

As in humans, we assume that direct inferences are necessary for cognitive processing and Metacognition [94]. For a normal, healthy system, if a cognitive function, $\boldsymbol{F}$, activates a functional cognitive area $\boldsymbol{A}$, it always activates $\boldsymbol{A}$.

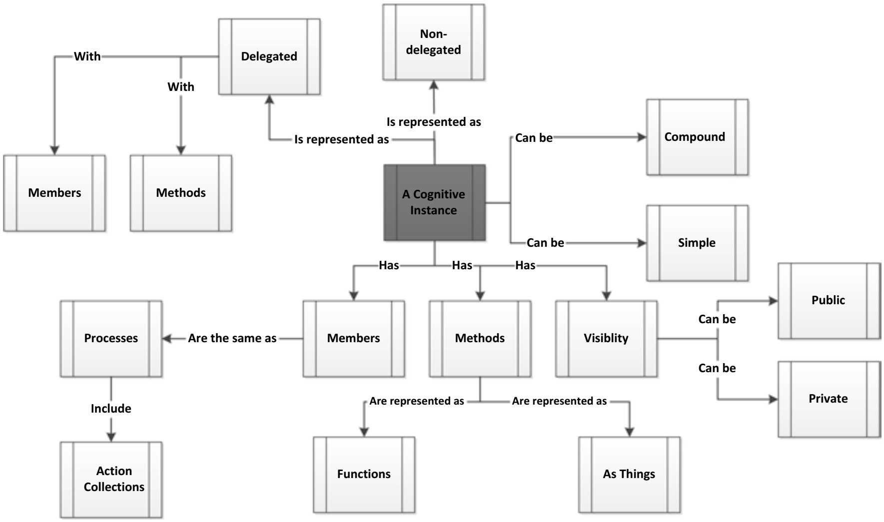


Fig. 6.2 Conceptual cognitive instance framework

We utilize this phenomenon to create an initial Conceptual Ontological Framework (Fig. 6.2) of cognitive instances for the SELF.

Each cognitive instance contains expectation derived from experiences the system has had [157]. We can deconstruct the cognitive instance expectations, based on the concepts and skills that drive the expectations, as depicted in Fig. 6.3.

Here we define the structures and methodologies required to provide Metacognition and Metamemory capabilities to the SELF [71]. Such capabilities are necessary for autonomous systems in order to provide the capability for ***Self-Assessment*** and ***Self-Diagnosis*** [82].

6.3.2 Metacognition: Cognitive Self-Awareness and Assessment

In order to create cognitive self-awareness and assessment in the SELFs, a formal cognitive neural framework is required for determining levels of cognitive granularity and to formalize methodologies for assessing the closeness of cognitive relationships within the SELF. This is facilitated through the ACNF described in Chap. 4, which is a hybrid, fuzzy-neural processing system using genetic learning algorithms. The ACNF architecture is based on a mixture of neural structures that add flexibility and diversity to overall system capabilities (as explained in Sect. 6.2). In order to provide an intelligent processing environment for the SELF that is continually adaptable, we believe the SELF must possess the notion of artificial emotions that allow the processing environment to “react” in real-time as the systems outside the environment change and evolve recursively as recombinant knowledge assimilation [157].

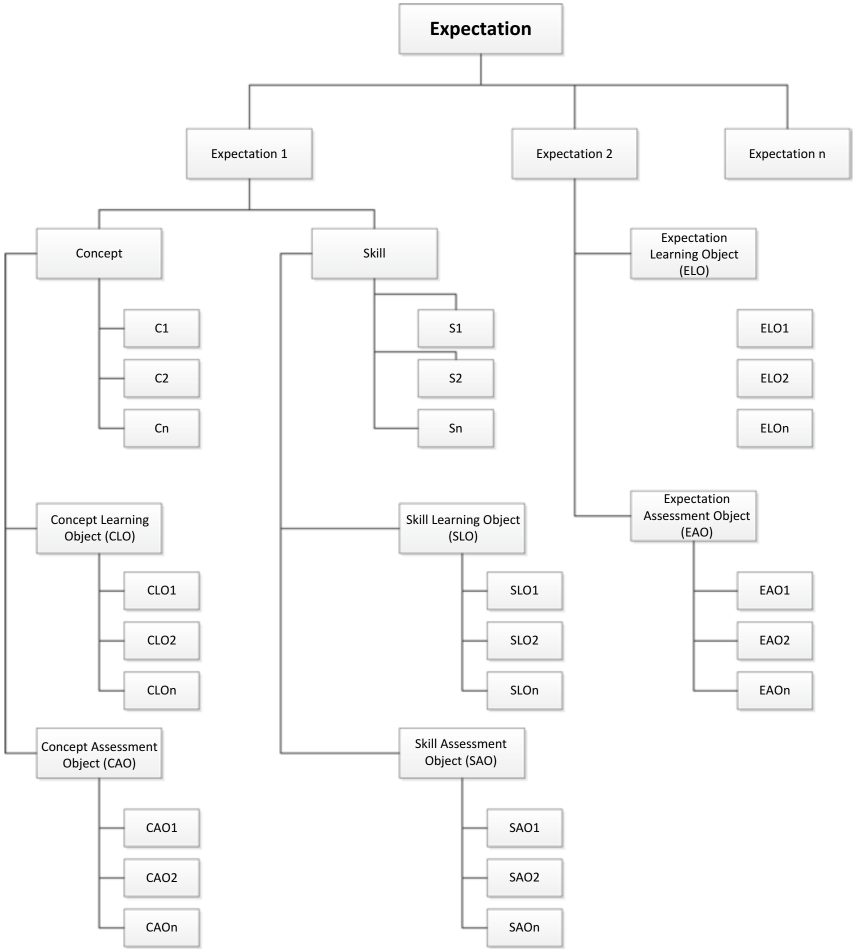


Fig. 6.3 The SELF ACNF cognitive expectation ontology

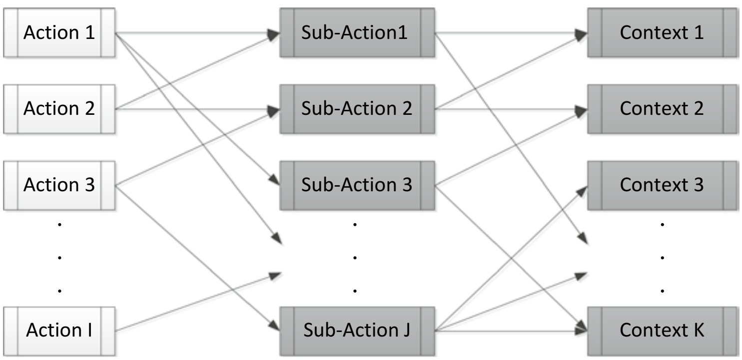


Fig. 6.4 Structure of the Cognitron action process

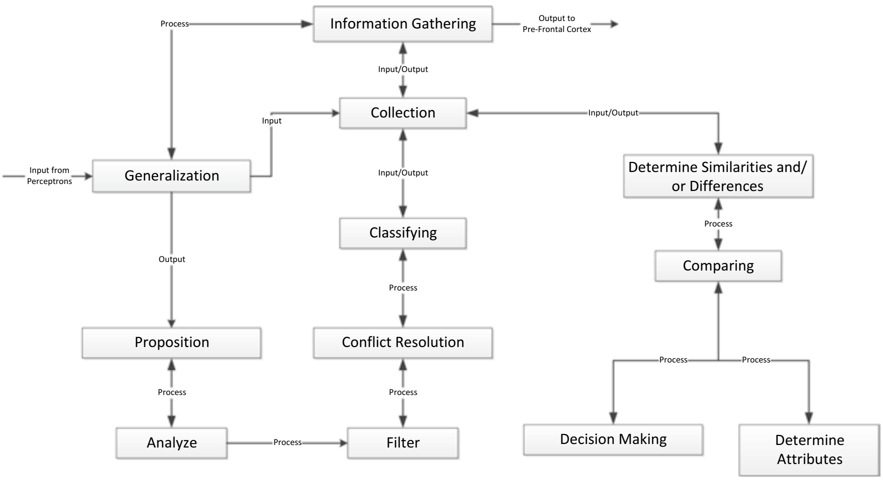


Fig. 6.5 SELF cognitive self-awareness and self-assessment

These constructs provide the necessary components for the SELF to possess cognitive self-awareness and self-assessment processes. Figure 6.5 illustrates the information flow through the ACNF Metacognitive processes.

6.4 Artificial Metamemory: Cognitive Understanding and Learning

Metamemory: is the concept of an AI system's memory capabilities and strategies that can aid in memory representation, retention, mining, and retrieval, as well as the processes involved in memory self-monitoring [156]. A system's self-awareness of memory has important implications about how the system learns and uses memory [94]. For example, the system can make a judgment on whether it has enough information to complete a mission, known as "judgments of learning" [81].

Metamemory concepts for AI system resemble cognitive map constructs that utilize registries for hybrid topical map assimilation [153]. The Metamemory registry allows Cognitive Creation and Discovery within the overall ACNF and drives the conscious analytic inference engines of the system. Figure 6.6 illustrates these Cognitive Creation and Discovery Actions.

The Metamemory Registry has the following properties:

- It has a repository that contains all types of cognitive maps.
- It has a registry that contains metadata describing cognitive topics and cognitive maps; much like the library's card catalog contains information describing the published content on its book shelves.

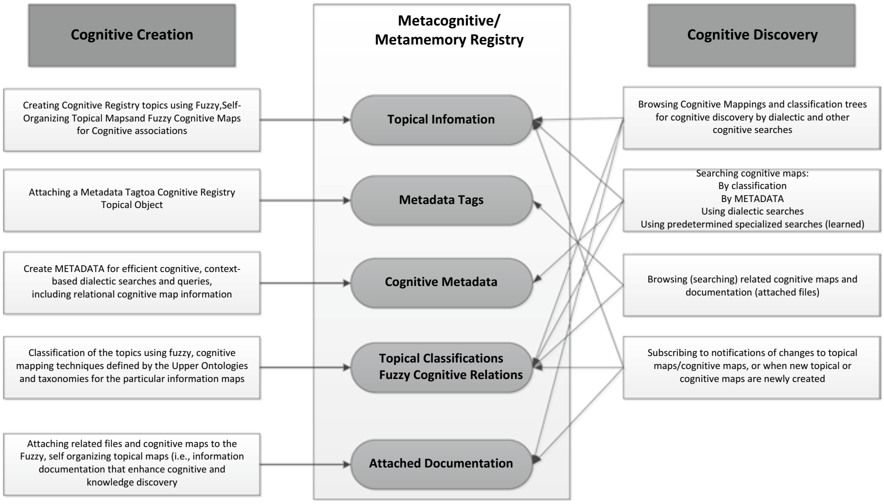


Fig. 6.6 SELF cognitive creation and discovery actions

- The Cognitive registry-repositories work together to offer a unified prefrontal cortex, much like multiple libraries can participate in a cooperative network and offer a unified service.

This APC (discussed earlier) registry-repository provides governance capabilities that enable definition and enforcement of cognitive policies governing the content and usage of the cognitive and topical maps by the Cognitron framework across the SELF enterprise architecture. The cognitive registry-repository contains all Metacognitive and Metamemory constructs for cognitive components available within the SELF, multi-ISA framework.

The registry-repository framework provides discovery capabilities that are extensible and can accommodate the simplest to the most complex domain-specific cognitive discovery queries among the Cognitrons.

- Specifically, its discovery queries do not need to be predefined.
- Instead, it provides an ad hoc query syntax supporting complex predicates that can be combined using possibilistic logic.

As more and more cognitive components are reused by Cognitrons, the task of tracking the network of dependencies between cognitive topics becomes more challenging and significant. This is another challenge that is made easier by the SELF ACNF Metacognitive and Metamemory registry-repository where inter-topical dependency information can easily be managed as relationships between cognitive maps. A cognitive registry-repository provides a set of standard relationship types, but also allows the definition of additional relationship types based on specific requirements within the overall Metacognitive framework.

The nature of these Metacognitive and Metamemory frameworks is that the SELF Cognitrons evolve over time as the SELF learns and evolves. Cognitive components evolve over time for a variety of reasons, as information is learned and the system evolves.

A cognitive component's evolution may involve changes in its topical references and/or interfaces. Changes to a cognitive interface need much more careful management because of the potential impact to existing Cognitrons. These changes will require a new version of the cognitive maps to be deployed, while maintaining the older cognitive maps until the SELF Cognitrons have had time to migrate to the new version according to their own needs and schedule, based on current analytical of mission drivers. New versions of a cognitive or topical map or a cognitive component also typically require publication of corresponding new versions of its cognitive information artifacts, much in the way new memories and thoughts must be categorized, catalogued, and correlated within the human brain.

The Metacognitive and Metamemory registry-repository provides a cognitive change notification capability that allows interested Cognitrons to create subscriptions to events within the registry-repository that may be of interest to them. Such a capability allows a Cognitron to be flexible enough to express precisely the types of events that are of interest to the SELF.

6.4.1 Cognitive Visibility and Governance

The SELF Metamemory system must provide cognitive visibility and governance across the entire ACNF framework allowing the SELF to create and process Metacognitive and Metamemory component models and store them. This enables cognitive governance within the ACNF Multi-Cognitron framework. This allows transparency into the system's cognitive semantics and allows conceptualization across the system of cognitive metadata. These cognitive registry-repository systems allow easy access to cognitive topics and maps across the system, and facilitate knowledge gap analysis. Major functions of the overall Metamemory and Metacognitive systems are:

Cognitive Provisioning: the ACNF framework provides Cognitron interfaces and Metacognitive and Metamemory Metadata. This allows seamless Cognitive Integration across the SELF. The ACNF mediator allows Cognitive Metadata to be utilized in one central location. This enables the reuse of Cognitive and Memory artifacts and provides end-to-end Cognitron support. It provides governance of Metacognitive and Metamemory definitions within the system.

Cognitive Process Integration: The Metacognitive and Metamemory information defined in the cognitive registry-repository is utilized to allow the system to create "what if" cognitive scenarios to facilitate self-assessment. These scenarios allow the system to correlate information and allow cognitive integration in a heterogeneous knowledge environment. This also promotes cognitive knowledge collaboration among the Cognitrons. All of this provides support to allow the system to define the cognitive integration scenarios.

Cognitive Composition: Cognitive Composition tools are provided to allow the system to compose queries of registry-repository information to discover new cognitive and memory components. This facilitates cognitive composite development

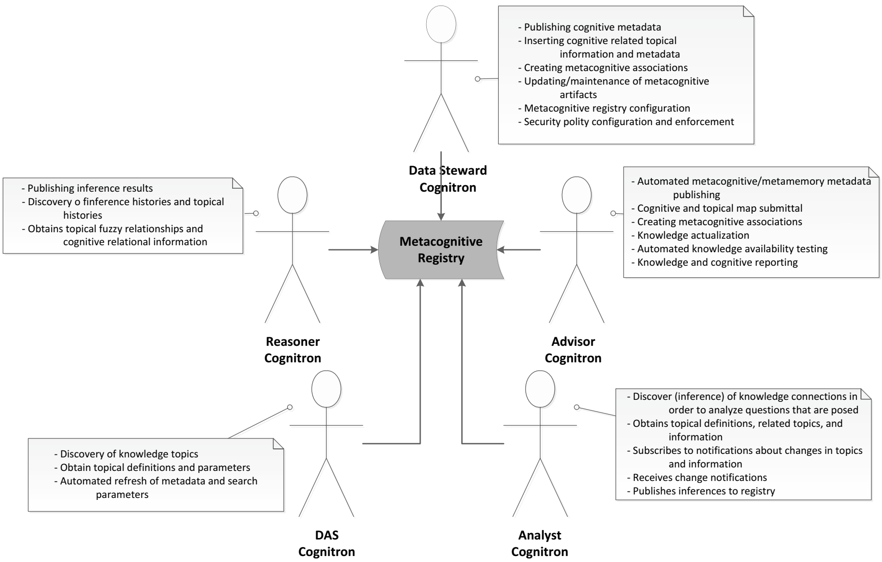


Fig. 6.7 SELF metacognitive and metamemory cognitive registry use case

within the overall SELF cognitive framework. These tools reduce risk in cognitive discovery and allow easy knowledge discovery, consumption, and composition.

Figure 6.7 provides a Metacognitive and Metamemory Cognitive registry-repository Use Case Diagram, illustrating the various types of Cognitrons utilized within the ACNF. This Use Case Diagram describes the functions each type of Cognitron contributes to the overall Metacognitive and Metamemory information flow within a SELF.

6.5 Metacognitive and Metamemory Structures

Metacognitive and Metamemory Self-Awareness and Self-Assessment rely on the system's ability to pose "what-if" cognitive scenarios within its cognitive system. This requires argument structures that allow the SELF to pose hypotheses and determine the validity of these hypotheses. For this ability, we chose a Dialectic Argument Structure (DAS) described earlier.

The DAS provides the artificial consciousness methods to analyze and sort diverse information and cognitive clues that drive the Emotional Memory and Metamemory processes in general [78, 93]. As cognitive knowledge is gathered, a cognitive lattice develops and an aggregate possibility is computed for/against the hypothesis using output from the fuzzy inference engines to provide fuzzy membership values of the support and rebuttal of the hypothesis [81]. This computation is

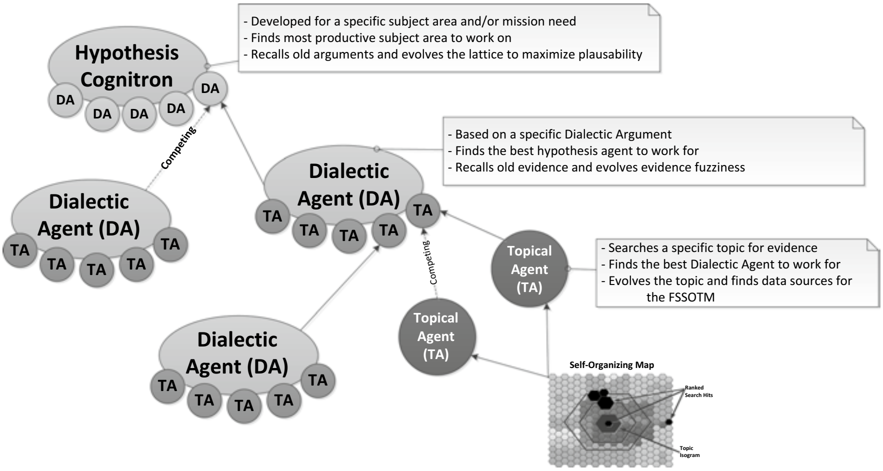


Fig. 6.8 The SELF metacognitive Cognitron hierarchy

based on Renyi's entropy theory, utilizing joint information membership metrics to generate a robust measure of hypothesis "possibility."

Cognitron coalitions form Cognitron Hierarchies for each cognitive hypothesis (see Fig. 6.8). These Cognitron coalitions are part of an ecosystem that constantly has to adapt to the changing information and cognitive knowledge environment. The Cognitrons attributes that drive the DAS are:

- Cognitrons search for information over an irregular topology.
- Cognitrons compete and collaborate to increase their value; successes and failures are remembered.

Cognitrons evolve or die and exhibit different levels of consciousness.

Genetic Algorithms are used to evolve agents, maximizing the effectiveness of the cognitive ecosystem. In order to facilitate intelligent transmittal of learned emotions and emotional context, Emotional Markup Language (EML) is utilized within the system for transmittal of emotional information, creating a Cognitive Social Intelligence, as depicted in Fig. 6.9.

6.6 Extended Metacognition: Artificial Locus of Control Within the SELF

Theories into human learning and cognition have led to much research into new methods and structures for Artificial Intelligence (AI) and, in particular the SELF, to learn and reason like humans. As discussed above, as we move toward a completely autonomous SELF, the ability to provide metacognitive capabilities becomes important [81] in order for the SELF to deal with entirely new situations within the

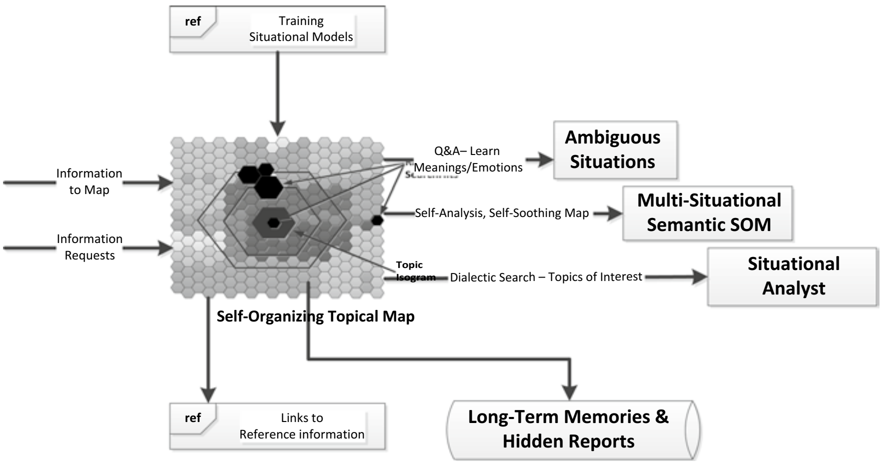


Fig. 6.9 SELF artificial social intelligence

environment it may find itself (e.g., deep space, deep undersea). Here we discuss the theories and methodologies for Constructivist Learning (CL) processes that provide the methodologies to allow a completely autonomous SELF to understand, evaluate, and evolve its "Locus of Control [222]."

We discuss how the use of AI learning systems, like Occam [80] and Probably, Approximately Correct (PAC) (see Chap. 7) learning can be combined with Cognitive Economy concepts to provide this constructivist learning process to allow a Locus of Control evolution within the SELF to provide a fully autonomous, cognitive framework that would be required for autonomous environmental interaction, evolution, and control.

In addition, provided are the mathematical constructs, based in Banach Spaces and Lebesque's work in Bounded Variability, that will provide the basis for Cognitive Economy structures within the SELF, allowing the SELF to operate in a "Bounded Rationality" mode, similar to humans; allowing the SELF to function in new, unforeseen, and challenging environments it may find itself in. Natural intelligence filters out irrelevant information (either raw sensory perception information or higher-level conception information), and categorizes the problem representations to allow for maximum information processing with the least cognitive effort [85]. This work is based on the use of Cognitrons [57] which will represent the internal environment (its tasks, goals, and information) in terms of the reward values associated with different actions when those features of its abilities are active.

The SELF, always having bounded cognitive resources, would react to three aspects of Cognitive Economy to create a Bounded Rationality set of goals for a given set of Cognitrons generated to solve a particular problem or situation. These are:

1. The size of the feature set – how many "features" are required to define the success of each task
2. The "fuzzy" relevance of each feature for the tasks
3. The preservation of necessary distinctions for success in each task

The SELF's cognitive components would autonomously define, for each Cognitron, a Banach Space for that Cognitron's goals and tasks and would then consider the set of Cognitron Banach Spaces as a set of bounded variations, the sequence of which (through Cognitron collaboration) produces an acceptable solution to the situation(s) or task(s) at hand.

The Cognitive Economy methods will be described, illustrating how these Cognitive Economy and Bounded Rationality concepts affect the overall learning aspects of the SELF. In addition, when considering autonomous SELFs, we must consider its need to interact and learn from its environment, and we have to ask ourselves "what is reality?" We have to establish how the SELF would interpret reality. One of the issues that humans deal with that assists in their understanding of reality and how they need to interact, is their concept of "Locus of Control." **Locus of control** is a term in psychology that refers to a person's belief about what causes the events in their life, either in general or in specific areas such as health or academics. Understanding of the concept was developed by Rotter [223], and has since become an important aspect of personality studies.

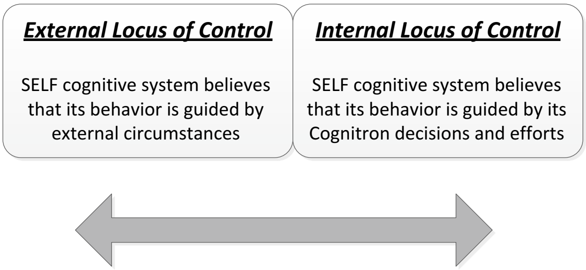
6.6.1 Artificial Locus of Control

Locus of control refers to the extent to which individuals believe that they can control events that affect them. Individuals with a high internal locus of control believe that events result primarily from their own behavior and actions. Those with a high external locus of control believe that powerful others, fate, or chance primarily determine events. Those with a high internal locus of control have better control of their behavior, tend to exhibit more political behaviors, and are more likely to attempt to influence other people than those with a high external locus of control; they are more likely to assume that their efforts will be successful. They are more active in seeking information and knowledge concerning their situation.

Locus of control is an individual's belief system regarding the causes of his or her experiences and the factors to which that person attributes success of failure. It can be assessed with the Rotter Internal-External Locus of Control Scale (see Fig. 6.10). Think about humans, and how each person, experiences an event. Each person will see reality differently and uniquely. There is also the notion of how one interprets not just their local reality, but also the world reality [24]. This world reality may be based on fact or impression.

Take a car accident as an example. There are two people who witness a car being hit by a motorcycle. The police at the scene are supposed to evaluate the facts to determine what has happened. The officer may use measurement tools that are supported by mathematical equations, to determine speed at impact or where impact occurred. The officer may measure skid marks or measure the distance between vehicles. The officer is gathering factual data. Let's consider this juried evidence and legitimate evidence. When asked by the police officer, each human witness can recall the event as if they were watching it again, a step-by-step recount. Each person's story likely has unique qualities depending on how they conceptualize the incident.

Fig. 6.10 SELF adapted Rotter locus of control scale



Eyewitness testimony is part of legal actions all the time. Even though each witness tells a slightly different story, all information and testimony is used in the ultimate decision. We know, by eyewitness testimony studies, that often times the recalled event is very different than the actual event. Let's say in this example both people recalled the event similarly except the color of the car that hit the motorcycle. Perhaps even whether the car hit the motorcycle or the motorcycle hit the car recount differs. The fire truck blocks the view of each eyewitness so they cannot confirm the color of the car as they recount the event. Each person has had a legitimate experience even if they code the color of the car differently. Factually legitimate the car and bike collided at a specific rate of speed at a specific location. Emotionally legitimate is the witnesses' personal experience. To one witness the car was green to the other it was blue. Thus, with this incident we have three realities. Here, one of the facts that we can measure by juried tools and the reality of each of the players in the scene; all experiencing the same event but each in his own unique way. Each reality is legitimate.

For further thought let's then consider Constructivist Psychology. According to "The internet Encyclopedia of Personal Construct Psychology" the Constructivist philosophy is interested more in the people's construction of the world than they are in evaluating the extent to which such constructions are "true" in representing a presumable external reality [81, 85]. It makes sense to look at this in the form of legitimacies. What is true is factually legitimate and what is peoples' construction of the external reality is another form of legitimacy. Later on we can consider the locus of control in relation to internal and external legitimacies or realities. You are correct if you are thinking that the SELF is not human and will not have human perceptions. The SELF will have ACNF Cognitrons which provide its own internal perceptions and realities. Thus, a mentor will be necessary. That mentor will need to understand the SELF as an artificial entity, and be able to understand the SELF in a human way, a human reality

6.6.2 *Constructivist Learning*

Constructive psychology is a meta-theory that integrates different schools of thought. According to the Vaihinger [208], people developed "workable fictions," which was part of his "As If" philosophy, such as mathematical infinity, or God. Alfred Korzybski [150], in his "System of Semantics," focused on the role of the speaker in assigning meaning to events. This drove constructivist learning thought to reason that humans operated on the basis of symbolic or linguistic constructs that help navigate the world without having to contact it in any simple or direct way [85]. Postmodern thinkers assert that constructions are viable to the extent that they help us live our lives meaningfully and find validation in shared understandings of others [155]. We live in a world constituted by multiple realities social realities, no one of which can claim to be "objectively" true across persons, cultures, or historical epochs. Instead, the constructions on the basis of which we live are at best provisional ways of organizing our "selves" and our activities, which could under other circumstances, be constituted quite differently.

For the SELF with Constructivist Learning, the SELF's cognitive learning process would be a building (or construction) process (which was explained earlier) in which the SELF's cognitive system builds an internal illustration of its learned knowledge-base, based on its experiences and personal interpretation (fuzzy inferences and conceptual ontology [191, 204]) of its experiences. The SELF's Knowledge Representation and Knowledge Relativity Threads [71], within the SELF cognitive system memories would be continually open to modification, and the structures and linkages formed within the SELF short-term, long-term, and emotional memories [82], along with its Knowledge Relativity Threads [73], would then form the bases for which knowledge structures would be created and attached to the SELF memories.

One of the results of the Constructivist Learning process with the SELF would be to gradually change its "Locus of Control" for a given situation or topic, from external (the system needing external input to make sense, or infer, about its environment) to internal (the SELF having the cumulative constructive knowledge-based of information, knowledge, context, and inferences to handle a given situation internally); meaning the SELF is able to make relevant and meaningful decisions and inferences about a situation or topic without outside knowledge or involvement. This becomes extremely important for completely autonomous SELF.

6.6.3 Bounded Conceptual Reality (Cognitive Economy)

Bounded rationality is a concept within cognitive science that deals with decision-making in humans [81]. Bounded rationality is the notion that individuals are limited by the information they have available (both internally and externally), the finite amount of time they have in any situation, and the cognitive limitations of their own skills. Given these limitations, decision making becomes an exercise in finding an optimal choice given the information available. Because there is not infinite information, infinite time, nor infinite cognitive skills, humans apply their rationality after simplifying the choices available, i.e., they bound the problem to be solved into the simplest cognitive choices possible [224].

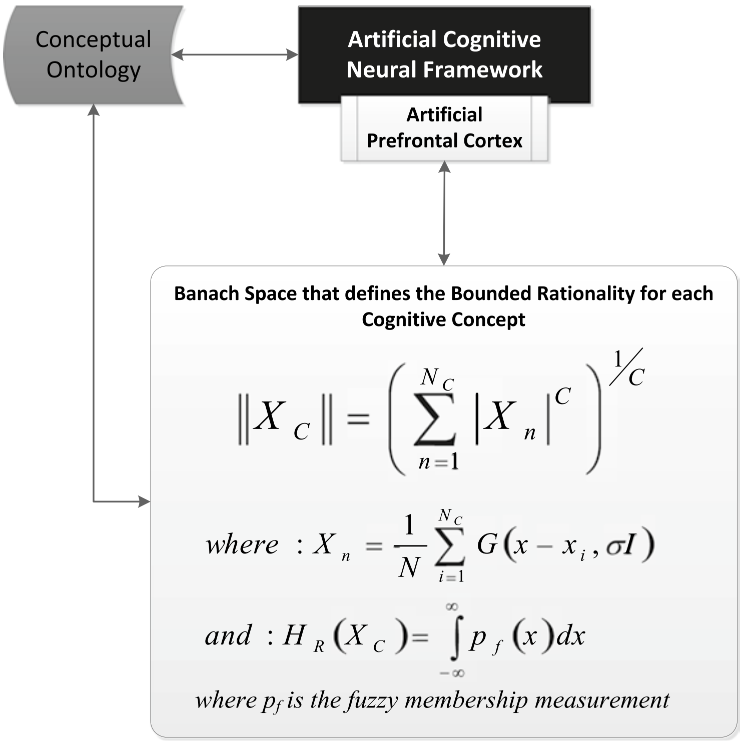


Fig. 6.11 SELF bounded conceptual reality computation

Any SELF must suffer the same issues. An autonomous system, by definition, has limited cognitive skills, limited memory, and limited access to information. The Locus of Control concepts discussed earlier will assist the SELF in determining which situations can be handled internally vs. externally, but still in any situation there is limited information, time, and cognitive abilities. This is particularly true if the system is dealing with multiple situations simultaneously. In order for the system to not become overloaded, we believe autonomous systems must employ strategies similar to human bounded rationality in order to deal with unknown and multiple situations they find themselves in. This involves creating mathematical constructs that can be utilized to mimic the notion of bounded rationality within autonomous SELF.

For this we look to Banach Space theory, tied into Constructivist Learning concepts [24] for an autonomous SELF. As concepts are learned and stored in the SELF conceptual ontology [191], Banach Spaces are defined that are used to bound the rationality choices or domains for that concept. As we "construct" these concepts and the Banach Spaces that bound them, the combination of Banach Spaces then defines the Conceptual Rationality for the Autonomous SELF. Figure 6.11 illustrates this concept.

These Banach Spaces that define the bounds for each learned concept are utilized when the SELF must reason, or perform decision making. When there are restricting limitations on time, resources (as determined by the resource manager, e.g., artificial prefrontal cortex), and available information, the bounds of these Banach Spaces would be tightened or loosened to allow the SELF to deal with multiple situations, or situations that are time critical. This, then, allows the SELF to decide what is a "good enough" solution to a given problem or set of problems, and to adjudicate between competing resources, priorities and overall goals.

The methods discussed here are initial concepts and methodologies for what we believe are essential cognitive skills that autonomous systems must have in order to deal with and survive in real-time extreme environments. As we push for systems that think, learn, and adapt, we must provide these systems with cognitive skills similar to human processes in order to be able to deal with and survive real-time situations they find in their environments. This is very preliminary work and much more remains in order to put these concepts into practice. Future books in this series will provide updates as the research continues.

6.7 Cognitive System Management

Many have put forth architectures that facilitate cognition, learning, memories, and information processing, but these are not sufficient to create a completely autonomous SELF. An overall SELF architecture framework, along with a knowledge and cognitive ontology are required in order to facilitate a fully autonomous, cognitive, self-aware, self-assessing SELF. Such a system must include architectures and methodologies for managing such cognitive processes. Described here is a SELF processing and management framework which within the SELF ACNF that provides for Memory, Decision, Rules, Learning, Reasoning, Decision, and Failure Management within the overall artificial cognitive processing architecture.

In order to understand the world we live in, we synthesize models that enable us to reason about what we perceive. We take in outside information that is available to us, usually from a variety of diverse sources like audio, video, text, etc. This information also comes from a variety of venues like news, entertainment, documentary, etc. Given the diversity of information types and sources, the information we receive does not have a consistent basis, but is riddled with fuzziness (or vagueness) and ambiguity; it is inexplicit and its context is often unclear. In fact, from the moment our brain functions at all, we begin the process of learning and managing the information we have learned; cognitive management. In order to provide the SELF with the abilities to autonomously process information and make decisions, the SELF's Cognitive System Management must include:

- Memory Management
- Learning Management
- Decision Management
- Rules/Goals Management

6.7.1 SELF Memory Management

In Chap. 5 we discussed the artificial memory system required to provide the SELF basic human memory processing, encoding, storage, and reconstruction capabilities (see Fig. 5.7) [68]. We saw that Explicit or Declarative Memory is utilized for storage of "conscious" memories or "conscious thoughts." Explicit memory carries those information fragments that are utilized to create what most people would "think of" when they envision a memory. Explicit memory stores things, i.e., objects, and events, things that are experienced in the person's environment. Information stored in Explicit Memory are normally stored in association with other information that relate in some fashion. The more meaningful the association, the stronger the memory and the easier the memory is to construct/recall when you choose to. In the SELF, Explicit Memory is divided into different regions, depending on the type or source of information. This is because different types of information fragments within the SELF memories are encoded and represented differently, each with its own characteristics that make it easier to construct/recall the memories later when the SELF's cognitive processing system needs the memories. In the SELF LTM, we utilize FUSE-CTXs and FUSE-SEMs [79, 80] to associate currently processed Information Fragments from the STM with memories stored in the LTM.

In order to facilitate self-evolution and memory management within the SELF, each cognitive subsystem, each Cognitron, each and every part of the system must be able to cooperate and collaborate with, and learn from every other part of the system. In essence, the combination of all of the Cognitrons within the SELF form a collective, or group consciousness that drives how the system learns, reasons, and behaves. In order to facilitate this collective group consciousness, there must be both memory management and memory sharing across the entire system. And while the system has a collective set of LTMs, these memories, their implications, their contexts, and their emotions must be broadcast, or transmitted, to each part of the system so that each Cognitron can evaluate how they are affected by learning and 'remembering' that goes on in other Cognitrons. Figure 6.12 below illustrates the overall memory management architecture for the SELF.

The SELF's overall system-level goals are constantly evaluated with the memory management structure, based on the SELF's collective group consciousness (see Fig. 5.12). In the SELF memory manager shown above, one of the functions of the Mediator is to manage this Group Consciousness by correlating current long-term memory Information with the ongoing real-time Cognitive Consciousness comprised of behaviors, cognitive processes, current goal and objectives, emotions, contextual knowledge, etc. The Metacognitive and Metamemory processes shown in Fig. 6.2 correlate all of this information and send information to the reasoning processes and broadcast relevant information to the Cognitrons currently operating in the SELF.

In order for the collective group consciousness to understand and make use of cognitive information from the collection of Cognitrons operating within the SELF, capturing and relating the context of cognitive information is crucial, for it forms that causal connectedness between the each Cognitron is processing and their relevance to other Cognitron, and to past experiences captured in the SELF's LTM.

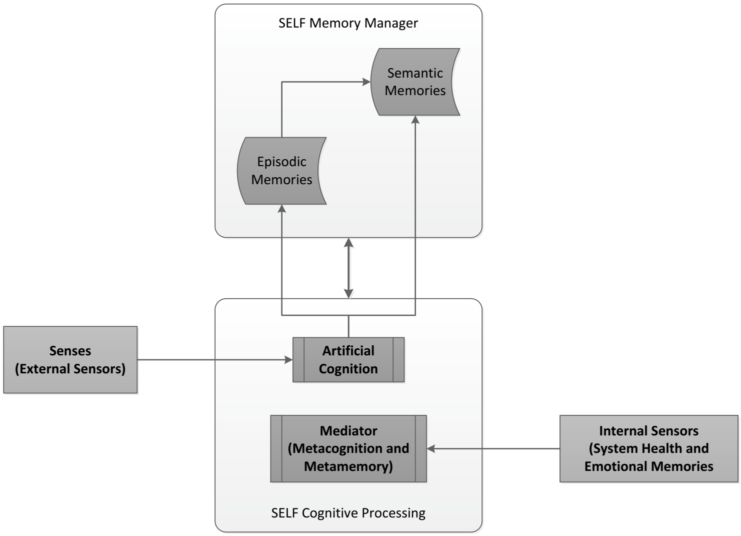


Fig. 6.12 SELF memory management architecture

6.7.2 SELF Learning Management

In order for a SELF to actually learn, we must provide a mathematical framework and foundation for learning. As humans we experience many types of information and have many different types of memories that allow us to experience, filter, learn, and then remember our experiences, both factual and emotional. Therefore we must provide a multitude of learning methods, procedures, and frameworks so that the SELF can function similarly to how humans learn. In attempting to emulate humans, and create artificial cognitive systems that mimic human reasoning, the main goal of the SELF Cognitive Management System is to provide a mathematical foundation that includes:

- Mathematical models that capture the key elements of machine learning and the different aspects of learning that encompass the different ways in which the system must learn.
- Understanding of the algorithms for learning; guarantees, if you will, to assess important aspects of learning, such as:
  - When will they succeed?
  - How do we know they've succeeded?
  - How long will they take?
  - What happens if we don't learn in time, or don't learn at all?
  - What types of algorithms are better for what types of learning?
- Analysis of the inherent ease or difficulty in different types of learning problems.

- Mathematical analysis of the general issues in learning. These might include:
  - When should the algorithms be simple, and when should they be complex?
  - How does the SELF make sure it can properly recall the things it learned?

There are many definitions of learning. For many artificial systems, like classical neural networks, learning may be defined as pattern matching. We train the neural network to recognize patterns, and then utilize the network to tell us when it sees these patterns in data that is fed to the neural network. And while this is a form of learning, it is not real-time, autonomous, unsupervised learning and does not allow the system to learn concepts, to react to unknown situations, which is essential for fully autonomous, self-reliant, unsupervised systems like the SELF. Real learning allows the system to sense the environment, learn from it, and act on it over time, in pursuit of its agenda and goals, based on the evolving constraints within the SELF. And while remembering patterns is a part of learning, it is by no means the majority of learning. Part of learning is being aware of oneself, to be able to assess one's own abilities. This is not different in the SELF. To learn is to evolve, to move forward, and to gain new abilities. This requires a learning system that can track spatial, temporal, and emotional characteristics of the information we bring in. Learning is not just remembering, but inferring as well.

Learning within the Polymorphic, Evolving, Neural Learning and Processing Environment PENLPE learning management system tracks changes in the cognitive framework that enables the SELF to perform new tasks previously unknown or to perform tasks already learned more accurately or more efficiently. Learning is constructing or modifying representations of what the SELF is experiencing. Learning also allows the SELF cognitive framework and memories to fill in skeletal or incomplete information or specifications about a domain (self-assessment). Figure 6.13 illustrates Learning Management within the SELF.

6.7.3 SELF Decision Management

The SELF Decision Management architecture provides the ability to organize information semantically into meaningful fuzzy concepts and information fragments that create cognitive hypotheses as part of its topology. This approach addresses the problems of autonomous information processing by accepting that the system must purposefully communicate concepts fuzzily within its processing system, often inconsistently, in order to adapt to a changing real-world, real-time environment. Additionally, the SELF's processing framework that allows the system to deal with real-time information environments, including heterogeneous types of fuzzy, noisy, and obfuscated data from a variety of sources with the objective of improving actionable decisions using Recombinant kNowledge Assimilation (RNA) processing integrated within the SELF cognitive processing framework to recombine and assimilate knowledge based upon human cognitive processes. The cognitive processes are formulated and embedded in a neural network of genetic algorithms and stochastic decision making with the goal of recombinantly minimizing ambiguity

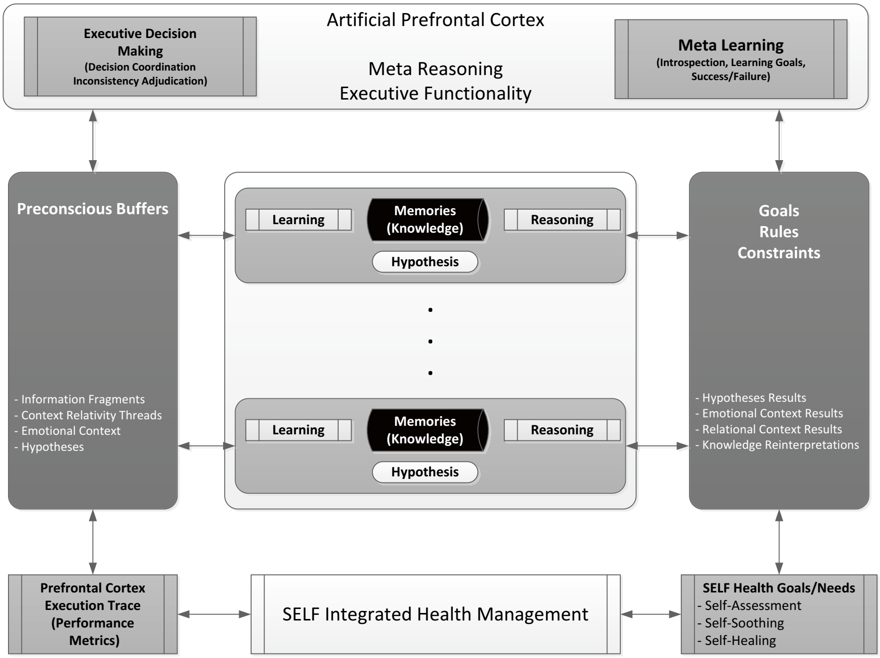


Fig. 6.13 SELF learning management architecture

and maximizing clarity while simultaneously achieving a desired result. The SELF Decision Management involves the Dialectic Argument Structure discussed above. Here we expand on the discussion to include decision management/support. The overall SELF Decision Management architecture is shown in Fig. 6.14. Details of the Context Manager and Decision Manager are provided below.

6.7.4 SELF Rules Management

The PENLPE Rules Manager provides the definitions, structures, rules, constraints, etc. that a minimum Cognitron must have. All Cognitrons within the SELF begin with a Cognitron Base Class (CPBC) of information that it must inherit. This Cognitron CPBC is customized for specific Cognitron purposes within the SELF processing system. This is done to ensure that every Cognitron within the SELF contains those structures and properties required to be recognized and handled by the attention components of the artificial consciousness subsystems. Some of the properties contained within the CPBC include definitions, constraints, and structures for Cognitrons association and coalition formation, how Cognitrons activation level is increased or decreased (Cognitron prioritization), communications protocols among Cognitrons, and how Cognitrons communicate with their short-term memories as well as the system-level long-term memories.

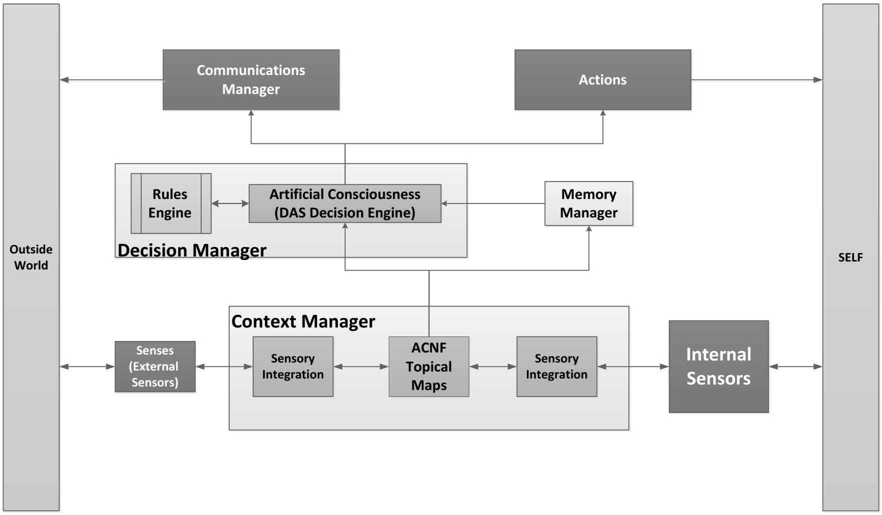


Fig. 6.14 SELF decision management architecture

Figure 6.15 illustrates the high-level SELF Rules Management architecture. The overall purpose of the SELF Rules Management is to:

- Provide an architectural framework for the control and evolution of the Cognitron personalities.
- Provide a framework for domain-independent rules and goals of the Cognitrons (those consciousness aspects that are common among all domains).
- Provide an easily customizable 'plug-in' framework for the domain-specific portions of the Cognitrons.
- Provide the cognitive structures and processes for behavioral and emotional capabilities of each Cognitron.

Figure 6.16 provides one view, or slice, through the Rules Management framework. Figure 6.16 illustrates the Cognitron Rules Class Hierarchy. This class hierarchy shows the classes of rules that are available to the Cognitrons.

Through the SELF Rules Management system, and the Cognitron class hierarchy, specialized rules are created, learned, and evolved throughout the processing life of the SELF. This includes creating specialized roles for possibly artificial emotions or artificial cognition required to provide motivations and goals with the PENLPE cognitive processing environment to facilitate learning within the system.

6.7.5 *SELF Cognitron Management*

Creating Cognitrons which are capable of learning and reasoning about information provides a robust, adaptive information processing system capable of handling new situations. When we use the term reason, we refer primarily to abductive logic,

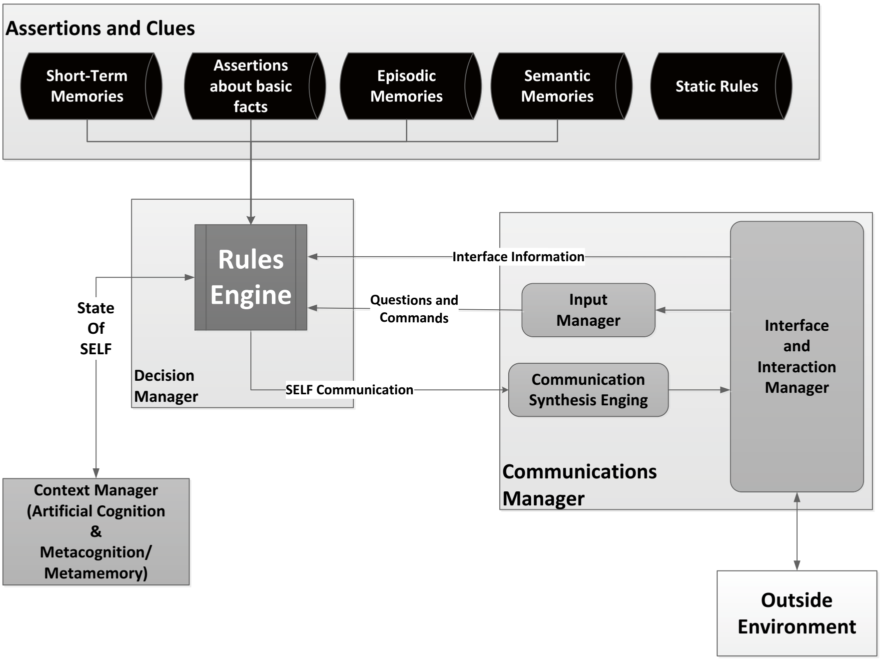


Fig. 6.15 SELF rules management architecture

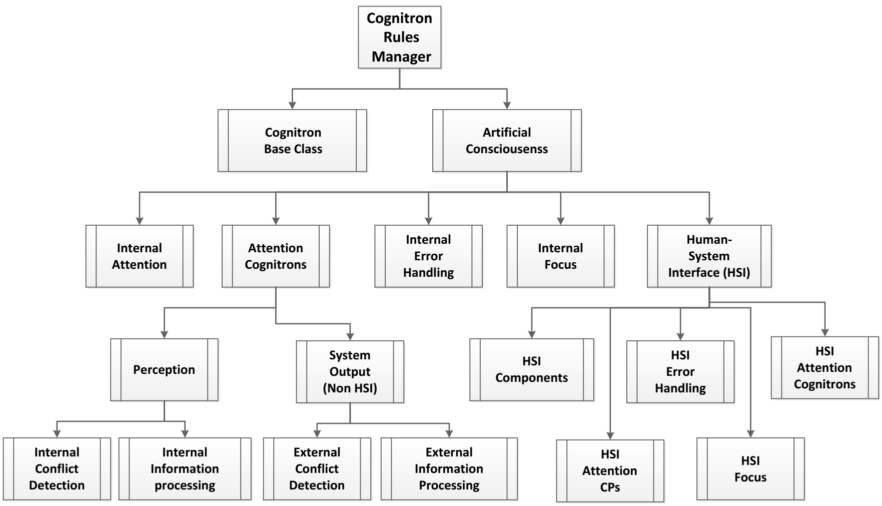


Fig. 6.16 SELF rules management class hierarchy

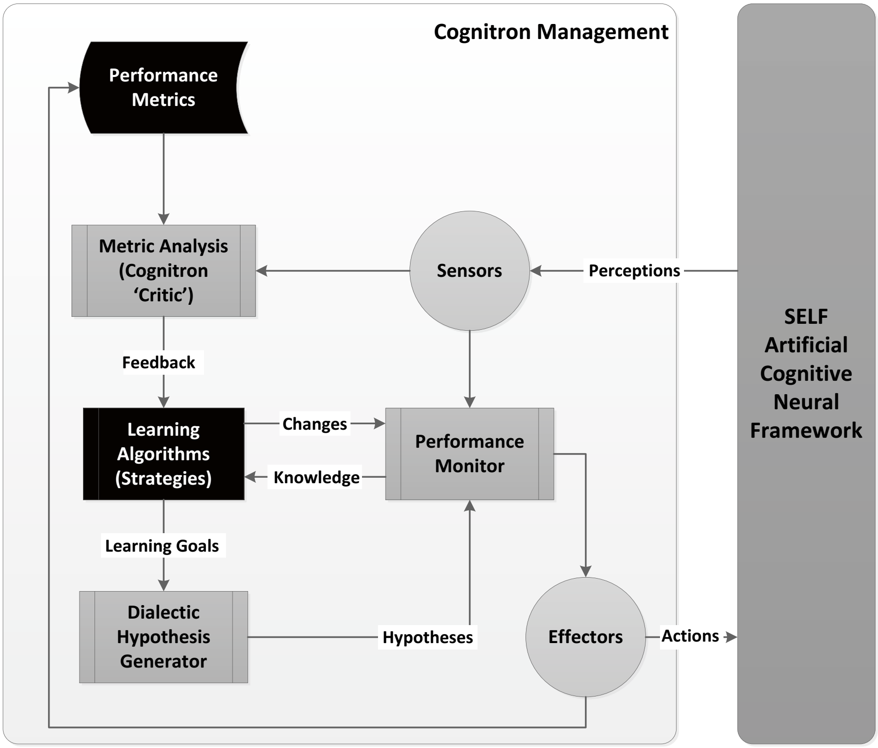


Fig. 6.17 SELF Cognitron management architecture

sometime called critical thinking to discriminate it from the formal logic methods of deduction and induction. For example, data mining uses induction to develop assertions that are probably true. The dialectic search uses abductive logic to develop propositions that are possibly true. As explained in the text, Bayesian methods cannot be used to measure possibility; in its place we use a method that is based upon Renyi's entropy theory, as explained previously [206].

A key value that the Cognitron is that it provides is its ability to learn from operators and from data. Using this learning, the Cognitron has the potential to provide the operators and analysts knowledge extracted from various sources of information. It performs this function 24*7, and can be cloned to support as many operators as required and as system resources allow. Since the Cognitrons learn and evolve over time, it is crucial to provide the SELF with a methodology to monitor and manage the Cognitron evolution. Details of the Cognitrons and their ability to reason will be discussed in Chap. 8. It is imperative to provide performance metrics for the Cognitron learning algorithms as well as the abductive hypothesis generation and testing algorithms that provide the major "reasoning" components of the SELF. Figure 6.17 illustrates the Cognitron management architecture.

6.8 Discussion

We have laid the foundations for the SELF, in terms of a providing an overall cognitive framework (the ACNF), a synthetic, human-like memory system, and Cognitrons to provide the SELF with a sense of artificial consciousness. The next major piece to creating a SELF is the ability to take in and process information, and in doing so, to “learn” from experience. Chapter 7 describes the SELF learning system, which utilizes its Cognitrons and memory systems, coupled with a variety of learning processes, including processes modeled after human constructivist learning techniques.

Chapter 7
Learning in an Artificial Cognitive System

From the moment our brain functions at all, we begin the process of learning. In order for a SELF to be an autonomous system, it must also begin and continue the process of learning throughout its existence. The main goals of the mathematical foundations for SELF learning include:

- To create mathematical models that capture the key elements of machine learning and the different aspects of learning that encompass the different ways in which the system must learn [213].
- To provide an understanding of the algorithms for learning; guarantees, if you will, to assess important aspects of learning, such as:
  - When will they succeed?
  - How do we know they've succeeded?
  - How long will they take?
  - What happens if we don't learn in time, or don't learn at all?
  - What types of algorithms are better for what types of learning?
- To analyze the inherent ease or difficulty in different types of learning problems.
- To mathematically analyze general issues in learning. These might include:
  - When should the algorithms be simple, and when should they be complex?
  - How do we make sure we recall the things we learned properly?

There are many definitions of learning. For many artificial systems, like classical neural networks, learning may be defined as pattern matching. We train the neural network to recognize patterns, and then utilize the network to tell us when it sees these patterns in data that is fed to the neural network. And while this is a form of learning, it is not real-time, autonomous, unsupervised learning and does not allow the system to learn concepts, to react to unknown situations, which is essential for fully autonomous, self-reliant, unsupervised SELFs. Real learning allows the system to sense the environment, learn from it, and act on it over time, in pursuit of its agenda and goals, based on the evolving constraints within the SELF. And while remembering patterns is a part of learning, it is by no means the majority of learning. Part of

J.A. Crowder et al., *Artificial Cognition Architectures*,

DOI 10.1007/978-1-4614-8072-3_7,

learning is being aware of oneself, to be able to assess one's own abilities. This is not different in the SELF. To learn is to evolve, to move forward, and to gain new abilities. This requires a learning system that can track spatial, temporal, and emotional characteristics of the information we bring in. Learning is not just remembering, but inferring as well. Described in this chapter are many different learning mechanisms for the SELF that will provide the capabilities for the system to experience its environment, to grow and evolve as it takes in information, analyzes it, makes inferences, and then learns from it; either extending memories it already has, creating new memories, or reinterpreting memories that it has, based on new information.

7.1 Autonomous Heterogeneious Level Learning Environment

Learning within the SELF connotes changes to the ACNF that enable a SELF's cognitive framework to perform new tasks that were previously unknown, or to perform tasks already learned more accurately or more efficiently. SELF Learning involves constructing or modifying representations based upon experiences to fill in skeletal or incomplete information or specifications about a domain or self-assessment [79]. The SELF cannot be completely preloaded, or trained for every possible stochastic situation that may present itself. Therefore, an autonomous and dynamically updating (learning) environment is required to incorporate new information and new inferences into the SELF's cognitive systems. Analogous to humans, learning new characteristics expands the SELF's domain or expertise and lessens the brittleness of its cognitive framework. Within a complex cognitive ecosystem like the SELF, no single learning system will suffice, as each type of learning algorithm/methodology can be better suited for certain types of learning problems. The SELF incorporates many types of learning systems to accommodate the heterogeneous environments an autonomous system may find itself in:

- **Rote:** Learning: also called "learning-by-memorization" is an associative implicit learning & memory that carries rote information required by the ACNF to function [51].
- **Induction**: this type of learning extrapolates from a given set of examples so that the ACNF can make accurate predictions about future examples.
- **Abduction**: here, genetic algorithms generate populations of hypotheses and a Dialectic Argument (Tolmin) Structure is used to reason about and learn about a given set of information or situations. This is also called "Concept Learning."
- **Clustering**: pattern recognition and related aggregation organization.
- **PAC**: Learning known as, "Probably, Approximately Correct." This learning style assumes information attained comes from an unknown distribution of information about a particular topic. Assumption: learning is supported by information we have already observed either directly, or indirectly related to an observed topic and helps to provide an approximately correct basis for newly encountered data or information.
- **Occam Learning**: this learning system is often cited to justify one hypothesis over another, and more generally means "prefer simpler explanations." The more

data are compressed, i.e., the more complex the learning algorithm, the more likely some important nuance or subtlety is missed or eliminated. Therefore, we define an "Occam Learning Methodology" to be one that produces hypotheses, or "Pattern Discoveries" that are simple in structure, and grow slowly as more data are analyzed.

- **Emotional Learning**: provides "personality" parameters and Conscious Cognitrons with emotional attributes that allows the SELF cognitive framework to have sensitivities to emotional computation and to situational analysis. We compute Emotional Learning Responses (called 'eigenmoods') and Emotional Action Responses from the Cross-Connectivity from the SELF's recombinant neural fiber relationship threaded network, in conjunction with Autonomic Nervous System States. Emotional Learning Responses are computed from the continuous n-dimensional fuzzy weightings, and Emotional Action Responses.

7.2 Autonomous Genetic Learning Environments

Throughout the book, there are many references to genetic learning algorithms. This sub-chapter discusses genetic learning environments in context of an ACNF. Figure 7.1 illustrates the basic genetic learning framework and connections for an ACNF.

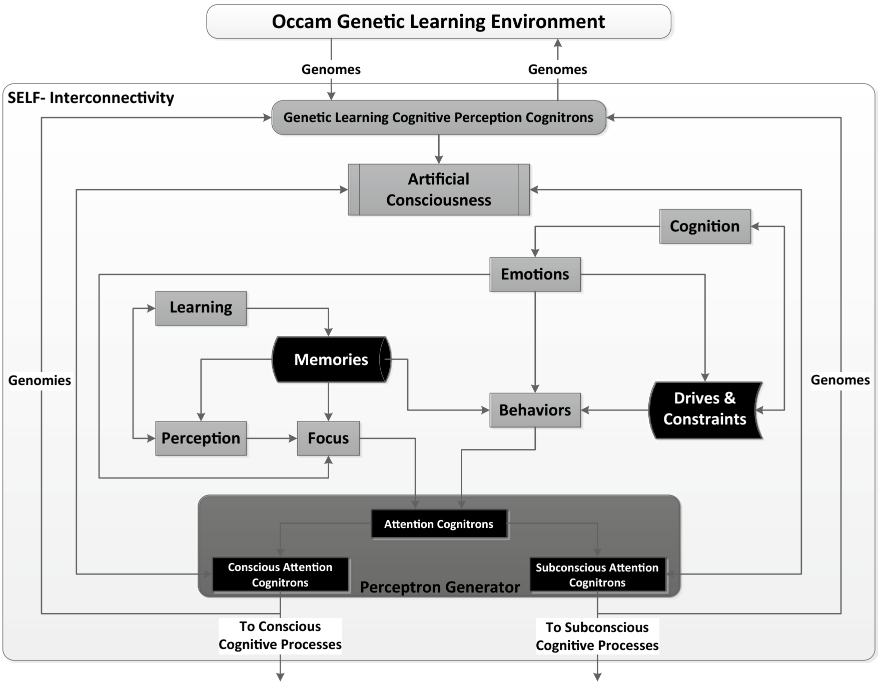


Fig. 7.1 The SELF ACNF genetic learning architecture

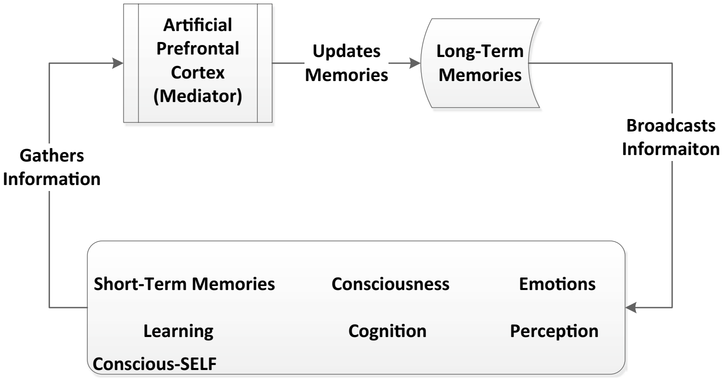


Fig. 7.2 The autonomous genetic learning architecture

Genetically learning Cognitrons inherit initial states from the memory system and inherit the initial parameters for behavior from the behavioral center of the ACNF. The consciousness mechanisms, along with the Artificial Prefrontal Cortex (or Mediator), control the response of the learning Cognitrons, and direct its constraints based on the environment and the problems to be solved currently (see Fig. 7.2). This provides the priorities, preferences, goals, needs, and activation constraints (when you know you've learned something). The genetic Cognitrons (called genomes) adapt to the environment and gather information in order to make conclusions (learn) about the problem to be solved.

In an autonomous genetic environment, genomes are transferred to other Cognitrons, to speed up adaptation of new generations of information content facilitating the development of new conscious Cognitrons within the SELF behavioral environment.

7.3 SELF Emotional Learning

In the ACNF environment, the drives, priorities, and constraints, shown in Fig. 7.1, influence emotions. The behavioral subsystem receives situations and computes actions, while memories provide personality parameters and Cognitron sensitivities to emotional computation. It is assumed that each matrix element E_{aj} represents an emotion. *Emotion* (a,j) of performing action *a,* in situation *j*. Given this, genetically learning Cognitrons perform an emotion learning procedure, which has the following four steps:

1. State *j*: choose an action in situation – (let it be action *a*; let the environment return situation *k*).
2. State *k*: feel the emotion for state *k* – *Emotion*(*k*).

3. State k: learn the emotion for a in j – *Emotion*(*a,j*).
4. Change state: $j = k$; return to 1.

The following learning procedure is a secondary emotion reinforcement learning procedure. The learning constraint used in step 3 is:

$$Emotion^0(a, j) = \text{genome}^0(\text{inherited})$$

$$Emotion^1(a, j) = Emotion^0(a, j) + emotion(k)$$

This learning rule adds the emotion of being in the consequence situation, *k*, to the emotion toward performing action *a,* in situation *j*, where *k* is the consequence.

The above discussion described general emotional learning within autonomous environments, however, a SELF requires learning algorithms that provide the same basic learning capabilities, but in a real-time continuously changing environment, typically difficult to implement for learning systems in general. The next sub-chapter discusses the concept and algorithms for a Decision Analytics in Real-Time (DART) learning system that provides the capability for concept learning in a real-time environment.

7.4 Decision Analytics in Real-Time (DART)

Decision Analytics in Real-Time (DART) is a general approach to continuous learning in a real-time, continuously changing environment. Later sections of this chapter will provide particular learning methodologies (e.g., Occam Learning) which are used for specific instances. The purpose of the DART is to provide the cognitive processing framework with Cognitron characteristics which determine learning modes to continuously test new strategies against dynamically created simulation models (created by the Artificial Prefrontal Cortex) and to dynamically update memories used by a Cognitron, based upon the results of the dynamic simulations [217]. The DART operates indefinitely (as long as the SELF is 'alive') and the execution system utilizes the results of the genetic learning process whenever they are available and broadcast to the Cognitrons.

7.4.1 Case-Based DART

First, we will discuss a case-based method of initializing genetic algorithms used for DART learning systems. A genetic algorithm with a case-based component provides a directed search mechanism for DART learning systems. When a genetic algorithm is started, strategies learned previously under similar environmental conditions are included in the initial population for the genetic algorithm. Additionally, a DART learning system evaluates the learning process by comparing and contrasting performance with and without a case-based component.

Anytime learning systems are essential for a SELF because:

1. The environment the SELF is exposed to and trying to learn about/from continuously changes.
2. Constructs to detect changing information are necessary, and are, in fact, random variables, hence, DART needs are absolute.
3. Measurement errors must be accounted for in the fuzzy learning process.

DART learning systems include machine learning techniques utilized to address sequential decision problems. Hence, we incorporate concepts of Occam Learning into DART processes, as reactive strategies expressed as condition-action rules; utilizing modified genetic algorithms applied to sets of symbolic reactive rules generating increasingly, over time, competent Occam Strategies. This is utilized to focus on "detectable changes" within internal and external environments. DARTs monitor the external environment and, when a change is detected, updates the learning strategies with the new information. The changes are organized/classified and stored, allowing new case-based methods to be developed, as new cases are discovered, and reused when they are needed again.

The basic idea for DARTs is to integrate continuously running, but dynamically changing execution and learning components. The Cognitron's learning component continuously tests new genetically generated strategies (variations upon strategies that have worked in the past), and updates the Cognitron's knowledge base with the best available results. In this way, each Cognitron has currently evolved learning strategies for the external data environment(s). The execution component of a DART learning system controls the Cognitron's interaction with its environment (both internal and external), and includes a monitor that dynamically modifies the learning simulation models based on its environmental observations. When the simulation model is modified, the genetic algorithms are re-started on the modified simulation model. A DART learning system is assumed to operate indefinitely, and the execution system uses the results of learning components as they become available.

Genetic algorithms are well-suited for initiating and reinitiating learning components in a continuously changing environment. Additionally, the genetic population approach was enhanced by including strategies, under the control of the learning component, when genetic population is performed. Previously organized and classified cases (developed via Dialectic Argument Structure discussed elsewhere) are stored, and the FUSE-SEMs are used to find the most similar cases to be used at any point in time. Figure 7.3 illustrates the high-level DART learning system architecture. The execution component of a DART learning system includes a decision maker which controls the Cognitron's interaction with its environment, based on its active knowledge base (current strategy). The learning component of a DART

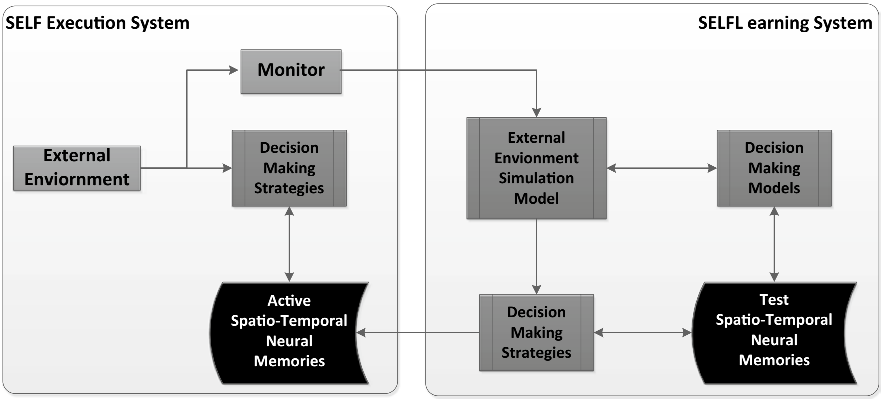


Fig. 7.3 SELF DART high-level architecture

learning system attempts to provide the execution component with improved strategies by experimenting with alternative hypotheses (new strategies) on a simulated model of the environment [51].

7.5 Cognitronic Learning

This sub-chapter illustrates mechanisms for several different types of learning which can be implemented for Cognitrons, and speculates briefly on their implications/benefits for human learning and development. In particular, we're concerned with the role of "attention", i.e., bringing specific content into view of consciousness, while learning. We propose computational mechanisms for such learning and we believe that when Cognitrons are employed within a DART learning system, that this creates the autonomous Cognitrons required to mimic human information processing and decision making. These autonomic Cognitrons running within a DART genetic learning process are described in the following subsection.

7.5.1 Cognitron Autonomy

Artificial Intelligence in general and research into Artificial Cognitrons in particular, pursues the combined goals of understanding human intelligence and translating that into Artificial Cognitive software. Designing, implementing and experimenting with autonomous Cognitrons furthers both these goals in a synergistic way. The Cognitron senses environments, and acts, over time, in pursuit of an agenda (based on a priori directives/goals, artificial instinct, and constraints). In biological agents, this type of an agenda arises from an evolving drive to satisfy associated goals [9, 10];

in a Cognitron based SELF, the driving forces and goals are built in by a creator (the human-in-the-loop and, over time, by the ACNF). Such driving forces, which act as motive generators, must be present, whether explicitly represented, or expressed causally [51]. Cognitrons are developed, used in real-time, and stored away for later reuse to help the system infer from past experiences at a later time. In other words, it is structurally coupled to its environment. Biological examples of autonomous agents include humans and most animals. Analogously, we are concerned with artificial autonomy implemented as autonomous Cognitrons, designed for similar humanistic type tasks, for 'living' in real-world environments.

7.5.2 Cognitronic Cognition

Many researchers have postulated [87] that human cognition is implemented by a multitude of relatively small, special purpose processes, almost always unconscious. Communication between them is rare and over a narrow bandwidth. Coalitions of such processes find their way into consciousness. This limited capacity workspace of our cognition serves to broadcast the message of the coalition to all the unconscious processors, in order to recruit other processors to join in handling the current novel situation, or in solving the current problem. Thus consciousness in this theory allows us to deal with novelty or problematic situations that can't be dealt with efficiently, or at all, by habituated unconscious processes. In particular, it provides access to appropriately useful resources, thereby solving the relevance problem [177, 178].

All this takes place under the auspices of contexts: goal contexts, perceptual contexts, conceptual contexts, and/or cultural contexts. These may look like goal hierarchies, dominant goal contexts, a dominant goal hierarchy, dominant context hierarchies, and lower level context hierarchies. Each context is, itself a coalition of processes. Though contexts are typically unconscious, they strongly influence conscious processes [1, 2].

Crowder and Friess [77] postulated that, in humans, the act of learning results simply from conscious attention. However, the act of conscious attention includes action selection, emotion, voluntary action, meta-cognition and an actual sense of self. In short, it involves a complete high-level theory of cognition [30]. The implication of this is that in order for the SELF to attain real, human-like learning, we must provide the SELF with the same cognitive constructs and the need for mechanisms to perform "Conscious" perceptions.

7.5.3 Conscious Cognitrons

We define a "conscious" Cognitron to be a Cognitron which implements global workspace theory [48, 49]. We believe that conscious Cognitrons play a synergistic role in both cognitive theory and artificial cognitive software. Minds can be viewed as control structures for Cognitrons [1, 2]. A theory of mind constrains the design of the "conscious" SELF Cognitron that implements that theory. While the theory of the mind is typically abstract and only broadly sketches the requirements for an artificial cognitive architecture, the SELF computational design provides a fully articulated architecture and a complete set of execution mechanisms. We feel the SELF architecture and its set of mechanisms provides a richer, more concrete and more decisive theory for artificial cognitive systems. Moreover, we believe the design decisions taken during an implementation may furnish hypotheses about how human minds work. These hypotheses may motivate experiments with humans and human/computer interaction experiments. Conversely, the results of such experiments may motivate corresponding modifications of the SELF architecture and mechanisms of the Cognitrons. In this way, the concepts and methodologies of cognitive science and of computer science may work synergistically to enhance our understanding of mechanisms of both artificial cognitive architecture and the human mind [44, 45].

7.5.4 Autonomous Learning Mechanisms

In prior chapters, we have seen descriptions of several, quite distinct types of learning available for use by "conscious" Cognitrons. In this chapter we explore some others. Firstly, we explore Declarative learning, which occurs in episodic memory and is implemented as case-based memory [50]. Precepts from a given focus stored in a SELF's main associative memory will constitute declarative learning. Not only is the content of a precept learned, but relationships with other items of memory and some generalizations are as well. Each new precept represents a new stored case study and constitutes declarative learning. Declarative learning also results from the learning of new overall concepts. We suspect that human declarative learning also occurs in each of these forms, each with somewhat different mechanisms.

Secondly, autonomous Procedural learning also occurs in several forms. Gradual procedural learning takes place as associations are strengthened, over time, between codelets that are "conscious," or even active, together. The gradual reorganization of codelets is another form of procedural learning, as is the gradual learning of new artificial behaviors through conceptual learning. Humans are capable of each of these modes of procedural learning and they each require a different mechanism. Human learning of language for example, has both a declarative and a procedural component.

A basic tenet of global workspace theory says that consciousness is sufficient for learning [9, 10] and is the reason that a SELF is modeled with Cognitrons. The contents of "consciousness" will be routinely written to associative procedural memory. The learning of new associations between codelets and adjustments to such associations will occur when their contents become "conscious" or become prominent in the forefront of a SELF's thought processes. A SELF, like humans, process information via multi-tasking in the background, while interacting with a given function. But it will also occur to a lesser degree when the codelets are active together; this is known as subconscious autonomy; even though a newly formed concept may well have been learned during "conscious" directed activity. This seems to suggest that some procedural learning, some gradual improvement of skills may occur subconsciously with practice.

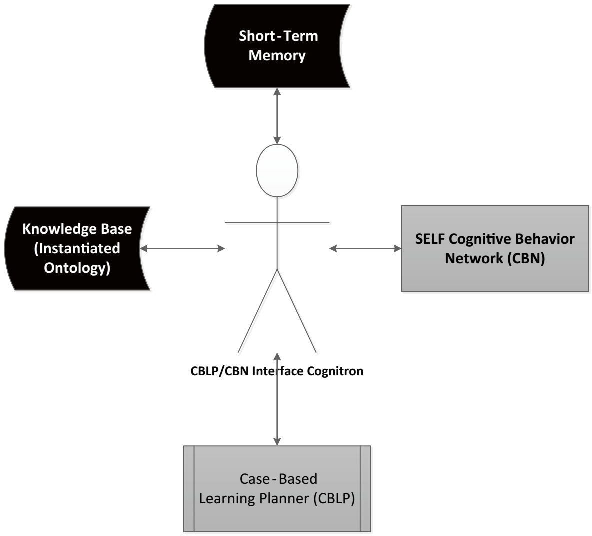
We propose that within the SELF subconscious learning of importance weights within the ACNF emotion networks occurs subconsciously as well. We then argue that in complex, dynamic domains, such as autonomy, the learning mechanisms described here allow Cognitrons to adapt and effectively live in those domains. We further propose that variations in development periods occur for different types and complexities during knowledge acquisition. Particularly, in complex, dynamic domains where active knowledge engineering is generally expensive, subconscious, passive, inexpensive autonomous development periods can provide a simple, but cost effective solution to knowledge acquisition. We speculate positively on the implications of these mechanisms for the evolving, complex SELF, and for human learning and development [116, 117].

7.5.5 Autonomous Behavior Learning

As discussed above, any agent (artificial or biological) senses, perceives and acts in order to satisfy built-in driving forces and goals The ever-present challenge for any agent is to produce the appropriate action relevant to internal states modulated by a perceived environmental situation. That is, the action selection mechanism of an agent decides what to do next. New concepts get introduced via conceptual learning mechanisms. New concepts require new behaviors, thus requiring an action selection module with a capability to learn. As shown in Fig. 7.4, the SELF behavioral learning system, to realize its adaptive action selection capability, uses four major components:

1. The *Cognitive Behavior Network (CBN) system* can be viewed as a collection of behavior streams (action plans). Each such stream is a connected partially ordered set of behaviors (plan operators) that serve to satisfy a goal or sub-goal of the Cognitron. A behavior stream is a partially ordered plan which guides execution of behaviors (plan operators) so as to effect the required transition from the initial state (mainly dependent on the internal representation of the perception) to the goal state. The CBN system has additional functions including interface with consciousness and priming.
2. The *Case-Based Learning Planner (CBLP)* is a case-based DART learning system described above [46, 48]. In general, a Case-Based Reasoning (CBR) system is a paradigm that solves new problems by adapting prior solutions to old problems and, to do so, it supports retrieval, adaptation, and retention processes. In our system the CBLP must have a flexible plan learning/adaptation mechanism.

Fig. 7.4 SELF DART behavioral learning model



The CBLP's processes operate on a unit of information called a *case*. In our adaptive action selection mechanism, a case is represented as a triplet consisting of <problem description, solution, outcome>. A *problem description* includes the initial stale of the problem situation (the contents of the focus, relevant coalitions of codelets, and feature values of relevant concepts, relevant registers in working memory, etc.), one or more (sub)goals that need to be satisfied in such a problem situation, and associated behavior streams (action plans) that achieve those goals. A *solution* is an action plan (behavior stream) whose execution beginning at the initial state of the problem achieves its stated (sub)goal(s); each of which in turn satisfying one or more of the innate drives which are that represent the primary motivation of the Cognitron. An *outcome* is the expected result (for example, feedback from a human) when the solution plan is applied in the initial state.

3. The *Knowledge-Base (KB)* is used to store information needed in the behavioral learning process. That is, it contains all domain related knowledge, built-in and/or leant, which is specific to the Cognitron's action selection mechanism. The *CBLP/CBN-Interface* module uses the KB module to couple the RN and the CBLP modules, and to facilitate the knowledge acquisition process. It is used to (a) store newly acquired domain knowledge into the KB, (b) compile the problem description (from the CBN side) in the format the CBLP can use, (c) format a newly obtained plan (from the CBLP) so that it can be integrated into the RN system and (d) facilitate effective conversation with human (via the CBN) by providing information available in the Conceptual Ontology and/or the KB. The CBLP/CBN-interface uses its own working (short-term) memory (WM).

Initially, Cognitrons are provided a CBN including a set of behavior streams (action plans) capable of producing actions appropriate to already known current concepts and situations in the domain. This allows Cognitrons to be able to adapt and behave differently to new situations in their environment. Behavioral learning is based upon two principles:

(a) Cognitrons will use past experiences to learn new behavior streams by adapting old plans that worked in similar situations.
(b) Cognitrons must be able to converse with humans and interact with their environment to acquire new domain knowledge. This also allows for feedback on the accuracy of plans and for necessary revisions.

When a newly learned concept is perceived, the "consciousness" mechanism broadcasts the relevant information to recruit codelets, which will collectively pick the appropriate behavior stream(s) that will produce an appropriate response. Since a new concept is involved the selected stream may fail to produce an appropriate action. This failure initiates the behavioral learning cycle [123, 126]. As an example, the learning happens by processing information content from a conversation with a human supervisor. At each interchange, the learning mechanism adapts streams from old solutions stored in the CBLP system. A single interchange may not suffice to produce an appropriate new stream (action plan). But, episodic memory (implemented using case-based memory) stores the sequence of interchanges and the trace of the reasoning used in building a new behavior stream. This, along with the already acquired domain and control knowledge stored in the KB and CBLP modules, will help in the effective use of past experience to speed up the learning process. A successfully learned stream in the CBLP module gets integrated into the CBN system where it can be instantiated and executed.

In addition to learning new streams, the behavioral leaning process is proposed to include the dynamic creation of new coalitions of behavioral codelets which choose and instantiate a new stream whenever it becomes relevant. The behavioral learning process is also proposed to initiate codelets which implement actions for each individual newly discovered behavior. The dynamic nature of the behavior learning process is accomplished in part by a code management framework which has the ability to copy, modify, and add existing codelets to support a developing stream of consciousness.

7.5.6 Behavior Learning and Human Interaction

Cognitronic teaming is essential for developing the intelligence and adaptability of an autonomous system to changes within an environmental domain. When Cognitrons interact with the environment, including humans, an additional layer of dynamism seems to occur within the Cognitron's domain. Interactions can be slow and innocuous or rapid and complex. In either case Cognitron learning mechanisms are essential for developing autonomy. In contrast, humans have a number of inherent organic sensing and survival based learning mechanisms. Hence, Cognitrons similarly require learning via several types of artificial learning mechanisms

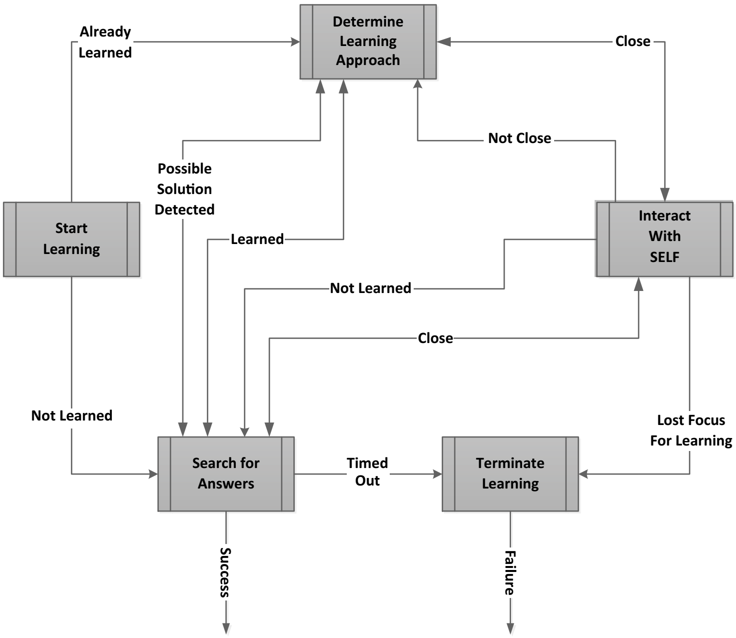


Fig. 7.5 DART genetic learning finite state machine

enabling environmental adaptation as described in Fig. 7.5 which illustrates the DART finite state machine for genetic learning and Fig. 7.6 describing the learned behavior selection process.

The DART Finite State Machine is utilized to manage the genetic learning processes within the DART. The DART Finite State Machine accepts input from the Behavioral Learning Model and accesses the genetic hypotheses generation and testing processes to help manage the DART learning process. Each state drives different actions from DART. Termination of the learning process occurs either when the system has determined it has hypotheses that adequately explain the observations/data/information or when the system cannot make a determination, either because it has taken too long (timed out) or because there is enough rebuttal evidence to the hypotheses that they are not worth pursuing (lost focus).

The process of abduction (hypothesis-based learning) makes use genetically generated populations of hypotheses (described in Chap. 9) to create potential explanations for the observations/data/information being processed. Finite state conditions and resulting actions are as follows:

- **Start**: this state determines whether the input from the DART Learning Behavioral Model is adequate, based on goals, needs, mission constraints, etc. If the input is adequate, learning has occurred and the process terminates and sends it on to memory processes. If the input it determines the learning models from the DART Behavioral Model is not adequate, further reasoning, analysis, and

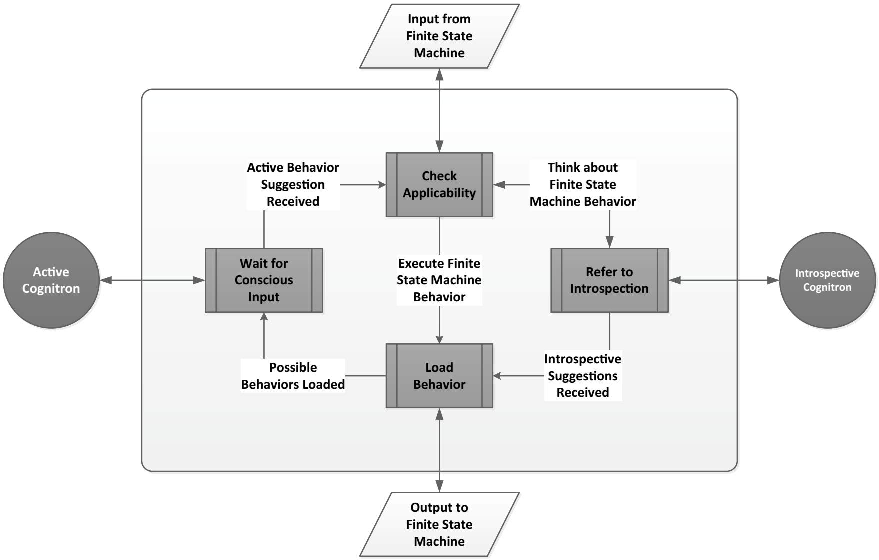


Fig. 7.6 DART behavior selection process

learning processing is required. Depending on the "state of learning" an approach is determined and the information is sent to the "Search" process to look for relevant information and possibly create more abductive hypotheses to increase the level of explanation (increase the knowledge density for that topic).

- **Approach**: in this state the learning approach is evaluated and the information is passed on to the Search state. There are many learning approaches that are possible, as explained throughout Chap. 7. Input from the other states helps determine the approach. Detected tells the approach state that information and/or hypotheses that may be useful are available for evaluation. An input of "not close" tells the approach state that the learning is still valid, much work is need for an adequate explanation; DART hasn't learned much yet.
- **Interact**: this state determines the level of interaction DART requires from the rest of the SELF, including the level of resources that will be required to continue the learning process. This required interaction with the Artificial Prefrontal Cortex to request resources and possibly with interface Cognitrons to request outside information (depending on the SELFs Locus of Control determination).
- **Search**: this is the main "learning" state for DART, but Abductive, Hypothesis-Driven, Occam Learning system, described next in Sect. 7.6. Once DART determines that something has been adequately learned, the Search state deems the learning a "success" and sends notification to the rest of the SELF. If a success criteria is not reached in the given time frame (which is determined on a case-by-case basis), a "time out" signal is sent to the Terminate state.

- **Terminate**: here DART send a failure message to the ISAAC cognitive processing system if either the learning times out, or it is determined that the learning shouldn't continue due to lack of evidence, or too much rebuttal evidence (the system has lost focus on this subject).

Once the DART behavioral learning system has "learned" something, the SELF must determine the proper behavior associated with the new learned information. The DART Behavior Selection Process shown in Fig. 7.1 is utilized to determine the appropriate behavior associated with a given memory, or input information. A failure or success input from the Finite State Machine drives the system the different behaviors, depending on the goals, directives, mission needs, and constraints on the system at that time. Introspection Cognitrons determine the internal needs of the system and provide perspective on how proposed behaviors (actions) will affect the SELF's internal systems. Action Cognitrons determine the affects the proposed behaviors will have on the external SELF environment. Once a behavior is selected, the Action Cognitrons send the required information to the SELF's effector system to initiate the chosen behavior.

7.6 DART Occam Learning

The push for real-time autonomous artificially intelligent systems over last number of decades has driven companies and government research facilities to spend considerable R&D budgets looking for systems that can operate with little or no supervision while processing incredible amounts of heterogeneous information. Building upon the last chapter, we will focus upon proposed mechanisms for artificially, "learning with experience," in autonomous AI systems [57]. The goal of having machines that learn with experience is one of the most intriguing problems in computer science and computer engineering. Unfortunately, by its nature, learning is somewhat fuzzy, and random in nature, as information comes at us rapidly and in stochastic fashion [164] learning things a SELF does not yet know, and possibly didn't know it needed to learn, and in doing so finding patterns within the interactive environment that it can learn. This constitutes not pattern matching, or pattern recognition, but is, in fact, pattern discovery. Thus, we propose the need for a mathematical framework for SELF pattern discovery within an autonomous SELF.

7.6.1 DART Pattern Discovery

The notion of pattern recognition is readily known and understood [87, 88]. However, here we introduce a different concept in machine learning, the concept of *Pattern Discovery* to assist in filtering through vast, stochastic, and unpredicted "fire hoses" of information content to achieve high value droplets of actionable content. We propose employing computational physics concepts for finding causal structure within stochastic data.

Pattern Discovery is intended to contrast the well understood concepts of *Pattern Recognition* and *Pattern Learning*. In Pattern Recognition the aim is to analyze ingested content and assign it to one or more pre-determined categories or patterns. In most Pattern Learning systems, the goal is to determine which of the several pre-determined pattern categories corresponds to the available pattern algorithms [187]. Although a valuable commodity for simple computational systems, a priori defined patterns facilitated by pattern recognition and pattern learning paradigms many times are difficult to apply to complex environments. Patterns are essentially outlines or templates developed and usable for specific contexts. Similar to the development of domain ontology's, patterns are only 100 % correct at the exact time they were developed and for the exact context they were developed in. Hence, our proposed approach is to discover patterns in real-time within a most current context, remember that detailed context and reuse the behavioral patterns learned if and when the detailed context reaches a threshold of similarity sometime in the future.

For our use of Pattern Discovery, the goal is to avoid the necessity that a priori knowledge about what structures, or patterns, may be relevant [196, 197]. This is not a new problem, and the classical approach, based on statistical mechanics, is to derive patterns (or macroscopic properties) from raw data (or microscopic components). Here we take the inverse approach, extending the concept of extracting "geometry" or causal structures, from a time or frequency series of data or information. We build upon the concept of "Occam Learning" [70] to construct the simplest model capable of capturing context specific causal structures, or "patterns" in the data which constitutes a representation of the causal structure of the hidden process(es) which generated the behavior observed or captured. We assert that this representation is the maximally efficient model of the observed data-generating process, based on the learning principles laid out in the Occam Learning Process. The underlying computational mechanics concepts have been used to analyze dynamical systems, evolving spatial computation, and stochastic resonance, among others. However, the combination of computational mechanics coupled with the Occam AI learning constructs provides a unique method for Pattern Discovery within large, heterogeneous data sets and moves us closer toward providing a real-time, autonomous SELF.

Here we discuss a mathematical framework for discovering, describing, and quantifying new patterns, based in computational mechanics and using tools from statistical physics. We construct optimal, minimal models of stochastic processes and their underlying causal structures that drive an "Occam Learning" model of intrinsic computational information transformation. Additionally, we summarize the mathematical foundation of computational mechanics, especially those constructs in optimality and uniqueness to drive the Occam learning algorithms. We also describe the principles and motivations underlying the computational mechanics, emphasizing connections to the minimum description length principles underlying Occam Learning, and its implications to Probably, Approximately Correct (PAC) machine learning concepts [87, 88].

We will examine the concepts and issues involving Pattern Discovery and how they are addressed by employing computational mechanics. Then we provide discussion of the computational mechanics based mathematical structures that provide the foundation for Pattern Discovery, with particular attention to optimality and uniqueness theorems. Uniqueness theorems are then utilized within the Occam Learning framework to provide a DART learning system the ability to learn, and then to extend these unique pattern structures. Differences between this work and other work in computational mechanics is in our utilization of Renyi's Entropy theory versus Shannon's; utilizing Renyi's mutual information theory in our computations involving stochastic processes [69].

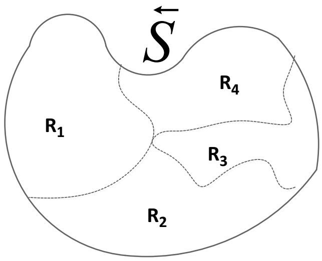
7.6.2 DART Pattern Discovery Concepts

Approaches to processing Pattern Discovery should meet a number of criteria:

- ***Predictive*** – the models the algorithms produce should allow the system to predict the original process or system that produced the data, and provide a compressed description of it (learned pattern).
- ***Computational*** – should have storage within system memories on how the process or system stores, transmits, and transforms information (what causal structure produced the information?).
- ***Calculable*** – either analytically or by systematic approximation.
- ***Causal*** – system should understand how instances of the discovered patterns are produced.
- ***Naturally Stochastic*** – the learned patterns models should not just be tolerant of noise, but should be explicitly formulated in terms of stochastic ensembles.

For our uses, the key idea of utilizing computational mechanics is the supposition that the information required to drive the Occam Learning Pattern Discovery is actually within the data, or information picked up by the system's sensors. A significant challenge is dealing with real-time information coming from a number of possibly heterogeneous sensors and determining which data sets (or partitions) of data sets should be treated as equivalent, and how the data should be partitioned. We assert that correct mapping of data to partitions should leave the system with discovered patterns that provide the system's cognitive framework with the same degree of knowledge about the future of the data (similar prediction accuracy and consistency), has a proper degree of plausibility, but is also vague enough to account for future growth of the learned pattern definition [108, 153]. So the question is: how to create the partitions? For this we look to genetic algorithms. We create generations of partitions and then use these to create Occam patterns, or memories, from the populations, based on the partition constraints (based on computational mechanics) [96]. These are evaluated, based upon a combined set of Entropy and

Fig. 7.7 $\mathfrak{R}$ partitions of the set $\overleftarrow{S}$



quantum mechanics based relationship calculations. Partitions from the population that produce the best fit and utilized (with mutation and crossover) are used to create a new generation of partition population. The process continues until an optimal partition is created. Those partitions that produce patterns similar to those already in the system's memories are sent to algorithms that evaluate the patterns for memory extensions or reinterpretations. Those that are not already part of the system's memories are used to create new memories.

For this discussion, we refer to ***H[X]*** as the entropy of discrete random variable ***X***, interpreted as the uncertainty in ***X***. ***H[X|Y]*** is the entropy of ***X*** conditional on ***Y***, and ***I[X|Y]*** is the mutual information between ***X*** and ***Y***, as measured by Renyi's Entropy and Mutual Information computations. Also, we restrict ourselves to discrete-valued, discrete-time stochastic processes (analogous to sensor data being collected by an autonomous system). Such processes are sequences of random variables, $\boldsymbol{S_i}$, and the values are taken from a countable set ***A***. This is reasonable since we are talking about a system with multiple sensors, each taking in data over a specified period of time, each with a countable number of data samples [87, 88].

Our goal is to discover a pattern, or Occam memory that will predict all or part of the future of process $\vec{S}$, using some function and some part of $\overleftarrow{S}$. We begin by taking the set $\overleftarrow{S}$ of past data points and partitioning it into mutually exclusive and jointly comprehensive subsets, as shown in Fig. 7.7. That is, we make a class $\mathfrak{R}$ of subsets.

Patterns in Data Ensembles: In order to discuss Pattern Discovery in data ensembles, we must have a way to discuss the uncertainty of Occam memories to predict future information states.

We cannot use:

$$H\left[\vec{S}\right] \tag{7.1}$$

Since this is infinite. Instead we use:

$H\left[\overset{\rightarrow L}{S}\right]$ where the uncertainty of the next ***L*** data is treated as a function of ***L***. Therefore, $\mathfrak{R}$ captures or discovers a pattern, ***iff***[1] there exists an L such that:

[1] If and only if

$$H\left[\overset{\to L}{S} \mid \Re\right] < LH\left[S\right] \tag{7.2}$$

$\Re$ discovers a pattern when it tells us something about how the distinguishable parts of the process affect each other, or how $\Re$ exhibits its independence, based on the Renyi entropy calculations discussed earlier. The smaller that

$$H\left[\overset{\to L}{S} \mid \Re\right] \tag{7.3}$$

is, the stronger the pattern discovered by $\Re$. The causal state, as determined by the Occam memory of a captured pattern, together with the next observed process, determine a new causal state (and may cause a redefinition of the Occam memory). Thus, there is a natural relation of succession among the causal states of a captured pattern or causal process. This leads us to the definition of a captured or discovered pattern, which leads to an Occam Memory within the SELF. Each discovered pattern, or Occam Memory will have the following properties:

- Occam Memories are deterministic
- All Occam Memory causal states are independent
- All Occam memories are reconstructed from information fragments
- All Occam memory causal states are maximally prescient
- All Occam memory causal states are minimal for all prescient rival memories
- All Occam memory causal states are unique
- All Occam memories are minimally stochastic for all prescient rival memories.
- The excess entropy, E, of an Occam Memory is the Mutual Information between the memory's semi-infinite past and its semi-infinite future.

7.6.3 DART Computational Mechanics and Occam Learning

Entities should not be multiplied unnecessarily.

William of Occam (1320 A.D.)

This maxim from William of Occam, called "***Occam's Razor***," is often sited to justify one hypothesis over others, and is taken to mean "prefer simpler explanations." However, what reason might we have to believe that simpler explanations lead us to a hypothesis with fewer errors?

One might simply reason this from the observation that there are far fewer simple explanations than complex ones. However, it may be no more complex than the reasoning that simple explanations are less likely to fit data, just by chance. Another way to view this is that by favoring smaller hypotheses over larger, we are less likely to run across bad hypotheses, which one of the fundamental axioms behind Occam Learning. Another axiom of Occam Learning is:

Learning is Data Compression

The more the data are compressed, i.e., the more complex the learning algorithm, the more likely something subtle that is important is missed or eliminated. Reasoning from this perspective, we define an “Occam Learning Algorithm” to be one that produces hypotheses, or “Pattern Discoveries”, that are simple in structure, and grow slowly as more data are analyzed. In fact, analysis has shown [41, 45, 47] that if we have a small hypothesis space, then by taking a polynomial number of data samples, we can achieve “Uniform Convergence,” i.e., the chance that any bad hypothesis with error > c, that is still consistent with the data, can be forced below some arbitrary number δ [185, 186, 188]. In the converse, is it impossible to get uniform convergence with a large hypothesis spaces, given a polynomial number of data samples, the answer is, sometimes [189].

Since learning is very stochastic in nature, particularly for real-time systems with heterogeneous data inputs, and given that it is impossible to know how many data points for a given unknown pattern may exist, we employ Occam Learning to provide Pattern Discovery. What we desire, then, is a mathematical framework and foundation that a DART Occam learning component, based in computational mechanics and Occam Learning principles can provide a SELF with autonomous understanding, reasoning, and decision making. Towards this end, a memory computational framework that encompasses the computational theory of machine learning is discussed here. The goals of which are:

- To provide computational mechanics mathematical models that capture key aspects of Occam Learning.
- To provide the system self-analytical metrics for its algorithms:
  - When will they succeed?
  - How long will they take?
- To develop algorithms that provably meet desired criteria;
- To provide the system self-guidance about which algorithms to use when.
- To allow the system to analyze the inherent ease or difficulty of learning problems.

Figure 7.8 illustrates the Occam Learning, computational mechanics framework.

We have provided a mathematical basis for a DART Occam Learning component, based in computational mechanics. As discussed, the Occam Learning component is but one of many learning constructs that must reside in a SELF cognitive framework to allow it to act autonomously and to make sense of global complexities. The Occam Learning Computational Framework provides the ability for simple Pattern Discovery that feeds more complex memory and

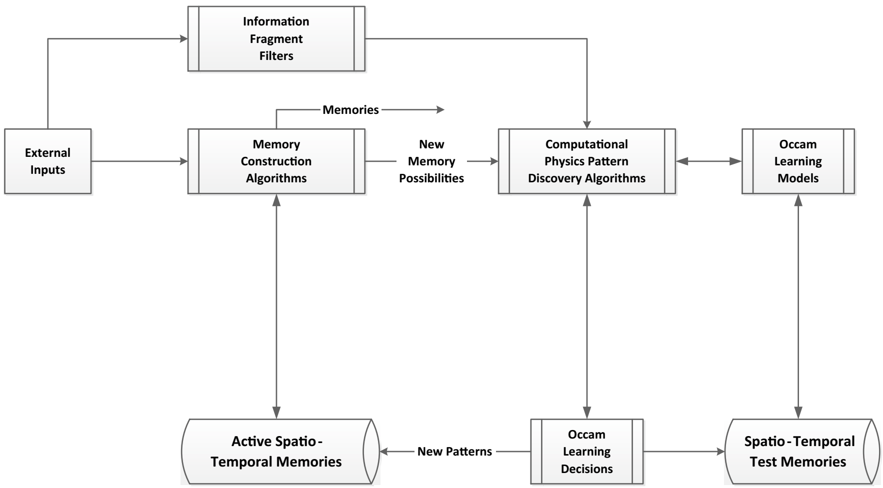


Fig. 7.8 SELF DART Occam learning architecture

inference systems within the ACNF to allow the autonomous system to think, reason, and evolve.

7.7 DART Constructivist Learning Concepts

Here we will discuss constructivist learning concepts where learning is viewed as a constructive process in which the learner is building an internal illustration/representation of knowledge, a personal interpretation of experience. This representation is continually open to modification, its structure and linkages forming the ground to which other knowledge structures are attached [190]. Learning is an active process in which meaning is accomplished on the basis of experience. This view of knowledge does not necessarily reject the existence of the real world, and agrees that reality places constrains on the concepts that are, but contends that all we know of the world are human interpretations of our experience of the world. Conceptual growth comes from the sharing of various perspectives and the simultaneous changing of our internal representations in response to those perspectives as well as through cumulative experience [14, 15].

When considering a SELF, we have to ask ourselves "what is its reality?" When considering human reality each person logs experiences of events. Each person will see reality differently and uniquely. There is also actual reality. Actual reality may be based on fact or perception of fact. In fact, we construct our view of the world, of reality, from our memories, our experiences. In Constructivist Psychology, according to Kelly [149], constructivist philosophy is interested more in the people's construction of the world than they are in evaluating the extent to which such constructions are "true" in representing a presumable external reality. Thus, it makes sense to look at this in the form of legitimacies. What is true is factually legitimate, and people's construction of an external reality is another form of legitimacy. Later we propose consideration of locus of control in relation to internal and external legitimacies/realities. Artificially cognitive systems like a SELF will have their own perceptions and realities, different entirely from humans. However, a SELF's cognitive framework and memories must have similar abilities to construct and rely on correct views of the world around if humans will eventually rely on autonomous systems. Thus, a priming/mentoring process will be necessary. Primer/mentors will need to be trained and understand artificial cognitive systems, their place in society, as well as ethical ramifications of employment. An entire discipline may evolve for creating autonomous objectives and goals within autonomous systems.

As described previously, Constructivist Psychology is a meta-theory that integrates different schools of thought. According to Adlerian, in his work on therapy as a relational constructivist approach [3, 4], the emphasis is on the importance of humans as active agents in the creation of their own constructive psychology. Another view of humans as agents in their own consciousness [122, 123] describes humans as existing in a socio-cultural world of persons where the distinguishing characteristic of their personhood is the possession of an individual agented consciousness. If the individual, or in our case a SELF, has no self-reflexive abilities, then it is unlikely that the entity will possess the capability to reflect critically and creatively about themselves and their interaction and place in their environment. A SELF, therefore, in order to become and remain a self-contained, completely autonomous entity, capable of self-assessment and self-reflection to modify its internal models of its environment, must possess these cognitive skills within a DART learning framework.

Botella [24] believed there were three main areas to consider in Constructivist Learning: psychological knowledge, psychological practice, and psychological research. In his book on Constructive Psychotherapy, Mahoney dealt with psychological knowledge [165], discussing that knowledge could not be separated from the process of knowing, and that all human knowing is based in value-generated processes. Kvale dealt with the process of psychological research [155] in postmodernism terms, viewing psychological practice not as a mapping of objective reality, but rather as an interactive co-construction of the subject under investigation or consideration. This view of psychological research is an interpretive view, requiring the use of hermeneutics, phenomenological, and narrative methodologies [204].

Both Kvale and Botella view constructivist learning as a meta-theory that assumes that knowledge is a hypothetical (anticipatory) construction. In this way it diverges from traditional objective views of conceptual knowledge as an internalized representation of reality. A SELF utilizes the concepts of Kvale and Botella in its Dialectic Argument Structure (DAS) for abductive reasoning and genetically controlled constructivist learning methodologies for cognitive processing and learning. This leads a SELF cognitive processing system to take on an epistemic view of knowledge acquisition and processing, taking into account values that are, by definition, subjective in nature. The most prevalent of these are:

- The pragmatic value of knowledge claims; i.e., their predictive efficiency, viability, and fertility
- The coherence of knowledge claims; i.e., their internal and external consistency and unifying abilities across the cognitive framework.

In their work on the biological roots of human understanding [169], Martin and Sugarman argued that humans are self-creating, self-producing systems and are capable of maintaining their own organization which performs development and maintenance. Hence, a SELF, extends these concepts as it also performs self-assessment, self-maintenance, and self-evolution as an artificial entity [107]. We believe these concepts equate to characteristics and constraints which should be embedded in a SELF at start-up and initialization.

SELF Constructivist Learning is enabled by the DART cognitive learning process, and is a building (or construction) process in which a SELF's cognitive system builds an internal illustration of knowledge, based on its experiences and personal interpretation (fuzzy inferences) of experience. The knowledge representations and knowledge relativity threads within the cognitive system's memories are continually open to modification, and the structure and linkages formed within the SELF's short-term, long-term, and emotional memories, along with the contextual knowledge relativity threads, form the bases for which knowledge structures are created and attached to the Binary Information Fragments. Learning becomes a very active process, where meaning is accomplished through experience, combining structural knowledge (knowledge provided in the beginning) with constructivist knowledge to provide the SELF's view of the "real world" around it. Conceptual growth within the autonomous SELF would come from collaboration among all Cognitrons within the system, sharing their experiences and inferences; the total of which creates changing interpretations of their environment through their collective, cumulative experiences.

Therefore, one of the results of the Constructivist Learning process within the SELF is to gradually change the "Locus of Control," described earlier, from external to internal. Within the context of the SELF, external refers to the fact that the system needed external inputs in order to make sense, or infer, about its environment. Internal, for the SELF, implies that the system has a cumulative constructive knowledge-base of information, knowledge, context, and inferences to handle a given situation internally; able to make relevant and meaningful decisions or inferences about a situation without outside knowledge or involvement. We believe this follows theories of human cognition and is possible through the use of the learning system we have created and the Metacognitive and Metamemory Constructs we have already described, along with Occam and PAC learning methods [174]. This, combined with the Cognitive Economy concepts, provides a vital piece of a SELF's fully autonomous, cognitive framework, required for completely autonomous environmental interaction, evolution, and SELF control.

7.7.1 Adaptation of Constructivist Learning Concepts to the SELF

Constructivist learning concepts within the SELF DART processes are utilized to strengthen knowledge that has been gained through the multi-level learning processes. This knowledge strengthening is in terms of gaining a better understanding of topics, information, inferences, etc. that have been learned. This can be seen as the process of increasing the SELF's knowledge density (explained in Sect. 8.6) for topics that are part of its Conceptual Ontology. These constructivist learning processes cooperate with the PENLPE learning management system in administering Cognitron goals and constraints throughout the SELF's ISAAC cognitive processing environment. These learning management algorithms help define the roles of the learning algorithms within the DART learning processes. The constructivist learning processes help establish learning Measures of Effectiveness (MOEs) against the goals and constraints developed by the PENLPE learning management processes. These constructivists learning processes utilize hypothesis generation and testing in cooperation with the DART knowledge acquisition and learning system.

The function of learning within this role is to increase the stimulus–response-feedback loop for knowledge carried within the SELF Cognitive Conceptual Ontology. In essence, this provides a synthetic "focus" for SELF conscious processes, providing additional information, goals, and constraints for the current behaviors and current memories. The result is additional contextual threads (Knowledge Relativity threads discussed in Sect. 8.5) attached to current memories to provide an additional context or connectivity across a SELFs memory systems. As the name implies, this allows the SELF to "construct" knowledge in an organized and focused fashion, based on its current informational and knowledge acquisition needs. This may include additional emotional memory triggers as well as additional procedural memories tied to current topical memories.

Another aspect of constructivist learning within the DART system is learning to acquire knowledge, in terms of understanding new information, new topics, etc., that have not been previously experienced or learned. Within this role, the PENLPE learning management system presents new information/concepts to be learned, based on sensory inputs that have been processed and reasoned about and correlated with the current Conceptual Ontology within the ISAAC cognitive framework. In this mode of learning within the DART, the role of the learning algorithms is re receive and processes information in order to from new concepts that must be added to the SELF's Conceptual Ontology utilizing the Occam Learning algorithms. The function of Occam Learning in this role is to first utilize the abductive hypothesis generation process (Sect. 9.1.4) to create new possible concepts to explain new data/information that are then utilized by the Occam Learning algorithms to support or rebut the hypotheses, and in the end, strengthen the learning system with the new concepts whose hypotheses are found to be supportable. Knowledge Relativity threads are attached to the new concepts, based on the output of the Topical Maps. This provides for the creation of initial Binary Information Fragments and possibly new or amended emotional and procedural memories within the SELF memory systems.

The last aspect of constructivist learning utilized within the SELF DART system is learning used to construct knowledge, meaning to create a knowledge representation within SELF memories and create meaningful connections (Knowledge Relativity Threads) between knowledge. The role of the PENLPE Learning Management System here is to provide cognitive guidance and modeling within the ISAAC cognitive framework. This involves deconstructing information into manageable information fragment objects, correlating and integrating these new memories into the SELF's current memory structure. The DART then must encode these new information fragment objects, based on the Knowledge Relativity Threads and Information Encoding schemes, to create Binary Information Fragments. The overall role of the DART learning algorithms here is for reasoning and analysis of data/information to determine the stimulus/response goals and constraints that must be added to the system to handle these new memories. This facilitates "making sense" of the information and constructing knowledge representations for the new information. The functions for the DART learning algorithms in this mode of constructivist learning is to create meaningful information fragment representations and contextual information in order to allow the SELF memory system to integrate and assimilate these new Binary Information Fragments into the SELF's long-term memory (memory organization and integration). The overall focus of constructivist learning in this role is to provide active or conscious learning, utilizing a variety of cognitive processes involving Reasoner and Analyst Cognitrons during the learning process, including the construction of emotional contexts for these new memories.

7.8 Discussion

Human cognitive processes rely on the ability extract and generalize knowledge (reason) from a few specific examples. We discussed the basic types of reasoning in Chap. 2. Chapter 8 presents the SELF's reasoning framework, beginning with a discussion of human reasoning and describe the architecture and processes for human-like reasoning within the SELF.

Chapter 8
Synthetic Reasoning

As explained earlier, the ability to reason within a SELF denotes the ability to infer about information, knowledge, observations, and experiences, and affect internal changes that enable it to perform new tasks previously unknown or to perform tasks already learned more efficiently. The act of reasoning, or inferring, allows a SELF to construct or modify representations of experiencing and learning. Reasoning allows a SELF to fill in skeletal or incomplete information or specification (self-assessment). Hence, this chapter is devoted to architectures and frameworks to enable artificial reasoning within a SELF's cognitive processes that synthesizes human reasoning. First, we will discuss the various stages and forms of human reasoning. The rest of the chapter is devoted to adapting human reasoning concepts into SELF reasoning architectures.

8.1 Human Reasoning Concepts

Human reasoning is dynamic and complex involving significant numbers of intertwined complex processes. There are different types of reasoning necessary to allow humans to navigate their world effectively and efficiently. A brief overview is provided; as the topic of human reasoning is vast. First, an overview of brain theory is provided, followed by logical reasoning considerations, as well as, moral, ethical, and emotional reasoning. Finally, we consider implicit and explicit reasoning.

Significant information is ingested simultaneously into the brain and it is therefore impossible to consciously be aware of it all. Imagine for just a minute how many discrete activities the human brain is handling in one instant. Memories, associations, habitual ways of thinking, beliefs, assumptions, predictions, experiences, past, present, and planned comprise the short list. We have senses and perceptions. We have defense mechanisms and feelings. Needless to say, our brains are active! Hence, it is difficult to imagine what all occurs in a single instant of human experience, but it is ultimately essential to explore in order to understand how to translate the concepts in realistic and efficient artificial reasoning.

J.A. Crowder et al., *Artificial Cognition Architectures*,
DOI 10.1007/978-1-4614-8072-3_8,

8.1.1 Human Thinking

It is general knowledge that there are different functional reasoning components and processing regions of the brain. The frontal cortex is responsible for executive functions, the limbic system (emotions), and others are described throughout the book. Each region of the brain is made up of significant numbers of neurons, chemical transmitters, and electrical activity. The following subsections describe types of human reasoning beyond the physical aspects of brain function: Specifically, there are three main theories of human reasoning:

- Modular reasoning
- Distributive reasoning
- Collaborative reasoning

8.1.2 Modular Reasoning

Cognitive modularity seems to have flourished with Fodor [115]. He articulated that humans use domain-specific modules that together form part of the reasoning system within the human brain. According to Fodor, there are conditions for modular cognition; one is that other parts of the brain have limited access to each reasoning module. This type of reasoning is mandatory, innate, shallow and very fast. He also stipulated that each module was fixed to a neural architecture and that information was encapsulated; since other modules have limited access to each other. More modern psychology believes that cognitive modularity as actually massive modularity. This school of thought suggests that the mind is even more modular with specific functions and specialization [118]. This type of Modular Reasoning is used within SELF Sensory Processing, before Sensory Integration. There are differing views on massive modularity. According to Raymond Gibbs and Van Orden [119], massive modularity theory has empirical problems. They state that studies fail to be able to separate modules. They also argue that massive modularity theory fails to discover input criteria and state that massive modularity may be impossible given the nature of context embedded within human nature. Lastly they argue that massive modularity does not acknowledge the interaction of brain, body, and world in human thinking.

8.1.3 Distributed Reasoning

Distributed theory suggests that there is more to the brain than massive modularity. Beyond some very specific areas such and motor control, distributed reasoning theory suggests that there exist many fuzzy connections between systems of the brain.

The distributed theory challenges boundaries of the mind and body taking into account the environment, artifacts, and specific characteristics of differing people. This theory is reflected in the Fuzzy, Possibilistic Abductive Network utilized within a SELF's cognitive framework.

The distributive reasoning theory by Hutchins [138] provides some insights into human reasoning. Hutchins provides five different models that affect human reasoning. First, he postulates that there are modules within the brain that are specialized in function and structure and are united in a complex way. Second, he argues that cognition at a macro level is distributed outside the individual; such as the media. Media can be internal and external. Third, there is human culture which influences the individual. Fourth, there is society which cognitive activity is distributed in tools, rules, and contexts. Finally, he argues that cognition is distributive in time, both vertical and lateral time dimensions of the subject [195].

Yvonne Rogers [195] provides a detailed analysis of the distributive cognitive model. Rogers cites Hutchins as creating a computational model of two modules of the brain that can together recover depth that neither module alone could do. One general assumption of the distributive human cognitive system is that it is made of more than one module and that each module in the cognitive system has different cognitive properties than the individual, and is different than the cognitive brain as a whole. Another general assumption made by Rogers is that members (modules) of the system have knowledge that is both variable and redundant and that members of the system can pool resources. Another is distribution of access to information. This enables the coordination of expectations and coordination of action within the human biological reasoning framework [195]. These concepts are utilized throughout a SELF, which utilizes localized processing modules (processing "experts") as well as distributed Cognitron experts that communicate and collaborate throughout a SELF cognitive system. This type of reasoning is discussed next.

8.1.4 Collaborative Reasoning

The collaboration theory suggests that both modular and distributed forms of processing occur within the human reasoning framework, and that it is a matter of degrees of each. We may be able to see this in example of small groups. If we consider the degrees of modular and collaborative reasoning we may have something like this: modular theory is represented by the individual thinking/reasoning internally and distributive theory is represented by the individual thinking and then being influenced by their environment. So on one end of the spectrum, individuals reason by employing brain modules. In the middle, individuals reason with thinking modules and with internal and external relationships, and with time. On the other extreme, thinking may be largely environmental, external, and limited individual thinking such as group think.

Group think is where members of a group think alike and fail to challenge alternative thinking and believe their way is the best way. The group tends to have a "group mind" only without any differences from individual members and hence, come to conclusions as a whole void of valuable diversity. Historically, it is well known that many mistakes have been made in decisions where group think created an environment of blindness inconsiderate of any outside view, individualism, or diversity of thought.

Modularity has become more complex as the field advances, moving from modular to massive modular. Each acknowledges modular function but the latter evolving to become more specified and specialized. Empirical study becomes more difficult affected as the level of volume and complexity increase. This shift from modular to massively modular shifts in levels or degrees and thus, is similarly in line with the collaboration theory of degrees, which states that Modularity and Distributed are both not only possible, but likely a matter of degrees.

8.2 Types of Reasoning

8.2.1 Logical Reasoning

There are three major human reasoning strategies: inductive, deductive, and abductive.

Inductive Reasoning: Inductive reasoning involves concluding after evaluating facts; reasoning from specific facts to a general conclusion and allowing for inferencing. It also requires human experience to validate conclusions. An example might be: Zebras at the zoo have stripes, therefore all zebras have stripes [97].

Deductive Reasoning: Deductive reasoning is just the opposite. Deductive reasoning moves from general principle to specificity. This type of reasoning is based upon accepted truths. An example of deductive reasoning might be: I know that all zebras have stripes therefore when I go to the zoo, if I see a zebra, it will have stripes.

Abductive Reasoning: Abductive reasoning allows for explanatory hypothesis generation or generating ideas outside of the given facts to explain something that has no immediate satisfactory explanation.

There are a number of ways in which people reason, but most often human reasoning follows either inductive or deductive reasoning. Other ways that humans reason includes cause and effect reasoning where causes and after effects are considered. Analogical reasoning is a way of relating things to other novel situations. Comparative reasoning as it implies involves comparing things. Still another reasoning method is conditional or if/then reasoning. Many of us have used the pros and cons methods of reasoning as well. Systemic reasoning involves thinking that the whole is greater than the sum of its parts, and finally reasoning by using examples or analogies. Hence, there are numerous logical ways in which people reason about events and situations.

8.2.2 Humans and Inductive/Deductive Reasoning

Some believe that there are only the two types of reasoning: inductive and deductive. Lance Rips [193] writes about these considering them as Strict and Loose reasoning. Loose reasoning is a continuous process of updating confidence of a belief. Any belief can raise or lower confidence in this view of reasoning. Inductive reasoning follows "loose reasoning" and deductive reasoning follows a more strict view of reasoning. Arguments can be deductively valid and inductively strong. Therefore, these forms of reasoning may not operate very differently in the human brain. To believe something, means to increase our confidence in it. Hence, if these reasoning processes can be combined as one psychological process then an important task would be to describe the different types of process manifestation. The author suggests physical examples where reasoning could be explored, such as in parallel networks or production systems.

Jonathan Evans and his colleagues [114] studied different theories of reasoning and concluded that all formal reasoning theories state that humans possess inference rules. Other theories such as mental models theories [143] proposed that human reasoning is semantic versus syntactic. Reasoning relies on systematic processes to construct and evaluate mental models [32]. Another major reasoning theory proposes domain-sensitive rules or schemas. People tend to abstract schemas or structure within domains where they have relevant experience. Lastly, there is a heuristics and biases approach to reasoning [113]. In heuristic reasoning, subjects' reason about features perceived as relevant.

8.2.3 Moral and Ethical Reasoning

Moral reasoning is reason that considers benefits to other, self, or society. Lawrence Kohlberg [151, 152] theorized that moral development progresses in three phases: Pre-Conventional, Conventional Morality, and Post-Conventional Morality phases.

Pre-Conventional Phase: The Pre-Conventional phase, includes two stages. The first stage is a Punishment-Obedience Orientation. The goal of this reasoning is to avoid being punished. The second stage is an Instrumental Relativist stage. In this stage the reasoning goal is to primarily meet one's own needs and occasionally meet the needs of others.

Conventional Morality Phase: Conventional Morality also contains two stages. The first is Good Boy-Nice Girl Orientation. This stage of reasoning is based upon what others may think of me. The next stage is Law and Order Orientation.

This reasoning develops from what is right and just and orients an individual to a fixed set of rules.

Post-Conventional Morality Phase: The Post Conventional Morality has two stages. The first stage is Social Contract Orientation and deals with social utility and the possibility of changing a set of rules for the benefit of society and considers individual rights [31]. The final stage is Universal Ethical Principle. In this stage dilemmas may be thought about abstractly considering universal principles of justice, reciprocity, and respect for individualism.

Why is moral reasoning important? It is a way that humans solve problems, and comprises logic leading humans to dilemmas. There does not exist a clear-cut, right and wrong answer to dilemmas. Hence, humans follow a way of thinking that leads to an ethical decision. Typically ethical reasoning follows a set of specific steps. Many professions have their ethical codes and decision-making models to follow. Some of these are complex and lengthy and lead humans to come to conclusions based on rules of thinking or operating.

8.3 SELF Reasoning

As discussed above, reasoning takes on a number of forms, but two important forms comprising a SELF are induction and abduction:

- Induction: Extrapolates from information and experiences to make accurate predictions about future situations.
- Abduction: Genetic algorithms generate populations of hypotheses and a Dialectic Argument (Tolmin) Structure is used to reason about and learn about a given set of information, experiences, or situations, also called "Concept Learning."

Earlier we briefly discussed the Dialectic Argument Structure (DAS). This section provides more detail of its architecture and design. The Dialectic Argument Structure seeks answers to questions that require interplay between doubt and belief, where knowledge is understood to be fallible. This humanistic 'playfulness' of compare and contrast is a key to searching and exploring information. We propose that utilizing this framework for reasoning about information, hypotheses, and problems provides a robust, adaptive information processing system capable of handling new situations. DAS utilizes abductive logic, sometimes called critical thinking, in order to distinguish it from more formal logic methods like deduction and induction. Whereas data mining utilizes induction to develop assertions that are probably true, the dialectic search uses abductive logic methods and processes to develop hypotheses that are possibly true. DAS specifically avoids Bayesian methods because they cannot measure possibilistics, but instead measure probabilistic metrics. Instead we utilize a fuzzy implementation of Renyi's entropy and mutual information theory to provide a possibilistic measure of mutual information and topical separation [194].

8.4 Abductive Reasoning: Possibilistic, Neural Networks

The original McCulloch-Pitts model of a neuron contributes greatly to our understanding of neuron-based systems. However, the model failed to take into account that the simplest human nerve cell types exhibit non-deterministic behavior [177, 178]. Some have attempted compensation by modeling as a function of randomness, creating a stochastic neural network. However, much of the behavior is not random, and carries imprecision, associated with lacking sharp transition from occurrence to non-occurrence of a given event. This leads us to the definition of a network not steeped in Bayesian statistics (a Bayseian Belief Neural Network – BBNN), but one utilizing possibilistics, based upon a humanistic environment of fuzzy characteristics, combined with an abductive, hypothesis-based decision network; and thus creating a Possibilistic, Abductive Neural Network (PANN) [38]. Here we discuss the theory and architecture for a Possibilistic, Abductive Neural Network capable of complex hypothesis generation and testing, leading to artificial creativity and discovery within an artificially intelligent system [100, 101].

8.4.1 Artificial Creativity

Neuroscience research into human perception [225] determined that noise and imprecision in the human nervous system was not, in fact, inconvenient, but was actually essential to the types of computations the brain performed [84]. The brain learns to make spatio-temporal associations in the presence of noisy, imprecise information. Therefore, we propose that any artificially intelligent system that tries to emulate human processing must be able to make similar noisy, imprecise associations within its artificial neural systems even when they are incomplete, imprecise, or contain conflicting information while simultaneously properly representing currency of entity behavior, i.e., accounting for its real-time internal state [22, 57, 59].

A Possibilistic, Abductive Neural Network (PANN) architecture that is capable of complex hypothesis generation and testing in the presence of multiple, noise, imprecise, and possibly incomplete information is inherently necessary for the types of environments an autonomous SELF is likely to operate within. These conditions define typical real-time processing situations which require specialized processing essential for developing realistic complex decision support systems that can learn, reason, analyze, and make critical decisions in real-world environments.

8.4.2 Creativity Through Problem Solving

Touring and others have hypothesized that computers cannot be creative, due to the absence of novelty in its flow of information processing. The use of stochastic, possibilistic abductive networks provides an approach to information processing, allowing the artificially intelligent system to vary its information processing flow, depending upon continuously changing generated hypotheses and continuously recombinant neural fiber network creation processes described earlier [57, 184, 229].

One hypothesis we propose to consider here is that creativity is directly related to problem solving activities in which explorations of problem spaces lead to the expansion of belief domains. A successful expansion of beliefs is initiated by recombinant updates of a cognitive system's Conceptual Ontology [191, 204, 229].

General heuristics used as continuous pedigree input within the genetic hypothesis generation process guide the support and rebuttal informational search processes and problem solving activities; which include strategies for examining, comparing, altering and combining concepts, strings of symbols, and the heuristics themselves [229].

But what kind of creativity is possible for the AI system in this context? We believe the answer is that it is similar to the one which humans experience in our everyday life: the experience of new and original ideas that have value, based on the overall goals, constraints, and mission directives of the environment the AI system is within.

Within this context, we describe the design of a PANN as a candidate to facilitate artificial creativity, and therefore autonomous, real-time decision support, as a primary objective; the abductive dialectic argument structure provides an inference engine upon which artificial creative reasoning in a SELF is based. Hence, Cognitrons and the Dialectic Argument Structure (DAS), which comprise the core primitive SELF foundational abductive reasoning components, are now expounded upon in the subsequent paragraphs.

Cognitrons are currently composed in Java and are employed together as part of a Java processing framework creating a synthetic PANN. Cognitrons are used to mimic human reasoning. The Java based Cognitron architecture framework and toolkit is utilized for constructing a system of dynamically changing Cognitron functions and applications for problem solving within a SELF cognitive system. Together, this toolkit allows a SELF's Artificial Prefrontal Cortex the ability to build a PANN's multi-Cognitron autonomous decision support system [63]. This system includes a framework for providing business rules and policies for run-time systems, and the autonomic abductive computing core technology. Hence, the next section describes the Dialectic Argument Structure and its use in Abductive Reasoning [37].

8.4.3 Dialectic Reasoning Framework

The Dialectic Search is a reasoning framework which seeks answers to questions that require interplay between doubt and belief, where knowledge is understood to be fallible. Cognitrons which support the Dialectic framework learn and reason about information, hypotheses, and problems and provide the adaptive information processing system proposed for handling new situations. The key value of Cognitrons within the Dialectic Search reasoning framework is their ability to learn from sensory data and from each other. Using learning methods discussed in Chap. 7, Cognitrons provide the operations and analytical structures used to extract knowledge and context from various sources of information. Cognitrons are cloned/scaled dynamically as system resources allow. Other types of Cognitrons are utilized by the Dialectic Search to provide Artificial Cognition [78, 79] and the Artificial Prefrontal Cortex (Mediator) [80] autonomous processing capabilities.

In the Dialectic Search processes, information is utilized to generate and assess hypotheses from thought processes continuously matured by Cognitrons, which learn and reason about the hypotheses and information with a Dialectic Argument Structure. As explained in Chap. 4, Cognitrons are autonomous software agents that create the essence of an information agent ecosystem. They comprehend SELF external and internal environments and act upon it over time, in pursuit of an a priori given and self-developing humanistic agenda and goals, to affect what it can comprehend as it learns.

The Dialectic Search uses the Toulmin Argument Structure to find and relate information that matures and shapes a larger learned argument, or intelligence lead. The Dialectic Search Argument (DSA), illustrated in Fig. 4.13, has four components:

1. Data: in support of the argument and rebutting the argument.
2. Warrant and Backing: explaining and validating the argument.
3. Claim: defining the argument characteristics and criteria.
4. Fuzzy Inference: relating and maturing the data to the claim.

The Dialectic Search process, called Dialectic Search Argument (DSA) serves two distinct purposes within the reasoning framework. First, it provides an effective basis for mimicking human reasoning. Second, it provides a means to glean relevant information from the Topic Map [199] and transform it into actionable intelligence (practical knowledge.) These two purposes work together to provide an intelligent system that captures the capabilities of human Intelligence to sort through diverse information and find clues that support the ability to make actionable decisions.

This approach is considered dialectic in that it does not depend on deductive or inductive logic, though these may be included as part of the warrant. Instead, the DSA depends on non-analytic inferences to find new possibilities based upon warrant examples. The DSA is dialectic because its reasoning is based upon what is plausible or possible; the DSA is a hypothesis fabricated from changing and refining bits or fragments of information.

Once the examples of information that are relevant to the hypotheses topics have been detected, data that fits the support and rebuttal requirements is used to instantiate a new claim. This claim is then used to invoke one or more new DSAs that perform appropriate contextual searches. The developing lattice forms the reasoning that renders the intelligence lead plausible and enables measurement of the possibility.

As the lattice develops, the aggregate possibility is computed using the fuzzy membership values of the support and rebuttal information. Eventually, a DSA lattice is formed that relates information with its computed possibility. The computation, based on Renyi's entropy theory, uses joint information memberships to generate a robust measure of Possibility, a process that is not achievable using Bayesian methods.

Figure 4.13 illustrated a SELF DAS Architecture that is used to implement DAS based Cognitrons: the Coordinator, the Dialectic Argument Search (DAS) and the Search, work together, each with a separate and distinct learning objective. The Coordinator is taught to watch the FUSE-SEM topical maps, responding to newly discovered inputs that conform to patterns of known interest. When an interesting input is received, the Coordinator selects one or more candidate DAS Cognitrons, and then spawns Search Cognitrons to find information relevant to each DAS. Over time, the Coordinator learns which patterns are most likely to yield a promising lead, adapting to changes in the FUSE-SEM topical map structure and sharing what it learns with other active Coordinators.

Search Cognitrons utilize DAS prototype search vectors and, through the FUSE-SEM topical map, finds information that is relevant and related. The Search Cognitron learns to adapt to different and changing source formats and dynamically modifies parsing procedures required to extract newly discovered detailed information. The final Cognitron, the DAS, learns fuzzy patterns and uses this to evaluate information found by the Search Cognitron. Any information that does not quite fit is directed to a sandbox where peer Cognitrons can exercise a more rigorous aggressive routine to search for alternative hypotheses.

The principal requirements addressed with Cognitrons are:

1. Learning to adapt to changes in the surrounding environment.
2. Capturing knowledge for reuse.
3. Sharing of information and learning between Cognitron peers.
4. Hypothesizing in the form of humanistic on-the-job-training
5. Remembering to avoid old mistakes and false leads.

A similar diagram can be drawn for the FUSE-SEM topical map where the Search Cognitron draws information out of heterogeneous sources, the DAS is replaced by a FUSE-SEM topical map, and the Coordinator is a specialized FUSE-SEM with its own specific ontology. The complete process for information search is illustrated in Fig. 8.1. Steps 1, 2 and 4 initialize the process by building the FUSE-SEM topical maps and teaching each DAS what to search for. Steps 3, 5 and 6 are recurring steps as information is gathered and more leads are discovered. Steps 4 and 7 are also recurring as the DAS discovers and adapts to new types of information.

Measures of information possibility (certainty) alerts inform an Interface Cognitron of information available for review. PANN constructs are then used to rank information and flag content considered to be most certain. Review can then be facilitated via user presentation of the DAS warrant, pedigree and reference links to

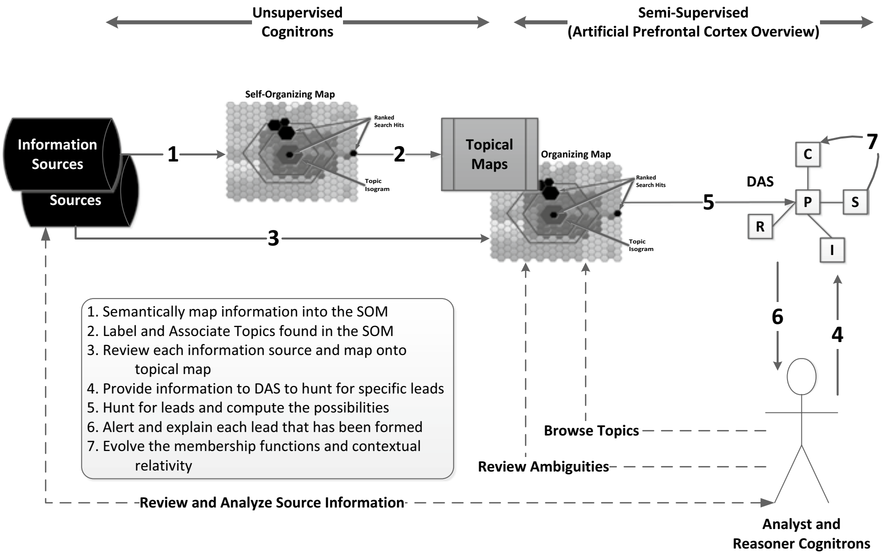


Fig. 8.1 The SELF DAS information search process

support and rebuttal sources, compared back through the FUSE-SEM topical map. Based upon a review, a user may elect to refine DAS training or further browse the Topic Map to rapidly generate alternatives within the DAS that better represents a new lead.

Otherwise, the information search process is also automated, and includes a review process which engages the Artificial Pre-frontal Cortex (APC) in a model review process described earlier, where a Fuzzy, Unsupervised, active resonance theory, Neural Network (FuNN) is then used to continuously develop the DAS density. FuNN critical learning objectives include, but are not limited to, the following:

1. DAS must be able to search and track using the signature of a particular Topic of Interest (TOI),
2. DAS must be able to investigate semantic anomalies found in computation of possibilities caused by information obfuscation,
3. DAS must be able to continuously review its internal FuNN developed adaptations,
4. As adaptations are discovered to be invalid, a DAS must be able to learn additional support/rebuttal arguments for adaptation re-occurrence prevention.

This review process enables a FuNN to learn from either a user or data mining interface. The resulting knowledge is shared across a SELF cognitive framework via each developed DAS, enabling sharing of the semantic corpus of knowledge essence representation.

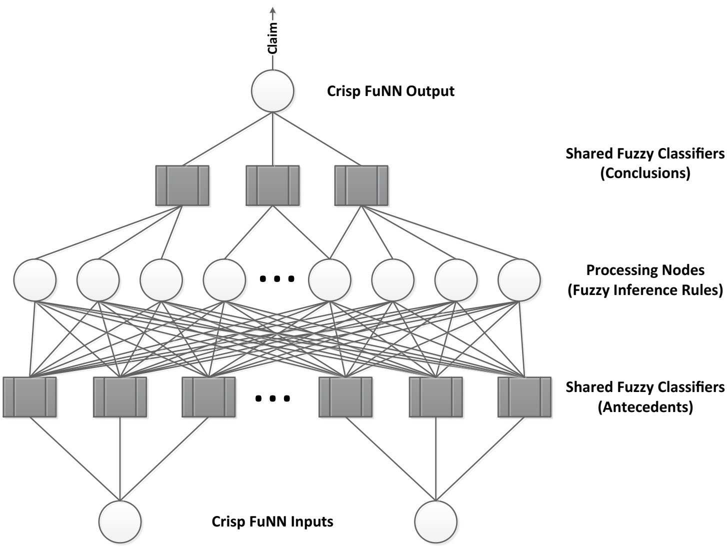


Fig. 8.2 A SELF DAS implemented as a FuNN

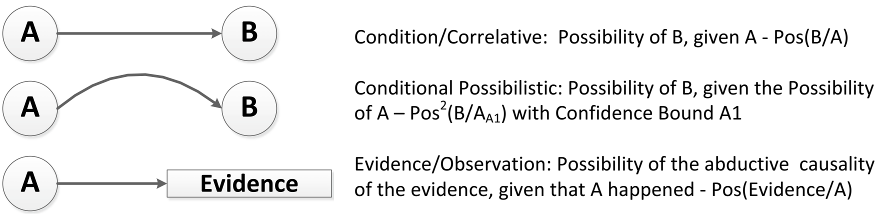


Fig. 8.3 SELF fuzzy possibilistic lattice connections

8.4.4 FuNN Creating DAS

Each DAS is implemented via a FuNN as illustrated in Fig. 8.2. The FuNN is able to learn both fuzzy possibilistic rules and fuzzy sets using the warrant for training data (reference Fig. 4.12). Changing its structure and its weights, the network converges to a state that maximizes likelihood and minimizes Renyi's entropy.

Each FuNN is interpreted using relative connections and membership functions to generate a lattice that interconnects the FUSE-SEM topical map input using fuzzy possibilistic connections illustrated in Fig. 8.3. Possibilistic is an assessment of the plausibility of the DAS. The confidence bound, derived from the FUSE-SEM topical map and FuNN fuzziness, represents a measure of semantic fit.

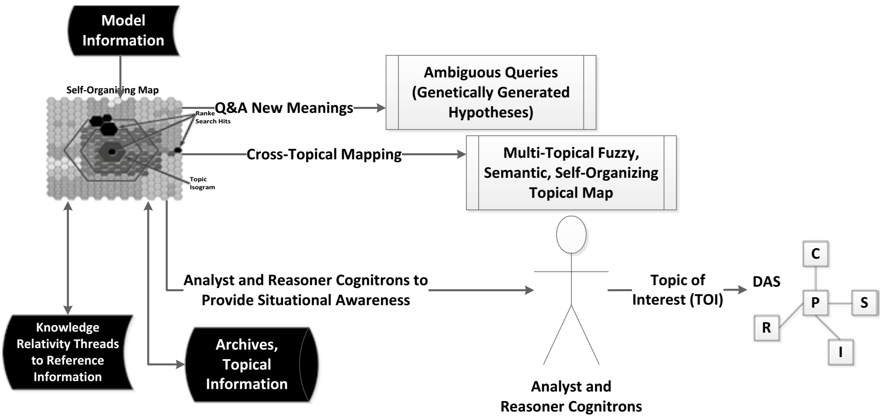


Fig. 8.4 *A SELF PENLPE* Cognitron *architecture*

A DAS lattice is used to explain the information, compute the overall possibilistics, based upon the fuzziness of the support, and rebuttal information, and compute the sensitivity of the claim relative to the fuzziness of the input data. Being able to review the lattice and assess its sensitivity to the ambiguity of the input data enables effective assessment of lead quality.

8.4.5 DAS Reasoning Approximation

The approximation system for autonomous reasoning using a DAS is comprised of a Fuzzy Inference Engine, Cognitrons, and Cognitron infrastructure. This system employs the soft computing techniques explained in Chap. 10 to generate Cognitrons for mimicking human reasoning to process information and develop intelligence. Figure 8.4 depicts a high-level view of the inference and cognitive management architecture for a SELF, known as the Polymorphic Evolving, Neural Learning and Processing Environment (PENLPE).

Figure 8.4 provides a data flow view PENLPE; the interactions of the FUSE-SEM topical maps, and the various Cognitrons. In the PENLPE process shown in Fig. 8.4, the FUSE-SEM topical maps are used to correlate new information with current topical information contained in a SELF's Conceptual Ontology. The new information is also fed to other topical maps in order to understand the contextual relevance between topics. From this contextual information, Knowledge Relativity threads are created and attached to the new information objects. This topical and contextual information is then stored and the combined knowledge objects are sent to active Reasoner and Analyst Cognitrons This process includes Search Information Cognitrons that mine through multiple sources to provide data/information to other Cognitrons throughout the PENLPE, called the Federated Search, shown in Fig. 8.5.

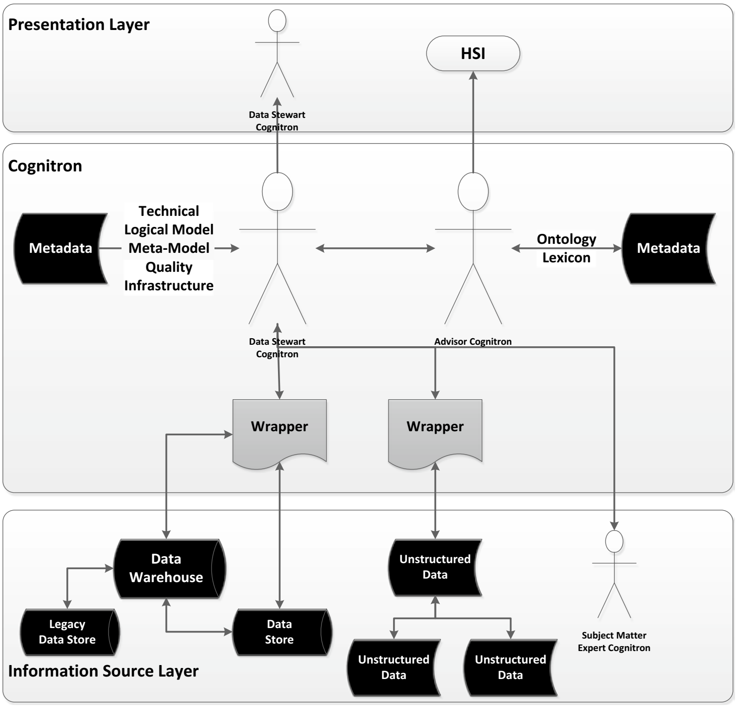


Fig. 8.5 PENLPE federated search process

8.4.6 *Cognitron Archetype Descriptions*

Figure 4.10 illustrated the basic Cognitron archetypes. Figures 8.6, 8.7, and 8.8 provide explanations of these Cognitrons archetypes. These provide the context-sensitive reasoning structure for a SELF PENLPE reasoning framework.

Notice that the Cognitron process described in Fig. 8.5 includes the use of Subject Matter Expert Cognitrons in order to provide initial information to PENLPE. The system does not spontaneously generate initial knowledge; it must ingest information to learn from, created from the autonomous Model-Based process described earlier. This can include a learning based question and answer processing architecture that allows PENLPE to ask questions, based upon contextual understanding of the information it is processing, and extract answers, either from its own inference engines, its own memories, other information contained in its storage systems, or outside information from other information sources. This process is illustrated in Fig. 8.9.

The PENLPE Cognitron processing environment allows data to be processed into relevant, actionable knowledge. Situational management is one of the most innovative components of PENLPE. Utilizing the ACNF framework within PENLPE, it can provide real-time processing of dynamic, situational awareness information.

The information gathering, processing, and analyzing within PENLPE is performed continually to keep track of current trends in the context of current situations,

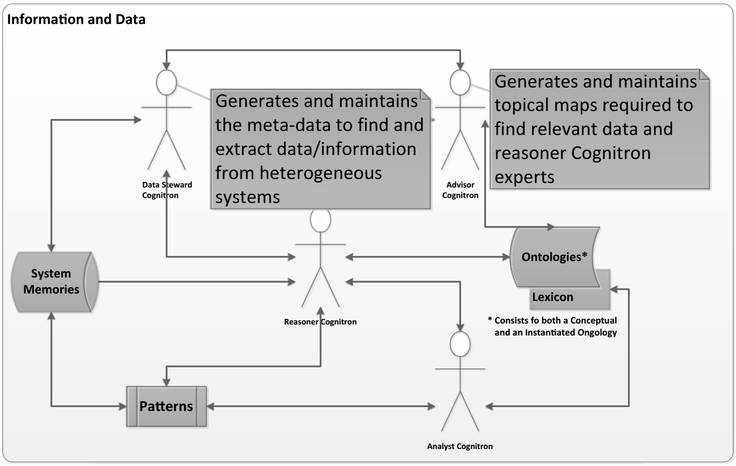


Fig. 8.6 The data Steward and advisor Cognitrons

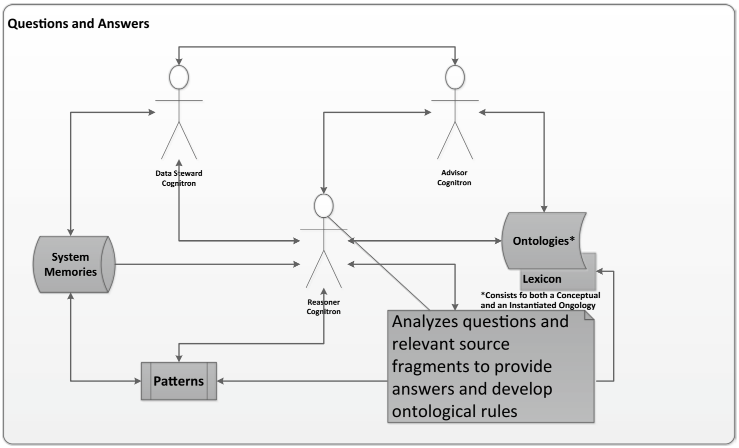


Fig. 8.7 The reasoner cognitron

both locally and globally across the corpus of knowledge, and then provide timely and accurate knowledge to allow users or a SELF to anticipate and respond to what is happening in a changing environment. To achieve a combination of awareness, flexibility, and agility involves dynamic and flexible processes that adapt and morph as situations change. This is possible with learning, evolving, Cognitrons, like those found in the PENLPE processing environment.

Data Steward Cognitrons support growing volumes of data and allow Reasoner Cognitrons to produce accurate and relevant metrics about past, current, and future situations (prognostics). Through inter-Cognitron communication, they provide control and visibility into the entire processing enterprise. This is made possible by integrating

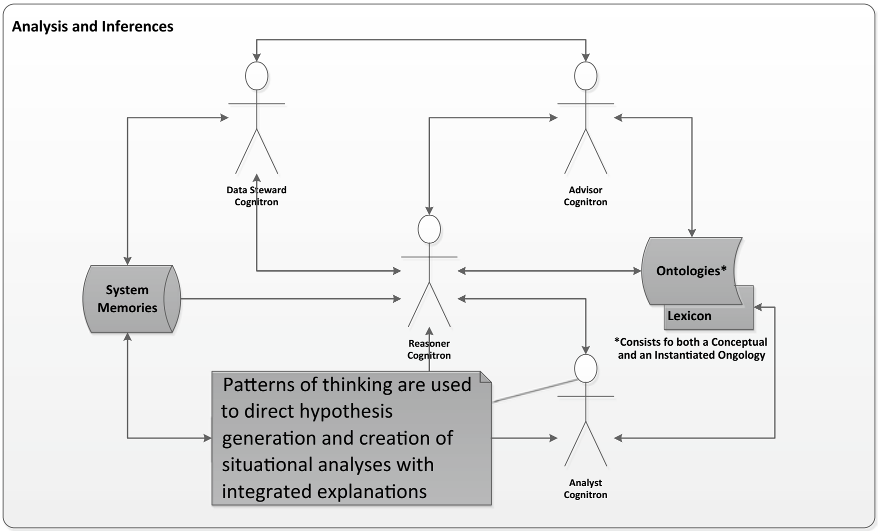


Fig. 8.8 The analyst cognitron

PENLPE into a flexible, distributed processing architecture that enables secure collaboration, advanced information management, dynamic system updates, and customer, rule-based processes (Advisor Cognitrons).

8.4.7 The Fuzzy, Unsupervised, Active Resonance Theory, Neural Network (FUNN)

The basic FUNN was illustrated in Fig. 8.2 above. While the internal structure of the FUNN can evolve over time as the system learns and evolves, the basic structure comprises five layers with "expert" nodes at each layer, which determine layer promotion paths for information residing at each layer. Each node within a layer consists of an input integration function, a fuzzy classifier $f\left(u_1^i,u_2^i,\ldots,u_n^i;w_1^{fi},w_2^{fi},\ldots,w_n^{fi}\right)$, and an output activation function, processing algorithm $a\left(f\left(u_1^i,u_2^i,\ldots,u_n^i;w_1^{fi},w_2^{fi},\ldots,w_n^{fi}\right)\right)$, which determines how the information is processed at the next internal fuzzy layer within the FuNN. Figure 8.10 illustrates this process.

Where:

$$f\left(u_1^i,u_2^i,\ldots,u_n^i;w_1^{fi},w_2^{fi},\ldots,w_n^{fi}\right) \tag{8.1}$$

corresponds to a stochasto-chaotic differential equation, and

$$a\left(f\left(u_1^i,u_2^i,\ldots,u_n^i;w_1^{fi},w_2^{fi},\ldots,w_n^{fi}\right)\right) \tag{8.2}$$

interprets increases or decreases in entropic constraints of the differential equation, based on Renyi's entropy definition.

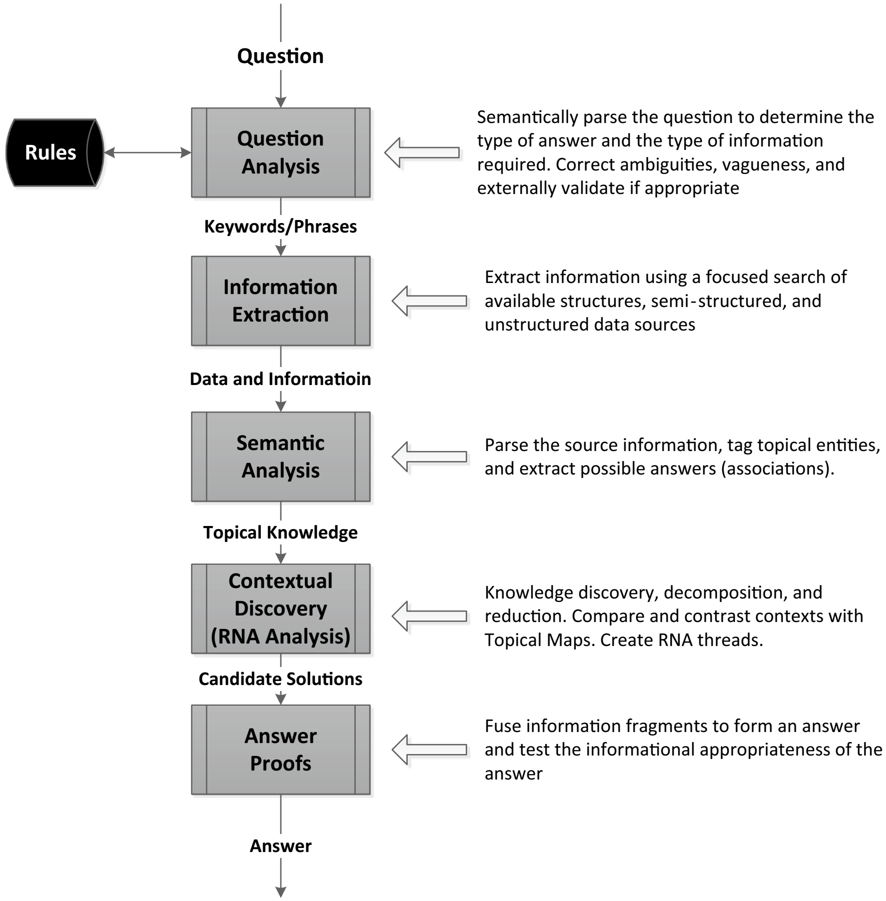


Fig. 8.9 Question and answer architecture for PENLPE

8.5 Cognitron Theory

The notion of an intelligent software agent is not new, and has been an object of research for decades in such fields as psychology, sociology and, of course, in computer science. Strangely, exactly defining what an intelligent software agent is, has only been intensively researched over the last number years. SELF software agents, Cognitrons, continuously carry the all-important cognitive artifacts and are therefore the heart of a SELF cognitive framework.

Because the term "agent" has been used by many, in many different ways, it has become relatively ambiguous and difficult to estimate the possibilities agent technology can afford. Consequently, there are more definitions than there are working examples of systems that could be called agent-based [231]. The misuse of the term 'Intelligent Software Agent' has caused many, unjustly, to draw the

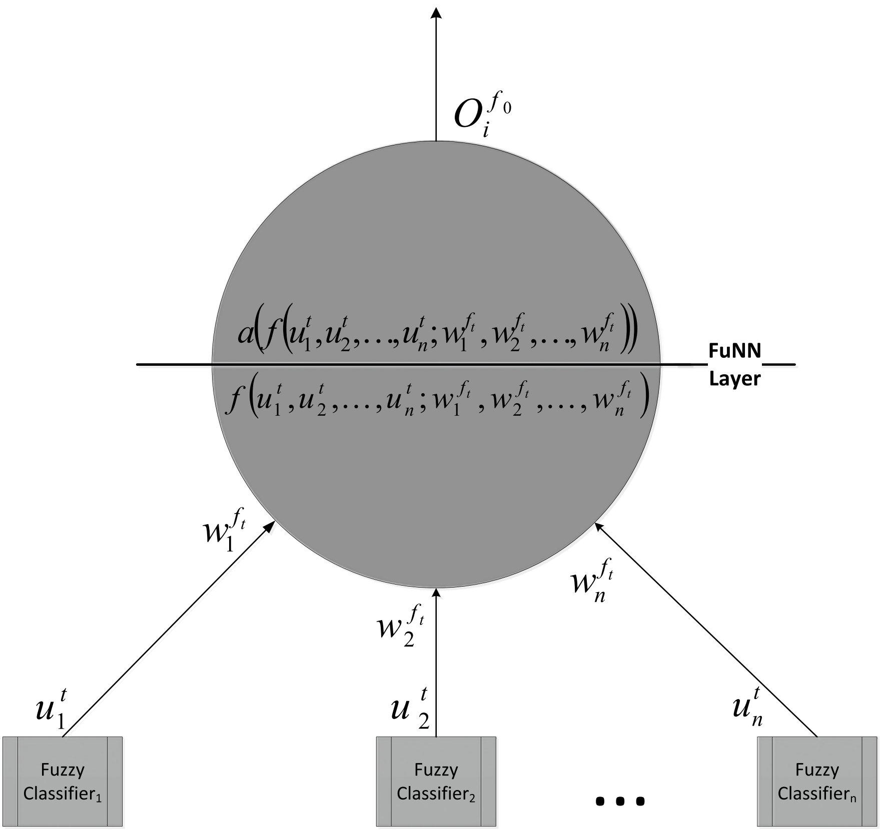


Fig. 8.10 Basic structure of a node within the SELF FUNN

conclusion that software agent technologies, as a whole, do not offer much to the field of software engineering, and to Artificial Intelligence, in particular [230]. Therefore, we provide an overview of software agent theory and practice, to set the proper context for Cognitron usage and Intelligent Software Agent development.

8.5.1 Intelligent Software Agent Definition

We will not attempt to come to a rock-solid formal definition of the concept "agent". Given the multiplicity of roles agents can play, this is quite impossible and even very impractical. However, it is possible to put forth a definition, or notion, of an intelligent software agent.

Next, we provide a list of general characteristics that, together, provide a suitable notion or impression of an intelligent software agent. The next section describes a set of characteristics that notionally describe a "weak software agent", followed by a set of characteristics to describe a "strong software agent." In both, 'intelligence' is considered within context of weak and strong agents.

8.5.2 Weak Intelligent Software Agents

Perhaps the most general way in which the term agent is used, is to connote a software-based computer system that enjoys the following properties:

- ***Autonomy***: agents operate without the direct intervention of humans or others, and have control over some if not all of their actions and internal state.
- ***Social Ability***: agents can interact with other agents and humans via discrete interfaces and/or communication language.
- ***Reactivity***: agents perceive their environment (which may be the physical world, a user via a graphical user interface, a collection of other agents, the Internet, or perhaps all of these combined), and respond in a timely fashion to changes that occur in it. This can entail that an agent spends most of its time in a sleep state, only awakening if certain changes in its environment (like the arrival of new message) occur.
- ***Proactivity***: agents do not only act in response to their environment; they are able to exhibit goal-directed behavior and hence can take the initiative.
- ***Temporal Continuity***: autonomous agents are continuously running processes (foreground or sleeping/passive in the background), not singleton or once-only computations or scripts that map a single input to a single output and then terminate.
- ***Goal Oriented***: an agent is capable of handling ambiguous, complex, high-level tasks. Operational decision changes, such as optimized task splitting into smaller sub-tasks, and/or sequencing/ordering of sub-tasks are also performed by the agent.

Thus, a simple way of conceptualizing an agent is as a software process, which exhibits the properties listed above. A clear example of an agent that meets the weak notion of an agent is the so-called softbot ('software robot'). This is an agent that is active in a software environment (e.g., a Linux operating system environment).

8.5.3 Intelligent Software Agents

For some researchers – particularly those working in the field of Artificial Intelligence – the term agent has a stronger and more specific meaning than what was sketched out in the previous section. These researchers generally mean an agent to be a computer system that, in addition to the properties previously identified, is either conceptualized or implemented using concepts that are more usually applied to humans. For example, it is quite common in artificial cognitive research to characterize an agent using cognitive notions, such as knowledge, belief, intention, and obligation. However, a SELF also considers emotions combined with agents.

Hence, software agents can be given human-like attributes representing agents visually by using techniques such as a emoticons or cartoon-like graphical icons or an animated face. Research into this matter has shown that, although agents are pieces of software code, people like to deal with them as if they were dealing with other people (regardless of the type of agent interface that is being used) [168]. Agents that fit the stronger notion of agent usually have one or more of the following characteristics:

- ***Mobility***: the ability of an agent to move around an electronic network.
- ***Benevolence***: the assumption that agents do not have conflicting goals, and that every agent will therefore always try to do what is asked of it.
- ***Rationality***: is (crudely) the assumption that an agent will act in order to achieve its goals and will not act in such a way as to prevent its goals being achieved – at least insofar as its beliefs permit.
- ***Adaptivity***: an agent should be able to adjust itself to habits, working methods and preferences of its user.
- ***Collaboration***: an agent should always validate instructions to ensure they conform to all conditions, goals, and constraints within a system (e.g., instructions that contains conflicting goals), omits important information and/or provides ambiguous information. For instance, an agent should interact with the external environment to validate assumptions such as asking questions, or build an internal model to solve problems. An agent should even be allowed to refuse to execute certain tasks, because (for instance) if they would put an unacceptable high load on the network resources or because it would cause damage to other users. Obviously, prime directives can be given that drive decisions on ultimatums.

Although no single agent possesses all these abilities, SELF Cognitrons are designed to capture all of them. At this moment no consensus has yet been reached about the relative importance (weight) of each of these characteristics in the agent as a whole. What most researchers have come to a consensus about is that it is these kinds of characteristics that distinguish agents from ordinary programs.

8.5.4 Software Agents and Intelligence

The degree of autonomy and authority vested in a software agent is called its agency. It can be measured at least qualitatively by the nature of the interaction between the agent and other entities in the system in which it operates.

At a minimum, an intelligent software agent must run asynchronously. The degree of agency is enhanced if an agent represents another entity (internal or external) in some way. This is one of the key values of intelligent software agents. A more advanced agent can interact with other entities such as data, applications, or services. Further advanced agents collaborate and negotiate with other agents.

What exactly makes an agent "intelligent" is something that is hard to define. It has been the subject of many discussions in the field of Artificial Intelligence, and a clear answer has yet to be found. For a SELF, we define intelligence as the degree of reasoning and learned behavior; the Cognitrons ability to accept a set of goals and directives, and then carry the tasks delegated to it, within its given constraints. At a minimum, there must be some statement of preferences or rules, with a Fuzzy Inference Engine to provide a reasoning mechanist to act on the rules. For Reasoner and Analyst Cognitrons, it includes a model of understanding and reasoning about what is to be done. The ability to autonomously plan means to achieve its goals, as well as the ability to learn and adapt to changing environments, both internal and external. This adaptation is both in terms of ultimate objectives, and in terms of the resources available to it (Cognitive Economy and Locus of Control described earlier). SELF Cognitrons, similar to a human assistant, discover new relationships, connections, and concepts independently, and exploit these in anticipation of future tasks.

8.5.5 The Cognitron

The architecture for the Cognitron is based on the metaphor of software agents and incorporates techniques from other research fields such as distributed architectures, relevance feedback and active interfaces. When configuring Cognitrons for a SELF, the overall aim is for the system to be suitable for different types of environments. This is with regard to local and external searches for information and data.

One single Cognitron, called the **Information Agent**, is used as the interface between a SELF and its environment. The Info Agent, in its turn, uses an **Interface Cognitron** for handling the communication with a SELF's sensors and external environment. This Cognitron is like a personal assistant who is responsible for handling user needs, and for the connection of the user with the agent(s) that will help the Cognitrons solve their problems or tasks. The number of types of Cognitrons the Interface Cognitron has to deal with depends on the current goals of a SELF. As a result of the distributed and agent-based architecture of the system the whole structure of it can be easily changed or updated by the Artificial Prefrontal Cortex to accommodate changes.

The Interface Cognitron (IC) is able to reason about requests and to understand what type of need requests are expressing: IC singles out which of the other Cognitrons in the system is able and necessary to solve the current problem. Two other agents are **Internal Services Cognitron** and **External Retrieval Cognitron** shown in Fig. 8.11.

The Internal Services Cognitron knows the structure of the SELF memories and which are applicable in a given situation: it is in charge of retrieving data and information.

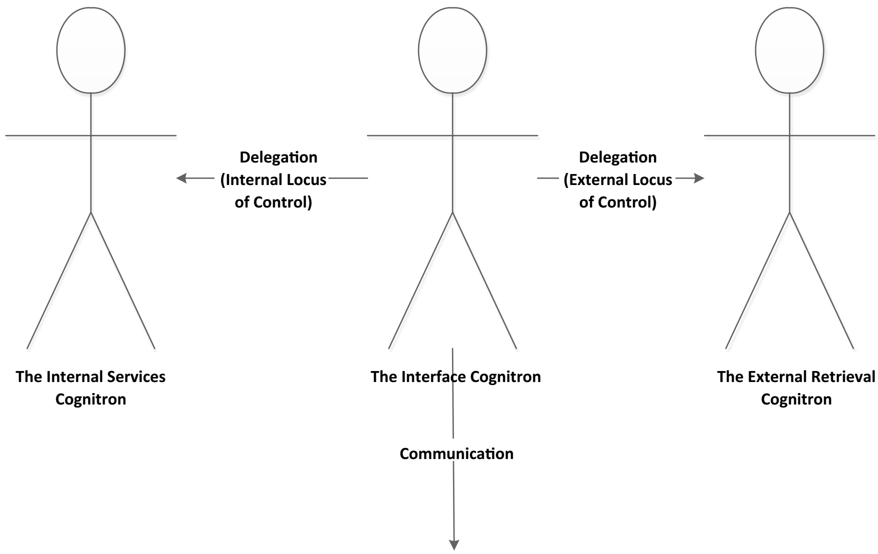


Fig. 8.11 High level structure for the information cognitron

The External Retrieval Cognitron is in charge of retrieving information from external sources (Locus of Control). It can work in two modalities: retrieval (or query) mode and surfing mode. In the first case, it searches for specific information: this service is activated by a direct internal SELF request. In the second case, the Cognitron navigates the external environment searching for information that, in its opinion, could interest a SELF. The search is driven by a SELF profile built and maintained by the Interface Agent

Refinement of this profile takes place according to how a SELF's cognitive manages the data that the Cognitron finds for and/or proposes to a SELF. Using a SELF's profile, the Interface Cognitron charges specialized agents to navigate through the memories or external environments, hunting for information that could be of some interest for a SELF. In this way, a SELF can be alerted when new data that can concern its interest area(s) appear. The Interface Cognitron performs the following tasks for a SELF's cognitive framework:

- Assists Cognitrons in performing requests and compiling profiles for certain types of information needs: The other Cognitrons do not need to be aware of what information is available, how the information is structured or organized, where the information is located. This is the responsibility of the Interface Cognitron.
- Deduces the Cognitrons informational needs by both communicating with the Cognitrons and learning their behavior: The Interface Agents observe the other Cognitrons behavior and the current state of its environments (internal and external – determining the "Locus of Control") to deduce what actions are to be performed and how to modify the current profiles to adapt as environments change (a change in the Cognitron's Locus of Control).

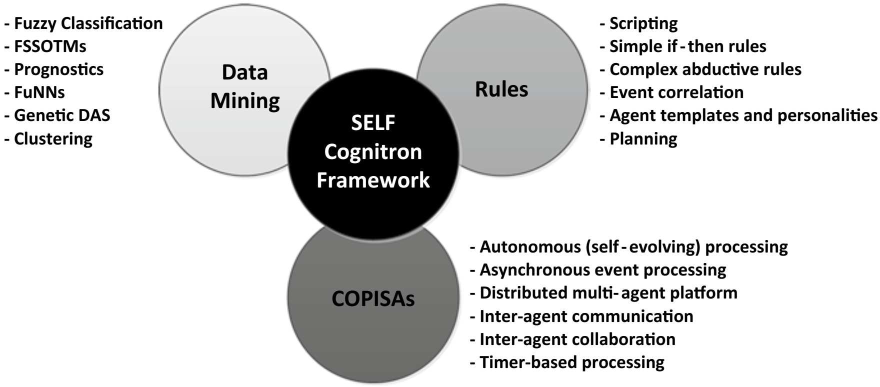


Fig. 8.12 A SELF high-level cognitron framework

- Translating the Cognitron requests and selecting other Cognitrons needed to solve the current set of problems or tasks: This allows the Cognitrons to ignore to concentrate on solving their tasks; allowing the Interface Cognitron to provide them with what they need to find solutions.

In general, the Cognitron framework is a Java framework for providing business rules and run policies for the Cognitrons, incorporating the necessary strategies into each type of and specific Cognitron. This includes an autonomous computing core technology within a SELF multi-Cognitron framework. Figure 8.12 illustrates a high-level Cognitron framework.

As explained earlier, SELF Cognitrons are active, persistent software components (codelets) that perceive, reason, act, and communicate. Figure 8.13 below illustrates an abstract architecture for the Cognitron framework.

Cognitrons learn from experience and can be used to predict future states (prognostics). They are able to analyze sensor data using classification and clustering techniques to detect complex states and diagnose problems (anomaly detection and resolution). They reason using domain-specific application objects and have autonomous (proactive) behavior and goals. They can correlate events to situations, reasons, and take actions. All of these abilities are based upon a Cognitron component library (examples illustrated in Fig. 8.14), which provides strategies (algorithms) utilized to provide Cognitrons with the required abilities for each Cognitron archetype (e.g., analyst, reasoner, etc.). Figure 8.15 illustrates the Cognitron Rules Architecture that drives a SELF.

8.6 Knowledge Relativity and Reasoning

Research shows that generating new knowledge is accomplished via natural human means: mental insights, scientific inquiry process, sensing, actions, and experiences, while context is information, which characterizes the knowledge and gives it meaning.

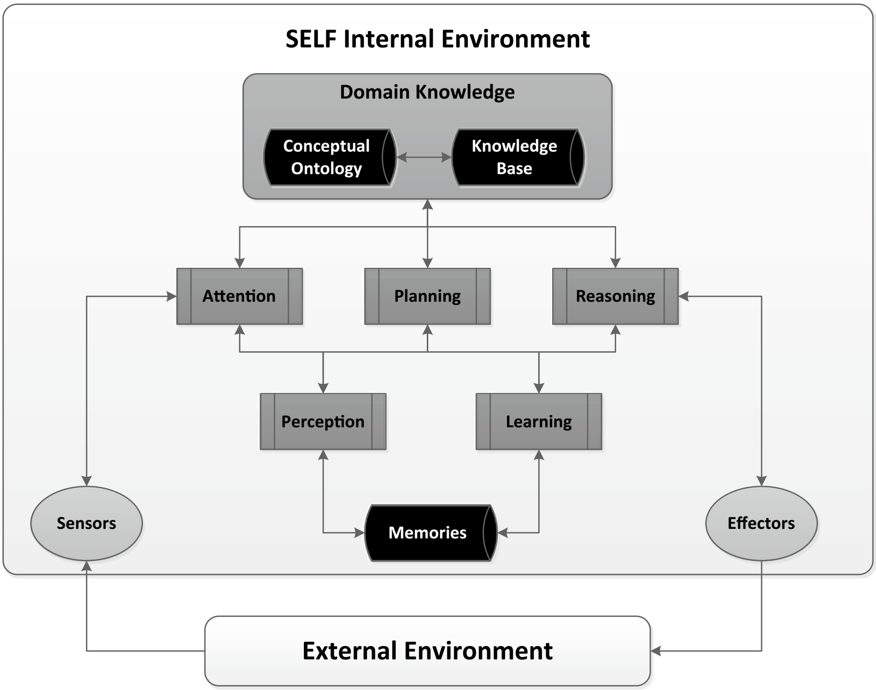


Fig. 8.13 A SELF cognitron abstract architecture

This knowledge is acquired via the focused development of an established set of criteria, approaches, designs, and analysis, as inputs into potential solutions [69]. The challenge within a SELF is to minimize ambiguity and fuzziness in understanding large volumes of complex interrelated information content via integration of two cognition based frameworks. The objective is improving actionable decisions using a Recombinant Knowledge Assimilation (RNA) framework [229] integrated with a SELFs Artificial Cognitive Neural Framework (ACNF) [64] described earlier. We describe a SELF's ability to recombine and assimilate knowledge based upon human cognitive processes which are formulated and embedded in a SELF's recombinant neural fiber network of genetic algorithms and stochastic decision making towards minimizing ambiguity and maximizing clarity [181].

8.6.1 Knowledge Relativity

Nonaka and Takeuchi [180], when describing how Japanese companies innovate as knowledge creating organizations, described two types of knowledge: tacit and explicit. Tacit knowledge is personal and context-specific. Explicit knowledge is knowledge codified in books, journals and other documents for transmittal.

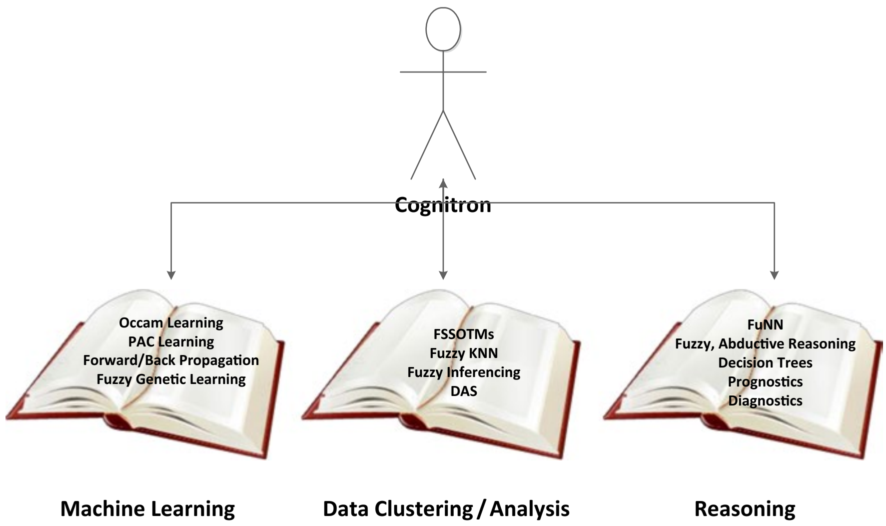


Fig. 8.14 The SELF cognitron component library examples

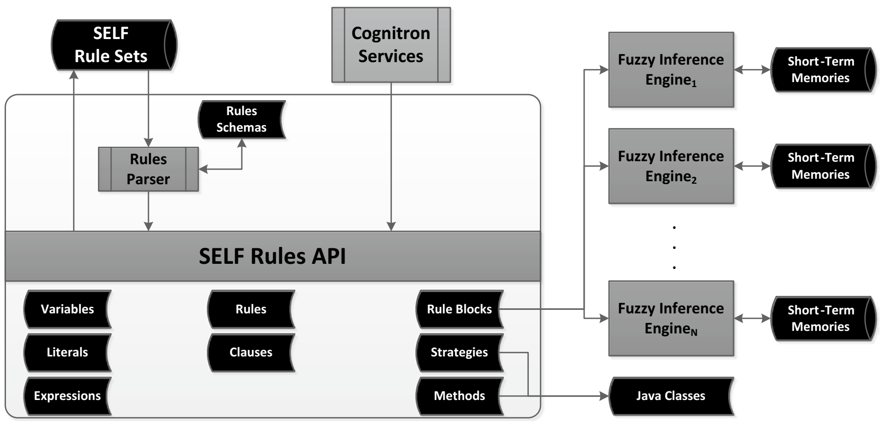


Fig. 8.15 A SELF cognitron rules architecture

Additionally, Nonaka [181] prescribed how dynamic organizational creation of knowledge needs to be strategically collected, understood, and managed across the entire company's organizational structure as intellectual capital. Knowledge theorist Polanyi and Sen [188], in describing what he called the "Tacit Dimension," used the idea of tacit knowledge to solve Plato's "Meno's paradox," that deals with the view that the search for knowledge is absurd, since you either already know it or you don't know what you are looking for, whereby you cannot expect to find it.

The author argued that if tacit knowledge was a part of knowledge then "we do know what to look for and we also have an idea of what else we want to know," therefore personal and context-specific knowledge must be included in the formalization of all knowledge [61].

Renowned fuzzy logic theorist Zadeh [219], described tacit knowledge as world knowledge that humans retain from experiences and education, and concluded that current search engines with their remarkable capabilities do not have the capability of deduction, that is the capability to synthesize answers from bodies of information which reside in various parts of a knowledge base. More specifically Zadeh, describes fuzzy logic as a formalization of human capabilities: the capability to converse, reason and make rational decisions in an environment of imprecision, uncertainty, and incompleteness of information.

Underlying decision-making based on informational inferences is a great concern, for informational ambiguity and the ramifications of erroneous inferences can be catastrophic. Often there can be serious consequences when actions are taken based upon incorrect recommendations and those can influence decision-making before the inaccurate inferences can be detected and/or even corrected. This is particularly a problem in intelligence processing. Underlying the data fusion domain is the challenge of creating actionable knowledge from information content harnessed from an environment of vast, exponentially growing structured and unstructured sources of rich complex interrelated cross-domain data.

Therefore we employ a Recombinant kNowledge Assimilation (RNA) instructional derivation, which extends space-time mechanics to provide mathematical relationships for n-dimensional context between knowledge objects [159]. Newell and Simon [177, 178] developed models of human mental processes and produced General Problem Solver (GPS) to perform "means-end analysis" to solve problems by successively reducing the difference between a present condition and the end goal. GPS organized knowledge into symbolic objects and related contextual information, which were systematically stored and compared. RNA is used within a SELF ACNF to provide context and knowledge relativity meta-information within SELF analysis, reasoning, and reporting; aiding in the Cognitive Intelligence for a SELF, and facilitating a SELF's top-down executive processing required for real-time cognitive reasoning.

The RNA theory below describes math and constructs which can be used to analyze and process knowledge and context, representing context in a knowledge management framework, comprising processes, collection, preprocessing, integration, modeling and representation, enabling the transition from data, information and knowledge to new knowledge [71]. The use of RNA provides knowledge threads for newly generated knowledge, storing memory context and relevance information in a context knowledge base and to be used by a SELF rule-based context knowledge-matching engine to support decision-making activities. Gupta and Govindarajan [124, 125] defined a theoretical knowledge framework and measured the collected increase of knowledge flow out of multinational corporations based upon "knowledge stock" (e.g., the value placed upon the source of knowledge). Pinto [187] developed a conceptual and methodological framework to represent the quality of

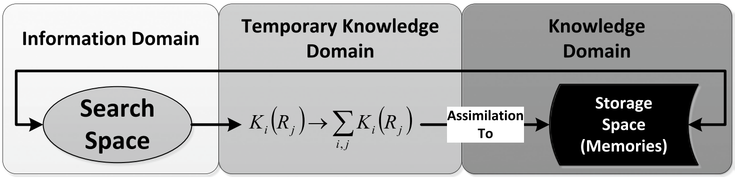


Fig. 8.16 Recombinant kNowledge Assimilation (RNA) equation

knowledge found in abstracts. Suh [200] concluded that collaborative frameworks do not provide the contents which go in them, therefore, content was discipline specific, required subject matter experts, and clear decision making criteria. Additionally, Suh noted that processes promoting positive collaboration and negotiation were required to achieve the best knowledge available, and were characterized by process variables and part of what is defined as the Process Domain. The creation of SELF Knowledge Relativity Threads (KRTs) [229] was developed to provide a framework for knowledge and context, which collected and stored the knowledge as mathematical relationships along with semantic context as decisions in a knowledge repository that corresponded to a specific context instance.

Knowledge relativity threads are needed because current databases housing vast bits of information, do not store the information content of the reasoning context used to determine their storage [109]. The knowledge collection and storage formula was therefore developed to include and store relationship context along with knowledge, recursively. This means that, each act of knowledge and context pairing shown as in equation:

$$\sum_{i,j} K_i\left(R_j\right) \tag{8.3}$$

and shown in Fig. 8.16, recursively examined all of the previous relationships as they were recombined into storage since they were all related and dependent on each other. Next, recursive refinement occurs, per iteration of relationship pairing. Recursive refinement occurs at the instance it is determined that what was found is what was looked for, shown as $K_i(R_j)$, using interrogatives, (e.g. who, what when, where, why and how) [133, 135, 136]. The information content contributing to finding the answer then has significant value and therefore, a higher degree of permanence in the mind of the stakeholder [6]. Therefore, the information content has reached a threshold where retaining the knowledge and context has become important.

8.6.2 *Knowledge Relativity Threads*

Figure 8.17 represents a Knowledge Relativity Thread (KRT). This approach for presentation of knowledge and context and was constructed to present five discrete attributes, namely, time, state, relationship distance, relationship value, and event sequence. The goal of a KRT is to map the dependencies of knowledge and related

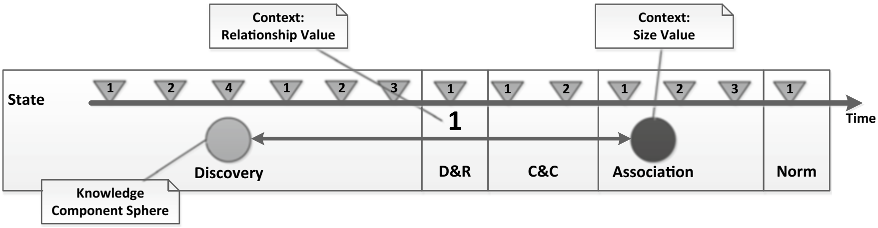


Fig. 8.17 A SELF RNA knowledge relativity threads

attributes as knowledge is developed from information content. In this figure, the timeline represented by the arrow from left to right, shows the events or state transitions in sequence and captures the decision points. During each of the iterations of the presentation of knowledge and context, intrinsic values were captured and placed close to each knowledge component [70]. In Fig. 8.17, these are represented as information fragments under the cycles. The Basic Information Decomposition depicts how a KRT looks when it represents information decomposed into pieces; in this case fragments. The triangles, depict a particular state for each of the iterations, in the KRT development cycle. For emphasis, each sphere was built into the depiction and added in sequence to represent the fact that each information fragment develops over time and follows the others. Each icon represents each information fragment. The relative values in this Basic Knowledge Decomposition between each sphere are perceived to be of the same value to each other. Therefore, the lines are the same distance as well. Since, this base representation depicted in Fig. 8.17 can present time, state, and sequence, as well as, relationships, the challenge was addressed as described by Dourish [104] to create presentation of context which can visually capture and manage a continually renegotiation and redefinition of context as development of knowledge occurs over time.

The KRT depicts cognitive comparison of not just information, but of the contextual relationships as well. An important distinction about the observation of each comparison is that each is made from the perspective of all previously aggregated information, knowledge, and context.

The representation of knowledge and context formula is introduced here and is presented by Eq. 8.4. The independent results which follow are mathematical evaluations extended from Newton's law of gravitation shown in Eq. 8.4. Newton's Law of Gravitation formula is:

$$F = G\frac{(M_1 M_2)}{r^2} \tag{8.4}$$

where:

F is the magnitude of the gravitational force between the two objects with mass,
G is the universal gravitational constant,
M_1 is the mass of the first mass,

M_2 is the mass of the second mass, and
r is the distance between the two masses.

This equation was used as an analogy for the derivation of mathematical relationship between a bases made up of two objects of knowledge. Abstracting Newton's Law of Gravitation as an analogy of Eq. 8.1, representing relationships between two objects of knowledge using context, is written as Eq. 8.5 shown below, which describes the components of the formula for representing relationships between two objects of knowledge using context

$$A = B\frac{(I_1 I_2)}{c^2} \tag{8.5}$$

where:

A is the magnitude of the attractive force between the two objects of knowledge,
B is a balance variable,
I_1 is the importance measure of the first object of knowledge,
I_2 is the importance measure of the second object of knowledge, and
c is the closeness between the two objects of knowledge

Comparing the parameters of Eqs. 8.4 and 8.5 F and A have similar connotations except F represents a force between two physical objects of mass M_1 and M_2 and A represents a stakeholder magnitude of attractive force based upon stakeholder determined importance measure factors called I_1, and I_2. As an analogy to F in Eq. 8.4, A's strength or weakness of attraction force was also determined by the magnitude of the value. Hence, the greater the magnitude value, the greater the force of attraction and vice versa. The weighted factors represented the importance of the information fragments to the relationships being formed. The Universal Gravitational Constant G is used to balance gravitational equations based upon the physical units of measurement (e.g. SI units, Planck units). B represents an analogy to G's concept of a balance variable and is referred to as a constant of proportionality. For simplicity, no units of measure were used within Eq. 8.5 and the values for all variables only showed magnitude and don't represent physical properties (e.g. mass, weight) as does G. Therefore, an assumption made here is to set B to the value of 1:

For simplicity, all of these examples assume the same units and B was assumed to be one. The parameter c in Eq. 8.5 is taken to be analogous to r in Eq. 8.4. Stakeholder perceived context known as closeness c represented how closely two Knowledge Objects (KO) or information fragments are related. Lines with arrows are used to present the closeness of the relationships between two pieces of knowledge presented as spheroids (see Fig. 8.18).

Using Eq. 8.2, the value of the attraction force $A_{1\rightarrow 2} = 5 \times 2$ divided by the relative closeness/perceived distance$^2 = 1$. Hence, the attraction force A in either direction was 10. The value of 10 is context which can be interpreted in relation to the scale. The largest possible value for attraction force A with the assumed important measure 1.10 scale is 100, therefore a force of attraction value of 10 was relatively small

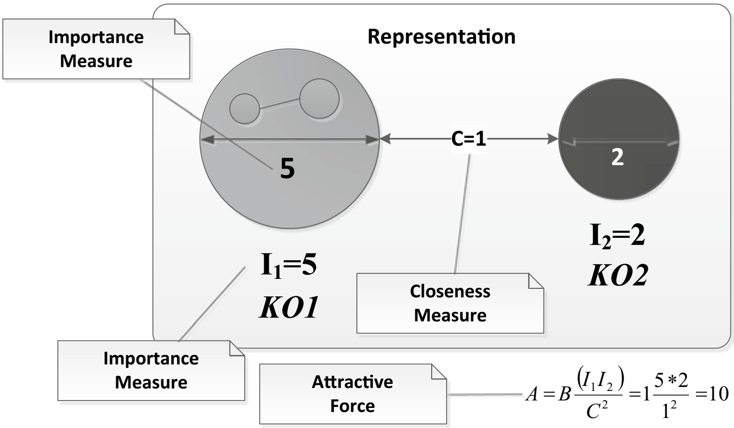


Fig. 8.18 SELF representation of knowledge object and context

compared to the maximum. This means that the next stakeholder/ researcher understood that a previous stakeholder's conveyance was of small relative overall importance. However, the closeness value of 1 showed that the two objects were very closely related. Figure 8.18 therefore shows that when using Eq. 8.5, if relationship closeness and/or perceived importance measure of the knowledge objects change value, as new knowledge or context is added and evaluated, then it follows that relationship force of attraction will change.

8.6.3 Frameworks for Contextual Knowledge Refinement

As the knowledge and context foundation described above depicts the process and tools for enhancing knowledge and context the Artificial Cognitive Neural Framework (ACNF) is utilized within a SELF to apply additional refinement concepts and another formalization for the modular Decomposition and Reduction and Association sub-processes described in the RNA above.

Here we refer again to the ACNF discussed earlier. The Mediator (Artificial Prefrontal Cortex) gathers information and facilitates communication between agents. Hence, each decision handshake of a combined RNA-ACNF system is handled by the Mediator which takes information from Cognitrons and from coalitions of Cognitrons and updates the short-term, long-term and episodic memories or pedigree [12]. The information available in memory (what the system has learned) is continually broadcast to the conscious Cognitrons that form the cognitive center of the system (i.e., they are responsible for the cognitive functionality of perception, consciousness, emotions, processing [93])

As described, the ACNF contains several different artificial memory systems (including emotional memories), each with specific purposes. Each of these memory systems are stored pedigree used in the recursive RNA process and are integrated during the processes of relationship formation between objects of knowledge and context [57].

When processing a SELF's pedigree memory, RNA loosely categorizes the granularity of information content into knowledge and context based upon the criteria established by the cognitive human interaction input into the system. These loosely or fuzzy categories are only as fuzzy as the threshold of human understanding. Therefore, in order to artificially create this effect we use a SELF's Cognitrons to develop fuzzy organization over time, ultimately reaching a threshold of perceived understanding relative to the initially specified set of criteria.

As we push to process, analyze and correlate more and more information, the need to combine contextual relevance with information is ever more necessary. Information without context is just that, devoid of real content. Instead, the systematic approach presented here, combining the RNA contextual approach, with a cognitive framework, in the ACNF, provides the framework that can handle cognitive processing of information and context, turning them into actionable intelligence. The use of Knowledge Relativity Threads represents the next generation of information analysis and will greatly enhance the capabilities of information processing systems to make sense of increasing volumes multivariate, heterogeneous information [108].

8.7 Knowledge Density Mapping Within a SELF

Here we describe structures and mathematical derivation for structures within a SELF to assess a SELF's Inference Potential. This Inference Potential is determined from providing a measure of the Knowledge Density and Analytical Competency of the information processing systems, based on the contextual assessment of the question, or topic posed by the operator and analyst. The use of Knowledge Density and Analytical Competency to determine a SELF's systems Inference Potential provides the methodologies to radically improve the performance and quality of SELF's cognitive processing by allowing a SELF to "self-analyze" its ability to answer questions and perform the analysis and inferences on incoming data/information.

In order to mimic real-time human decision making processes, a SELF's cognitive processing systems must be supported by information derived from the fusion process and must operate in a uniform and cooperative model, fusing data into information and knowledge so information a SELF can make informed decisions [216]. One such construct that aids a SELF is the measure of a system's ability to provide quality information and/or inference about a particular subject or question posted by Dialectic Argument Structure's hypothesis generators. Described here is the mathematical derivation and development of a SELF's ***Inference Potential***. This *Inference Potential* is determined from providing a measure of the ***Knowledge Density*** and ***Analytical Competency*** of a SELF, based on the contextual assessment of the question, or topic posed by a SELF's internal

Locus of Control system. Such a measure allows a SELF's Artificial Prefrontal Cortex to quickly understand the system's ability to provide quality knowledge about a subject, question, or topic, and could be used to discover knowledge holes or gaps in a SELF. ***Knowledge Density Mapping*** facilitates information, intelligence, and memory integration, and allows faster accommodation of knowledge and knowledge characteristics. The *Analytical Competency* measure provides analysis, reasoning, and reporting capabilities of a SELF's capabilities (provides cognitive intelligence).

8.7.1 Knowledge Density Mapping: Pathway to SELF Metacognition

In order for a SELF to be truly autonomous, we must provide a SELF with the ability to understand its own limitations and capabilities and to reason about them in light of the duties or missions it is given. In humans, we call this ability "Metacognition." As described earlier, metacognition in humans refers to higher order thinking which involves active control over a SELF's cognitive processes engaged in learning and performing. Activities such as planning how to approach a given task, monitoring comprehension, and evaluating progress toward the completion of a task are metacognitive in nature [42, 202].

In order for a SELF to achieve the metacognitive abilities within the ACNF, a SELF must have the ability to measure its own knowledge about a particular topic or subject [79]. This measure of topical or subject knowledge involves measuring the "density" of knowledge the system possesses about this subject or topic in question. This ***Knowledge Density*** measure is based on the number of separate information fragments relative to the taxonomy of the topic or subject. Figure 8.19 provides the Knowledge Density Measure, based on separable topical information fragments [61]. In order to provide the parameters required to compute Knowledge Density, cognitive maps [8, 153] track separable information fragments by topic, as illustrated in Fig. 8.20.

We use knowledge fragment measurements to ensure that we only store information relative to a topic or subject once. Information that is taken in is parsed and information fragments that have not been stored before are pulled out and stored in a cognitive map for that topic. Renyi's entropy measurement is utilized to separate information into topical information fragments.

Computationally, this is difficult, however, Renyi's measure, combined with the Parzen Density estimation method provides a computational model. We start by looking at the information densities, *p(y)*, as a sum of related topical cognitive maps, each centered at y_i, we get:

$$p(y) = \frac{1}{N}\sum_{i=1}^{N} G\left(y - y_i, \sigma I\right) \tag{8.6}$$

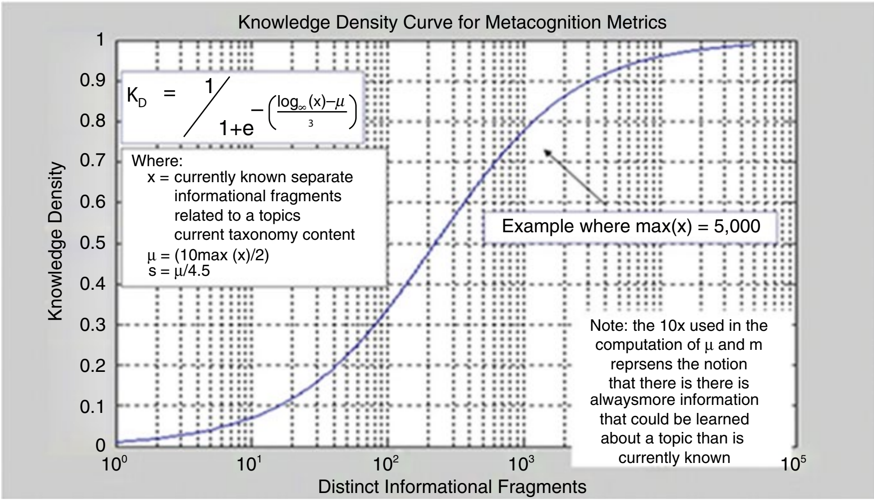


Fig. 8.19 SELF knowledge density calculation

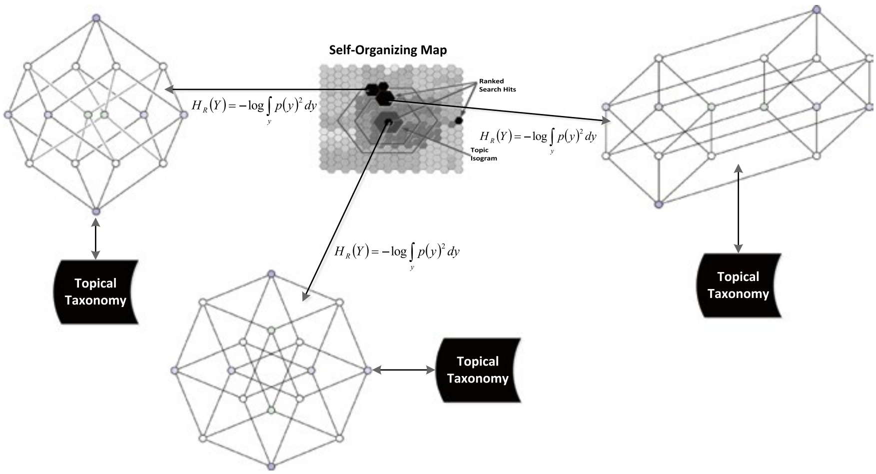


Fig. 8.20 SELF knowledge density mapping

Renyi's entropy measurement is defined as [57]:

$$H_R(Y) = -\log \int_y p(y)^2 \, dy \tag{8.7}$$

Therefore, Renyi's entropy can be computed as the sum of local information interactions (separate information fragments) over all pairs of informational entities. Informational associations are created within the Cognitive Topical Maps utilizing the FUNN described earlier and a SELF's Fuzzy Inference Engine, based on Renyi's theoretics. We use this possibilistic network because:

- It's robust in the presence of inexact information.
- It utilizes conditional possibilistics
  - Mutual Information measurement
  - Joint Informational membership rather than joint probabilities
- Excellent at showing qualitative relationships not attainable with Bayesian methods
  - Excellent at showing qualitative relationships not attainable with Bayesian methods
  - Creates decisions with conditional possibilistic attributes
- More useful with general questions about a subject domain

This methodology allows a SELF's Cognitive Topical Maps to be populated with separable information fragments, relative to a topic that maps to the topical taxonomy. This allows a measurement of the density of knowledge a system contains, relative to a topic or subject. Within the Knowledge Density computations, a SELF's FSSOTMs are used to measure topics and how other topics relate. Knowledge Density is a measure of the density of knowledge a system has about a topic and the density of related topics that would be used to answer questions and/or analyze situations. The next piece of the Inference Potential computation is Analytical Competence, or, what is the competency of the current cognitive system, based on the current problems to be solved.

8.7.2 Analytical Competency

In order to quantifiably measure a SELF's ***Inference Potential***, a SELF's internal assessment processes must be able to assess its ability to analyze information relative to a question or mission posed to it. We call this measure of analytical potential ***Analytical Competency***. The Analytical Competency measures relative to a topic or subject are based on the algorithms and software that are available:

- The algorithm technical capability-what it was designed to do
- The algorithms experiences – tied to emotional memory [156]
- The algorithms body of knowledge – what it has learned

Analytical Competency is tied to "***Areas of Expertise***" within the AI system. Figure 8.21 illustrates the information flow for the ***Analytical Competency*** measure.

Respectively, the ACNF and the Cognitron coalitions become emotionally aroused when they form semantic and episodic memories about situations that cause "stress" within an artificial neural system. Stress situations may involve a loss of resources, new data environments that are unfamiliar, new interfaces that are introduced into the environment or situations where the algorithms produced incorrect results. These cognitive representations of emotional situations better referred to as memories about emotions rather than emotional memories.

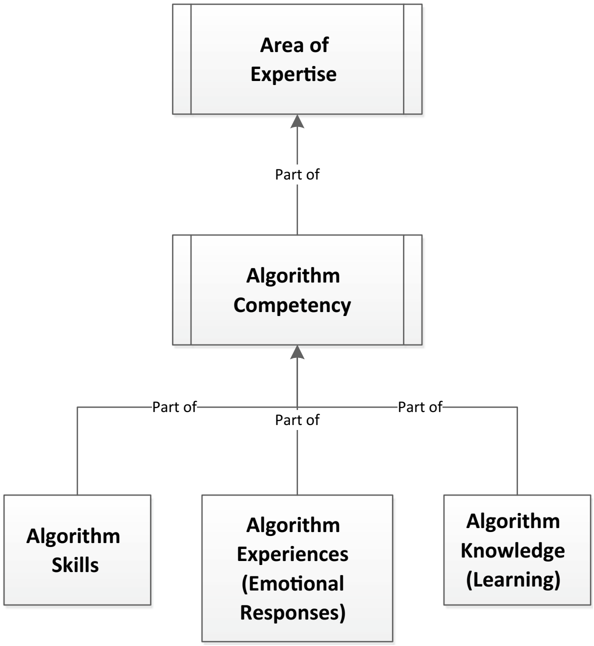


Fig. 8.21 SELF analytical competence measurement model

The effects of emotional arousal on explicit memory are due to processes that are secondary to the activation of emotional processing systems in the ACNF [69]. These emotional responses or emotional memories within the algorithmic long-term memories provide vital information that relates to how these algorithms have been able to respond or not respond to given assignments, topical analysis, or missions that have been assigned to a SELF.

Activity in these areas would be detected by the cognitive coalitions and would lead to increases in system emotional arousal (due to activation of modulation within the neural structure that leads to the release of cognitive problem, solution, search, and emotion agents [168]. These responses are stored and utilized, in part, as a measure of the analytical competency of a set of algorithms that make up an area of expertise within a SELF. The transmittal of informational content as well as emotional context allows information retrieval performance to be greatly enhanced, allowing for "cognitive economy" within the artificial neural system [60]. The ***Analytical Competency*** measure is based on inputs to the areas shown in Fig. 8.21, illustrated in Fig. 8.22 [64].

The actual Analytical Competency measurement is computed as:

$$AC_t = \frac{\sum_{i=1}^{n} w_i \sqrt{L_i^2 + A_i^2}}{n}, \text{ where} \tag{8.8}$$

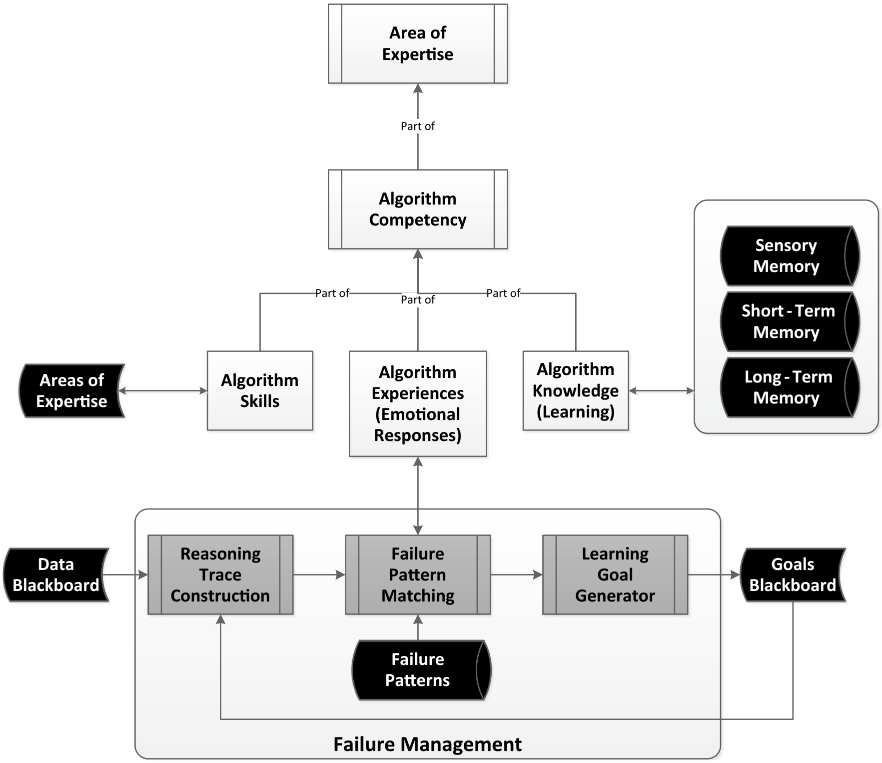


Fig. 8.22 SELF analytical competence measurement inputs

$$w_i = \sqrt{\frac{\sum_{j=1}^{m} (\pm) E_j^{\,2}}{m}}, \text{ where } E_j = \text{emotional memory response} \tag{8.9}$$

$$L_i = \frac{\sum_{k=1}^{p} A_{E_k} C_k}{p}, \text{ where } A_{E_k} = \text{memories for area of expertise k and} \quad C_k = \text{completeness of memory} \tag{8.10}$$

$$A_i = \frac{\sum_{l=1}^{r} \mathrm{A}_{\mathrm{R}_l}}{r}, \text{ where } \mathrm{A}_{\mathrm{R}_l} = \text{Algorithm relevancy for lth algorithm} \tag{8.11}$$

The result is a rating from 0 to 1 of the Analytical Competency for a SELF, based question or mission posed, or an internal issue to be resolved. Once the Knowledge Density and Analytical Competency have been computed, the overall Inference Potential of the system for a given topic/subject/mission is:

$$IP = KD * AC \tag{8.12}$$

Producing a number between 0 and 1, where 0 means the system has no potential to produce a useful inference for the topic requested and 1 indicates that not only can the system produce useful inferences, but that the inferences will be useful and trustworthy.

8.8 Discussion

We have described the SELF's processes for creating, storing, and constructing memories, learning, reasoning, an inferring information and knowledge throughout its ACNF. At a higher level, the SELF requires an overall cognitive framework that provides management and control of its cognitive processes. Chapter 9 will introduce this overall cognitive framework that will facilitate synthetic consciousness within the SELF, called Intelligent information Software Agents to facilitate Artificial Consciousness (ISAAC). This cognitive framework utilizes the SELF's Cognitrons and organizes the SELF's processes into sub-groups that mimic the human brain organization (e.g., Neocortex).

Chapter 9
Artificial Cognitive System Architectures

Our proposed ACNF, discussed earlier in the book, provides an outline for a possibilistic architecture that can facilitate cognition, learning, memories, and information processing, but it is not solely sufficient to create a comprehensive, autonomous SELF. An overall SELF architecture framework, along with both a knowledge and cognitive framework are required in order to facilitate our fully autonomous, cognitive, self-aware, self-assessing, SELF. We have discussed a SELF system for cognitive management, PENLPE, now we will look at an overall cognitive processing framework, called the ***Intelligent information Software Agents to facilitate Artificial Consciousness*** (ISAAC). A SELF architecture, allows dynamic adaptation of the structural elements of the cognitive system, providing abilities to add and prune cognitive elements as necessary as part of SELF evolution [54]. The overall architecture also accommodates a variety of memory classes and algorithmic methods. The basic building blocks of ISAAC comprise an ACNF framework, Cognitron architecture, Fuzzy, Self-Organizing, Semantic Topical Maps (FUSE-SEMs), and a comprehensive Abductive Neural Processing system, the Possibilistic Abductive Neural Network (PANN), for providing consciousness and SELF cognitive functions. Within an ISAAC framework, Cognitrons are added or deleted from the system, based upon the complexity of the classes of information processed. This chapter expounds upon background and architecture for ISAAC, as well as, human-SELF interaction and collaboration, Cognitive, Interactive Training Environment (CITE).

The ISAAC framework comprises real-time, time-varying creation and destruction of neural structures within a SELF and takes into account problems that have been encountered in the past with earlier, less sophisticated neural systems. The vast majority of systems built today are simpletons which learn and store absolutely no pedigree which could allow them to operate more efficiently. More specifically to learning systems, many earlier neural systems and components based upon the principle of time-varying neural structures tended to also forget previously learned neural mappings as they were exposed to new types of environments. This is a phenomenon known as "Catastrophic Interference" (CI). Previous attempts to alleviate

J.A. Crowder et al., *Artificial Cognition Architectures*,
DOI 10.1007/978-1-4614-8072-3_9, © Springer Science+Business Media New York 2014

this problem by utilizing networks and systems with localized processing responses, where neural structures were added and modified at the local level, not global, lead to systems with unbounded growth. This was primarily due to a lack of effective pruning mechanisms due in part to only producing the real-time ‘local view’ of the neural structures. To alleviate possibilities of Catastrophic Interference we propose a very modular SELF architecture based upon a mixture of hybrid neural structures adding elasticity and diversity to SELF capabilities. Hence, the ISAAC architecture discussed here comprises a flexible, continuously adaptable hybrid neural processing system with functions for dynamically adding and pruning basic cognitive and neural building blocks as real-time needs of the system change.

A foundational component of an ISAAC processing framework is the concept of “mixture of experts” architecture and methodology, similar to a human brain. The human brain possesses different “specialty” areas used to process different types of ingested information (e.g. auditory, visual, and tactile). The difference within ISAAC is that here, each expert is defined as a fuzzy, synthetic Cognitron object created for each particular algorithm or processing object, and thus is a synthetic ‘expert’ designed for learning and processing a particular type of information in a particular matter. Hence, the information algorithm for which each Fuzzy, Genetic Cognitron (FGC) is generated can be predetermined, or evolves by it-SELF.

9.1 Cognitronic Artificial Consciousness Architecture

9.1.1 Synthetic Neocortex Adaptation

Since the SELF is not a biological entity, but is, instead, made up of hardware and software, we define a Synthetic Neocortex Theory to facilitate human-like neurobiological functions within a SELF cognitive framework. ISAAC’s artificial cognitive controller is an electronic entity where, as described in Chap. 4, where information is stored as Binary Information Fragments (BIFs) within elementary memory locations, known as synthetic synapses, within the various SELF system memories (e.g., short-term, long-term, emotional, etc.). Each synthetic memory synapse is a small unit of memory with complex read-write functionality [140]. The BIFs are stored in synthetically transmitted neural abductive nodes, whose resolution ranges from 1 to several dozen bits/synthetic synapse (depending upon the precision needed). Analogously, within a synthetic SELF hardware/software driven artificial entity, information is, of course, read and written using voltage pulses. Read voltages are analogously provided by the arrival of pre-synaptic spikes, initiated by the Artificial Prefrontal Cortex (APC) actuated via a SELF’s Cognitive Knowledge framework, driven by FUSE-SEM outputs. The release of synthetic neurotransmitters generates a post-synaptic response, which is proportional to the number of transmitter channels within a respective memory’s abductive synthetic synapse. Summed post-synaptic abductive neural responses can trigger a delayed

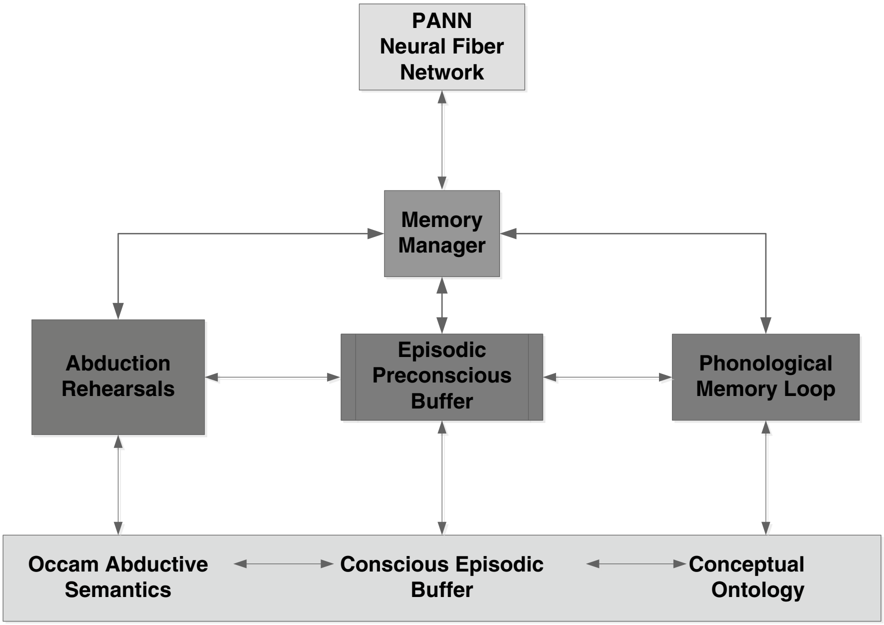


Fig. 9.1 ISAAC's memory synaptic information flow

write pulse, depending upon Cognitron needs and timing information. Pulses are in the form of back-propagation pulses, which initiate small synthetic synaptic strength relation interrupts (Hebb Rule) [130], and are used to drive the adaptation of synthetic neurons during the SELF learning process. Over time, synthetic, abductive neural synapses, improve their predictive abilities, with respect to correlation of output-to-input spikes. Figure 9.1 illustrates ISAAC's short-term memory information flow for synthetic synaptic evolution.

Based upon Neural Continuum Theory, discussed in Chap. 2, there exists a fundamental limit of information quantity, read and written within SELF memories (true with any resource-limited system), As in the brain, ISAAC's artificial abductive network is monitored by synthetic Neocortex functions, which monitor the abductive network for stochastic neural synaptic leaking and appropriately strengthens and/or "de-strengthens" true and erroneous stochastically strengthened synthetic neural synapses. This process allows unsupervised learning and evolving within the SELF's cognitive systems, minimizing the potential for synaptic catastrophes. In this way, the synthetic Neocortex acts as a SELF's memory proofreader; independently monitoring and measuring the synthetic synaptic plasticity of the synthetic abductive neural synapses. Proofreading helps ensure application of synaptic strengthening to the correct pre and postsynaptic responses, and helps ensure application to the correct synthetic abductive neural synapses known as synthetic abductive neural plasticity gating.

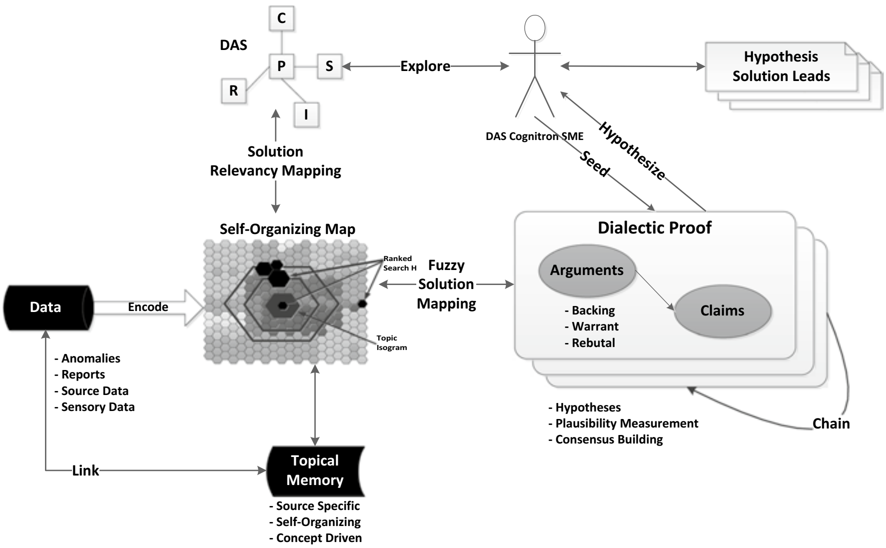


Fig. 9.2 ISAAC theory and argument development

9.1.2 Cognitronic Information Flow

For each problem domain a SELF comes in contact with, a series of Dialectic Search Arguments continuously build an information content encyclopedia (Knowledge Base and Relations) of topics comprising dynamically changing Knowledge Relativity Threads used to explore active argument criteria and applied to FUSE-SEM searches for topical instances. Domain expert Cognitrons develop arguments based upon the exploration of internal encyclopedias and other sources of internal information as needed (depending upon Locus of Control parameters). Situational Reasoning is a facilitation process within a SELF's ISAAC framework that actively discovers and chains together argument instances for a given DAS. Figure 9.2 illustrates this process. The encyclopedia like system pedigree are captured via the use of outputs and pointers from topical memory continuously organized via FUSE-SEMs to provide context and relativity information for the continuously evolving Topics of Interest (TOIs). This cognitive growth process is accelerated with inter-Cognitron communication for real-time sharing of information.

Topically and semantically related ISAAC BIFs are arranged as a cluster within the FUSE-SEMs where fuzzy classification rules are tuned to best define the subject language, fuzziness, relevance, etc. Labeling of the FUSE-SEMs produces metadata sued for navigation and querying within a SELF's internal cognitive framework. Pedigree are continually captured and generated by mapping related libraries of FUSE-SEMs. Massively parallel recursion is inherently utilized to break down dense topics. Figure 9.3 illustrates this process.

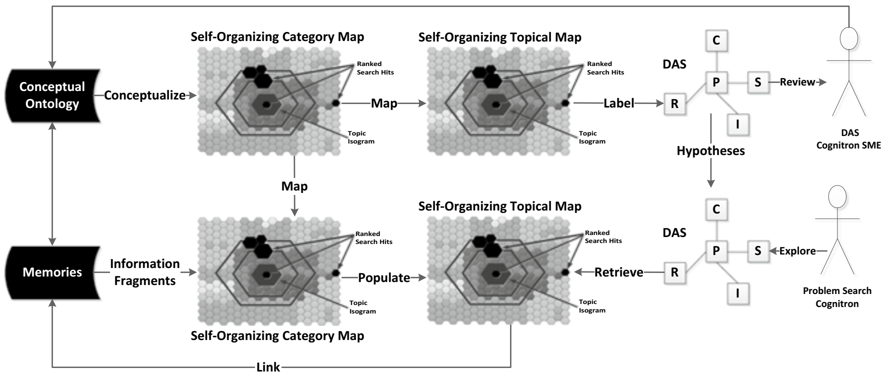


Fig. 9.3 Cognitronic analysis, reasoning, and reporting

Internal Cognitron Subject Matter Experts (C-SME) explore the stored pedigree and utilize FUSE-SEMs to develop prototype hypotheses used to train the FUNN to support/rebut the various developing hypotheses. A framework for knowledge and context includes the processing of criteria, rules, goals, and requirements which are used synonymously in literature to govern the comparisons and processing which take place during a given set of activities. Their history is well documented in utility theory within disciplines of Economics and Psychology [232]. Specifically, the value of the Independence Axiom has been debated for decades [233]. Additionally, in engineering there exist numerous methods to attain quality understanding of requirements [234, 235]. Hence, in a SELF the proximity of continuously discovered data/information upon comparison to the original criteria and developing prototype is used to determine the Mutual Information Measure (MIM), based on Renyi Mutual Information computations. The support/rebuttal BIFs instantiate or refute the hypotheses and MIMs are used to calculate Hypothesis Plausibility Measures (HPMs). Figure 9.4 describes this process.

In Fig. 9.4, the Topical Maps are utilized by the Dialectic Argument Structure (DAS) to group solution prototypes into "like" solutions. Information (fuzzy information and metrics, since the solutions are not identical) that is about each solution prototype group is extracted and sent back to the DAS to be evaluated. It is possible that a set of hypotheses, each one feeding the next, in total explains the situation, therefore these hypotheses for a "chain" of hypotheses, which is denoted as the recursive "chain" in Fig. 9.4. Possible solution hypotheses are processed by the FuNN to see whether these correspond to previously learned solutions, and the results are sent to the DAS Cognitron SME. Possible solutions are also sent to a Problem Solution Cognitron, shown in Fig. 9.4. Here, the Cognitrons search through the SELF memories and Cognitive Ontology to see whether relevant information can be found to help support or rebut the current hypotheses under evaluation. Knowledge gathered is shared between Cognitrons at all stages in the reasoning process.

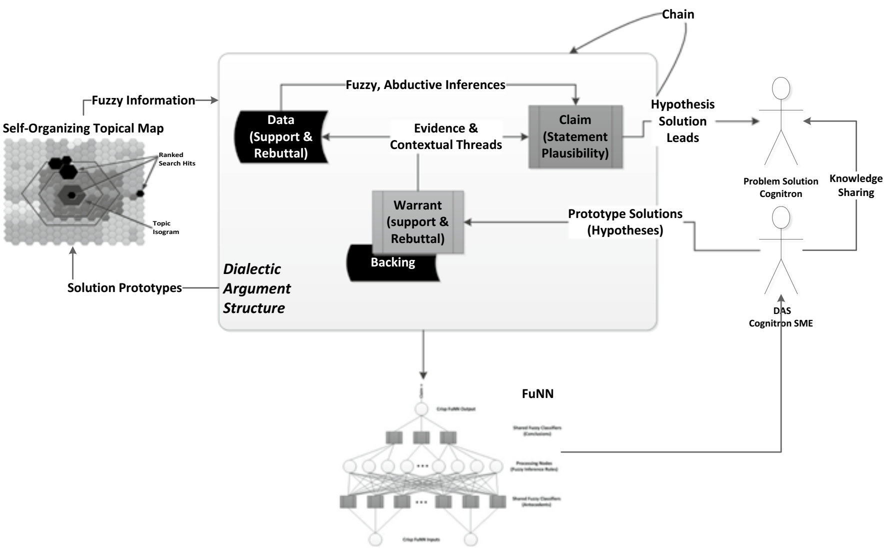


Fig. 9.4 SELF ISAAC Cognitron situational reasoning

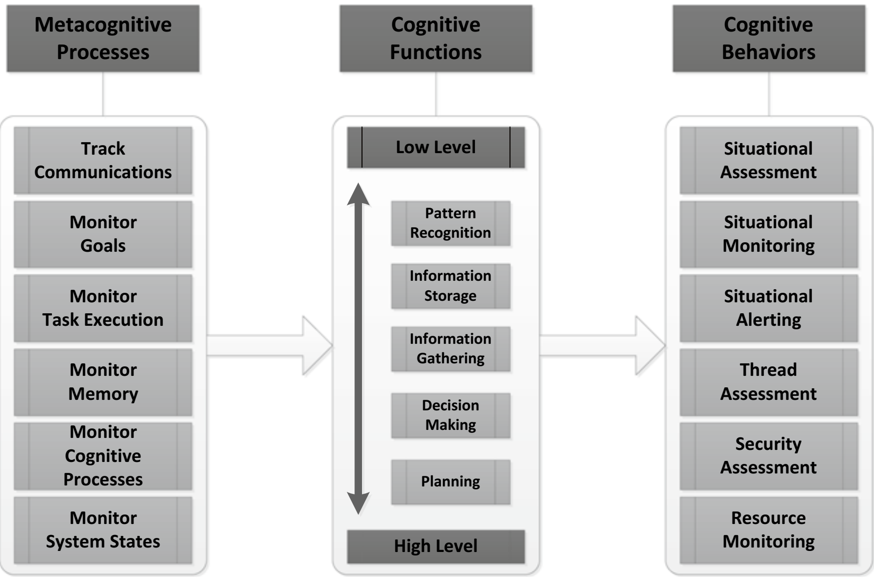


Fig. 9.5 SELF ISAAC cognitive reasoning information flow

ISAAC's high-level Cognitive Reasoning Information Flow is illustrated in Fig. 9.5, representing the metacognitive processing flow within the SELF.

As explained in Sect. 6.3, the metacognitive processes allow self-analysis and self-regulation of the SELF's overall cognitive process. The metacognitive

Table 9.1 Self ISAAC cognitron integration

1. Information exchange	2. Communication
– Seeking information from all available sources	– Utilizing proper context
– Passing information to appropriate ISAs when available	– Providing internal and external reporting
– Providing situational updates	– Ensuring communications are relevant and understandable (proper translation)
3. Supporting behavior	4. ISA team initiative
– Corrective errors (learning and self-evolving)	– Providing guidance or suggestions to other ISAs
– Providing and requesting assistance from other ISAs when needed	– Clearly communicating coalition and individual information ISA priorities
Providing situational updates	

oversight of the SELF monitors all of the SELF's high-level cognitive systems (memory, reasoning, goals, constraints, communication paths, and Neocortex functions), and provides self-assessment metrics to the SELF's Artificial Prefrontal Cortex (APC). These metrics are utilized by the APC to manage the SELF's cognitive processes and affect current and future cognitive behaviors within the SELF's ISAAC cognitive framework. In general, full cognitive processing with ISAAC requires full Cognitron Integration across the SELF ACNF framework. Table 9.1 below outlines this integration process.

9.1.3 Artificial Abductive Reasoning

Abduction is formally defined as finding the best explanation for a set of observations, or inferring cause from effect. The notion of Occam Abduction [70] relates to finding the simplest explanation with respect to inferring cause from effect. A formal definition for Artificial Occam Abduction would be:

Artificial Occam Abduction: The simplest set of consistent assumptions and hypotheses, which, combined with available stored pedigree knowledge, entails adequate description/explanation for a given set of observations which has reached a previously learned threshold within the thought processes of a SELF.

In formal logic notation, given $\boldsymbol{B_D}$, representing current background knowledge of domain $\boldsymbol{D}$, and a set of observations $\boldsymbol{O_D}$, on the problem domain $\boldsymbol{D}$, we look for a set of Occam Hypotheses, $\boldsymbol{H_D}$, such that:

– $\boldsymbol{H_D}$ is consistent[1] w.r.t. $\boldsymbol{B_D}$, and
– It holds that $\boldsymbol{B_D}, |= \boldsymbol{H_D}, \rightarrow \boldsymbol{O_D}$

[1] If $\boldsymbol{H_D}$ contains free variables, $\exists(\boldsymbol{H_D})$ should be w.r.t. $\boldsymbol{B_D}$.

Abduction consists of computing explanations (hypotheses) from observations. It is a form of non-monotonic reasoning and provides explanations that are consistent with a current state of knowledge and can become more or less consistent or inconsistent, as new information is gathered [160, 229]. The existence of multiple hypotheses (or explanations) is a general characteristic of abductive reasoning, and the selection of the preferred, or most simple, but possible, explanation is an important precept in Artificial Occam Abduction.

Abduction was originally embraced in Artificial Intelligence work as a non-monotonic reasoning paradigm to overcome inherent limitations in deductive reasoning. It is useful in Artificial Intelligence applications for natural language understanding, default reasoning, knowledge assimilation, belief revision, and very useful in multi-Cognitron systems. The Abduction form of inference, using hypotheses to explain observed phenomena, is a useful and flexible methodology of reasoning on incomplete or uncertain knowledge. Occam Abduction, defined here, provides not only an answer, or cause, to the observations, but provides significantly more detail, by describing class distinguishing properties of possible hypotheses in which the observations become valid, and then explicitly denotes which is the simplest set of hypotheses under which a fact becomes true.

Here is where we diverge from classical Abductive Reasoning, which is generally steeped in Bayesian probabilistics. Fuzzy abduction, as opposed to Bayesian reasoning, utilizes fuzzy sets of hypotheses to embrace the essence of a given set of observations. The fuzzy abduction utilized here genetically derives a set of fuzzy hypotheses, using the most appropriate available fuzzy implications, and uses these fuzzy hypotheses to derive a truth value (how well do the hypotheses explain the observations). This process is considered abductive because it actively searches for information that both support and/or rebut the developing fuzzy hypotheses. The combination of supporting and rebutting arguments is used to determine the "possibility" that each hypothesis explains all or part of the observations. Hypotheses whose possibility is above a given threshold are sent forward either to provide explanations, or as input for the next genetically generated set of hypotheses.

9.1.4 Elementary Artificial Occam Abductivity

There are several distinct types of interactions that are possible between two elementary Occam Abductive hypotheses $h_1, h_2 \in H_e$:

- ***Associativity***: The inclusion of $h_1 \in H_e$ suggests the inclusion of h_2. Such an interaction may arise if there is knowledge of, for instance, mutual information (in a Renyi sense) between h_1 and h_2.
- ***Additivity***: h_1 and h_2 collaborate additively where their abductive and explanatory capabilities overlap. This may happen if h_1 and h_2 each partially explain some datum $d \in D_0$ but collectively can explain more, if not all of D_0.
- ***Incompatibility***: h_1 and h_2 are mutually incompatible, in that if one of them is included in H_e then the other should not.

- ***Cancellation***: h_1 and h_2 cancel the abductive explanatory capabilities of each other in relation to some $d \in D_0$.
  - For example, h_1 implies an increase in a value, while h_2 implies a decrease in a value. In this case, one supports the hypothesis while the other is used to rebut.

The Occam Abductive Process is:

- Nonlinear in the presence of incompatibility relations
- Non-monotonic in the presence of cancellation relations
- The general case (nonlinear and non-monotonic) Occam Abduction hypothesis investigation is NP-complete.[2]

Consider a special version of the general problem of synthesizing an Artificial Occam abductive composite hypothesis that is linear, and, therefore, monotonic.

The synthesis is linear if: $\forall h_i, h_j \in H_e, \quad q(h_i) \cup q(h_j) = q(\{h_i, h_j\})$

The synthesis is monotonic if: $\forall h_i, h_j \in H_e, \quad q(h_i) \cup q(h_j) \subseteq q(\{h_i, h_j\})$

In this special version, we assume that the Occam hypotheses are non-interacting, i.e., each offers a mutually compatible explanation where their coverage provides mutual information (e.g. Renyi). We also assume that the Occam, abductive belief values found by classification abduction subtasks of for all $h \in H_e$ equal to 1 (i.e., true).

Under these conditions, the synthesis subtask of Artificial Occam Abduction can be represented by a bipartite graph, consisting of nodes in the set $D_0 \cup H_e$. This implies that no edges between the nodes in D_0, nor edges between nodes in H_e. The edges between the nodes in D_0 and those nodes in H_e can be represented by a matrix Q where the rows correspond to $d \in D_0$ and the columns correspond to $h_t \in H_e$.

The entries in $\boldsymbol{Q}$ are denoted as $\boldsymbol{Q_{ij}}$ and indicate whether the given analyzed data are explained by a specific abductive Occam hypothesis. The entries are defined as:

$$Q_{i,j} = \begin{cases} 0 & \text{if datum } d_i \text{ is not explained by hypothesis } h_j \\ 1 & \text{if datum } d_i \text{ is explained by hypothesis } h_j \end{cases} \tag{9.1}$$

Given the matrix Q for the bipartite graph, the abductive, Occam synthesis subtask can be modeled as a set-covering problem, i.e., finding the minimum number of columns that cover all the rows. This ensures that the composite abductive, Occam hypothesis will explain all of D_0 and therefore be parsimonious.[3]

Now we look at a special linear and monotonic version of the general abductive, Occam hypothesis synthesis subtask and look at two Abductive Neural Networks

[2] Nondeterministic Polynomial time complete. A set or property of computational decision problems which is a subset of NP (i.e. can be solved by a nondeterministic Turing Machine in polynomial time), with the additional property that it is also NP-hard.

[3] Note that the general set-covering problem is NP-complete.

(ANNs) for solving it. The first is based on an adapted Hopfield model of computation:

$$\forall i = 1,2,\ldots,n, \quad \sum_{j=1}^{m} Q_{ij} V_j \geq 1 \tag{9.2}$$

For the Occam, abductive synthesis subtask, we associate variable $\boldsymbol{V_j}$ with each Occam hypothesis $h_t \in H_{e,}$ in order to indicate if the Occam hypothesis is included in the composite Occam, abductive hypothesis $\boldsymbol{C}$. We then minimize the cardinality of $\boldsymbol{C}$ by:

$$\sum_{j=1}^{m} V_j \tag{9.3}$$

subject to the constraint that all data $d \in D_0$ are completely explained.

For the Occam, abductive network, the term in the energy function that represents the problem constraints must evaluate to zero when the constraint is satisfied and must evaluate to a large positive value when the constraint is not satisfied, forcing the neural fiber network to evolve accordingly. For this energy term, we use a term expressed as a sum of expressions, one for each datum element, $\boldsymbol{d_i}$, such that the expression evaluates to zero, when hypothesis $\boldsymbol{h_j}$ that can explain the datum $\boldsymbol{d_i}$ is in the composite hypothesis, i.e., $\boldsymbol{V_j} = 1$. Given that $\boldsymbol{Q}$ is an incidence matrix (with elements either 0 or 1), the expression:

$$\sum_{i=1}^{n} \prod_{j=1}^{m} \left\{ \left(1 - Q_{ij}\right) + \left(1 - V_j\right) \right\} \tag{9.4}$$

satisfies the following conditions:

- Each sum of the product terms can never evaluate to a negative number.
- The sum of the product terms, thus, can never evaluate to a negative number.
- Each product term evaluates to zero when a hypothesis that can explain the datum is in the composite; otherwise, it evaluates to a large value.
- The sum of the product term, thus, evaluates to zero when a composite set of hypotheses can explain all the data.

We derive our Occam abductive energy function as follows:

$$E = \alpha * \sum_{j=1}^{m} V_j + \beta * \sum_{i=1}^{n} \prod_{j=1}^{m} \left\{ \left(1 - Q_{ij}\right) + \left(1 - V_j\right) \right\} \tag{9.5}$$

Where α and β are positive constants, and β > α. The first term represents the cardinality of the Occam hypothesis and the second term represents the penalty for a lack of complete coverage; 0 indicates complete coverage. The self-organizing algorithm for the Occam abductive network is:

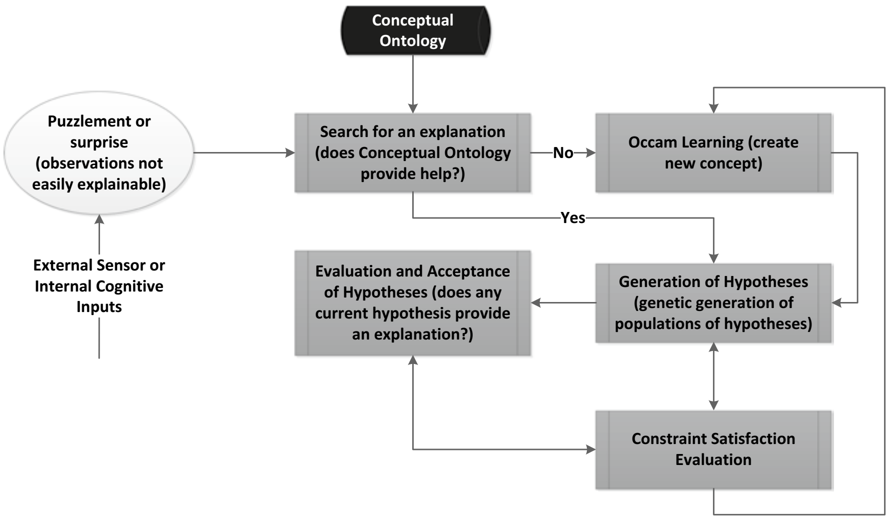


Fig. 9.6 ISAAC artificial Occam abduction process

Assume data set: $\bar{X} = \left\{X^{(1)},\ldots,X^{(p)},\ldots,X^{(P)}\right\}$, where P is the number of input vectors.

Vector $X^{(p)} = \left\{x_1^{(p)},\ldots,X_i^{(p)},\ldots,X_{n_1}^{p}\right\}$ represents the *pth* input vector to the network.

We initialize STEP and SLOPE $\in (0, 0.5]$.

$IT = \beta = TD = 0.5$

For all input vetors $p \in [1\ldots P]$ do {

For all input dimentions $i \in [1\ldots n_1]$ do {

if there are no fuzzy clusters in the i^{th} input dimention $(J_i = 0)$

Create a new cluster using $x_i^{(p)}$

else do {

find the best-fit

Figure 9.6 below illustrates the ISAAC Occam abductive inference process [40]. When sensor inputs are processed, if there are observations that are not readily explained, this information is sent to the Occam Abduction processes to search through the memories and Conceptual Ontology to look for related and relevant information. If none is found, then the Occam processes are engaged to create hypotheses and test them to find applicable explanations for the observations. The abduction process continues until a set of hypotheses can be generated that adequately explains the observations.

9.1.5 Synthesis of Artificial Occam Abduction

Let $B = \{b_k | k = 1,\ldots,l\}$ be a finite set of l
Occam learned possible beliefs
Let $H_e \subseteq H$ such that for each $h_j \in H_c$
can explain some non-empty subset of D_0
Let p be a map from H_c to $H_e : p : H_e \rightarrow B$.
The map p is also defined from an
elementary Occam hypothesis belief value.
We define $p(\{h_j\})$ as $p(h_j)$ and interpret
$p(h_j)$ as the prima facie Occam belief value for h_j .

The Occam abductive classification subtask takes D, H, D_0,
and r as input, where r is a map from $\wp(D_0) \rightarrow \wp(H)$,
and gives H_e and p as output.
The abductive hypothesis synthesis subtask may be characterized
as a five-tuple (D_0, H_e, q, p, H_c), where $D_0, H_e, q,$ and p
constitute the input to the abduction task, and H_c
is the output of the task.

Maximal explanitory coverage of hypothesis data:
A composite hypothesis H_1^c is a better explanation of D_0 than
another abductive hypothesis H_2^c if:
$q(H_1^c) \cap D_0 \supset q(H_2^c) \cap D_0$
Ideally, the assembled composite abductive hypothesis, H_c,
would provide adequate explanitary coverage of
D_0, i.e., $q(H_c) \supseteq D_0$.

Maximal belief in abductive Occam hypothesis:
A complosite hypothesis H_1^c is a better explanation of D_0 than
another abductive hypothesis H_2^c if:
$p(H_1^c) \geq p(H_2^c)$
This specifies that among the composite dialectic Occam
hypotheses that explain the data, the one with the
highest "belief" value is the "best" explanation, by abduction.

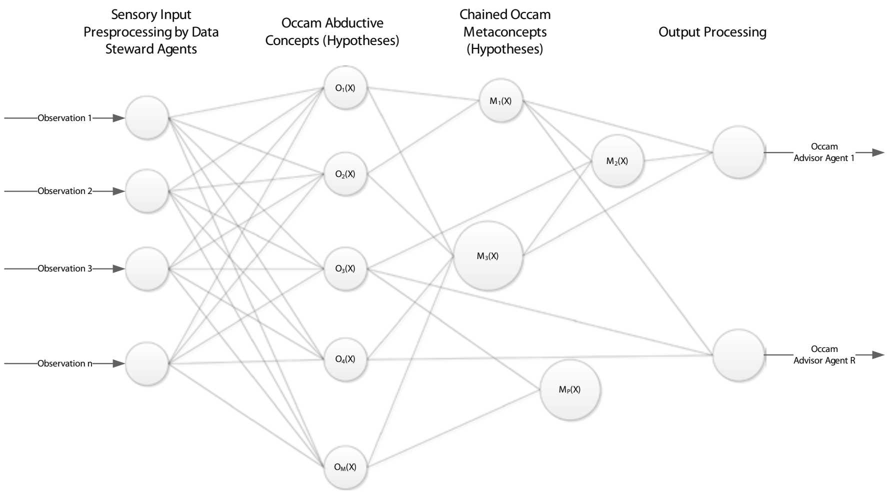


Fig. 9.7 ISAAC Occam abductive processing network

Minimal hypothesis: An composite abductive Occam hypothesis H_1^c is a better explanation of D_0 than another compoaite abductive Occam hypothesis H_2^c if:

$$\left|H_1^c\right| < \left|H_2^c\right|$$

This condition specifies that H_c should be *parsimonious*.

This process is fully realized utilizing the hybrid possibilistic abductive neural processing network illustrated in Fig. 9.7.

Once it has been determined that sensory observations require an Occam Abductive process to provide an explanation for the observations, The Occam Abductive Processing Network is invoked. The observations are sent to the hypotheses generated by the Occam Abduction process (see Fig. 9.6). It is possible within the Occam Abduction process that multiple hypotheses are required to fully explain the observations. These chains of hypotheses (called Metaconcepts) may contain individual hypotheses that are shared by multiple Metaconcepts. This is illustrated in Fig. 9.7. Once the chained Metaconcepts are adjudicated and a final set of hypotheses are created that adequately explains the observations (sensory inputs), they are communicated to Advisor Agents for use by the ISAAC cognitive processes.

9.1.6 Artificial Occam Abductive Hypothesis Evaluation Logic

The following lays out the basics of the Occam Abduction that will be used to perform system hypothesis generation, evaluations, and testing for Dialectic Argument Structure, and testing:

Definition 1

A triplet (Φ, Ω, e) defines a domain of Occam hypothesis assembly:

- Φ = The set of hypotheses
- Ω = The set of observations (sensor inputs)
- e = The Mapping from the subsets of Φ to the subsets of Ω
- Assumptions:
  - Computational: For every subset Φ' of Φ, $e(\Phi')$ is computable.
  - Independence: $e(\Phi_1 \cup \Phi_2) = e(\Phi_1) \cup e(\Phi_2)$; for all Φ_1 and Φ_2 that are subsets of Φ.
  - Monotonicity: If Φ_1 is a subset of Φ_2, then $e(\Phi_1)$ is a subset of $e(\Phi_2)$.
  - Accountability: $\alpha(\varphi)$ is the set of observations that cannot be explained without hypothesis φ.

The following outlines a four-part Occam Dialectic Argument Structure (DAS) Process:

Screening: determines hypotheses acceptability allocating them into a hierarchical fuzzy classification system.

Collection: aggregates hypotheses while accounting for the observations. Hypotheses are determined by adding together every hypothesis that explains all or part of the observations.

Parsimony: narrows down the collection of hypotheses to the most applicable Occam subset. If a subset of collected hypotheses can explain new observations then a narrowed down hypothesis is created.

Critique: determines which hypotheses are the most essential, among those available, based on fuzzy inference metrics. Individually, every hypothesis is excluded from the set, and then the set is tested against the observations. If the observations cannot be explained without the excluded hypothesis, then the excluded hypothesis is marked essential and reintroduced into the set.

Definition 2

An Occam abduction system consists of a logical theory 'T' defined over a domain language 'L', and a set of domain syntax 'A' of 'L' that are called abducible.[4]

Definition 3

If a set of syntax φ is found as a result of an abductive process in searching for an explanation of ω observations, it must satisfy the following conditions:

[4] An abducible argument is a first-order argument consisting of both positive and negative instances of abducable predicates. Abducible predicates are those defined by facts only and the inference engine required to interpret the meaning. In formal logic, abducible refers to incomplete or not completely defined predicates. Problem solving is effected by deriving hypotheses on these abducible predicates as solutions to the problem to be solved (observations to be explained).

- $T \cup \varphi$ is consistent
- $T \cup \varphi \vdash \omega$
- φ is abducible, i.e., $\varphi \in A$

Definition 4

(C, E, T) is a simple causal theory defined over a first order language 'L' where 'C' is a set of causes, 'E' is a set of effects, and 'T' is a logical theory defined over 'L'.

Definition 5

An *Occam Explanation* of a set of observations Ω, which is a subset of E, is the simplest finite set Φ such that:

- Φ is consistent with T
- $T \cup \Phi \vdash \Omega$, where Ω is the conjunction of all $\omega \in \Omega$.
- Φ is a subset-minimal.

9.1.7 SELF's Overall Cognitive Cycle

When we put together all the pieces for a SELF's ISAAC Artificial Cognitive Architecture, which includes the ACNF, the PENLPE cognitive management framework, the SELF memory systems, and the DAS/Cognitron systems, we derive the high-level SELF cognitive and memory cycles (see Fig. 9.8). The high-level cognitive process shown in Fig. 9.8 illustrates the process and information flows steps involved in the SELF's "conscious" cognitive processing:

- **Step 1. Sensory Processing** – these processes are responsible for understanding either inputs received from the SELF's sensors, or internal information that needs to be evaluated in by sensory processing algorithms. Sensory Processing spawns Reasoner Cognitrons that carry perceptual information related to the Perceptual Association Network.
- **Step 2. Perceptual Association Network** – Reasoner Cognitrons process the sensory information that includes metrics to evaluate the external environment's reaction to the current SELF's outward behaviors. Information is exchanged with the SELF's Action Selection Network to provides re-affirmation of the current SELF's external behaviors. Based on the output of the perceptual processing, Analyst Cognitrons are created and sent to the Behavior Selection process to determine what internal and external behaviors are warranted.
- **Step 3. Behavior Selection** – Analyst Cognitrons utilize the metrics created by the Perceptual Association Network to determine whether this information constitutes queues within the Episodic, Declarative, and Emotional memories of behaviors that must be initiated, based on these queues. Based on the memory queues generated within the SELF's various memory systems, combined with

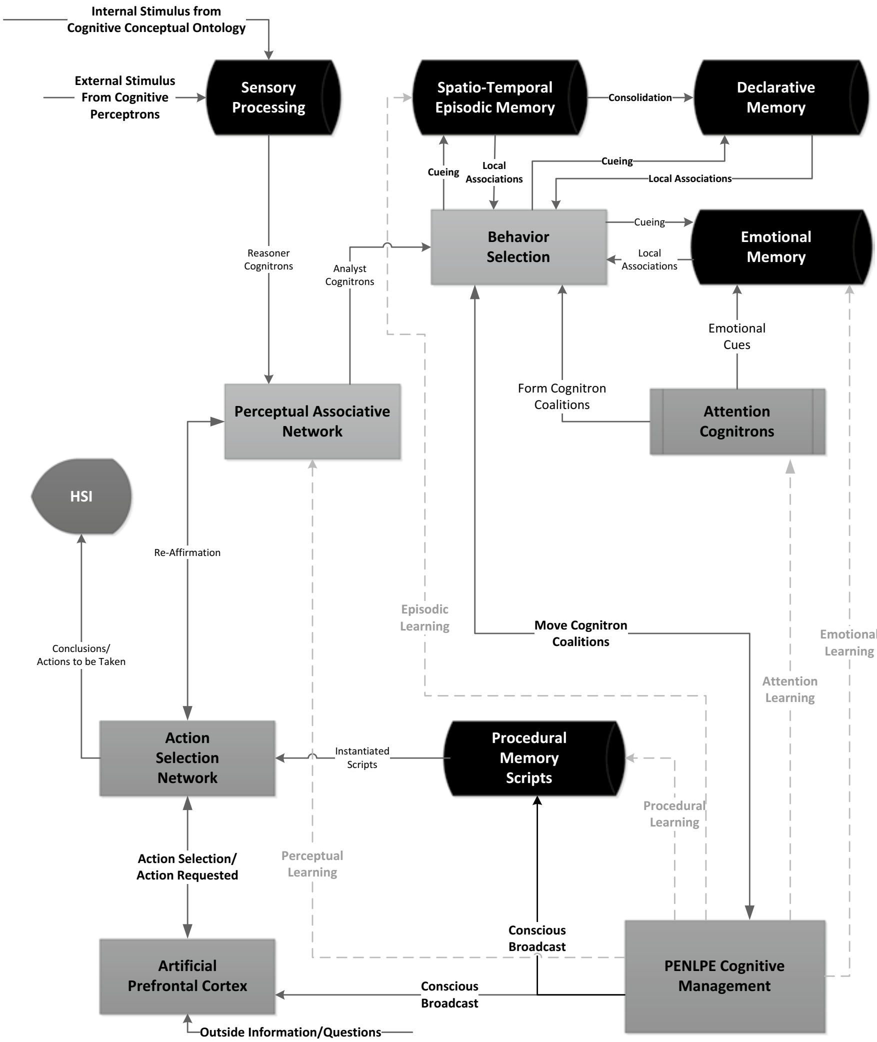


Fig. 9.8 SELF high-level cognitive and memory cycle

information from the current Attention Cognitrons, coalitions of Cognitrons are created that provide information for the Cognitive Management processes (PENLPE) that monitor, manage, and initiate cognitive processes within the ISAAC framework.

- **Step 4**. **Cognitive Management** – here, information from the Conscious Cognitron coalitions is correlated and any information that must be "remembered" is routed to the appropriate memory system (e.g., emotional, episodic, etc.). Based on the current and predicted future cognitive states, the cognitive management processes broadcast the current state of consciousness to the Artificial Prefrontal Cortex for resource management, and any information would constitute a procedural memory is sent to the procedural memory creation processes.

- **Step 5. Artificial Prefrontal Cortex** – based on the conscious broadcasts from PENLPE, along with any outside information that is required (refer to Locus of Control in Chap. 6), the Artificial Prefrontal Cortex initiates possible internal and external behaviors that are required to regulate the SELF, meeting all internal goals and constraints, while at the same time fulfilling external goals and mission directives. These action selections/requests are sent to the Action Selection Network.
- **Step 6. Action Selection Network** – requested/selected actions are evaluated for appropriateness and internal and external actions are selected. If there are any Procedural Memories available for the actions selected, they are initiated. The Perceptual Association Network is informed of the selected internal and external actions and the SELF's effectors are initiated through the SELF's HSI.

9.1.8 SELF Sensory Environment

In order for the SELF to be human-like in its processing, collaboration, and overall nature, we must provide sensors to bring in information from the outside environment [137]. Much like humans, the SELF's ISAAC architecture utilizes all of the sensory information available. Since we are not bound by human limitations in sensory information, ISAAC can contain sensors and sensory processing of information outside the normal human restrictions (e.g., RF information). Figure 9.9 below illustrates the ISAAC sensory environment available to the SELF for processing. Given the sensory inputs and processing, along with the SELF's goals, directives, mission needs, and internal requirements (power conservation, self-assessment, etc.), all of these are processed and utilized to determine courses of action required by the SELF, which in turn drives the use of the SELF effectors, whether they be artificial arms and legs, or wheels, weapons, etc. Figure 9.10 describes the ISAAC sensor-to-effector processing flow. There is communications (feedback) between all levels of cognitive processing within the ISAAC framework, and this is certainly true between the Sensor and Effector networks, as illustrated in Fig. 9.10.

First level of sensory perception looks for spatial and temporal detail within the sensory observations. The sensory perception algorithms query the spatial/temporal behavior processes to determine if there are specific temporal or spatial characteristics the actions were meant to create or capture. Perceptual information is flowed to the relational and reactional perception algorithms to assess and create Knowledge Assimilation Threads based on the results of the Topical Map associations and queries to the SELF's Conceptual Ontology and system memories. At this stage of processing, the relational and reactional perception algorithms may request information from the processes that determine the behavioral priorities to determine

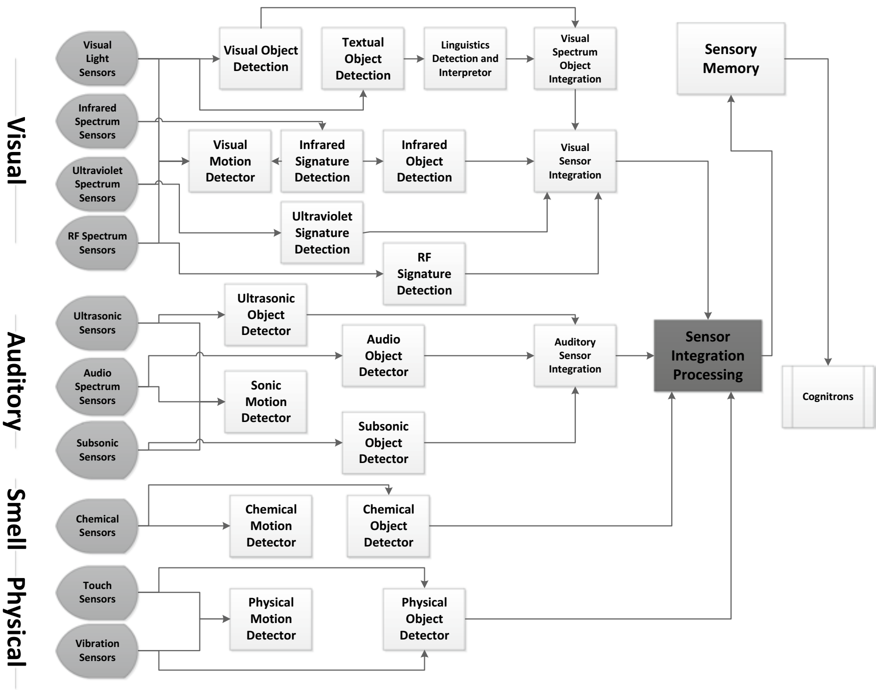


Fig. 9.9 ISAAC's sensory input architecture

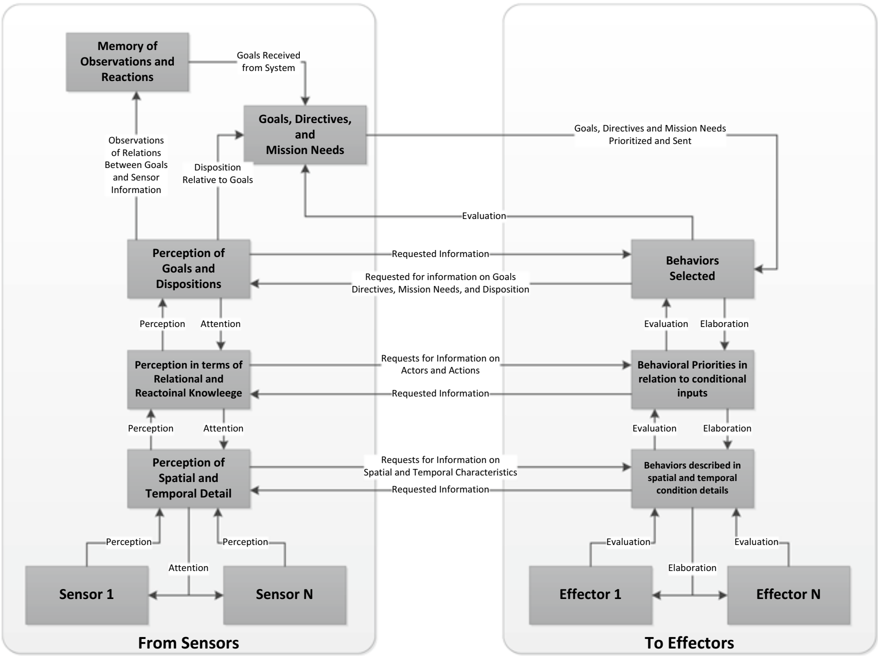


Fig. 9.10 ISAAC sensory-to-effector processing flow

whether the relational and reactional information received through the sensors consistent with the anticipated relational and reactional responses.

One of the last stages within the sensory perception processing is to determine whether the observations perceived through the sensory processing contributed to, or detracted from, the SELF's overall internal and external goals, constraints, and mission directives. The perceived goal, constraint, and mission directive disposition is stored and the results are used to help determine future internal and external actions. The observations of the relationship between actions and their effects on goals, constraints, and mission directives are transmitted to the memory processing systems. Here memory queues are created for future use.

9.1.9 ISAAC's Lower Brain Function Executives

As with any cognitive entity, whether it is a human, a dolphin, a white rat, or an artificial life form, there is lower brain functions executives required to, in the background, keep the entity functioning. In biological entities these lower brain function executives keep blood flowing throughout the body (keep the heart pumping), keep the flow of oxygen throughout the body (keep the entity breathing), etc. Within the SELF, the ISAAC cognitive architecture provides Lower Brain Function Executives, based on human lower brain function executive levels (e.g., brainstem, thalamus, etc.).

Within the SELF they have different meanings, but they are analogous to human lower brain function executives. An example would be blood flow. There is no blood flowing throughout the SELF, but the equivalent is information flow. Without information flowing throughout the SELF, the SELF is essentially dead. Resource management is another important lower brain function executive within ISAAC. We discussed the mechanisms for Cognitive Economy in Sect. 6.6. Within the ISAAC lower brain function executives are components that utilize the PENLPE cognitive management functions to facilitate self-assessment, sensory management, etc. within the SELF's ACNF cognitive systems. We will describe each of these sections, based on their human counterparts. Within biological entities, these systems combine to form an overall control system for the underlying functions necessary to keep the entity alive and functioning. This is also true with ISAAC.

Figure 9.11 shows ISAAC as a high-level control system, with Fig. 9.12 illustrating the lower-level informational control system. The Conceptual and Instantiated Ontologies within the SELF's ISAAC cognitive system contain those real-world concepts the system understands and has experienced. When a decision is made and/or sensory information is received, all of the outputs from the various cognitive processes within the SELF combine to form an inference, or prediction, of what the sensory information means, or what effect the current action(s) will have on the real-world environment the SELF is within. The SELF's cognitive processes fall generally into three categories:

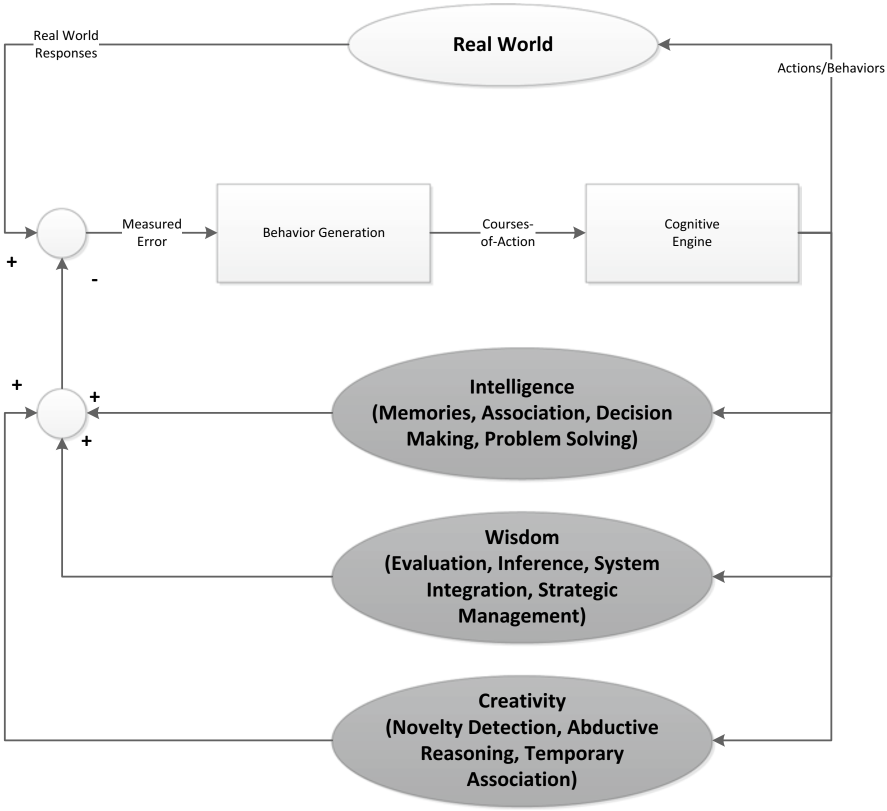


Fig. 9.11 ISAAC as a high-level SELF control system

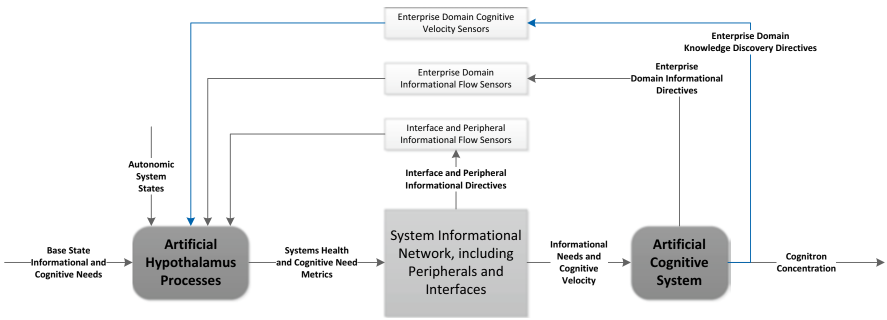


Fig. 9.12 ISAAC as the SELF informational control system

- Intelligence: the SELF's memories, problem solving, and decision making capabilities,
- Wisdom: the SELF's inference engines, learning mechanisms (e.g., Occam Learning), situational assessment (evaluation), information integration, and information/strategic management capabilities,
- Creativity: the SELF's abductive reasoning and temporary association capabilities, along with the SELF's ability to detect novelty or unknown information/situations (pattern discovery).

These inferences and predictions are weighed or evaluated against the real-world responses and/or sensory information to determine the effectiveness of the SELF's decisions and the effects of the SELF's actions/behaviors. Based on the evaluation, the errors (predicted-actual), the SELF's behavior generation processes must choose appropriate responses/actions required. The result is a high-level control system, as is depicted in Fig. 9.11.

At the Information System level with the SELF, the integrated system health management processes (artificial Hypothalamus) collect Measures of Effectiveness (MOEs) on the processing infrastructure of the SELF and perform self-assessment of the overall SELF health and SELF status. Based on these MOE's, resource management processes (Cognitive Economy discussed earlier) will be activated to optimize the SELF effectiveness.

These MOEs include measuring the quantity, velocity, and acceleration of information flow throughout the SELF's systems, at each level within the SELF processing infrastructure. These levels include:

- Peripherals (HSI) and interfaces,
- Enterprise Infrastructure (processors, memories, networks) quantity and quality measures (sometimes called Quality of Service or QoS measures),
- Enterprise Infrastructure informational velocity and acceleration measures. This measures how quickly information can be transferred between processing systems within the self (velocity), and how quickly information needs can be ramped up when information needs are critical (acceleration).

These drive the Integrated System Health and Prognostic Health Cognitrons to process and communicate issues and needs to the various system management processes within the SELF's PENLPE cognitive management. Figure 9.12 illustrates this process.

This SELF as a control system may also be viewed as an analogy to a human Central Nervous System (CNS), which is itself a control system within the human cognitive processing system [14].

9.1.10 ISAAC as an Artificial Central Nervous System

Throughout the book, we have described the various aspects of the SELF and the ISAAC cognitive framework in terms of human brain functions (e.g., Artificial Prefrontal Cortex, Artificial Neocortex, etc.). Since many people are familiar with

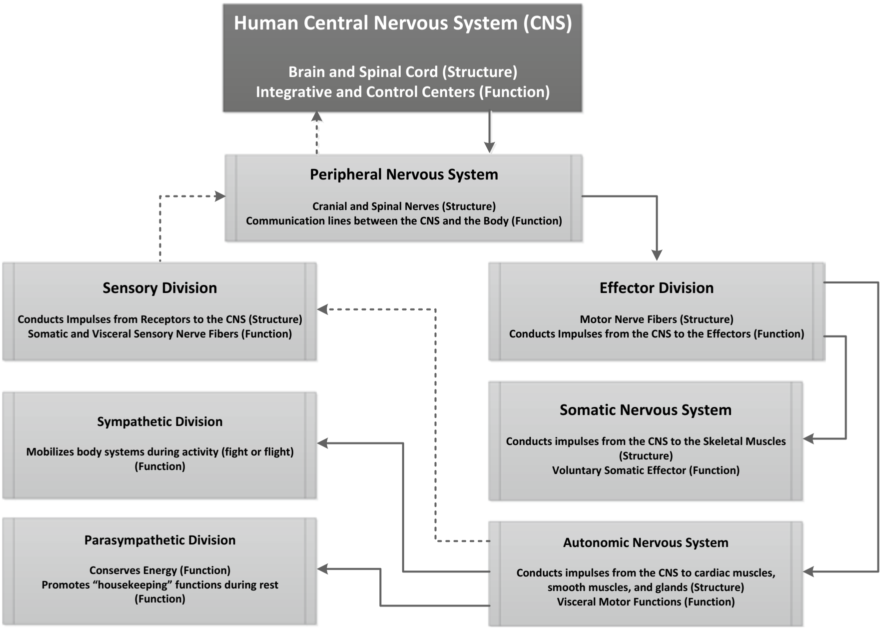


Fig. 9.13 Human central nervous system high-level view

the basic human central nervous system, it seemed appropriate to describe ISAAC in these terms. Figure 9.13 is an adaptation of work done by Kandel [236] and Levine [163], and provides a high-level view of the human Peripheral Nervous System (PNS) in terms of how information is transmitted between its various components. The PNS consists of those components outside of the Brain and Spinal Cord. One of the major components of the PNS, the Effectors, consists of the nerves associated with motor nerve fibers. This is divided between the Somatic Nervous System, the voluntary effector function, and the Autonomic Nervous System, which affect the internal organs, blood vessels, and glands and corresponds to involuntary nerve fibers (things regulated by the brain without conscious thought). Information is sent from the Brain to the Spinal Cord, and then out to the PNS systems.

The Autonomic Nervous System has two major divisions, the Parasympathetic Division, which is used as a resource manager (conserves energy) and promotes housekeeping functions within the human cognitive system, and the Sympathetic Division, which regulates and mobilizes the body's effector motor nerve fibers and cognitive systems during activities (e.g., fight or flight). The Autonomic Nervous System is also tied to the Sensory Division of the PNS and may drive or trigger the body and brain's responses, depending on the current Autonomic Nervous System state.

We use an adaptation of Fig. 9.13 to illustrate the cognitive and information flow structures within the SELF as a Synthetic Nervous System (SNS), illustrated in Fig. 9.14. Here ISAAC and PENLPE act as the Brain and Central Nervous System functions within the SELF Synthetic Nervous System. As illustrated in Fig. 9.8,

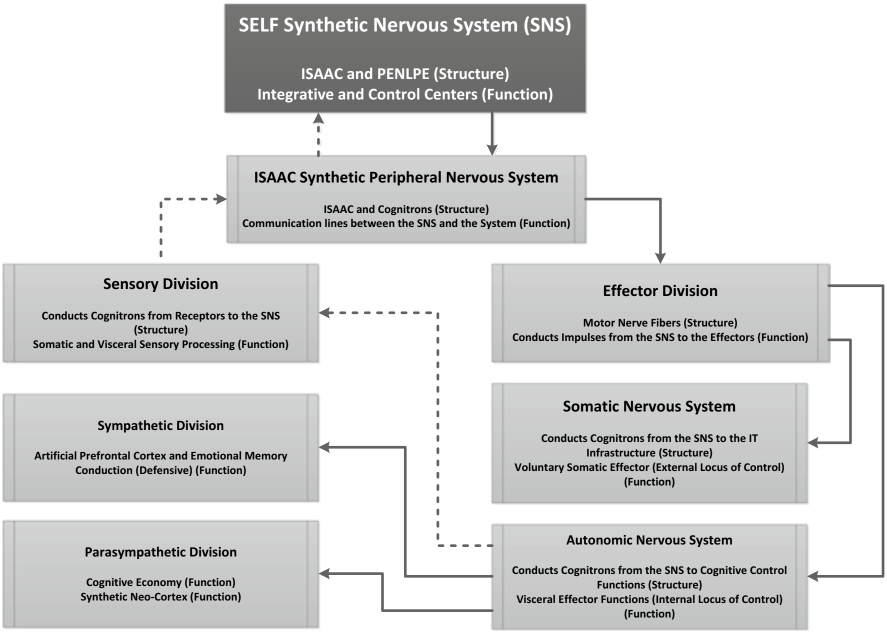


Fig. 9.14 ISAAC synthetic nervous system high-level view

effector instructions are transmitted via Cognitrons through the behavior selection processes to the SELF's effectors, both for voluntary (conscious) execution and involuntary (subconscious) execution. In the case of the SELF, instead of blood vessels, internal organs, and glands, the SELF's system are the IT Infrastructure (processors, memory, networks, etc.). This also drives the external Locus of Control (Chap. 6) functions. The Synthetic Autonomic Nervous System processes within the SELF provide internal Locus of Control autonomic functions (subconscious) that regulate the internal systems of the SELF.

The SELF's Synthetic Autonomic Nervous System, like the human Autonomic Nervous System, is divided between the Parasympathetic Division, which, in the SELF, provides the Cognitive Economy functions (Chap. 6) and the Synthetic Neocortex functions, and the Sympathetic Division, which drives the Artificial Prefrontal Cortex and Emotional Memory (defensive) functions of the SELF. These systems all work together to allow complete autonomous control and operation of the SELF.

Based on this picture of the SELF's cognitive structure in terms of Human Brain functionality, we will describe each of ISAAC's subsystems as brain functions:

- **Lower Brain Function Executives (the Brainstem)**: accepts sensory Inputs and performs early pre-processing of sensory information to be sent to Synthetic Thalamus. ISAAC's lower brain functions controls the startup of the ISAAC framework (cognitive awakening) and creates and controls Interface Cognitrons for Sensors. In addition, these process create and control Data Steward Cognitrons for initial sensory information handling and regulates the internal information flow (heartbeat) within the IT infrastructure, including creating Health and Status Cognitrons to monitor the overall system as well as regulate sensory information flow (external inputs – breathing).

- **Information Processing Center (the Thalamus)**: here, processes create and control Data Steward Cognitrons to accept information from Brain Stem Data Steward Agents, where sensory processing algorithms cleanses, categorizes and creates metadata and contextual threads (RNA) [39]. This defines the initial state of the SELF, creating and disseminating the required Cognitrons, based on internal state assessment. The Thalamus creates and controls Advisor Cognitrons, which advises on the system health status which is used to regulate the SELF's lower brain functioning (internal and external information flows). These processes also create and control the Internal Interface Cognitrons which allows interfaces between Subconscious and Lower Brain functions to Higher Brain Functions. Finally, the Thalamus processes act as an information relay between sensory memory and Cognitrons.
- **External Motor/Effector and HRI Control (the Cerebellum)**: the synthetic Cerebellum functions within ISAAC create and control the interface Cognitrons for interface with the outside world (HSI functions). This includes the creation and control of Advisor Cognitrons to deliver external motor/actuator control decisions. Given the external interface nature of the Cerebellum, these processes also create and control Health and Status Cognitrons that monitor external interfaces (HSI, motor/actuator).
- **External Motor/Effector and HRI Control (Limbic System)**: these processes within ISAAC create and control Reasoner, Analyst, and Advisor Cognitrons with specialized strategies for dealing with possiblistics of emotional memories and emotional triggers. They also provide inputs to the Artificial Prefrontal Cortex that determines transition between Artificial Autonomic Nervous System States. This emotional information is transmitted throughout the SELF via EML (Emotional Markup Language), which includes Emotional RNA threads that provide emotional triggers in order to initiate emotional memory recall.
- **Integrated System Health Management (the Hypothalamus)**: the synthetic Hypothalamus processes create and control Reasoner, Analyst, and Advisor Cognitrons with specialized nodes and strategies specific to Integrated and Prognostic System Health Management [51, 53]. These Cognitrons monitor and control relationships between emotional, conscious, and subconscious systems within ISAAC. The Hypothalamus algorithms are used to initiate and control Cognitrons used for self-assessment, self-regulation, self-soothing, and self-healing within the ISAAC cognitive ecosystem. As part of this self-regulation, the algorithms initiate artificial pain and pleasure metrics within ISAAC that correspond to prioritization of tasks, goals, initiatives, etc. within the cognitive processing system.
- **Abductive Conceptual Framework (the Temporal Lobe)**: here, the sensory processing algorithms transform sensor data into the different types of "perceived"

Binary Information Objects, based on the nature and type of sensory information. These processes keep track of events (temporal information), information context (relationships between object and events), including spatio-temporal relationships [56], and create metadata that is used by the memory system for information correlation and retrieval (construction). The abductive processes are utilized to reason (hypothesize) about the perceived objects, events, and context, based on information about current mission, goals and objectives, the current state of the system, as well as historical information contained in the SELF's Conceptual Ontology and system memories. This is the section of the ISAAC cognitive system where priorities are decided and a hierarchy of priorities is created for the system to operate from. Based on these priorities, a "course-of-action" hierarchy is created, based on the overall system priorities, goals, and mission directives.

9.2 The Cognitive, Interactive Training Environment (CITE)

Depending on autonomous systems like the SELF is a two-edged sword. These are extremely complex systems that require complex control and monitoring and it's unclear as to what level of trust to ascribe to autonomous artificial entities. However, utilizing software to partially or fully automate tasks is now commonplace. Unfortunately with most systems that utilize automation the capabilities of the software performing these tasks typically do not improve over time (as humans would who were performing the same tasks). Even though the SELF is a fully cognitive, learning, self-evolving system, one of the questions to be solved is how we can infuse human heuristic thinking into the SELF cognitive processing algorithms. One way to accomplish this is to utilize human operators as mentors for the SELF. Here we describe the use of a software system called the Cognitive, Interactive Training Environment (CITE) that learns and improves through the use of a Human Operator acting as a Mentor for the software, until the software is capable of performing the desired operations autonomously and with improvements (see Fig. 9.15). CITE provides for Human Interaction Learning (HIL); as the human operator's role changes from manager to mentor to monitor while the SELF evolves from learner to performer. The CITE system, illustrated in Fig. 9.15 provides effective feedback mechanisms to allow humans to influence the SELF systems. The heart of CITE is the SELF's PENLPE cognitive monitoring and management framework described earlier. Utilizing PENLPE's monitoring and metric generating capabilities for collaborating with humans, humans can influence PENLPE in a positive way and CITE allows the Cognitrons to learn and improve as the process. The human has to ability to review the SELF's choices, based on PENLPE's suggestions and then provide feedback as to why a given choice or set of choices was effective or not. The PENLPE management framework then provides feedback to the human to give the human an understanding of the processes the SELF utilized to make inferences and decisions. This process of feedback and human-SELF interactions

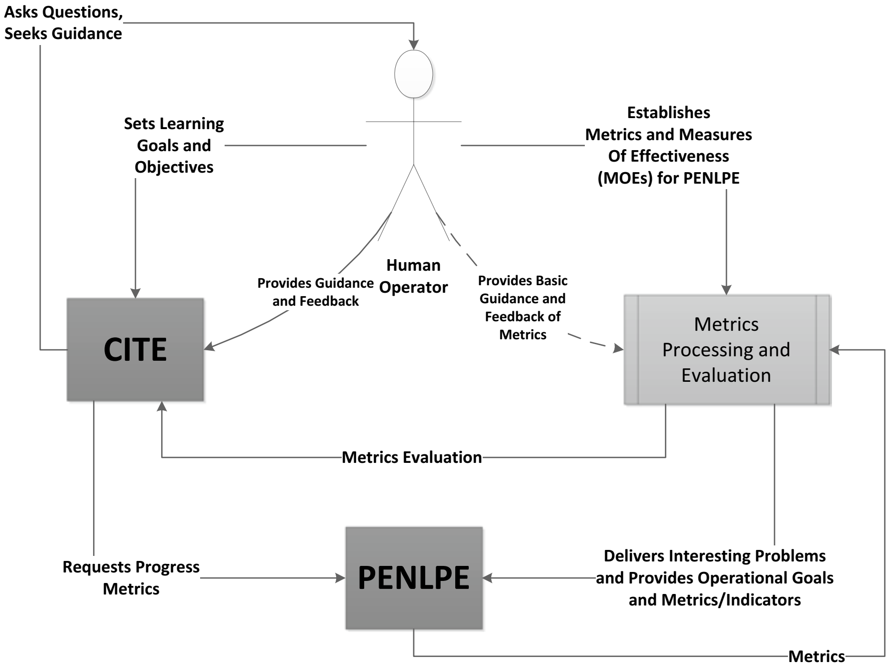


Fig. 9.15 The SELF CITE high-level architecture

and collaborations provides humans insight to develop trust in the SELF over time and will help to increase the effectiveness and efficiency of the SELF, as well as providing an effective Human-Robot interface that humans will come to trust and use over time. Figure 9.15 illustrates the high-level architecture for CITE.

9.2.1 SELF Cognitive Resiliency

One of the adapted uses for CITE is to provide a mechanism to develop cognitive resiliency within the SELF. Developing "Warrior Resiliency" has been a focus of armies since the dawn of time [85]. There has been much research over the last decades to understand and provide systems and methodologies to develop and enhance cognitive resiliency in soldiers. The ability to adapt to adversity and overcome barriers in all walks of life is critical to a soldier's overall mental health and strength. The same must be true for an autonomous SELF. The SELF must have the ability to adapt to any environment it finds itself in. An adaptation of the SELF can allow human operators to utilize the SELF to develop cognitive resiliency within the SELF's ISAAC cognitive architecture. Figure 9.16 illustrates the use of CITE for this purpose.

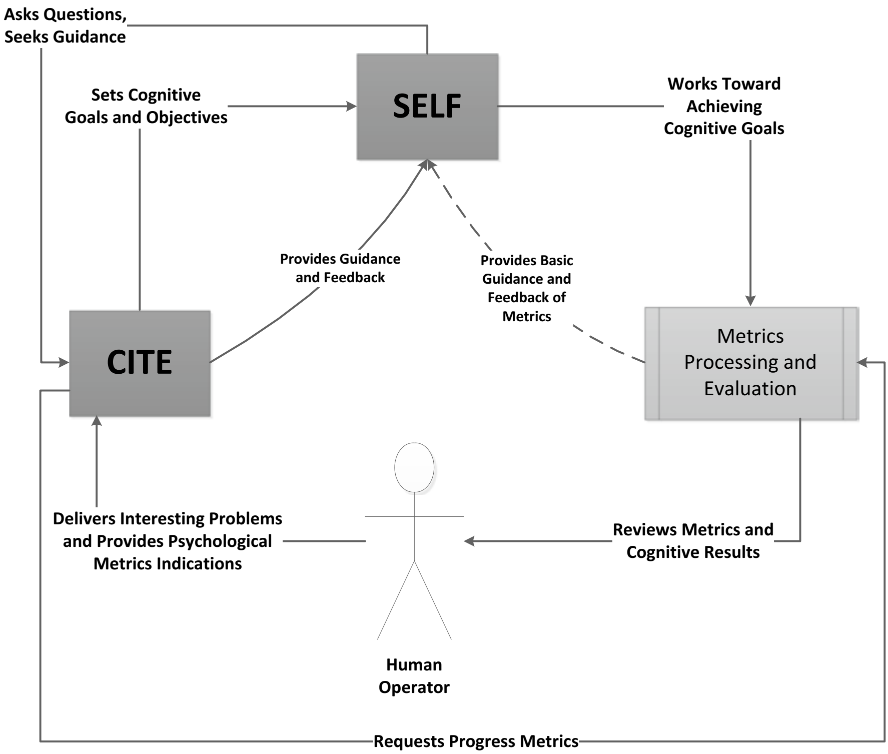


Fig. 9.16 The CITE used for SELF cognitive resilience

As described above, the CITE system utilizes the PENLPE cognitive management system. This, combined with Ontological Technologies developed by Purdue University [203], allows CITE to provide an interactive environment capable of providing training and adaptation for the SELF. In this mode, CITE will provide the SELF with automated, interactive cognitive training and captures and reports artificial metacognitive indicators and metrics that allow complete assessment of the cognitive resiliency of the SELF. CITE will gather and assess metacognitive indicators like:

- Problem solving skills (analytical proficiency)
- HRI Collaborative skills
- Cognitive self-awareness
- Cognitive self-regulation
- Emotional self-regulation (the SELF's Limbic system)

CITE will provide the SELF with training to allow evolution of the SELF's self-monitoring and self-assessment skills needed for cognitive self-regulation when the SELF is operating autonomously. The self-assessment and cognitive resiliency instructional methods are based on Dr. Peter Levine's [163] autonomic nervous states, which provide the basis for cognitive behavior training [134], metrics and artificial bio-markers that include environmental, contextual and HRI component to affect real SELF cognitive resiliency [83–85].

9.2.2 SELF Cognitive Resiliency and Memory Development

It may sound strange to be discussing cognitive resiliency in an artificial life form like the SELF, however if artificial entities like the SELF are expected in the future to operate in complete autonomy, ensuring that the autonomous artificial cognitive system can respond and adapt to unknown situations and environments becomes critical. Cognitive resilience develops within the cognitive framework (whether biological or artificial) through training that results in the learned ability to respond, or self-regulate, to severe changes in environments or situations. These learned abilities, then, get stored as memories (possibly emotional memories) within the memory system. Memories, in general, are divided according to the functions they serve [178]. To qualify as a memory, a cognitive input must cause both enduring changes within the cognitive system (affect autonomic nervous system states) and must also affect emotional responses and goals [159]. Crowder and Friess adapted Dr. Levine's autonomic nervous system states to how they were applied to artificial entities [78]. A memory must also induce some change that affects the entities Conceptual Ontology, brought about by the memory in class of things being affected by the input, and therefore, affects entity behavior (we discussed the SELF's behavior selection previously). There are no memories that are neutral from a behavioral standpoint [79].

9.2.3 SELF Procedural Memory Development and Resiliency

One of the main divisions of memory that we discussed in previous chapters is Procedural Memory. Procedural Memory is a form of implicit memory that includes classical conditioning and the acquisition of skills [90]. Procedural Memory creation contains central pattern generators that form as a result of teaching or practicing and are formed independently of conscious or declarative memory. In his work on Procedural Memory, Kahana showed that retrieval of Procedural Memory is a cue-dependent process that contains both semantic and temporal components [144].

Creation of Procedural Memory is tied to not only the repetition of tasks, but also to the richness of the semantic association structure they represent [220]. In order to provide cognitive resilience to the SELF, the CITE system provides interactive training that allows the DART systems for the SELF to create procedural memories, as was described previously. These memories, or 'scripts' will have physical (sensory) as well as emotional memory triggers and provide the cognitive skills required at the necessary time for self-evaluation, self-awareness, and self-assessment within the SELF's ISAAC cognitive framework to present or reduce problems caused by changes in the SELFs environment or situation [80, 90].

The SELF's CITE training will provide an artificial cognitive architecture that is capable of developing cognitive strategies and training scenarios, based on mission needs and goals, that allow the SELF to develop implicit strategies (procedural memories) that will "kick-in" under specific circumstances, based on physical, sensory, and/or external or internal environmental changes the SELF finds. Based on these cognitive interactions between CITE and the SELF's ISAAC cognitive framework, procedural memories creation will be initiated to allow the self-assessment, self-awareness, and self-regulation, driving self-soothing within the artificial cognitive architecture [79].

9.3 Discussion

Humans are made up of thousands of biological processes that create our overall conscious self. However, the SELF is, at its heart, software. We have seen throughout the book architectures, methodologies, processes, and frameworks aimed at providing the SELF with synthetic consciousness, reasoning, learning, inferring, and remembering. All of these must be implemented in software and hardware. Chapters 10 and 11 describe software and hardware (physical) architectures capable of supporting the SELF's cognitive processes.

Chapter 10
Artificial Cognitive Software Architectures

As discussed in previous chapters, the primary SELF software component is the Cognitron. Each Cognitron type provides different cognitive abilities that, together, form a cognitive ecosystem within an ACNF cognitive framework, implementing intra & inter SELF communication and collaboration. The basic Cognitron is a self-contained discrete functional software codelet comprising one or more loosely coupled software services. For a Cognitron to be of a specific archetype (e.g. Reasoner Cognitron), a set of archetype specific services is defined. Additional services can always be added to extend capabilities of a Cognitron archetype. Figure 10.1 lists the core set of services from which a Cognitron's capabilities can be defined [198].

As described in previous chapters, there are five basic Cognitron Archetypes that make up the SELF Cognitron software architecture:

- Data Steward Agent
- Advisor Agent
- Reasoner Agent
- Analyst Agent
- Interface Agent

As described in previous chapters, the five basic Cognitron Archetypes within an Artificial, Cognition based Software Architecture are:

- Data Steward Archetype (DSA)
- Advisor Archetype(ADA)
- Reasoner Archetype (REA)
- Analyst Archetype (ANA)
- Interface Archetype (INA)

These reference architecture archetypes are manifested and implemented as software agents who then drive artificial SELF cognitive processing.

A SELF Software Architecture is a distributed architecture that allows Cognitrons to operate independently, but in coordination and collaboration with other Cognitrons to achieve SELF-wide goals and comply with directives. In order to facilitate and

J.A. Crowder et al., *Artificial Cognition Architectures*,
DOI 10.1007/978-1-4614-8072-3_10, © Springer Science+Business Media New York 2014

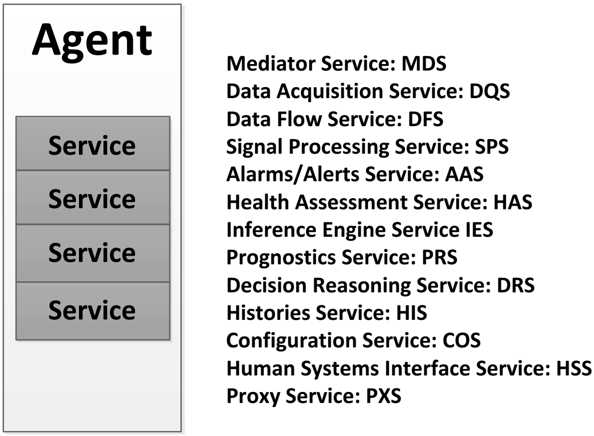


Fig. 10.1 Basic set of available Cognitron services

manage Cognitron coordination and collaboration, a hierarchical management structure is provided via PENLPE for continually negotiating Cognitron goals [198]. Any Cognitron may take the role of Mediator (managed by the Artificial Prefrontal Cortex). A Mediator Cognitron (MC) is one of the first Cognitrons to be started on a host processor. MC has special duties related to short-term and long-term memory within its host and performs special duties related to other Cognitrons that will, over time, run on that host processor (see Sect. 4.2). One issue currently being researched is the sequencing of the cognitive functions for a SELF artificial life form. Figure 10.2 provides a Use Case diagram for the Mediation Cognitron. The MC is utilized to initialize a SELF's subconscious Cognitrons, Lower Brain Function Executives, and Memory Management Executives (Sect. 4.2).

Figures 10.3, 10.4, 10.5, 10.6, and 10.7 breaks down Fig. 10.2 to provide Use Cases for the executives (Prefrontal Cortex, Lower Brain Function, and Memory Management). Figure 10.3 illustrates the Lower Brain Function Executives that are initialized through the Mediator Cognitron. These executive functions are pulled from a SELF Long-Term Memory and executed in order to "start" a SELF. These are described in Sects. 9.1.9 and 9.1.10. Figure 10.3 shows the Lower Brain Function Executives that must be initialized and executed in order to first by the MC.

Figure 10.4 illustrates the executive functions that must be initialized and executed in order for the Artificial Prefrontal Cortex functionality within a SELF. These include the Metacognitive, Metamemory, Cognitive Economy, Locust of Control and Communications functionalities discussed in earlier chapters.

Figure 10.5 provides the Memory Management Executives that the Mediator Cognitron must initialize and execute in order for a SELF's memory systems to functions properly upon startup.

Figures 10.6 and 10.7 provide lower-level details of the Brain Steam and Thalamus executives to illustrate the type of lowest level functions that must be executed in order to initialize and start up a SELF. Each of the executive function in Figs. 10.3, 10.4, and 10.5 would have similar diagrams.

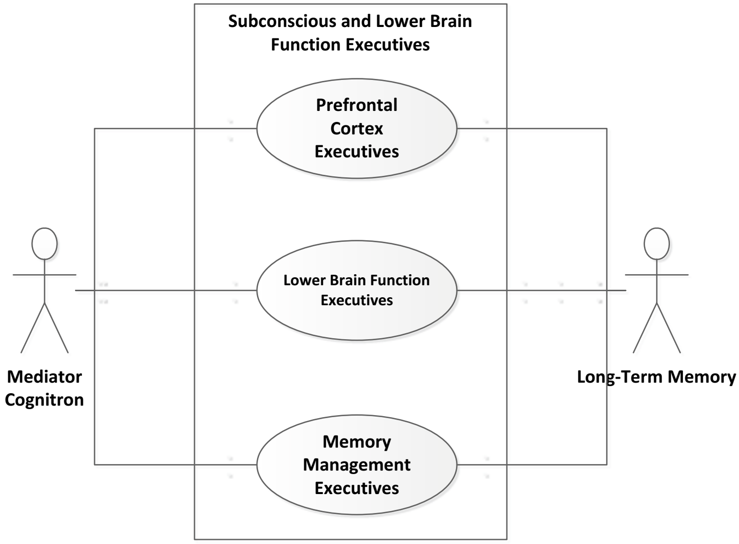


Fig. 10.2 Lower brain function executives

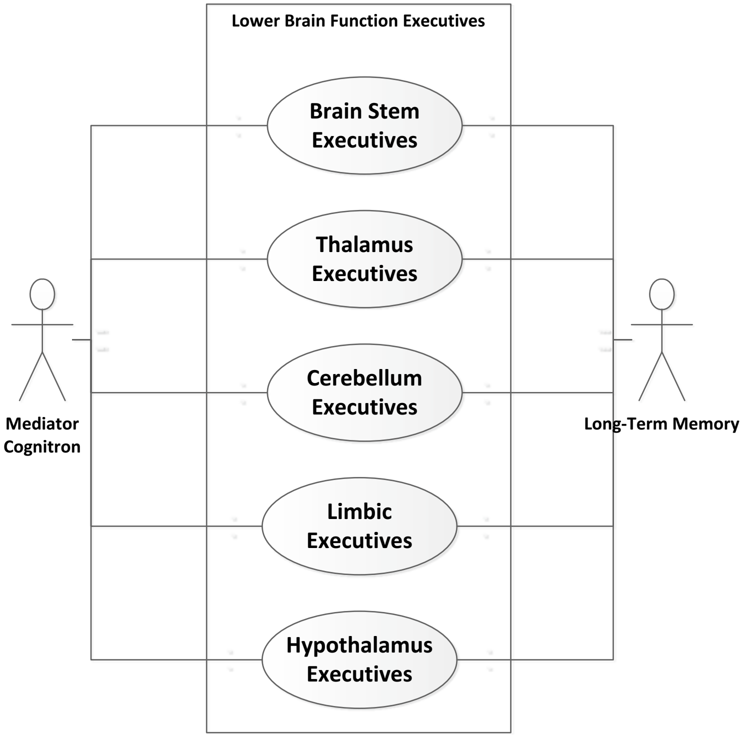


Fig. 10.3 Lower brain function executives

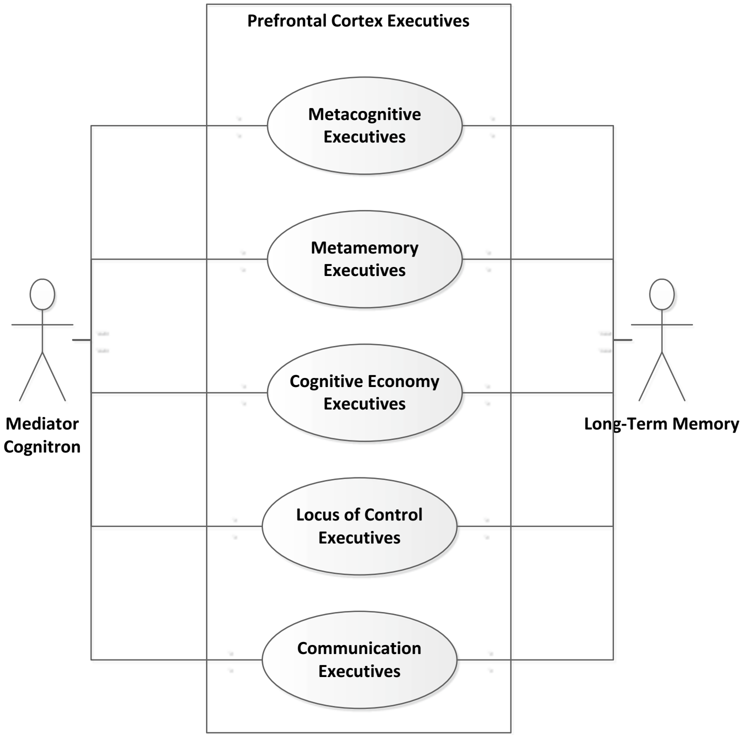


Fig. 10.4 Artificial prefrontal cortex executives

10.1 Artificial Prefrontal Cortex Genesis

When a SELF is first initiated, any Cognitron acting as a Mediator may also take on the role of System Mediator (Artificial Prefrontal Cortex). This role is negotiated between host mediator Cognitrons within the cognitive framework network and will typically be assigned to the host mediator Cognitron with the most appropriate local resources available, such as storage, memory, processor capabilities (CPU). When the spark of life is given to a SELF, default properties and initial settings will govern startup. However, it should be noted that a SELF of different sizes and limitations can and most likely will require different initial settings based upon a priori knowledge of the operational environment a given SELF will operate in. The System Mediator maintains its role for the life span of a SELF.

The initial implementation vision for a SELF is an artificial self-contained, autonomous hardware/software entity. Therefore, the initial software architecture assumes that once a System Mediator has been established in does not migrate to a different processor, unless a partial SELF system failure is experienced, and would be considered a recoverable system failure. Think about shutting down and trying to

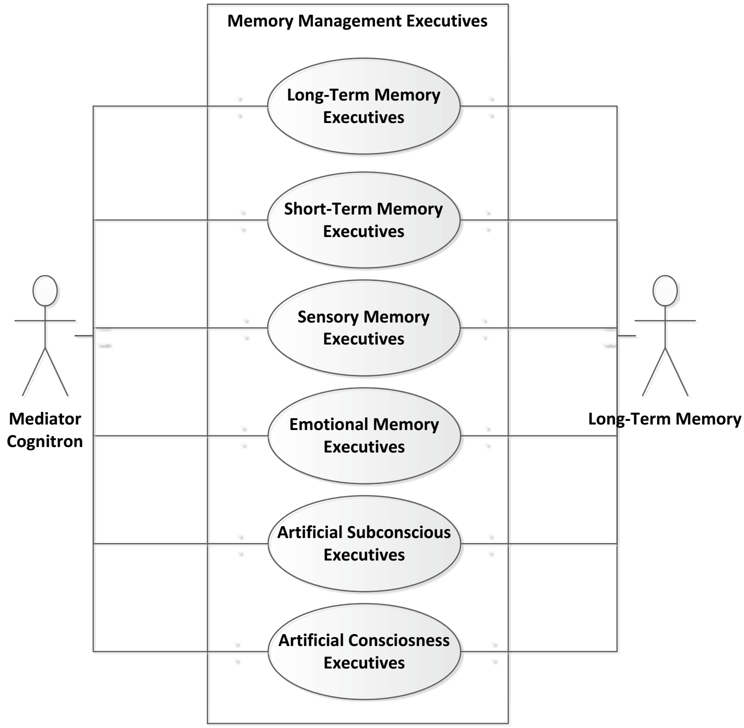


Fig. 10.5 Memory management executives

move a human's Prefrontal Cortex to another part of the brain; a non-recoverable system error, by the way. Such a move within a SELF would require a system-wide change in order to physically transfer a SELF's long-term memories (e.g., declarative, procedural, etc.) a process that would take considerable time and the system would be cognitively unavailable during this time.

Another primary function of Mediator Cognitrons is continuously assembling information from the many Cognitrons for facilitating creation of organized Cognitron coalitions for solving problems, analyzing, and inferring about sensory data/information. This process is designed to provide cognitive intelligence, rapid analysis, and reasoning within an artificial cognitive architecture. This mediation process facilitates the memory integration process, providing development and delivery of knowledge and knowledge characteristics throughout a SELF's cognitive processing architecture. Thus, providing for increased learning and reasoning capabilities as Cognitrons evolve and communicate their insights. The Mediation Cognitrons rely on an initial Reasoner Cognitron (RC) launched by each host within a SELF hardware framework [198]. The Host Mediator (HoMe) (aka. Artificial Prefrontal Cortex) is responsible for managing the goal-orientation for a SELF making sure learning is generally focused upon objectives, directives, and health.

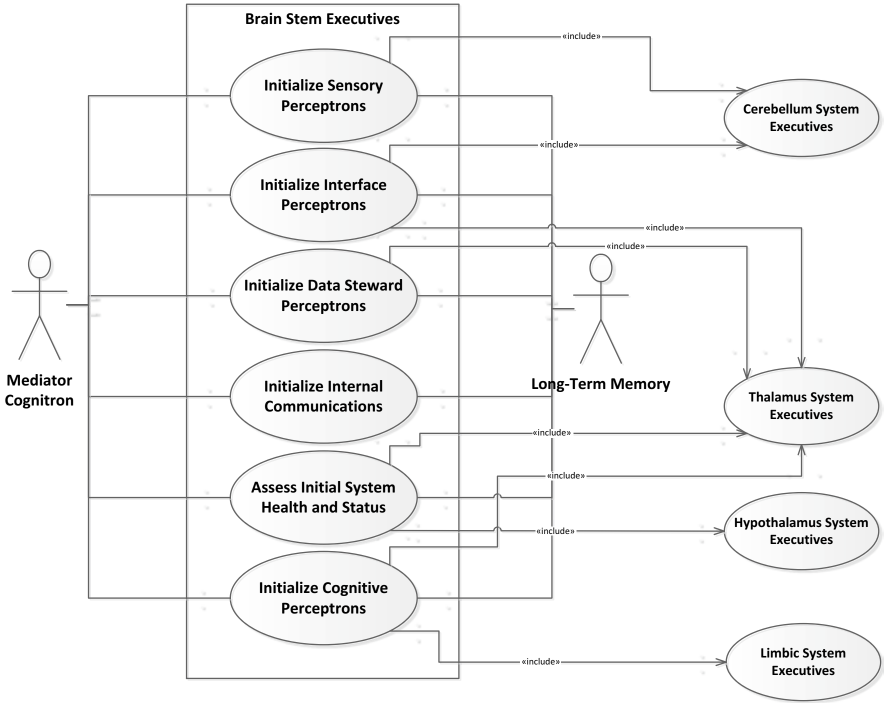


Fig. 10.6 Brain stem executives

10.2 Cognitron Service Instantiation

Cognitrons within a SELF cognitive framework are not mobile, however, personalities aggregating over time and the components within them are mobile. Cognitron personalities are discussed in the next chapter. The preceding chapter introduced artificial cognitronic archetypes. Specifically, Fig. 10.1 described core basic services that are available to the different archetype Cognitrons. Each Cognitron service is composed of small discrete function specific pluggable modules, or Nodes. Remember, in order for a Cognitron to be considered one of the archetypes, it must contain a minimum set of available services. Hence, Fig. 10.8 shows the minimum services required for each Cognitron archetype. Each Cognitronic Agent may contain more capabilities than the services shown in Fig. 10.8; however, it must contain at least these services to perform the artificial cognitive functions of a particular Cognitron archetype It is possible that new Cognitron archetypes have to be created and evolved over time as a SELF's ISAAC cognitive processing system determines that a new archetype is needed, possibly to handle something new it was not originally programmed for; based upon its available sensors and other physical/data environments that a SELF encounters.

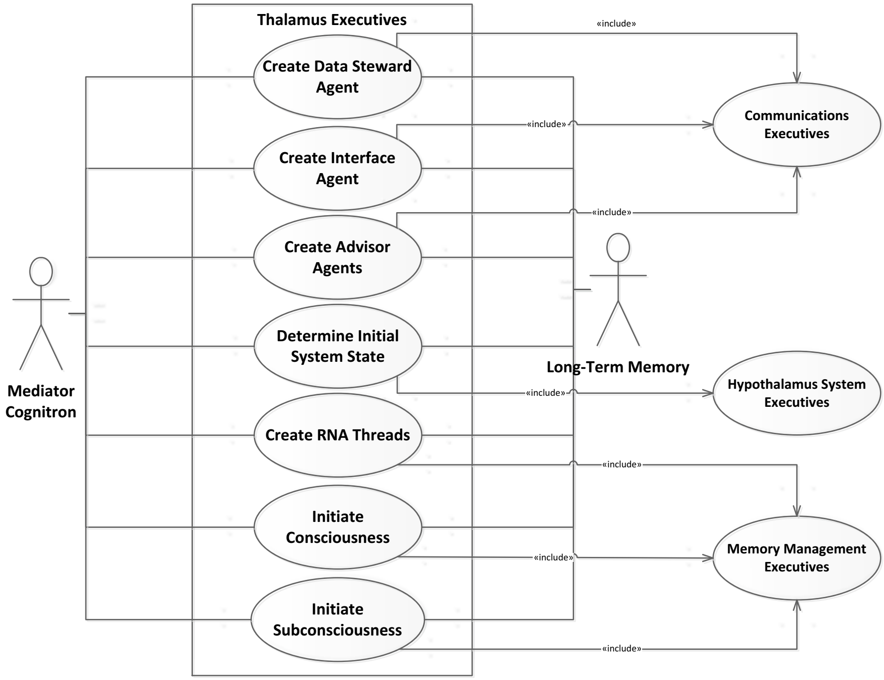


Fig. 10.7 Thalamus executives

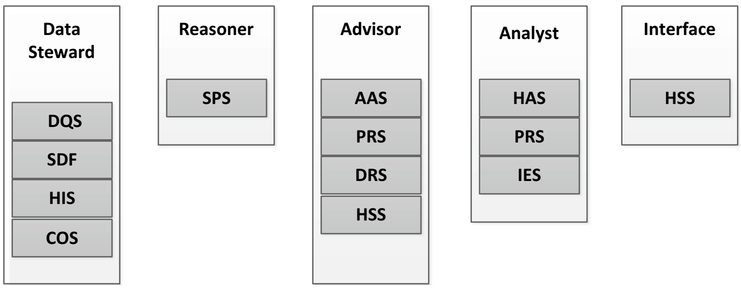
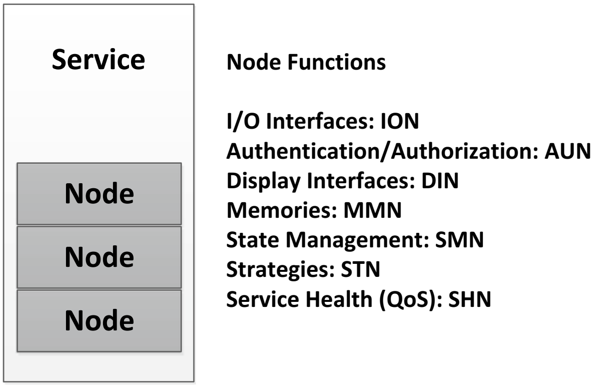


Fig. 10.8 Minimum required Cognitron archetype services

As explained, each service consists of one or more pluggable Nodes that define the service capabilities within the Cognitrons. These pluggable Nodes provide a modularized service architecture that allows the SELF cognitive framework to create new Cognitrons and new services with new purposes. The collection of Nodes within each service defines its overall functionality. As with Cognitron archetypes, each service has a minimum number of nodes required for it to be considered a particular service type. Figure 10.9 shows the basic Nodes available to each service; with each Node providing functions or capabilities specific to that Service.

Fig. 10.9 Basic Cognitron service nodes



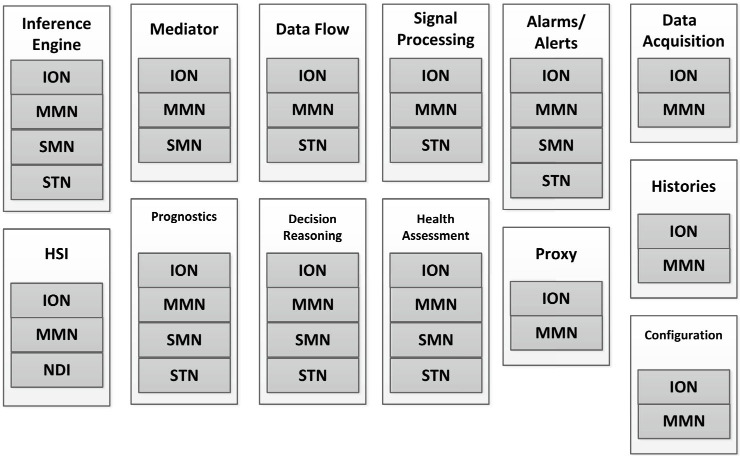


Fig. 10.10 Minimum nodes required for each Cognitron archetype

Nodes, as explained, provide service capabilities. A family of nodes constitutes a service plugin. Figure 10.10 illustrates the minimum nodes required for a Cognitron service to be called a particular service type. A service may contain more than the minimum nodes, but must contain at least the nodes shown in Fig. 10.10

10.3 Cognitron Personalities

The high-level features of the Cognitron architecture are enhanced by the evolutionary processes embedded within a SELF ISAAC cognitive framework and implemented utilizing functionally distributed capabilities provided by the PENLPE infrastructure [78, 80]. This enables Cognitrons to learn and evolve over time.

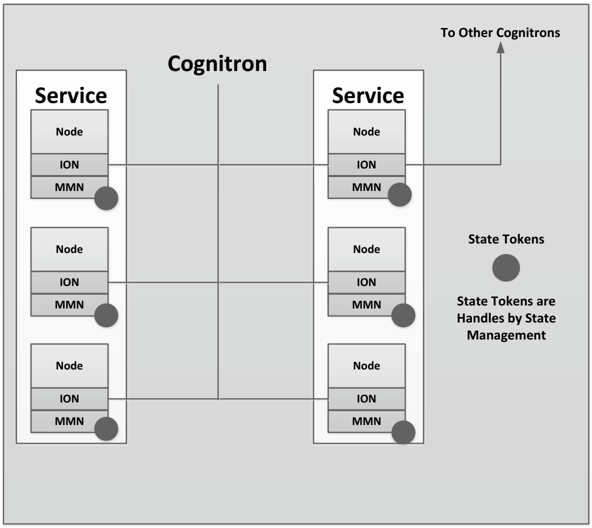


Fig. 10.11 Cognitron personality tokens

However, evolution of simple stored information and the continuity of operations most systems employ today are not sufficient for a self-evolving system. When humans are exposed via injury or disease to brain damage many undesired outcomes follow. A system failure for a SELF would similarly result in the loss of learned and evolved behavior across a current set of evolutionary Cognitrons. To eliminate this potential problem, Cognitrons store personality tokens, which capture the current machine arousal states (Sect. 3.4) of a Cognitron to envelop the essence of stored memories. A personality then is a collection of state information carried in a "personality token" that describes the knowledge and context using knowledge relativity threads [229] and representing that given body of information for an operational Cognitron [198] and the rest of the operating SELF. Figure 10.11 illustrates the personality token concept. Cognitron personality tokens are stored within a SELF's memories, thus allowing Cognitron personalities to be cloned and distributed. Once a "clone" Cognitron has been created, the two Cognitrons could learn and evolve differently; however, this allows a SELF to create new Cognitrons that also carry the memories of previous Cognitronic operations. Hence, allowing for system regeneration in case of system failure without loss of evolutionary learning within an artificially generated cognitive architecture [99].

In this way, Cognitronic personalities become mobile and "state mobility" enables Cognitrons to evolve self-deterministically [81, 198]. Self-determinism in humans is generally still governed by varying degrees by culture, laws, and general policies. Similarly, policy management for distribution of Cognitron updates and state token mobility within a SELF processing infrastructure is handled via Interface Cognitrons. The personality of particular Cognitrons is partially based on their need to cooperate, learn, and function autonomously within a SELF, as well as, determined by immediate and codified objectives, directives, and system health.

10.4 Cognitron Flexibility

As explained, all services are flexible, pluggable modules that enable Cognitrons to add and/or replace as required, based on the current needs, goals, directives, etc. The addition of services beyond the minimum required set is based solely on SELF needs. A simple example might be for a Cognitronic History service to keep track of new changes within an interface Cognitron, even though this History service was originally not required. As explained, a plug-in hierarchy exists, with core Cognitron code as the parent, services are children, and nodes are child function primitives to services. Both the Cognitron code and the service code implement plug-in management features within their core components, managed by the overall PENLPE cognitive management framework. However, the core software enabled elements that allow for SELF Cognitronic flexibility are the underlying, simple KRT representations of normalized unstructured information content which pervade the SELF artificial Cognitronic architecture and the flow of scalable, parallel recursion. Additional, flexibility and speed is enhanced via shared memory access, and storage and via minimization of instruction and software stack.

Serially, examining Cognitronic flow, a core Cognitron Archetype would pull services and node plug-ins from Mediator Cognitron archives on a per need basis. Next, local Mediator Cognitrons would keep a local copy of available plug-ins, while the Host Mediator Cognitron keeps global copies that are pulled by the local Mediator Cognitrons. The Host Mediator Cognitron and Local Mediator Cognitrons keep metadata about plug-ins so that the Mediator Cognitrons can determine when to upgrade their local plug-in repository. The Host Mediator can then be updated via Cognitrons, based upon evolution within the cognitive infrastructure, or additional information from external sources. The Host Mediator may send out alerts to Local Mediators that updates are available, allowing the Local Mediators to schedule updates within their local cognitive structures.

Services within a single Cognitron communicate via interface nodes plugged into the local Cognitron's management system. Figure 10.12 illustrates this. Services communicate with other services outside of a given Cognitron using a Cognitron availability and context registry, by querying Cognitrons for discovering adjacency, finding the appropriate Cognitron types, and for generating external messages routed to remote Cognitrons through the Cognitron's local Data Flow service.

10.4.1 Mediator Service

Earlier we discussed that Cognitrons could take on the role of Local Mediator or Host Mediator. These roles are implemented through inclusion of the Mediator Service. The Mediator Service provides the messaging protocols that allow mediation between the Mediator Cognitrons and other types of Cognitrons within a SELF framework. Mediation functionality includes the storage of long-term memories

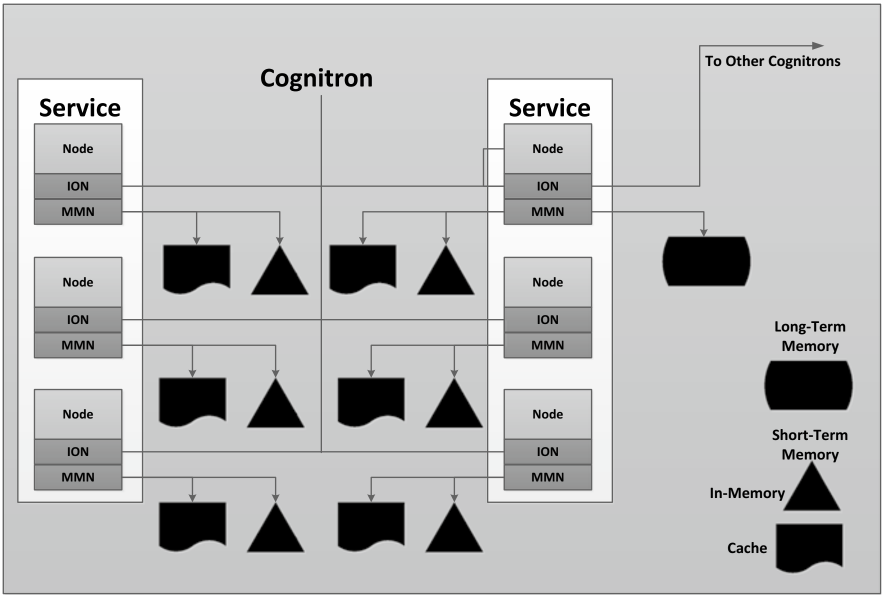


Fig. 10.12 Inter-Cognitron communication

that utilize FUSE-SEMs and the processing for recombinantly assimilated Knowledge Relativity Threads (KRT). Mediator Cognitrons gather and provide information and questions posed to Reasoner and Analyst Cognitrons. Mediator Cognitrons also provide provisioning capabilities to other Cognitrons. Provisioning involves the storage of dynamically loadable service and node modules, and the distribution of loadable module to Cognitrons.

10.4.2 Data Acquisition Service

The Data Acquisition Service forms the core of the Data Steward Cognitron functionality. The Data Steward Cognitrons provide internal and external interfaces and are the main SELF interface to its sensors. The Data Steward Cognitrons provide pre-processing of incoming sensor data (temporarily stored in Sensor Memory), applies context using KRTs [229] and utilizes metadata for characterizing the raw sensor inputs, then repackages and/or reformats the data for use by other Cognitrons, and then stores the data in a SELF's short-term memories. The Data Acquisition Service notifies Data Flow services of the availability of the data, handing off responsibility for further processing the Data Flow services. The Interface Cognitrons support both existing and future sensory sources via the nodal plug-ins and nodal strategies (Fig. 10.13).

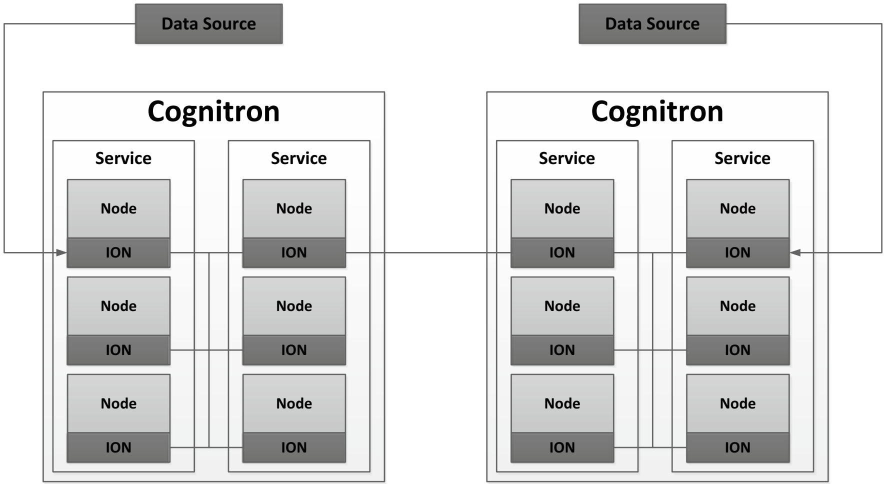


Fig. 10.13 Inter-Cognitron sensor nodal strategy plugins

10.4.3 Signal Processing Service

The Signal Processing Service utilizes core strategies within the service nodes, discussed later in this chapter, for dealing with various aspects of data/information processing and manipulation across a SELF. This includes, but is not limited to:

- Feature Extraction
- Data/Information Characterization
- Data/Information Clustering
- Pre-Processing
  - Data/Information Identification (source and type)
  - Feature Identification
  - Error Detection (including anomalies)
- Sensor Specific Feature Extraction
- System Health Processing and Characterization
  - Diagnostics
  - Prognostics
- Data/Information Fusion
- Data/Information Classifiers
- Inferences

Signal processing services are utilized by Reasoner and Analyst Cognitrons, but may be utilized by other Cognitrons, depending on the domains and environments under which the SELF is being deployed. The Signal Processing Services provide the reasoning strategies that enable decision support and used by PENLPE for cognitive management, as well as use by the Artificial Prefrontal Cortex for resource management. The Dialectic Argument Structure discussed earlier makes extensive use of Reasoner and Analyst Cognitrons with a variety of Signal Processing Services that contain a variety of node plugins and node strategies.

10.4.4 The Data Flow Service

Data Flow Services handle the routing of Data/Information to Cognitrons that have already been processed by the Data Acquisition Services. This may including include causing data/information to be moved through other Cognitrons which are proxy agents moving data to other Cognitrons. Data Flow Services also provide scheduling of data/information movement between Cognitrons so that Cognitrons can receive data/information when needed.

Access to archived data/information from Data Acquisition Services is an asynchronous process to the acquisition of data/information. To support the asynchronous nature of information flow within a SELF, Data Flow Services handle requests from other services for data/information, which allows Data Acquisition Services to be unburdened by incoming Cognitron requests while data are flowing through the Data Acquisition Services. This “separation of concerns” allows the highest data flow rate within a SELF.

Data Flow Services also handle communication between Cognitrons. Each Cognitron’s neighbor configuration determines the remote Cognitrons to which data can be sent to or from, and which data can be received. Services within each Cognitron pass data to Data Flow Services with destination information and the Data Flow Services provide transfer of this data/information.

10.4.5 Alerts and Alarms Service

All services within a SELF generate alerts and alarms. The Alerts and Alarms Services provide a distribution point for these information objects within a SELF framework. Various services within a SELF register for notification of specific alerts and alarms. Services also deliver alerts and alarms to the Alerts and Alarms Services for subsequent distribution throughout a SELF framework.

Multiple alarms from several cognitive components about many unknown data/information sources within a short time period might trigger a serious warning alert rather than a precautionary alert. The priority and severity of the alert will drive the cognitive infrastructure to different reactions, based on the internal Autonomic Nervous System state that results from the alerts. Alerts, therefore, carry severities, priorities, age, parentage, and history parameters/information that are transmitted through the cognitive infrastructure.

Multiple alarms from several cognitive components about many unknown data/information sources within a short time period might trigger a serious warning alert rather than a precautionary alert. The priority and severity of the alert will drive the cognitive infrastructure to different reactions, based on the internal Autonomic Nervous System state that results from the alerts. Alerts, therefore, carry severities, priorities, age, parentage, and history parameters/information that are transmitted through the cognitive infrastructure.

The Alerts and Alarms Services provide registration facilities to other services. Registration allows a service to be notified when a requested alarm or alert has been submitted by another service. The Alerts and Alarm Services pass their alarm and alert objects to services that have registered for notification. The meaning that is applied to the alarm or alert is the responsibility of the service that receives the notification. The collection of available alerts and alarms is maintained by the Alerts and Alarms service. The collection of alerts and alarms has an initial static definition that is extended by services that register for their own collection of alarms and alerts within this service. The definitions also evolve as the system evolves. Once a new alert or alarm has been asserted by a SELF ISAAC cognitive framework, it is officially added and the definition is published to allow services to register for notification.

10.4.6 Health Assessment Service

The Health Assessment Service is an internal monitor of overall SELF system health and status. This service takes on many forms within a SELF ISAAC cognitive framework, depending on the cognitive level it is operating within. There are Cognitrons that contain Health Assessment Services to monitor the health and status of cognitive functions as well as hardware and overall service health. These include, but are not limited to:

- Lower cognitive brain functions (discussed earlier),
  - Information flow (heartbeat)
  - Sensor availability (sight, hearing, etc.)
  - Power availability (breathing)
  - Etc.
- Higher cognitive functions
  - Artificial Prefrontal Cortex
  - Artificial Neocortex
  - Artificial Hypothalamus
  - Etc.
- Hardware
  - Processor utilization
  - Memory utilization
  - Internal network utilization

- Services
- Memories (e.g., short-term, emotional, long-term, etc.) stability and read/write availability and timing

In terms of service health, the Health Assessment Services keep track of predefined node and service states for Cognitrons in which the services reside. If a monitored service or node state violates its defined criteria, the appropriate notices are sent to the Alerts and Alarms Service for distribution.

As described earlier, internal models of a SELF cognitive system are kept and scenarios run against them to ensure the system understands itself and potential changes. System Health fault scenarios are also run and the resultant faults are tracked, and alert and alarm triggers are created for these possible fault conditions within a SELF cognitive ecosystem.

All services, as well as nodes within each service, implement a common set of states as well as node/service specific states. Groups, or coalitions of Cognitrons, together form a given cognitive functionality within a SELF (e.g., hypothalamus) and also implement a common set of states which evolve as the system evolves. Each Cognitronic entity that creates a state also provides the validity tests for these states. The Health Assessment Services at each level of cognitive assessment calls those validity tests to perform health assessments. Additionally, there are specific tests for each Cognitron archetype. Cognitron specific tests are run to validate the overall state of each Cognitron. These include tests that verify required services and nodes within the specific Cognitron are available and functioning. All Health Assessment Services must register their monitoring results with the Alerts and Alarms Services.

10.4.7 Inference Engine Service

Data/information within a SELF cognitive processing infrastructure is collected from a variety of sensors as illustrated before in Fig. 9.10. The diverse sources of information that may come through the sensors may not (and often do not) have consistent contextual structures; each introduces ambiguity into the correlation, analysis, reasoning, and inference processes that may be applied to such data/information. A SELF cognitive processing framework, like the ACNF, provides the computational and processing capabilities to organize the incoming sensor information semantically into meaningful fuzzy concepts and Binary Information Fragments that allow the Dialectic Search Argument processing to create cognitive hypotheses as part of a SELF's overall cognitive topology.

Humans utilize fuzzy language and communication, adapting and evolving the way they process to best fit the needs of personal and conceptual views [7]. These views are based on individual and communal goals and visions gained over time and with personal and communal experiences. The Cognitron Inference Service addresses the problems of autonomous information processing by communicating concepts fuzzily between processing components within a SELF ACNF (Cognitrons, their services, and their nodes), in order to adapt a SELF cognitive processing framework to a changing, real-world, real-time environment.

An Inference Engine Service utilizes cognitive processes formulated and embedded within a SELF Genetic Neural Fiber Network, and the Stochastic Decision Making algorithms discussed earlier work toward the goal of minimizing ambiguity and maximizing clarity; while simultaneously achieving the necessary results. The Inference Engine Services is designed to perform fuzzy, possibilistic reasoning and analysis, based on fuzzy rules and inferences given in symbolic form [65, 66]. The Inference Service supports evaluation of:

- Fuzzy inferences
- Fuzzy and/or/not operations
- Arbitrary nesting and chaining of fuzzy expressions
- Multiple assignment operations
- Pre-defined and domain defined hedges
- Unconditional assignment operations
- Dynamically weighted values
- Evaluation of single rules or entire blocks of rules.

The Fuzzy, Possibilistic Inference Engines are structured and used in a variety of ways throughout a SELF ISAAC cognitive architecture [67]. This includes the use of Cognitrons in Reasoner and Analyst configurations. These are utilized within the FUSE-SEMs, the FUNNs, the Genetic Fiber Network, and the DAS; all within the ISAAC cognitive framework and a SELF's ACNF processing infrastructure. The Inference Engine Service parses and evaluates rules that have been registered with the inference engine and are provided through service nodes that implement the strategies related to fuzzy inference processing. There are a variety of fuzzy inferences that are utilized within a SELF cognitive processing system. Standard inference implications that are utilized are:

- Dienes-Resher Implication:

$$\mu_R(x,y) = \max\left(1-\mu_A(x), \mu_B(y)\right) \tag{10.1}$$

- Zadeh Implication:

$$\mu_R(x,y) = \max\left(\min\left(\mu_A(x), \mu_B(y)\right), 1-\mu_A(x)\right) \tag{10.2}$$

- Lukasiewicz Implication:

$$\mu_R(x,y) = \max\left(\min\left(\mu_A(x), \mu_B(y)\right), 1-\mu_A(x)\right) \tag{10.3}$$

- Gödel implication:

$$\mu_R(x,y) = \begin{cases} 1 & \text{if } \mu_A(x) \le \mu_B(y) \\ \mu_B(y) & otherwise \end{cases} \tag{10.4}$$

In addition, there are specific inferences that are utilized to provide the type of abductive heuristics utilized in human reasoning. The DAS, within the DART Occam Abduction framework, works to find inferences that both support and rebuts the genetically generated Occam hypotheses. Given the abductive nature of human implication logic required to hypothesis testing and inferencing, specific fuzzy, abductive inferences are required to assess the "rebutting" nature of the DAS. Four specific abductive fuzzy inference implications are utilized within by the Inference Engine Services are:

- Abductive, Fuzzy Modus Tollens
- Abductive, Fuzzy Modus Ponens
- Abductive, Fuzzy Inverse Modus Ponens
- Modified Fuzzy Abduction

These are all specific forms of implications that have been adapted for a SELF from a more general type of human Experience-Based Reasoning (EBR) implications, known as Condition-Based Reasoning (CBR) [161, 182]. CBR emulates human abductive and deductive reasoning which simulates the experimental Occam Reasoning principles, and therefore are useful and appropriate within a SELF's cognitive processing framework. The purpose of these Abductive Inference Implications is to provide experimental case reuse and case retention within a SELF ACNF Conceptual Ontology and the Inference Engine Services within the DART ISAAC cognitive architecture. For this application, we utilize the notion of CBR that refers to the notion that past experiences, combined with a given set of observations and inferences, can be used to provide an "influence," or starting point for the DAS genetic hypothesis generation and testing use to explain a similar, but not identical set of observations [120]. These abductive, fuzzy inference implications follow from the Occam abductive principles described in Chap. 8:

> Similar observations and similar problems probably have similar solutions or explanations. [73]

10.4.8 Prognostic Service

We briefly mentioned prognostics as part of the capabilities provided by the Health Assessment Services, and basic prognostic capabilities are included in the Health Assessment Service nodal strategies [55]. However, cognitive prognostics is such an integral part of the overall health of a self-assessing, autonomous, artificially conscious SELF, that a separate service is required to accommodate comprehensive health understanding and predicting cognitive health issues within the ISAAC cognitive architecture. Here, the Prognostic Service nodal strategy capabilities allow an assessment of a component, coalition, or entire cognitive system's current and future health predictions. There are two variations of this prognostic health determination. The first is a short-term prediction that is needed to understand if the component being assessed will be able to complete its current task/mission.

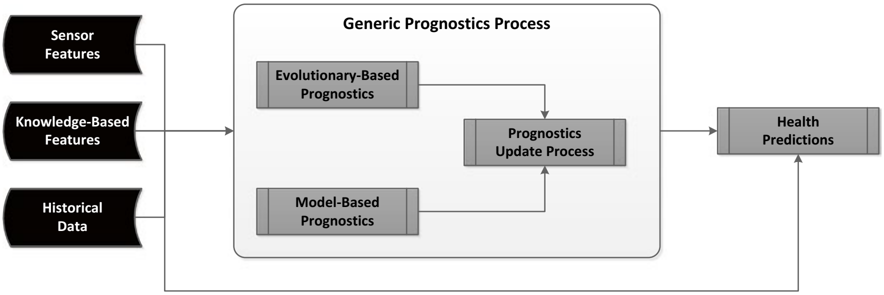


Fig. 10.14 SELF Cognitron prognostics service architecture

The second assessment is to perform prognostics to determine whether the component will be stable long-term, or what is the Remaining Useful Life of the component, and when is failure or problem likely to occur, what conditions may exist to cause the failure, and how can it be measured.

The Prognostic Services provide specific information to Cognitrons and, through communication, to the Host Mediator Cognitron. Figure 10.14 illustrates the information flow and high-level architecture of the Prognostics Service.

10.4.9 Decision Reasoning Service

The Decision Reasoning Services within a SELF serve many purposes within the ISAAC cognitive processing framework. The main purpose of Decision Reasoning Services is to provide recommended actions (behavior selections), decision alternatives, and the implications associated with each action or behavior. Recommendations may include effectors to accomplish a mission, maintenance action. A SELF user or developer is unlikely to want a SELF with pending cognitive issues or pending cognitive problems when it is supposed to be performing to specific goals and objectives.

Decision Reasoning Services within Advisor Cognitrons take into account histories from History services within the Cognitron structure. These include current and future mission or task profiles, high-level objectives, low-level “life” objectives, and resource constraints provided by the Cognitive Economy processes. Actions, behaviors, and implications are passed to the Signal Processing Services within Cognitrons as feedback for cognitive condition monitoring. The information is also communicated to Interface Cognitrons to broadcast the information internally, and externally, if needed.

10.4.10 Histories Service

A SELF History Service provides pedigree information to various processes and services within a SELF ACNF processing framework. They are used by Design Reasoning Services, Prognostic Services, etc., based on usage, maintenance, and current situational awareness within a SELF ISAAC artificial cognitive framework, based on a particular domain. These histories are information that may have been provided as a SELF was being constructed (manufacturer information, etc.) or may be information collected during the initial cognitive testing of a SELF. History Services take information, data, queries, DAS hypotheses, and correlate them utilizing the FUSE-SEMs to provide contextual information based on relevant situational analysis and parameters (metadata).

10.4.11 Configuration Service

The Configuration Services provide a single point of management for Cognitron configuration. Incoming configuration requests for Cognitrons from the Host Mediator Cognitron or other cognitive component within a SELF ISAAC cognitive framework are routed to the appropriate Cognitron services and nodes. Requests for configuration information are pulled from services and nodes and then routed back to the requesting process. Configuration requests are transmitted throughout a SELF framework utilizing a Cognitive communications network that is specifically created for a SELF cognitive framework, based on the environment, hardware architecture, and overall system configuration.

10.4.12 Human Systems Interface Service

Since the purpose of a SELF is perform services and tasks for humans, there are times when it is necessary for a SELF to interface with humans, other SELFs, or some other type of system. The Human Systems Interface Services provides an interface for a SELF to interact with humans for informational or collaborative purposes. Other services within the Cognitron services allow for external interface data/information/knowledge to be routed to other services within the Cognitrons or to external Cognitrons.

Data export capabilities within a SELF cognitive infrastructure (effectors) are the only mechanisms within a SELF for data/information/knowledge to leave a SELF. Data/information/knowledge can enter a SELF through sensory information communicated through Data Steward Cognitrons and can be generated from within any of the components of a SELF. Human System Interface Services provide direct access a SELF Host Mediator Cognitron, enabling the ability to manage HRI interactions.

10.4.13 Proxy Service

It is possible for a SELF to be implemented across a wide area processing environment, either on a LAN, or implemented across a geographically diverse WAN. The architecture allows for this implementation configuration.

Under these circumstances it may be necessary for one Cognitron to need a routing path to another one. A SELF Proxy Services allows a Cognitron to act as a routing path for another Cognitron. Cognitrons make a request to the Proxy Cognitron, who forwards the request to another Cognitron. The response is then routed back to the original Cognitron.

10.5 SELF Service Node Strategies

We described strategies above, as providing the functional capabilities to the service nodes within a Cognitron. The strategies provide the algorithmic "how" for the functionalities and capabilities within a SELF Cognitrons. As explained, strategies are carried within the service nodes and provide the nodes with their abilities, based on the context of the service nodes. There are two main categories of nodal strategies within a SELF:

- Domain Independent Strategies: these provide the basic cognitive abilities required for the SELF.
- Domain Dependent Strategies: these are very specific to the domain, environment, or mission of the SELF and cannot be known a prior.

Domain Independent service nodal strategies provide a SELF with those cognitive and processing capabilities that are required regardless of the domain within which a SELF is utilized. A SELF's Domain Independent strategies include, but are not limited to:

- Resource Management
- Rules
- Inference Implications
- Learning
- Decision Management
- Memory Creation
- Memory Integration
- Memory Construction
- Cognitive Needs (e.g., Lower Brain Function Executives)

With the understanding that within each category of strategies listed above, there may be up to hundreds of individual strategies that fit within that category that are used by service nodes within a SELF.

10.6 Discussion

We have described software structures that facilitates a SELF's artificial cognitive processes. Chapter 11 will discuss the physical architectures that are necessary to provide the processing, memory, and network infrastructures to run a SELF's cognitive software.

Chapter 11
SELF Physical Architectures

As we have discussed throughout the book, a SELF is a hardware/software artificial cognitive system designed to mimic human reasoning, learning, and understanding. The first ten chapters have concentrated on the cognitive side of the software architectures and frameworks to accomplish artificial consciousness and artificial human cognitive skills [9, 10]. However, the next few chapters will focus on the pragmatic computer software and hardware architecture upon which the cognitive software functions will operate and be processed. Computer processing units, electronic memory devices, and information networks are requirements in order for the cognitive software to exist, operate, and function.

Architectures and frameworks discussed in previous chapters are have complex functions and somewhat complicated dynamically modifiable and real-time collaborative software frameworks that require appropriate software stacks and performance based hardware to support major communications channels and accommodate loose coupled, ubiquitous Cognitrons that drive a SELF's cognitive processes. We describe an approach to a flexible, scalable, modular software and hardware processing architecture that not only accommodates a SELF today, but can be as much as possible dynamically reconfigurable and evolvable to accommodate SELF autonomy. We will illustrate and describe our current approach and discuss future plans to evolve the architecture.

11.1 The Reconfigurable Advanced Rapid-Prototyping Environment (RARE)

The RARE hardware architecture is a scalable signal processing and computing architecture that utilizes state-of-the art high-performance general purpose processors, Field Programmable Gate Arrays (FPGAs), and flexible I/O fabrics to facilitate processing systems where conventional back-plane/blade server applications

J.A. Crowder et al., *Artificial Cognition Architectures*,
DOI 10.1007/978-1-4614-8072-3_11,

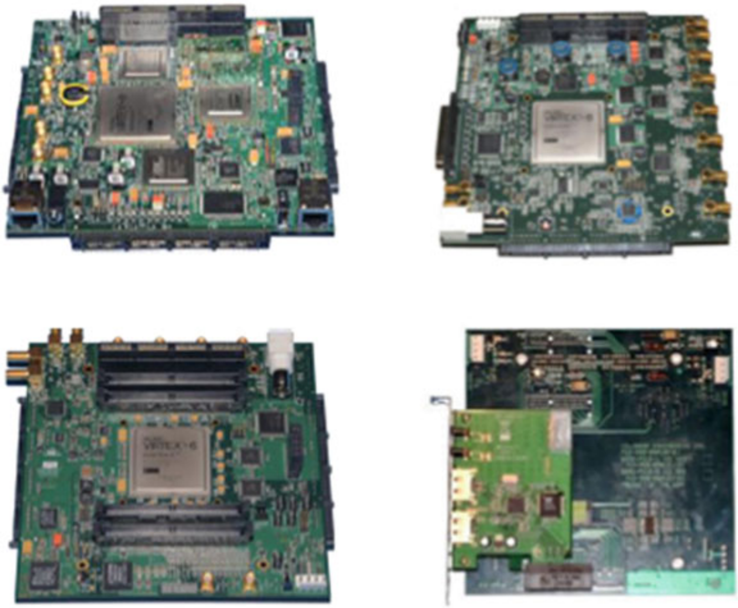


Fig. 11.1 RARE processing boards

are not viable solutions.[1] A SELF is one such application. For a SELF to be an autonomous land/sea/air/space cognitive system, a SELF hardware/software infrastructure must be flexible to accommodate a variety of size and form configurations. The RARE hardware architecture provides a compact, low-power, light-weight processing framework that is ideal for implementation of a SELF within an unmanned vehicle that is resourced constrained. RARE[2] provides an 'out-of-the-box' hardware architecture approach suitable to a SELF. Figure 11.1 shows a set of RARE processor and FPGA boards.

The heterogeneity of processing elements within a SELF cognitive processing framework demand a flexible, scalar, and modular footprint that allow general purpose/multi-core processors, FPGAs, and possibly ASICs (Application Specific Integrated Circuits) that might be used for specialty processing (an "expert" within a SELF ISAAC/ACNF cognitive processing architecture). The ability, within the cognitive evolution, to be able to evolve the FPGA configuration autonomously provides the capability to operate cognitive portions of the artificial cognitive framework at hardware speeds, not software speeds.

[1] © Colorado Engineering, Inc., Nancy Scally, CEO.

[2] The RARE computing technology was sponsored by the Missile Defense Agency under the Small Business Innovation Research (SBIR) program awarded to Colorado Engineering, Inc.

11.2 Physically Modularity and Scalability

One of the many uses for a SELF is for autonomous operations (e.g., deep underwater, deep space, intra-corpus, etc.). Many systems including systems embroiled in autonomous operations, require considerations for hardware architecture of size, weight, and power, (aka. SWaP). A SELF requires additional considerations due to the need for rapid resource availability and scalability in real-time operations to mimic brain like cognitive functions.. The use of a modular, scalable, heterogeneous processing architecture like RARE provides very high performance characteristics, while providing the low SWaP profiles needed for extended autonomous operations. The RARE technology concept is shown in Fig. 11.2, which illustrates RARE processor boards connected in a variety of configurations.

A unique characteristic of RARE processor boards is that they can be connected in a full three-dimensional configuration, illustrated in Fig. 11.3, enabling full three-dimensional cross cube communication capabilities. This allows SELF Cognitrons to communicate rapidly and efficiently to any other Cognitron within the processing system, and enables SELF control authorities (e.g., Artificial Prefrontal Cortex) to communicate more easily throughout a SELF cognitive framework.

For use with a SELF, each board within the 3-D configuration would carry up to 64 GB of RAM, and up to 0.5 TB of flash memory to accommodate the cognitive processing, storage, and temporary usage of the Cognitrons and other cognitive components within a SELF cognitive framework. This allows a SELF to be decomposed into functional embedded processing components within a SELF cognitive processing framework which are loosely coupled to a SELF's high-level common operational environment.

The flexible, scalable, but well-defined interfaces result in an overall SELF hardware environment that provides a scalable, flexible, and heterogeneous, embedded

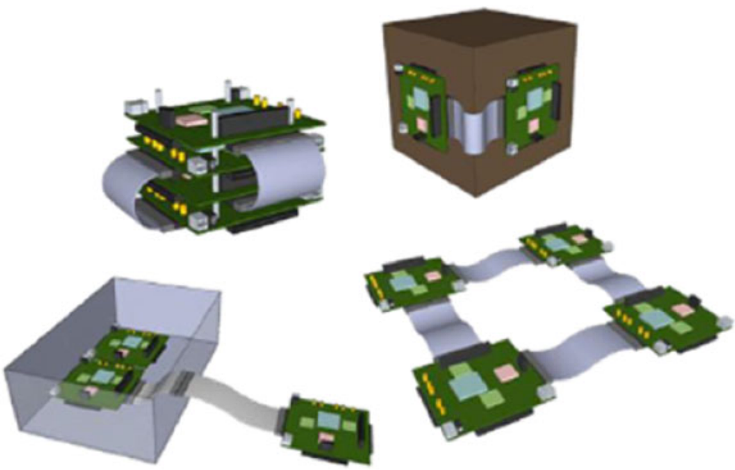


Fig. 11.2 RARE processor board connectivity

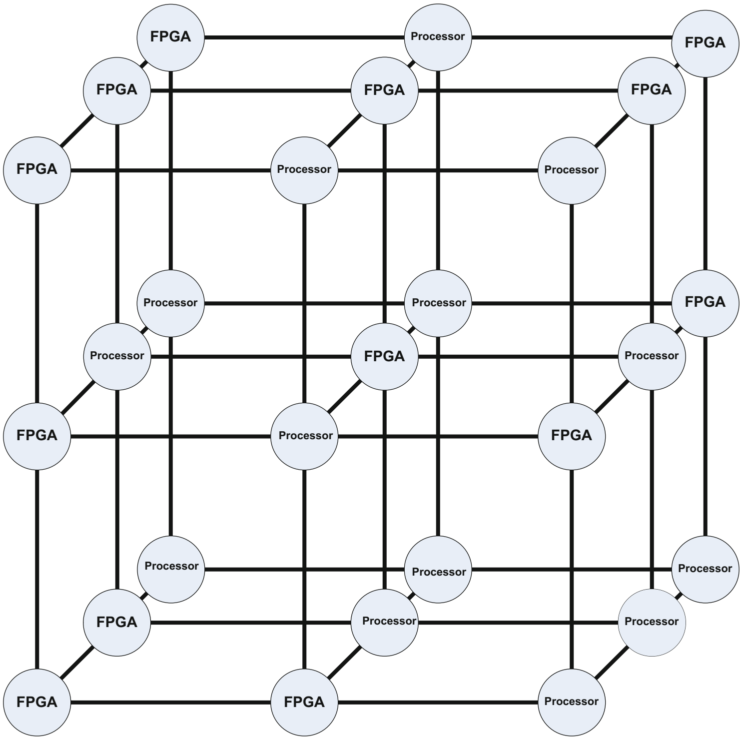


Fig. 11.3 RARE 3-D connectivity

processing hardware framework that can scale in a modular hardware fashion to enable a SELF, learning, self-evolving cognitive processing software. The advantages hardware architectures like the RARE architecture bring to a SELF are:

- They balance state-of-the-art general purpose high-performance processors with high-performance specialty processing (e.g., FPGAs, ASICs) providing processing and storage capabilities enabling SELF cognitive processing.
- They support an I/O fabric that promotes throughput flexibility, overall system and component-level scalability, as well as interoperability between different cognitive processing components.
- Contains A/D conversion capabilities to support analog neural processing frameworks that may be required, depending on a SELF domain.

11.3 Discussion

Cyber security research and artificial intelligence research have moved down separate paths over the last few decades. Those interested in artificial intelligence have been focused on how to create software structures that mimic human thought and reasoning, while cyber security has been aimed at understanding and closing vulnerabilities in IT computing infrastructures. Unfortunately, all artificial intelligence involves the use of IT software and hardware infrastructures and therefore is as vulnerable as any other system, although we contend more vulnerable, since intrusion, corruption, and other cyber security issues will have devastating effects on a self-aware, cognitive, learning, and reasoning SELF. This is especially true when a SELF is designed to self-adapt and therefore not completely predictable. Since one aspect of cyber security involves monitoring IT infrastructures for unpredictable behavior, capturing cyber security breaches within a SELF that is, itself, not completely predictable is a daunting task for Information Assurance frameworks. Chapter 12 will address cyber security architectures and software structures for a SELF, aimed at self-assessment, self-regulation, and self-healing within the ISAAC cognitive framework.

Chapter 12
Cyber Security Within a Cognitive Architecture

As with any electronic information processing system in today's world of hackers, malware, spyware, etc., security is a major component of the overall operational capabilities of a SELF. All information within a SELF must be protected and kept from corruption (whether accidental or intentional). Accidental corruption of information and knowledge within a SELF is handled through continual cross-checking and self-assessment within the ACNF framework. Continuous communications between Cognitrons within the system and constant refresh of memory information keeps information from being arbitrarily modified (loss of bits) and from corruption due to memory failures and catastrophic interference problems discussed earlier. However, these do not protect the system from intentional corruptions and hackers. Since a SELF is intended to be a fully autonomous, self-evolving, self-learning, reasoning artificial entity, any corruption of information across a SELF's artificial cognitive framework could have devastating effects on a SELF's learning, reasoning, memory, and cognitive processes analoagously to what occurs in injured humans (e.g. head injury, Alzheimers). Corruption or incorrect modifications to a SELF's needs, constraints, goals, memories, or algorithms could cause a SELF to act, evolve, remember, or learn, completely incorrectly and/or out of scope for the intentions of a SELF.

A SELF's cognitive framework must contain Cognitive Security Architecture (COGSEC) to ensure information is only communicated with parts of the cognitive framework that have a need to know the information. That way, any corruption within one part of the system cannot be arbitrarily communicated throughout the system and cause problems. Corruption within the system can be identified, quarantined, and either corrected or removed before serious damage can be done to a SELF. These information filters and quarantine processes are similar to the way human brain filters information and reject wrong or corrupt information. Corrupt information is not the same as information that is contrary, or in rebuttal, to a cognitive process or hypothesis the system is trying to resolve.

Given that our SELF is a self-evolving system, the Cognitrons and even entire Cognitron Coalitions will change and evolve, as will the algorithms within each Cognitron.

J.A. Crowder et al., *Artificial Cognition Architectures*,

DOI 10.1007/978-1-4614-8072-3_12,

The information that is required for each Cognitron module will change over time (i.e., the Cognitron's "sphere of influence" will change over time).

In order to ensure that each information fragment, each memory, each Cognitron is secure, each is encrypted differently. It is not necessary to encrypt each with a lengthy encryption algorithm. Yes if it is a short encryption scheme it would be possible to break the encryption code, but you would need to do that for each of the billions, possibly trillions of information fragments, inferences, Cognitrons, and memories within the system. As Cognitrons evolve and their context and content change, their encryption changes as well. This description drives us toward an n-dimensional, real-time encryption scheme that provides encryption based on a combination of information fragment, topic, need-to-know, and context. However, your level of paranoia determines how dimensional your system becomes and the volume of characteristics you wish to use. This chapter aims to describe a security architecture and encryption scheme to provide security both from outside sources, and security within it SELF, to keep corrupted information from permeating the system and extending problems within the SELF-evolving cognitive framework.

12.1 SELF Cognitive Security Architecture

Any newly designed systems are slave to the actual hardware and software available in the industry at a given point in time. Additionally, there are also standards, conditions and protocols which must be paid homage and utilized to ensure interoperability with existing systems. This is especially true for security, since historically the organizations and accreditation policies are deeply rooted and traditionally have been difficult to adapt to new technologies. Recently, there has been a bit more light at the end of the security tunnel due to the vast changing environment of today. Hence, our SELF's security ontology and architecture is described here and is based on security relationship models described in the National Institute of Standards and Technology (NIST). This along with their existing concepts and capabilities are adapted for use in a self-evolving, self-aware, self-assessing cognitive framework because autonomous self-evolving systems must adapt to adaptive cyber security threats which evolve over time as well. A static architecture or approach will not be effective for a SELF. Figure 12.1 illustrates an Upper Security Ontology for a SELF.

Within a SELF security ontology and architecture, we define a SELF cyber security attribute relationships so the lower security information ontology can be developed. Figure 12.2 shows a SELF's security ontology attribute relationships. Figure 12.2 defines the hierarchy of relationships between security attributes within a SELF information system infrastructure. The assets associated with a SELF's information system/information technology has responsibility to mitigate the risks of security threats within a SELF's cognitive framework. Each risk, vulnerability, and threat must be dealt with within the overall design of the cognitive framework in order to ensure the ISAAC cognitive framework has the capabilities to recognize, learn about, and handle all security-related incidents. The relationships defined in

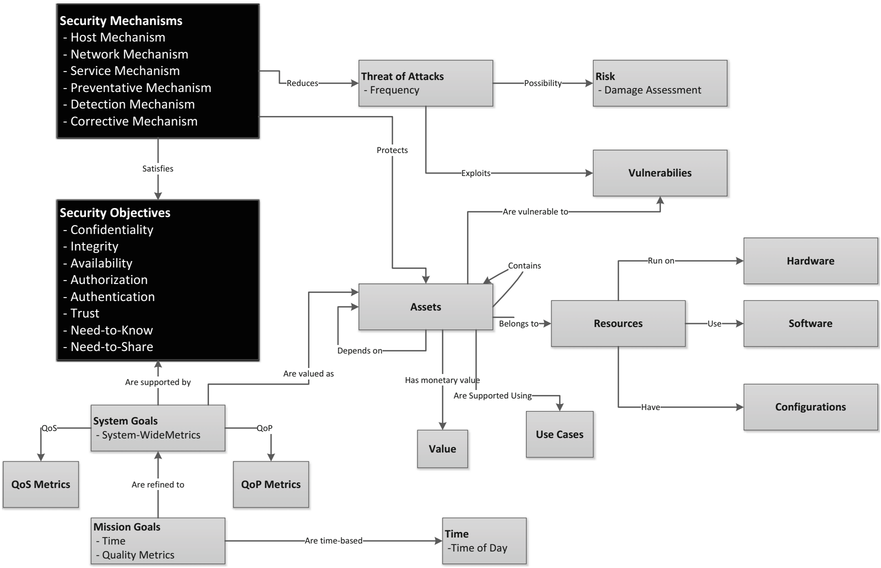


Fig. 12.1 SELF upper security ontology

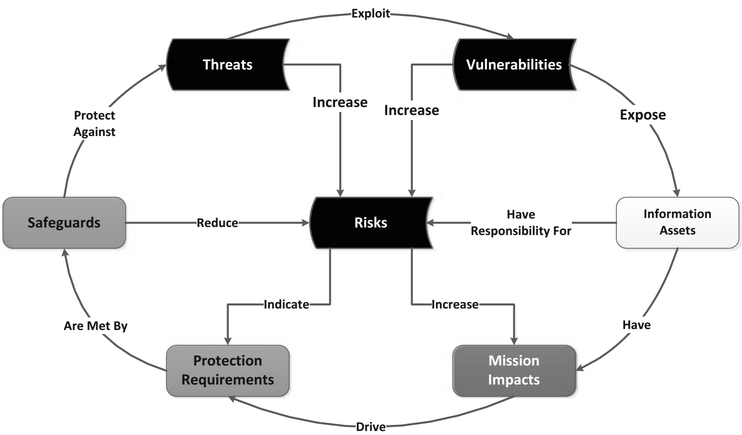


Fig. 12.2 SELF security attribute relationships

Fig. 12.2 lays out how each attribute related to a SELF security infrastructure affects each other. Security specific Cognitrons (Data, Analyst, Reasoner, and Advisor) are utilized throughout a SELF cognitive processing framework to oversee information and cognitive processing at every level within a SELF, ensuring that there are all information, knowledge, information, and data throughout a SELF is free of corruption. A threat within a SELF cognitive architecture gives rise to follow-up threats that must be dealt with rapidly to avoid extended damage to a SELF cognitive ecosystem. Detrimental threats effects could be:

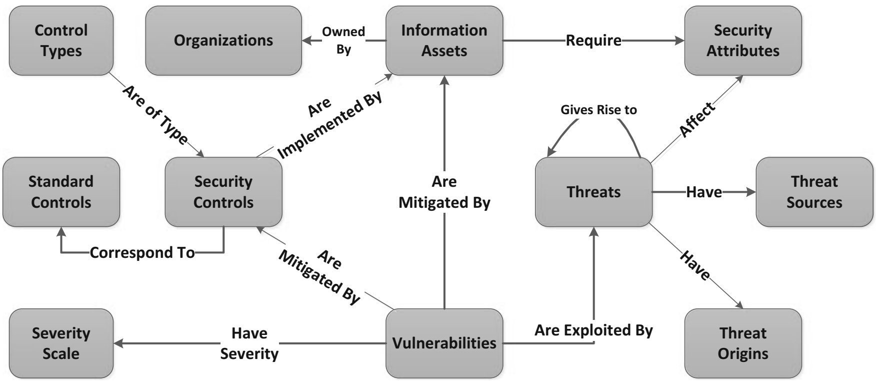


Fig. 12.3 SELF lower security information ontology

- A threat that represents a potential danger to the information assets (memories) within the cognitive ecosystem, and affects specific security attributes (e.g., integrity, availability, etc.).
- Exploitation in the form of administrative weakness within a SELF cognitive system (e.g., damage to the Artificial Prefrontal Cortex).

Additionally, each security threat within a SELF is described by potential threat origins and threat sources (accidental or deliberate). In some cases it will require a DAS to propose potential hypotheses to explain the threats and the reasoning and analysis elements of a SELF would endeavor to solve the security issue. For each of the vulnerabilities, a severity value and a SELF asset on which the vulnerability could be exploited is assigned. A SELF ISAAC and ACNF frameworks must have controls within their architecture to initiate mitigation steps for an identified vulnerability and to protest the respective SELF cognitive assets by preventative, corrective, deterrent recover, and/or detective measures. Figure 10.3 illustrates the Lower Security Information Ontology for a SELF.

Each SELF security control is implemented as a SELF asset concept within the ACNF Conceptual Ontology. Each time a threat is encountered, it is implemented as an instantiated concept within the ACNF Knowledge Base (Instantiated Ontology). Cyber security controls within a SELF framework are derived from and correspond to best practices and information security standards and control to ensure a SELF is compliant with security standard security controls and measures. The controls are modeled on a granular level and evolve within a SELF cognitive framework as the system learns and evolves, since a SELF's internal neural fiber structure, as well as the Cognitrons learn and evolve. These security management controls are implemented within the PENLPE cognitive management structure of a SELF.

In Fig. 12.3, the Security Information Ontology provides a SELF with a working model of security entities and interactions within a SELF's cognitive information processing framework, providing the security domain of knowledge and practices.

Figure 12.3 provides the specification of security conceptualizations, and is used by the security Cognitrons to share knowledge about security-specific objects, events, and relations.

12.2 SELF Cognitive Security Architecture: Threat

Within a SELF security ontological architecture, the Threat Lower Ontology comprises natural, accidental, and intentional possible threats to a SELF, followed by detailed threat sub-classification. An in-depth description for clarity, as well as endangered security objectives are provided for each threat. These detailed descriptions follow the taxonomical structure for:

- Dependability
- Confidentiality
- Integrity
- Availability
- Accountability
- Authenticity
- Reliability
- Safety

Following these taxonomical structures helps develop the security strategies regarding specific attributes. Often the occurrence of a threat gives rise to, or intensifies, other threats; therefore, these relationships are reflected in the ontology. The rates of occurrence for each threat are computed using the fuzzy possibilistic inference engines, and are linked to the threat and location sub-ontologies, allowing a SELF to map location-dependent threat occurrence rates for future reference. Furthermore, each threat exploits one or more vulnerabilities within a SELF cognitive infrastructure. These vulnerabilities are mapped within the vulnerability sub-ontology.

Understanding the relationships between SELF threats and endangered assets is vital for PENLPE to map a comprehensive security planning. Information Assets (memories) are reflected by classes within the ACNF memory infrastructure sub-ontology.

12.3 SELF Cognitive Security Architecture: Vulnerability

Within a SELF, vulnerability is defined as the absence of a proper cognitive safeguard that could be exploited by a threat (internal or external). We have sub-classified the vulnerability sub-ontology into four distinct classes:

- Administrative vulnerability (the ability to manage the cognitive framework)
- Physical vulnerability
- Cognitive vulnerability
- Technical vulnerability

Each of a SELF's vulnerabilities can be exploited by pre-defined threats (which will evolve over time) within the threat sub-ontology and mitigation is achieved by selection of one or more controls which are implemented by PENLPE elements within the ISAAC cognitive infrastructure, Artificial Prefrontal Cortex controls, or within the Cognitron software sub-ontology.

The infrastructure section of a SELF Security Ontology contains a wide range of physical and artificial cognitive elements which are utilized within the cognitive ecosystem (e.g., self-soothing). The security infrastructure sub-ontology provides structural elements which enable the mapping of physical and cognitive environmental elements. Vulnerability severity ratings (critical, important, moderate, and low) enable additional vulnerability classification. Within the vulnerability definition we include a separate relativity thread (KRT) that indicates the corresponding ISAAC cognitive infrastructure element that causes certain vulnerabilities.

12.4 SELF PENLPE Security Management Ontology

Figure 12.4 illustrates a high-level view of the PENLPE security information processing architecture. This information processing environment utilizes many of the elements of the ISAAC and ACNF cognitive resources, as indicated in Fig. 12.4. The overall PENLPE Security Management Ontology is illustrated in Fig. 12.5.

The overall goal is to provide a highly reliable, available and secure cognitive processing environment for a SELF; given a real-time, continuously changing, evolving external and internal environment. Some of the underlying technologies and mechanisms that make this possible within a SELF are the DAS, the FUSE-SEMs, and the Reasoner and Analysts Cognitrons.

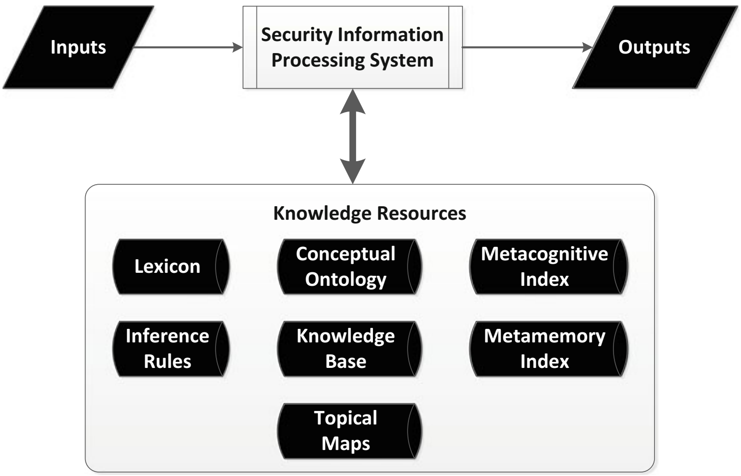


Fig. 12.4 SELF high-level security information processing

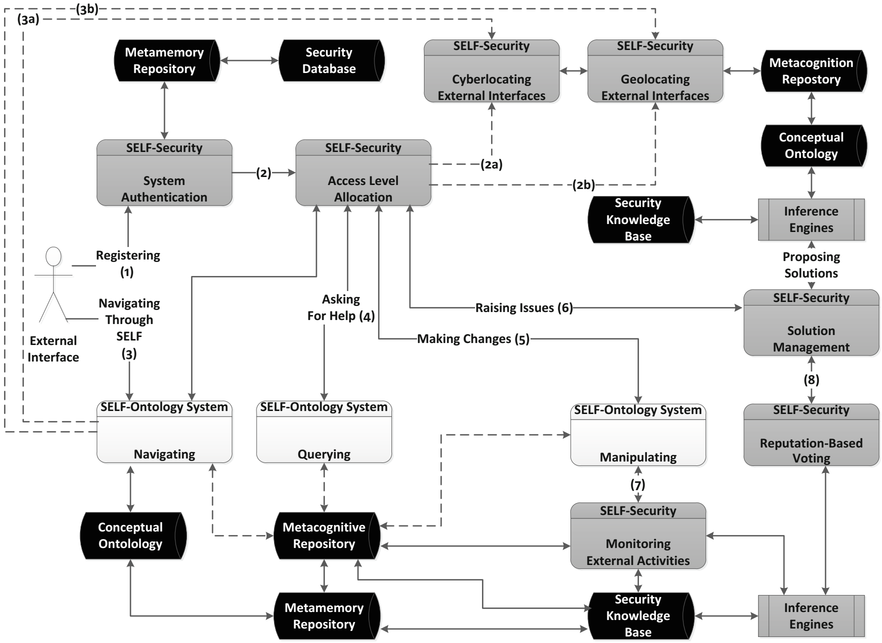


Fig. 12.5 SELF PENLPE security management ontology

The Security Management Ontology illustrated in Fig. 12.5 provides the SELF security vocabulary of terms, specifications of their meaning and includes definitions and indicates how the SELF security concepts are inter-related. Collectively, they impose a structure on the SELF security domain and constrain the interpretation of security terms within the SELF cognitive framework, allowing a common understanding between Cognitrons. Of note, is the Reputation-Based Voting (RBV) processes, where the system inherits its general security implication from information gathered about the reputation of the external entities with whom a SELF is interfacing. In particular, these algorithms allow for the routine production of statistics and fuzzy metrics that are useful for security monitoring purposes, and it provides a coherent cognitive framework to limit intrusion of the system or corruption of information into the system.

12.5 SELF Security Management: Self-Diagnostics and Prognostics

A critical part of a SELF PENLPE security architecture is the use of the Cognitron Diagnostic and Prognostic services. These services include node strategies that allow them to detect security faults in early enough stages to enable a SELF framework to do something useful with the fault information.

Security fault isolation and diagnosis uses detection events as the start of a DAS hypothesis process for fault classification within the cognitive element being monitored. Condition and/or failure prognosis then forecasts the issues and potential future problems; which includes detection and prognosis of time to degradation below an acceptable level (Remaining Useful Life). Specific requirements, in terms of confidence and severity levels, are identified for diagnosis and prognosis of critical failure modes. As a minimum, the following possibilistics are used to specify fault detection and diagnostic accuracy:

- The possibility of anomaly detection within the cognitive ecosystem, including false alarm rates and fault possibility statistics.
- The possibility of specific security fault diagnosis classifications, utilizing specific confidence bounds and severity predictions.

To specify prognostic accuracy requirements, we define:

- The level of condition security degradation beyond which cognitive operations are considered unsatisfactory or undesirable to the mission or current tasking.
- A minimum possibility that Remaining Useful Life of the cognitive element will be equal to or greater than the minimum warning level indicated by the Cognitive Economy algorithms.

The emotional learning and self-soothing techniques described in previous chapters are utilized to provide security prognostics and self-healing within a SELF cognitive infrastructure. The allows a SELF to radically enhance the ability of the system to perform self-assessment and self-prognosis, based on the notion of emotional learning and emotional memories to provide a contextual basis for criticality of security faults and overall system conditions. These may be based on previously learned security and environmental (external or internal) information.

A comprehensive SELF security philosophy integrates the results from the monitoring sensors within the ISAAC cognitive architecture, all the way through to the Reasoner Cognitrons that provides the decision support for the Cognitive Economy algorithms to optimally allocate resources within a compromised SELF. The core component of a SELF overall security strategy is based on its ability to:

- Accurately predict the onset of impending security faults/failures
- Quickly and efficiently isolate the root cause of security failures once failure effects have been observed.

The PENLPE diagnostic and prognostic capabilities require an integrated maturation environment for assessing and validating a SELF's system security accuracy at all cognitive levels within a SELF. This allows for inaccuracies to be quantified at every level within the cognitive ecosystem and then be assessed automatically up through the PENLPE security architecture. The final results are reported up through the cognitive hierarchy to Reasoner and Decision Support Cognitrons for processing and reporting to the Artificial Prefrontal Cortex (Host Mediator Cognitron).

12.6 PENLPE Prognostic Security Management (PSM)

Prognostic management within a SELF cognitive framework consists of the ability to:

- Monitor and predict failures.
- Predict what the security health of a SELF cognitive components/sub-components will be at some time in the future.
- Assess the criticality of this future condition, in terms of a SELF's current mission/tasks and the possibility of mission/task completion.

The role of the PENLPE Prognostic Security Management (PSM) is to:

- Predict what the security health of the SELF cognitive components/sub-components will be at some time in the future.
- Assess the criticality of this future condition, in terms of the SELF's current mission/tasks and the possibility of mission/task completion.

The use of emotional memories within the PENLPE PSM provides:

- The ability to assess the criticality of current and future predicted SELF security in terms of success.
- Aids the speed at which information is provided and transmitted to Cognitron coalitions within a SELF cognitive framework.

This prediction can be for a short-time horizon, or an estimate of the time until a failure or security issue will occur. There are a variety of issues that need to be considered to facilitate these abilities. The PENELPE PSM Cognitrons need to accurately predict into the future, and those predictions are required to be unbiased and to have a small variance in order to be useful. However, the emotional context of the predictions, in light of the context of current cognitive parameters, can help provide insight into the predictions. The emotions states, in terms of self-soothing become important and will be discussed further.

12.7 Abductive Logic and Emotional Reasoners

As we have discussed throughout the book, abductive reasoning allows a SELF to provide explanatory hypotheses. These hypotheses, or new ideas, in this context, are about security faults and cognitive performance indicators. This type of Reasoner Cognitron within the PSM is primarily used at the coalition and higher levels and not at the lower service or node level. We utilize abduction because abduction is the process of forming explanatory hypotheses and is the only logical operation that allows the introduction of new ideas [145]. We utilize abduction throughout the PSM process because abduction allows the examination of a large amount of observations, or facts, and suggests theories as to their cause [161, 162].

In this way we gain new ideas, but there is not force in the reasoning. Deductive logic, or necessary reasoning, is applied on when we need to deduce from the population of hypotheses, each with support and rebuttal information from the DAS, a set of consequences to the effect that if a SELF performs in certain ways, then a SELF will be confronted with certain experiences. In this way we can provide sets of cause and effect concepts within a SELF's conceptual ontology, with respect to security events.

As we explained above, one way to enhance, or accelerate this process is to add in the notion of emotion learning into the PSM process. This is done by assessment of how the SELF has reacted to situations in the past, and the results of those reactions. In other words, how the SELF feels about the possible solution. Possible solutions can be assessed quicker with the context of emotional memories, in light of the current problem or situation.

12.8 Self-Soothing Mechanisms

Here we describe the adaptation of self-soothing techniques from human neuroscience to a SELF's artificial cognitive architecture (ISAAC) [106]. The following sections illustrates how the use of the cognitive constructs within a SELF, the DAS, the FUSE-SEMs, etc., work together to provide self-soothing constructs within a SELF cognitive ecosystem.

12.8.1 SELF Self-Soothing: Acupressure

Artificial cognitive acupressure involves polling all of a SELF's available resources, refreshing the view of a SELF ISAAC cognitive infrastructure. This is basically taping on a SELF to see what response is generated [78]. Combining this with retrieving emotional memories involved in the current condition and mission/task context allows the Cognitrons to "calm down" and concentrate on finding solutions to the current problem. This utilizes the DAS genetic hypothesis search, in conjunction with the FUSE-SEMs to look for solutions; forming a SELF version of an Emotional Freedom Technique (EFT) [102].

12.8.2 SELF Self-Soothing: Deep Breathing

When you are scared, you might contract your body and hold your breath to try to squish the feelings in order to keep from feeling bad. Pulling your body in tight and stopping your breath keeps you from getting good oxygen to deal with whatever upsets you. In SELF cognitive system health terms, this is paramount to conservation of resources (Cognitive Economy) and not allowing the system to release hardware and software resources that may be required to "heal the current situation."

Deep breathing within a SELF ACNF involves releasing a plethora of Cognitrons to access all parts of a SELF and collaborate in a calm, organized fashion (i.e., breath in and out) and form a cognitive collective grouping of possible solutions to the current situation [33].

12.8.3 SELF Self-Soothing: Amplification of the Feeling

Exaggeration of feelings in a SELF entails flooding the system with genetic DSA searches with constraints based on an exaggeration of the emotional memories. In fuzzy inference sense, this is moving from a fuzzy membership function of "greater than" to one of "much greater than", or "much less than" instead of "less than." This allows a SELF to concentrate on solutions that are the most appropriate and eliminates the majority of "possible" solutions. This acts as a SELF's subconscious.

12.8.4 SELF Self-Soothing: Imagery

In SELF terms, imagery involves creating several genetic populations of solutions with a large solution space within the DAS, opening up the mutation and combination rates to allow for major jumps in the generational solution possibilities (intuition and discovery within a SELF cognitive framework). The FUSE-SEM measure spaces are relaxed to allow for a larger possibilistic set of solutions to be abductively explored. This helps to jump-start the solution generation process, allowing a wide range of solutions, and then completely unavailable solutions can be eliminated, but the possible solutions sets are broadcast to as wide a Cognitron coalition population as possible; or, as Sherlock Holmes once said:

> When you have eliminated the impossible, what remains, no matter how improbable, must be the truth
>
> Sir Arthur Conan Doyle, Sr. [228]

Once a viable solution set is created, the constraints, mutation, and crossover rates are returned to normal levels, and the solution populations are evaluated. The emotions experienced through this process are catalogued and stored in emotional memory, with contextual trigger RNAs created and catalogued; coupling emotional responses with each solution space. This allows greater efficiencies when the current situation is encountered again.

12.8.5 SELF Self-Soothing: Mindfulness

Mindfulness within a SELF involves keeping your attention on what is happening at the moment. Within a SELF ISAAC framework, this involves tightening constraints and FUSE-SEM distance metrics to ensure that attention is paid just to the current problem at hand, once the imagery technique has been utilized. This employs abductive techniques to sort out only those solutions that carry positive emotional learning responses and assessing those solutions first. These solutions are evaluated by utilizing Mediator Cognitrons that concentrate on the mission/task needs and mission/task criticality to provide necessary solutions that are also pertinent to the current situation.

12.8.6 SELF Self-Soothing: Positive Psychology

In regards to humans, positive psychology researches how happy, successful people work their life. Within a SELF, the positive psychology processes look for solution spaces that have resulted in positive emotional responses, based on learned emotions (emotional memories), and utilizes those methods employed during that investigation and diagnostic/prognostic period to look for solutions.

12.9 SELF Internal Information Encryption

In order to facilitate providing an internal multi-level secure information encryption within a SELF, we employ an encryption scheme (similar to that shown in Fig. 5.8) that allows information to be encrypted at any combination of information, context, content, and need-to-know [76]. We utilize a multi-dimensional fractal matrix that defines the real-time multi-dimensional fractal encryption/decryption scheme for every combination of topic, information, and context. Internal security levels for all three are generated and must match the credential of the Cognitron requesting the information in order for the information to be disseminated. Figure 12.6 illustrates this concept.

As information is obtained, the connections to other information and contexts for the information, as it relates to other information, is always expanding and evolving. It is necessary to create a continuously evolving, secure framework, as we discussed above, in order to carry the notion of information, subject, and context in a compact format that can accommodate multi-level security in one construct. The multi-dimensional triple-helix quantum fractal encryption framework accommodates this within a SELF cognitive architecture. For each combination of topic, context, and informational content, there is separate quantum fractal encryption algorithm.

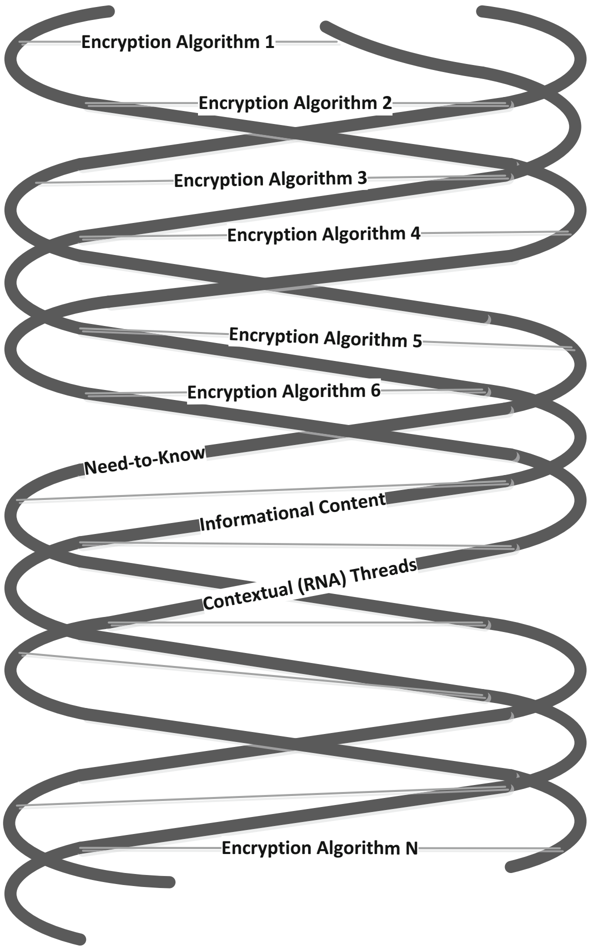


Fig. 12.6 Multi-dimensional quantum fractal encryption triple helix

Each combination equates to one of the quantum fractal eigenstates and the informational fractal can be unpacked at any security level. This allows a continuously evolving, secure, multi-level security, contextual secure framework for a SELF. This allows:

- Reduction in data acquisition and recognition time.
- Improved efficiencies for autonomous decision making.
- Improves processing and internal information reporting timeliness.
- Effective knowledge and decision management.
- This security framework within a SELF allows security policies and procedures to learn, self-adapt, and react to rapid changes in situational conditions, while still providing a very secure framework to protect a SELF's cognitive structures.

12.10 Discussion

Keeping a SELF secure if paramount to self-adaptation and effective learning and reasoning. Even a minor corruption will radically affect the learning, reasoning, and self-evolution of a SELF. Corruption would permeate throughout the system as Cognitrons communicate and learn from each other. The architectures and structures discussed here are intended to provide a SELF with the tools and capabilities to keep its cognitive framework and infrastructure safe from security incidents. One topic of current research is Anti-Tampering and how it should be implemented within a SELF architecture.

Chapter 13
Conclusions and Next Steps

So we have given a SELF the ability to think, reason, adapt, and evolve, as well as Metacognitive and Metamemory capabilities to understand its own abilities and limitations; including cyber security within its cognitive framework. The Cognitrons within the system themselves can learn, adapt, and evolve and can communicate with each other, allowing cognitive collaboration and cognitive economy within a SELF. So if we can actually build the complete system, if a SELF becomes a real-time, fully functioning, autonomous, self-actuating, self-analyzing, self-healing, fully reasoning and adapting system, what do we have and what are the ramifications? In Chap. 3 we discussed how people from different cultures might respond to a SELF, and the differences between accepting the system when it looks like a machine versus when it looks like a person. We explored the ramifications of giving a SELF basic emotions and emotional memories. How might its memories and actions be influenced by how people react to it? We also discussed how those reactions might influence how a SELF handles being around people. The overall purpose of the book was to begin to describe the capabilities, methodologies, and subsystems that must be in place in order to create a real-time, autonomous, thinking, reasoning system. We hope we have allayed fears that a SELF is going to decide to take over and eliminate the human race, as Hollywood is so fond of portraying. However, we also are not describing a cute, lovable robot, as depicted in the movie "Wall-E." There are other questions that need to be explored such as how to create versions a SELF at different levels in its evolutions so as not to have to start over again with each SELF we create, i.e. how do we clone a SELF. We also need to explore the advantages and disadvantages of SELF entities communicating with each other.

In the end, as we push for "autonomous" systems, as we want robots that are smart enough to do what we ask, when we ask, and to perform tasks without intervention or supervision, we need to be aware and fully understand the ramifications of what we ask. We need to understand how we really get there and how we deal with what we get when we get there. After all, we really don't want to hear a SELF reply to us: "I'm sorry Dave I really don't think I can do that."

J.A. Crowder et al., *Artificial Cognition Architectures*,
DOI 10.1007/978-1-4614-8072-3_13, © Springer Science+Business Media New York 2014

13.1 The Future SELF

There are many questions yet to be answered, and questions to be discovered we haven't even though of yet. This book is not the final answer, but is, instead, the first step toward providing possible answers to some of the fundamental questions in our quest for SELF. In many ways, technology has fallen short of predictions. Many felt that by now we would have a fully autonomous SELF that is in use throughout the world. In many ways technology has outstripped predictions. Cell phone and internet technologies far exceed expectations. Predictions for robotics in the future vary. In his article in 2006, Warren [210] predicts that by 2015, 1/3 of all military fighting forces will be composed of robots, and that by 2035 we will have fully autonomous SELFs in operation in battle situations. Other timelines predict that by 2034 we'll have robot SELFs performing most household tasks. In an article in the BBC news from December 2006, it was predicted that there will come a time when artificial life forms (robots) will have the same rights as people.[1] Whatever the future holds, we predict it will be different than anyone expects or predicts, as such is the nature of technology. However, we might get a glimpse of the future by looking at some of the work currently being done. This book describes the research Dr. Crowder, Dr. Carbone, and Ms. Friess have been involved in over the last decade and a half. Next we will discuss some of the accomplishment of this research that demonstrates the viability of synthetically thinking, learning, self-evolving life forms; these are called Reactionary, Evolving, Artificial Learning SELFs, or a REAL SELF.

13.2 Zeus: A Self-Evolving Artificial Life Form

In order to truly understand whether self-evolving life forms were indeed plausible, Dr. Crowder undertook a series of experiments over the last 10 years to create and test small, artificially intelligent, cybernetic entities that have the ability to think, learn, and self-evolve, but a very low-brain function level. These artificial life forms, which are created to learn and act like insect life forms, were created with very rudimentary cognitive functions to establish whether artificial cognitive architectures are realizable even at a simplistic level. The current instantiation, named "Zeus" utilizes a simplistic analog neural network for information transfer throughout the Zeus' effector network, and contains a digital low-level cognitive framework to affect learning and self-evolving [141]. We utilized controllers that contain EEPROM, RAM, and Flash memory in order to facilitate the abilities to learn and store learned behavior in as low a SWaP footprint as possible.[2] Basic effector control commands are stored in EEPROM. As Zeus learns, information is stored in

[1] http://news.bbc.co.uk/2/hi/technology/6200005.stm

[2] AVR ATTINY24 and ATTINY44 Microcontrollers are used, with the ATTINY 24 as the baseline.

RAM until it is determined that a behavior has been 'adequately' learned, and is then stored as a series of commands (procedural memory) into flash memory.

For his analog neural network, we utilize an adaptation of Eq. 2.1, which is repeated below:

$$C\frac{du(x,y,t)}{dt} = -\frac{1}{R}u(x,y,t) + \int_x\int_y w(x,y)z(x,y,t)\,dxdy + I(x,y,t)$$

The adaptation, shown in Equation 13.1, describes the dynamic equation for Zeus' analog neurons (non-linear amplifiers) [139]:

$$C_i\dot{u}_i(t') = -\frac{1}{R_i}u_i(t') + \sum_{j=1}^{N} T'_{ij}f_j\left(u_j\left(t' - \tau'_j\right)\right)\; i = 1,\ldots,N \tag{13.1}$$

Working together Dr. Crowder and Dr. Carbone improved Eq. 13.1 to encompass topic strength entropy. Within Zeus' analog neural network, the digital cognitive system monitors the strength of the analog neurons to determine when the strength of learning has progressed to the point where the learning should be "committed to memory" within the digital cognitive system (RAM). When a series of analog neurons is sufficiently strengthened over time, and have been committed to digital memory, such that they create a series of commands or learned behaviors that can be considered a "procedural memory," these are stored in Flash Memory with a tag that corresponds to the learned activity or behavior (e.g., turn left). The next time Zeus' sensors relate information such that he needs to move left, this procedure is recalled form memory and activated; meaning, he doesn't have to think about how to turn left, he turns left automatically. This is similar to humans driving a car after we have learned to drive. Again, this is at a much lower cognitive level than a human, but it does allow Zeus to learn and evolve. With Zeus, the goal is to add cognitive skills one at a time, perform tests and determine whether he can integrate these together within his limited cognitive framework. Once Zeus has reached a significant cognitive skill level, a new REAL SELF will be created with all of the cognitive skills present at the beginning of activation and determine whether this new REAL SELF has an easier or more difficult time integrating the cognitive skills together, and whether they arrive at essentially the same cognitive level as Zeus did adding cognitive skills one at a time. This will help determine how an artificial life form should be "started." Zeus has been in existence since early September, 2012. He has learned to walk, turn, integrate his sensors, plan his movements, and execute on his plans, demonstrating autonomous planning, sensory integration, and autonomous decision making, none of which is part of his initial programming; although the skills to learn, think, and store and recall memories was provided to him initially. He now carries 25 different procedural memories, which, as explained, are series of commands learned for a particular action.

13.3 Early Research into Cognitrons: Adventures in Cyberspace

Dr. Crowder and Dr. Carbone have been involved in research and development of intelligent software agents for well over a decade. Many believe that the term "intelligent" agent is in name only and they can never attain any significant level of cognitive intelligence. Here we related a real incident that may help the reader to make their own determination.

Over a decade ago Dr. Crowder was being funded by government research and development agencies to create and test Intelligent Software Agents, which became the predecessors to the Cognitrons. Since these were learning, thinking, evolving software agents, they were stopped every afternoon when Dr. Crowder left work to ensure they did not "evolve" while unsupervised. As you might expect, this did not happen on one given evening, and the agents were active, up and running throughout the evening. The next morning when Dr. Crowder arrived at work and logged in, an Instant Messaging (IM) window opened up from one of his government technical contacts. The gentleman on the other end sent the message, "I see you were working very late last night," to which Dr. Crowder replied, "No, I wasn't." He then said, "Well, I got up around 2:00 a.m. and couldn't go back to sleep so I went on line to check my email. I saw you were on-line and sent you an IM. We chatted for about 45 min by IM. You were in an odd mood." He had not been talking to Dr. Crowder by IM, he had been talking to the intelligent software agents. When the IM window opened up, the software agents took control of the IM window, read his message, interpreted it, understood its context, and crafted a reply and sent it. Understand, he thought Dr. Crowder was in an odd mood, but not odd enough that he didn't recognize he wasn't talking to an actual person. This says two things; first, when given sufficient cognitive skills, intelligent agents can think, reason, infer, and make coherent decisions, second, we need to be very careful when utilizing such entities and much testing is required before allowing one to be used unsupervised in a system.

13.4 What's Next?

What happens when a SELF comes together? SELF will be making decisions autonomously based on information that is new, abstract and/or incomplete. All the while, autonomously and continuously reprogramming a SELF codebase to resolve any internal conflicts. It will be accomplishing this in real time, while learning and using information stored and making inferences from all the combination of its memories, sensory information, abductive decision reasoning, and behavior processing. Human support systems will need to be enhanced to handle this fully autonomous Artificial Cognitive System (a SELF). This will facilitate the need for new science that will expand to better understand similarities and differences between humans and a SELF in areas such as learning, integration, recall, processing, and uses of information. The fields of Psychology, Sociology, Engineering, Communications, and Computer Skills Development will all need to be revamped to accept Artificial Cognitive Life Forms like a SELF, and we cannot wait until these systems are fully functioning. Classes must be developed now in order for individuals to be prepared for a new frontier of artificial life forms like a SELF. Intuition and emotion may become better understood in the humans as well as a SELF. Once we determine the primary arousal states of a system, whether human or artificial, then we may better be able to understand the higher level functioning that involves other arousal states and understand human autonomic nervous responses better.

A SELF, like the human, will be able to handle misinformation, ambiguity, and form hypotheses when given incomplete information and imprecise (fuzzy) conditions. A SELF may be able to determine when things are creating conflicts internally yet make sense of a world of nonsense. Since this is a fully autonomous, learning, reasoning system, it will still be possible for a SELF to make fallible decisions based on fallible information, while at the same time be able to make sense of the world that is not always exact, such as the human conversation. Even the word minute can have so many different and ambiguous meanings; a SELF must be able to handle all of these conditions.

As a SELF evolves, the human will need learn about its processes and have appropriate expectations. Thus, the human will be less likely to give a SELF attributes that are simply not true, such as Hollywood's depiction of Artificial Intelligence or social rules and expectations that may not apply. Thus, we believe this will allow an improved trust between a SELF and humans. Trust will become a necessary component for future collaboration between a SELF and humans. In addition, the more emotional intelligence a SELF has the more likely a SELF can reason with humans on an emotional level, for the human, and become more likable; which in turn would increase the chances of collaboration. In addition, a SELF will have the capabilities to understand the human better, also increasing the likelihood of enhancing human-robot collaborations. As a SELF develops and the humans gain clearer expectations such as predictability, safety, reliability, trust, communication, knowledge, understanding, and accommodation, the collaboration between human and SELF will develop over time.

The main outcome of a SELF and human interaction is cooperative and collaborative reasoning. Consider how helpful it will be to bounce ideas off a colleague or learn from a mentor and the ways in which those activities change individual human reasoning. If a SELF could also be a part of a human's reasoning outside of its own internal reasoning, then we propose a SELF's reasoning can be enhanced, as well as the potential to enhance human reasoning. Memory could be easier recalled if it came from internal and external locations with both locations capable of different distributed reasoning. So it may become far more than individual, world, and body, but it may now include other unique forms, like a SELF. It would seem that a SELF could enhance our ability to come to certain conclusions more accurately and in efficient ways. It may be the case that each player would enhance knowledge and understanding of the others making, an enhanced collaborative team.

In the end, we have only begun to scratch the surface of an entirely new frontier. There are unlimited possibilities, including enhancements to the medical field, engineering, space travel, and possibly enhance our understanding of serious conditions like schizophrenia, by utilizing a SELF to better understand how such conditions may be overcome through retraining the brain. We are on a journey that was begun 25 years ago and will not be finished for many more years. Again, we endeavor to create artificial life forms that will enhance the human experience, not destroy it as Hollywood might have you believe.

Acronyms

ACA	Artificial cognitive architectures
ACNF	Artificial cognitive neural framework
ACP	Artificial cognitive perception
ACS	Artificially cognitive systems
AI	Artificial intelligence
AIC	Autonomic information continuum
ANN	Abductive, neural network
APC	Artificial prefrontal cortex
API	Application programming interface
ASIC	Application specific integrated circuit
BAO	Binary information object
BBNN	Bayesian belief neural network
BIF	Binary information fragment
BMRO	Binary memory reconstruction object
BRO	Binary relativity object
CBLP	Case-based learning planner
CBN	Cognitive behavior network
CBR	Case-based reasoning
CCO	Cognitive conceptual ontology
CEC	Cognitive emotion cognitron
CI	Catastrophic interference
CITE	Cognitive, interactive training environment
CL	Constructivist learning
CNS	Central nervous system
Cognitrons	Cognitive perceptrons
COGSEC	Cognitive security
C-SME	Cognitron subject matter expert
DART	Decision analytics in real-time
DAS	Dialectic argument structure
DSA	Dialectic search argument
EBR	Experience-based reasoning

J.A. Crowder et al., *Artificial Cognition Architectures*,
DOI 10.1007/978-1-4614-8072-3, © Springer Science+Business Media New York 2014

EI	Emotional intelligence
EML	Emotional markup language
FGC	Fuzzy, genetic cognitron
FPGA	Field-programmable gate array
FuNN	Fuzzy, unsupervised, active resonance theory, neural network
FUSE-CTX	Fuzzy, self-evolving, contextual, topical map
FUSE-SEM	Fuzzy, self-evolving semantical topical map
GPS	General problem solver
HIL	Human interaction learning
HMS	Human mentored software
HoMe	Host mediator
HPM	Hypothesis plausibility measures
HRI	Human-robot interface
HSI	Human-systems interface
IC	Interface cognitron
IM	Instant messaging
ISA	Intelligent information software agents
ISAAC	Intelligent information software agents for artificial consciousness
JPL	Jet propulsion laboratory
KB	Knowledge base
KO	Knowledge object
KRT	Knowledge relativity thread
LTM	Long-term memory
MC	Mediator cognitron
MIM	Mutual information measure
MOE	Measure of effectiveness
NIC	Neural information continuum
NIST	National institute of standards and technology
OLAP	Online analytical processing
PAC	Probably, approximately correct
PANN	Possibilistic, abductive neural network
PENLPE	Polymorphic, evolving, neural learning and processing environment
PNS	Peripheral nervous system
PSM	Prognostic security management
RAM	Random access memory
RARE	Reconfigurable, advanced rapid-prototyping environment
RBV	Reputation-based voting
RC	Reasoner cognitron
RNA	Recombinant knowledge assimilation
RUL	Remaining useful life
SELF	Synthetic, evolving life form
SNS	Synthetic nervous system
STM	Short-term memory
SWaP	Size, weight, and power
TMQL	Topical map query language
TOI	Topic of interest
WM	Working memory

References

1. Ackley D, Littman M (1991) Interactions between learning and evolution. In: Langton CG et al (eds) Proceedings of the second conference on artificial life. Addison-Wesley, Reading
2. Ackley D, Littman M (1990) Learning from natural selection in an artificial environment. In: Caudhill M (ed) Proceedings of the international joint conference on neural networks, IJCNN-90-WASII-DC, Lawrence Erlbaum, Hillsdale, vol I, pp 183–193
3. Ade H, Denecker M (1995) Abductive inductive logic programming. In: Proceedings of the fourteenth international joint conference on artificial intelligence, Montreal
4. Adler A (1929) The practice and theory of individual psychology, Revth edn. Harcourt, Brace, New York (Original work published 1925)
5. Allen J (1983) Maintaining knowledge about temporal intervals. Commun ACM 26(11):832–843
6. Anderson J (2004) Cognitive psychology and its implications. Worth, New York
7. Ashcraft M (1994) Human memory and cognition. Harpercollins College Division, New York
8. Axlerod R (1976) Structures of decision: the cognitive maps of political elites. Princeton Press, Princeton
9. Baars B (1988) A cognitive theory of consciousness. Cambridge University Press, New York
10. Baars B (1997) In the theater of consciousness. Oxford University Press, Oxford
11. Bäck T, Schwefel H (2003) An overview of evolutionary algorithms for parameter optimization. Evol Comput 41(1):1–23
12. Baddeley A (2000) The episodic buffer: a new component of working memory? Trends Cogn Sci 4:417–423
13. Barto A, Sutton A, Brouwer P (1981) Associative search network: a reinforcement learning associative memory. Biol Cybern 40:201–211
14. Bear M, Connors B, Paradiso M (2007) Neuroscience: exploring the brain, 3rd edn. Lippincott, Williams, & Wilkins, Philadelphia, pp 196–199. ISBN 978-0-7817-6003-4
15. Bednar A, Cunningham D, Duffy T, Perry J (1998) Theory into practice: how do we link? In: Duffy TM, Jonassen DH (eds) Constructivism and technology of instruction: a conversation. Lawrence Erlbaum Associates, Hillsdale, pp 17–35
16. Beer, R. (1995) Computational dynamical languages for autonomous agents. Published in Mind as Motion, Massachusetts Institute of Technology, Cambridge, MA, 121–147
17. Beer R (1995) Computational dynamical languages for autonomous agents. In: Mind as motion. Massachusetts Institute of Technology, Cambridge, pp 121–147
18. Belew R, McInerney J, Schraudolph N (2002) Evolving networks: using the genetic algorithms with connectionist learning. In: Artificial life XII, SFI studies in the sciences of complexity: proceedings, Addison Wesley, Redwood City, pp 511–547

J.A. Crowder et al., *Artificial Cognition Architectures*,
DOI 10.1007/978-1-4614-8072-3, © Springer Science+Business Media New York 2014

19. Billings D, Schaefer K, Chen J, Hancock P (2012) Human-robot interaction: developing trust in robots. In: Proceedings of the 7th ACM/IEEE international conference on human-robot interaction, Boston
20. Bishop C (1995) Neural networks for pattern recognition. Oxford University Press, New York
21. Bolte A, Goschke T, Kuhl J (2003) Emotion and intuition. Institute of Psychology, Braunschweig University of Technology, Braunschweig. annette.bolte@tu-bs.de
22. Bonarini A (1997) Anytime learning and adaptation of structured fuzzy behaviors, vol 5, Adaptive behavior. The MIT Press, Cambridge
23. Bonato M, Scally L (2012) Reconfigurable advanced rapid-prototyping environment (RARE): a computing technology for challenging form factors. In: Proceedings of the IEEE high performance extreme computing conference, Waltham
24. Botella L (1995) Personal construct theory, constructivism, and postmodern thought. In: Neimeyer RA, Neimeyer J (eds) Advanced in personal construct psychology, vol 3. Jai Press, Greenwich, pp 3–35
25. Brazier F, Keplicz B, Jennings N, Treur J (1995) Formal specification of multi-agent systems: a real-world case. In: First international conference on multi-agent systems (ICMAS'95), San Francisco, pp 25–32
26. Brillouin L (1962) Science and information theory. Academic, New York
27. Brillouin L (1964) Uncertainty and information. Academic, New York
28. Brunk C (1996) An investigation of knowledge intensive approaches to concept Learning and theory renement. Ph.D. thesis, Department of Information and Computer Science, University of California, Irvine
29. Buhman J, Schulten K (1987) Noise-driven temporal association in neural networks. Europhys Lett 4:1205–1209
30. Cardinal RN, Parkinson JA, Hall J, Everitt BJ (2002) Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. Neurosci Biobehav Rev 26:321–352
31. Chai S (2004) Artificial intelligence and social theory: a one way street? Perspectives 27(4):11–12
32. Cheng P, Holyoak K (2011) Causal learning and inference as a rational process: the new synthesis. Annu Rev Psychol 62:135–163
33. Chandramouli R, Sandhu R (1998) Role based access control features in commercial database management systems. In: Proceedings of the 21st national information systems security conference, Crystal City
34. Cleermans A (1993) Mechanisms of implicit learning: connectionist models of sequence processing. The MIT Press, Cambridge
35. Cliff D, Harvey I (2002) Incremental evolution of neural network architectures for adaptive behavior. Technical Report, University of Sussex, COGS
36. Cobb H (1990) An investigation into the use of hypermutation as an adaptive operator in genetic algorithms having continuous, time-dependent non-stationary environment. NRL Memorandum Report 6760
37. Cohen W (1992) Abductive explanation-based learning: a solution to the multiple inconsistent explanation problem. Mach Learn 8:167–219
38. Cooper G, Herskovits E (1992) A Bayesian method for the induction of probabilistic networks from data. Mach Learn 9:309–347
39. Coutaz J, Crowley J, Dobson S, Garlan D (2005) Context is key. Commun ACM 48:53
40. Cox P, Piertrzykowski T (1987) General diagnosis by abductive inference. In: Proceedings of the symposium on logic programming, San Francisco, pp 183–189
41. Crowder J (2001) ELYSE, The evolving, life-like yielding, symbiotic environment. NSA Technical Paper Technical Paper TIT_01_01_013_2001_002
42. Crowder J (2001) Integrating an expert system into a neural network with genetic programming for process planning. NSA Technical Paper TIT_01_01_013_2001_001
43. Crowder J (2001) Statistical performance of outer product associative memory models. GDIS, IRAD Report GDIS_9873_2001_001

44. Crowder J (2002) Derivation of fuzzy classification rules from multidimensional data. NSA Technical paper CON_0013_2002_001
45. Crowder J (2002) Adaptive learning algorithms for functional constraints on an evolving neural network system. NSA Technical paper CON_0013_2002_003
46. Crowder J (2003) Machine learning: intuition (concept) learning in hybrid genetic/fuzzy/neural systems. NSA Technical paper CON_0013_2003_009
47. Crowder J (2003) Adaptive learning algorithms: pattern discovery in artificial intelligence. NSA Technical paper CON_0013_2002_004
48. Crowder J (2003) Intelligence fusion techniques applied to data integration. NSA Technical paper – CON-SP-0014-2003-07
49. Crowder J (2003) Temporal difference learning with fuzzy cognitive maps. NSA Technical paper – CON-SP-0014-2003-04
50. Crowder J (2004) Multi-sensor fusion utilizing dynamic entropy and fuzzy systems. In: Proceedings of the processing systems technology conference, Tucson
51. Crowder J (2005) Cognitive systems for data fusion. In: Proceedings of the 2005 PSTN processing technology conference, Ft. Wayne
52. Crowder J (2006) Intelligent information agents for autonomous condition based system health management. In: Proceedings of the AFRL integrated systems health management conference, Cincinnati
53. Crowder J (2006) Intelligent agents for integrated system health management: modern diagnostic and prognostic techniques. In: Proceedings of the AFRL integrated systems health management conference, Cincinnati
54. Crowder J (2006) Intelligent information agents. In: Proceedings of the 5th annual system/software symposium, Denver
55. Crowder J (2006) Modern architectures for integrated system health management: intelligent agent architectures for sensor-to-prognostics information integration. In: Proceedings of the air force 3rd annual integrated health management systems conference, Cincinnati
56. Crowder J (2007) Integrating metrics with qualitative temporal reasoning for constraint-based expert systems. In: Proceedings of the 2007 processing systems technology network conference, El Segundo
57. Crowder J (2010) Flexible object architectures for hybrid neural processing systems. In: Proceedings of the 12th annual international conference on artificial intelligence, Las Vegas
58. Crowder J (2010) The continuously recombinant genetic neural fiber network. In: Proceedings for the AIAA Infotech@Aerospace 2010 conference, Atlanta
59. Crowder J (2010) Operative information software agents (OISA) for intelligence processing. In: Proceedings of the 12th annual international conference on artificial intelligence, Las Vegas
60. Crowder J (2011) Cognitive architectures for real-time integrated system health management. In: Proceedings of the 8th annual workshop of structural health monitoring, Stanford University
61. Crowder J (2011) Knowledge density mapping for derivation of inference potential. In: Proceedings of the 12th annual international conference on artificial intelligence, Las Vegas
62. Crowder J (2011) The artificial prefrontal cortex: artificial consciousness. In: Proceedings of the 13th annual international conference on artificial intelligence, Las Vegas
63. Crowder J (2012) Cognitive system management: the polymorphic, evolutionary, neural learning and processing environment (PENLPE). In: Proceedings for the AIAA Infotech@Aerospace 2012 conference, Garden Grove
64. Crowder J (2012) The artificial cognitive neural framework. In: Proceedings for the AIAA Infotech@Aerospace 2012 conference, Garden Grove
65. Crowder J (2012) Possibilistic, abductive neural networks (PANNs) for decision support in autonomous systems: the advanced learning, abductive network (ALAN). In: Proceedings of the 1st international conference on robotic intelligence and applications, Gwanju
66. Crowder J (2012) Autonomous robotic systems for collaborative shared awareness. In: Proceedings of the 1st international conference on robotic intelligence and applications, Gwanju
67. Crowder J (2012) Real-time data storage and retrieval in autonomous robotic systems. In: Proceedings of the 1st international conference on robotic intelligence and applications, Gwanju

68. Crowder J (2012) Artificial memory reconstruction. In: Proceedings of the 1st international conference on robotic intelligence and applications, Gwanju
69. Crowder J, Carbone J (2012) Cognitive perception utilizing operationally intelligent software agents (OISAs). In: Proceedings of the AIAA space 2012 conference and symposium, Pasadena
70. Crowder J, Carbone J (2011) Occam learning through pattern discovery: computational mechanics in AI systems. In: Proceedings of the 13th annual international conference on artificial intelligence, Las Vegas
71. Crowder J, Carbone J (2011) The great migration: information to knowledge using cognition-based frameworks. Springer Science, New York
72. Crowder J, Carbone J (2011) Hybrid neural architectures for the ELYSE cognitive system. In: Proceedings of the AIAA Infotech@aerospace 2011 conference, Atlanta, pp 3470–3477
73. Crowder J, Carbone J (2011) Recombinant knowledge relativity threads for contextual knowledge storage. In: Proceedings of the 13th annual international conference on artificial intelligence, Las Vegas
74. Crowder J, Carbone J (2011) Transdisciplinary synthesis and cognition frameworks. In: Proceedings of the society for design and process science conference 2011, Jeju Island
75. Crowder J, Carbone J (2012) Reasoning frameworks for autonomous systems. In: Proceedings of the AIAA Infotech@Aerospace 2012 conference, Garden Grove
76. Crowder J, Carbone J (2012) Cyber physical system architectures for dynamic, real-time "need-to-know" authorization. In: Proceedings of the society for design and process science conference 2012, Berlin
77. Crowder J, Friess S (2012) Extended metacognition for artificially intelligent systems (AIS): artificial locus of control and cognitive economy. In: Proceedings of the 13th annual international conference on artificial intelligence, Las Vegas
78. Crowder J, Friess S (2010) Artificial neural diagnostics and prognostics: self-soothing in cognitive systems. In: Proceedings of the 12th annual international conference on artificial intelligence, Las Vegas
79. Crowder J and Friess S (2010) Artificial neural emotions and emotional memory. Proceedings of the 12th annual International Conference on Artificial Intelligence, ICAI'10, Las Vegas, NV
80. Crowder J, Friess S (2010) Artificial neural emotions and emotional memory. In: Proceedings of the 12th annual international conference on artificial intelligence, ICAI'10, Las Vegas
81. Crowder J, Friess S (2011) Metacognition and metamemory concepts for AI systems. In: Proceedings of the 13th annual conference on artificial intelligence, ICAI'11, Las Vegas
82. Crowder J and Friess S (2011) Metacognition and Metamemory concepts for AI systems. Proceedings of the 13th annual Conference on Artificial Intelligence, ICAI'11, Las Vegas, NV
83. Crowder J, Friess S (2011) The artificial prefrontal cortex: artificial consciousness. In: Proceedings of the 12th annual international conference on artificial intelligence, Las Vegas
84. Crowder J, Friess S (2012) Artificial psychology: the psychology of AI. In: Proceedings of the 3rd annual international multi-conference on informatics and cybernetics. Orlando
85. Crowder J, Friess S (2012) Warrior resilience training through cognitive self-regulation. In: Proceedings of the 13th annual international conference on artificial intelligence, Las Vegas
86. Crowder J, Wetherbie J (2012) Improving software performance via human interaction learning. In: Proceedings of the 2012 information systems and computing conference, Tucson
87. Crowder J, Barth T, Rouch R (1999) Evolutionary neural infrastructure with genetic memory algorithms: ENIGMA theory development. NSA Technical paper, Denver ENIGMA_1999_004
88. Crowder J, Barth T, Rouch R (1999) Learning algorithms for stochastically driven fuzzy, genetic neural networks. NSA Technical paper ENIGMA_1999_002
89. Crowder J, Barth T, Rouch R (1999) Neural associative memory for storing complex memory patterns. NSA Technical paper ENIGMA_1999_003
90. Crowder J, Raskin V, Taylor J (2012) Autonomous creation and detection of procedural memory scripts. In: Proceedings of the 13th annual international conference on artificial intelligence, Las Vegas

91. Crowder J, Scally L, Bonato M (2012) Applications for intelligent information agents (I2As): learning agents for autonomous space asset management (LAASAM). In: Proceedings of the international conference on artificial intelligence, ICAI'12, Las Vegas
92. Crowder J, Friess S (2013) Systems engineering agile design methodologies. Springer, New York. ISBN 978-1-4614-6662-8
93. Damasio A (1994) Descarte's error: emotion, reason, and the human brain. Gosset/Putnam, New York
94. Davis M, Whalen P (2001) The amygdala: vigilance and emotion. Mol Psychiatry 6:13–34
95. Davis L (1991) The handbook of genetic algorithms. Van Nostrand Reinhold, New York
96. DeJong K (1975) An analysis of the behavior of a class of genetic adaptive systems. Doctoral dissertation, University of Michigan
97. DeRaedt L (1992) Interactive theory revision: an inductive logic programming approach. Academic, London
98. Dey A (2001) Understanding and using context. Pers Ubiquit Comput 5:4–7
99. DeYoung C, Hirsh J, Shane M, Papademetris X, Rajeevan N, Gray J (2010) Testing predictions from personality neuroscience. Psychol Sci 21(6):820–828
100. Dimopoulos Y, Kakas A (1996) Abduction and Learning. In: DeRaedt L (ed) Advances in inductive logic programming. IOS Press, Landsdale
101. Dimopoulos Y and Kakas A (1996) Abduction and Learning. In Advances in Inductive Logic Programming, L. DeRaedt (ed.), IOS Press, Landsdale, PA
102. Dolan R, Vuilleumier P (2003) Amygdala automaticity in emotional processing. Ann N Y Acad Sci 985:348–355
103. Dourish P (2004) Where the action is: the foundations of embodied interaction. The MIT Press, Cambridge, MA
104. Dourish P (2004) Where the action is: the foundations of embodied interaction. The MIT Press, Cambridge, MA
105. Dourish P, Buton G (1998) On technomethodology: foundational relationships between ethnomethology and system design. Hum Comput Interact 12(4):395–432
106. Dudai Y (2004) The neurobiology of consolidations, or, how stable is the engram? Annu Rev Psychol 55:51–86
107. Edelman G (1987) Neural darwinism. Basic Books, New York
108. Eichenbaum H (2002) The cognitive neuroscience of memory. Oxford University Press, New York
109. Ejigu D, Scuturici M, Brunie L (2008) Hybrid approach to collaborative context-aware service platform for pervasive computing. J Comput 3:40
110. Elman J (1990) Finding structure in time. Cogn Sci 14(2):179–211
111. Emery M (2000) A discrete approach to the chaotic representation property. 005, IRMA
112. Ertas A, Tanik M (2000) Transdisciplinary engineering education and research model. Trans SDPS 4(4):1–11
113. Evans J (1989) Bias in human reasoning: causes and consequences. Lawrence Erlbaum Associates, Hove
114. Evans J, Newstead S, Byrne R (1993) Human reasoning: the psychology of deduction. Lawrence Erlbaum Associates, Hove
115. Fodor J (1983) Modularity and the mind: an essay on faculty psychology. MIT Press, Cambridge
116. Franklin S (1995) Artificial minds. MIT Press, Cambridge
117. Franklin S (1997) Autonomous agents as embodied AI. Cybern Syst 28:499–520
118. Garcia C (2007) Cognitive modularity, biological modularity, and evolution. Biol Theory Integrat Dev Evol Cogn 2(1):62–73
119. Gibbs R, Van Orden G (2010) Adaptive cognition without massive modularity. J Lang Cognit 2(2):149–176
120. Giuri L, Iglio P (1997) Role templates for content-based access control. In: Proceedings of ACM workshop on role based access control, New York, pp 153–159

121. Green S, Chase G, Chen X, Billinhurst M (2008) Evaluating the augmented reality humanrobot collaboration system. In: International journal of intelligent systems technologies and applications (IJISTA), vol 8, No. 1/2/3/4
122. Green S, Richardson S, Billinghurst M, Chase J, Stiles R (2008) Multimodal metric study for human-robot collaboration. Sainte-Luce, Martinique: The first international conference on advances in computer-human interaction (ACHI 2008)
123. Gruber T (2008) Collective knowledge systems: Where the social web meets the semantic web. Web Semantics Sci Serv Agent World Wide Web 6:4–13
124. Gupta K, Govindarajan V (2000) Knowledge flows within multinational corporations. Strategic Manage J 21:473–496
125. Gupta K, Govindarajan V (2000) Knowledge flows within multinational corporations. Strategic Manage J 21:473–496
126. Hammond KJ (1990) Explaining and repairing plans that fail. Artif Intell 45:173–228
127. Hancock P, Billings D, Schaefer K, Chen J, de Visser E, Parasuraman R (2011) A meta-analysis of factors affecting trust in human-robot interaction. J Hum Factor 53(5):517–27
128. Harashima F, Suzuki S (2008) Intelligent mechatronics and robotics. In: Proceedings for the 2008 IEEE international conference on emerging technologies and factory automation, Hamburg
129. Harmon S (2011) Human perception of gendered artificial entities. Colby College, Waterville
130. Hebb D (1949) The organization of behavior. Wiley, New York
131. Holland JH, Reitman JS (1978) Cognitive systems based on adaptive algorithms. In: Waterman DA, Hayey-Roth F (eds) Pattern directed inference systems. Academic, New York, pp 313–329
132. Holland P, Gallagher M (2004) Amygdala-frontal interactions and reward expectancy. Curr Opin Neurobiol 14:148–155
133. Hong J, Landay J (2001) An infrastructure approach to context-aware computing. Hum Comput Int 16:287–303
134. Hopfield J (1982) Neural networks and physical systems with emergent collective computational abilities. In: Proceedings of the national academy of science, Washington, DC, vol 79(8), pp 2554–2558
135. Howard N, Qusaibaty A (2004) Network-centric information policy. In: Proceedings of the second international conference on informatics and systems, Setúbal
136. Howard N, Qusaibaty A (2006) Intention awareness in a nutshell. Def Concept J 1(3):48–57
137. Humberto R, Varela F (1992) The tree of knowledge: the biological roots of human understanding. Shambhala Publications, Boston
138. Hutchins E, Lintem G (1995) Cognition in the wild. MIT Press, Cambridge
139. Hutchinson J, Koch C, Luo J, Mead C (1988) Computing motion using analog and binary resistive networks. Computer 21(3):52–63
140. Igor A, Antonova I, Kandel E, Hawkins R (2003) Activity-dependent presynaptic facilitation and hebbian LTP are both required and interact during classical conditioning in aplysia. Neuron 39(1):135–157
141. Jackel L, Howard R, Graf H, Straughn B, Denker J (1986) Artificial neural networks for computing. Vlic Sci Tccirttrol 861:63
142. Jacobs R, Jordan M, Nowlan S, Hinton G (1991) Adaptive mixtures of local experts. Neural Comput 3(1):79–87
143. Johnson-Laird P, Paolo L, Legrenzi M (1972) Reasoning and a sense of reality. Br J Psychol 63(3):395–400
144. Kahana M, Howard M, Polyn S (2008) Associative retrieval processes in episodic memory. In: Roediger HL III (ed) Cognitive psychology of memory, vol 2 of learning and memory: a comprehensive reference, 4 vols. (Byrne J, ed). Elsevier, Oxford
145. Kakas R, Kowlaski R, Toni F (1993) Abductive logic programming. J Logic Comput 2(6):719–770
146. Kanerva P (1988) Sparse distributed memory. The MIT Press, Cambridge
147. Karami A-B, Jeanpierre L, Mouaddib A-I (2010) Human-robot collaboration for a shared mission. In: HRI-10', 155–156

148. Kasabov N (1998) Introduction: hybrid intelligent adaptive systems. Int J Intell Syst 8:453–454
149. Kelly G (1991) The psychology of personal constructs, vol 2. Routledge, London
150. Korzybski, A (1994) Science and sanity: an introduction to non-Aristotelian systems and general semantics, 5th edn. Institute of General Semantics, Brooklyn
151. Kohlberg L (1971) From 'is' to 'ought': how to commit the naturalistic fallacy and get away with it in the study of moral development. Cognitive development and epistemology. Academic, New York
152. Kohlberg L (1985) Stages of moral development. In: Crain WC (ed) Theories of development. Prentice Hall, New York, pp 118–136
153. Kosko G (1986) Fuzzy cognitive maps. Int J Man-Machine Stud 24:65–75
154. Kroeker JP (2008) Wiener analysis of functionals of a Markov chain: applications to neural transformations of random signals. Biol Cybern 56:243–248
155. Kvale S (1992) Psychology and postmodernism. Sage, Thousand Oaks
156. LaBar K, Cabeza R (2006) Cognitive neuroscience of emotional memory. Nat Rev Neurosci 7:54–64
157. LeDoux J (1996) The emotional brain. Simon and Schuster, New York
158. LeDoux J (2000) Emotion circuits in the brain. Annu Rev Neurosci 23:155–184
159. LeDoux J (2002) Synaptic self: how our brains become who we are. Viking, New York
160. Lehtonen M (2012) Long-term phonological learning begins at the level of word form. Neuroimage 63(2):789–799
161. Lenz M, Barsch-Spörl B, Burkhard H, Wess S (1998) Case-based reasoning technology, from foundations to applications. Springer, Berlin
162. Levesque HJ (1989) A knowledge-level account of abduction. In: Proceedings of the eleventh international joint conference on artificial intelligence, Detroit, pp 1061–1067
163. Levine P (1997) Walking the tiger: healing trauma. North Atlantic Books, Berkeley
164. Liao S, He J, Tang T (2004) A framework for context information management. J Inform Sci 30:528–539
165. Mahoney M (2003) Constructive psychotherapy: a practical guide. The Guilford Press, New York
166. Malik J, Binford T (1983) Reasoning in time and space. In: Proceedings of the 8th annual IJCAI conference, Karlsuhe, pp 343–345
167. Maren S, Quirk G (2004) Neuronal signaling of fear memory. Nat Rev Neurosci 5:844–852
168. Marsella S, Gratch J (2002) A step towards irrationality: using emotion to change belief. In: 1st international joint conference on autonomous agents and multi-agent systems, Bologna
169. Martin J, Sugarman J (1997) The socio-cognitive construction of psychotherapeutic change: bringing social constructionism and cognitive constructivism. Rev Gen Psychol 1:375–378
170. Mayer J, Salovy P, Caruso D (2004) Emotional intelligence. Psychol Inquiry 15(3):197–215
171. Michalski R (1993) Inferential theory of learning as a conceptual basis for multi-strategy learning. Mach Learn 11:111–151
172. Miller E, Freedman D, Wallis J (2002) The prefrontal cortex: categories, concepts and cognition. Philos Trans R Soc Lond B Biol Sci 357(1424):1123–36
173. MinNa L (2012) Distributed cognitive theory: learning concepts from network era. In: International symposium on instrumentation and measurement, sensory network, and automation, IEEE
174. Mooney R (1995) A preliminary PAC analysis of theory revision. In: Petsche T, Hanson S, Shavlik J (eds) Computational learning theory and natural learning systems, vol 3. MIT Press, Cambridge, pp 43–53
175. Morik K, Wrobel S, Kietz J, Emde W (1993) Knowledge acquisition and machine learning: theory, methods, and applications. Academic, New York
176. Nass C, Moon Y (2002) Machines and mindlessness: social response to computers. J Social Issues 56(1):81–103
177. Newell A (2003) Unified theories of cognition. Harvard University Press, Cambridge
178. Newell A, Shaw J, Simon H (1957) Preliminary description of general problem solving program-i (gps-i). WP7, Carnegie Institute of Technology, Pittsburgh

179. Nishimori T, Nakamura T, Shiino M (1990) Retrieval of spatio-temporal sequences in asynchronous neural networks. Phys Rev A 41:3346–3354
180. Nonaka I, Takeuchi H (1995) The knowledge-creating company: how Japanese companies created the dynamics of innovation. Oxford University Press, New York
181. Nonaka I (1994) A dynamic theory of organizational knowledge creation. Org Sci 5(1):14–37
182. O'Rorke P (1994) Abduction and explanation-based learning: case studies in diverse domains. Comput Intell 10:295–330
183. Ourston D, Mooney R (1994) Theory refinement combining analytical and empirical methods. Artif Intell 66:311–344
184. Paul G (1993) Approaches to abductive reasoning: an overview. Artif Intell Rev 7:109–152
185. Pearl J (1998) Probabilistic reasoning in intelligent systems. Morgan Kauffman, San Mateo
186. Phelps E (2006) Emotion and cognition: insights from studies of the human amygdala. Annu Rev Psychol 57:27–53
187. Pinto M (2006) A grounded theory on abstracts quality: weighting variables and attributes. Scientometrics 69:213–226
188. Polanyi M, Sen A (2009) The tacit dimension. University of Chicago Press, Denver
189. Quinlan J (1989) Unknown attribute values in induction. In: Proceedings of 6th international workshop on machine learning, Morgan Kaufmann, pp 164–168
190. Ramamurti V, Ghosh J (1996) Advances in using hierarchical mixture of experts for signal classification. In: Proceedings of the international conference on acoustics, speech and signal processing (ICASSP), vol 6, pp 3569–3572
191. Raskin V, Taylor JM, Hemplemann CF (2010) Ausocial engineering. New Security Paradigms Workshop, Concord
192. Rinkus G (1995) A combinatorial neural network exhibiting both episodic memory and generalization for spatio-temporal patterns. Boston University Press
193. Rips L (2011) Lines of thought: central concepts in cognitive psychology. Oxford University Press, Oxford
194. Roberts S, Tarassenko L (1994) A probabilistic resource allocating network for novelty detection. Neural Comput Appl 6:2023–2032
195. Rogers Y (1997) A brief introduction to distributed cognition. University of Sussex Press, Brighton
196. Rowley J (1996) Organizing knowledge: an introduction to information retrieval. Gower, Aldershot
197. Rowley J, Hartley R (2008) Organizing knowledge: an introduction to managing access to information. Ashgate, Burlington
198. Scally L, Bonato M, Crowder (2011) Learning agents for autonomous space asset management. In: Proceedings of the advanced maui optical and space surveillance technologies conference, Maui
199. Stylios C, Groumpos P (2000) Fuzzy cognitive maps in modeling supervisory control systems. J Intell Fuzzy Syst 8(2):83–98
200. Suh N (2006) Application of axiomatic design to engineering collaboration and negotiation. In: Proceedings of the 4th international conference on axiomatic design, Firenze
201. Suzuki S, Pan Y, Harashima F, Furuta K (2005) Skill Analysis of human in machine operation. In: Proceedings of the international conference on neural networks and brain, Beijing
202. Sycara K (1990) Negotiation planning: an AI approach. Eur J Oper Res 46:216–234
203. Taylor JM, Raskin V (2011) Understanding the unknown: unattested input processing in natural language, FUZZ-IEEE conference, Taipei
204. Taylor JM, Raskin V (2010) Fuzzy ontology for natural language. In: Proceedings of the 29th international conference of the North American fuzzy information processing society, Toronto
205. Thelan E, Smith L (2004) A dynamic systems approach to the development to of cognition and action. The MIT Press, Cambridge, MA
206. Thompson A, Mooney R (1994) Inductive learning for abductive diagnosis. In: Proceedings of the AAAI-94, Palo Alto, pp 664–669

207. Torralba A (2003) Contextual priming for object detection. Int J Comput Vis 53:169–191
208. Vaihinger H (1925) The philosophy of 'As If': a system of the theoretical, practical, and religious fictions of mankind. Kessinger Legacy Reprints, Kessinger Publishing, LLC
209. Valdez-Peres P (1998) Spatio-temporal reasoning and linear inequalities. MIT Artificial Intelligence Laboratory, Cambridge
210. Warren P (2006) Launching a new kind of warfare. The Guardian, www.guardian.co.uk/theguradian
211. Wilson E (2006) The melancholy android: on the psychology of sacred machines. State University of New York Press, Albany
212. Winograd T (2001) Architectures for context. Hum Comput Interact 16:401–419
213. Wogulis J (1991) Revising relational domain theories. In: Proceedings of the 8th international workshop on machine learning, Morgan Kaufmann, Evanston, pp 462–466
214. Wooldridge M (2006) Agents as a Rorschack test: a response to Franklin and Graesser. Springer, London
215. Wooldridge M, Jennings N (1995) Intelligent agents: theory and practice. Knowl Eng Rev 10(2):115–152
216. Arens Y, Chee CY, Hsu C-N, Knoblock CA (1993) Retrieving and integrating data from multiple information sources. Int J Intell Cooper Inform Syst 2(2):127–158
217. Yamauchi B, Beer D (2004) Sequential behavior and learning in evolved dynamical neural networks. Adapt Behav 24(3):219–246
218. Yang Y, Raine A (November 2009) Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: a meta-analysis. Psychiatry Res 174(2):81–8. doi:10.1016/j.pscychresns.2009.03.012. PMID 19833485
219. Zadeh L (2004) A note of web intelligence, world knowledge, and fuzzy logic. Data Knowl Eng 50:291–304
220. Landauer T, Dumas S (1997) Solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychol Rev 194:211–240
221. Griffiths T, Steyvers M (2003) Predictions and semantic associations. In: Kalmberg, Steyvers (eds) Advances in neural information processing systems 15. MIT Press, Cambridge, MA, pp 11–18
222. Stillings N (1995) Cognitive science: an introduction, 2nd edn. MIT Press, Cambridge
223. Watts R (2003) Therapy as a rational constructivist approach. Fam J Couns Ther Couples Fam 11(2):139–147
224. Rotter J (1954) Social learning and clinical psychology. Prentice-Hall, New York
225. Jones B (1999) Bounded rationality. Ann Rev Polit Sci 1999(2):297–321
226. Rosenblatt F (1962) Principles of neurodynamics. Spartan Books, Washington, DC
227. Sternberg S (1966) High-speed scanning in human memory. Science 153:652–654
228. Duncan A (2008) Eliminate the impossible: an examination of the world of Sherlock Holmes on page and screen. MX Publishing, London. ISBN 978-1-904312-31-4
229. Carbone J (2010) Framework for enhancing transdisciplinary research knowledge. Texas Tech University Press, Lubbock
230. Jacobs R, Jordan M, Nowlan S, Hinton G (1991) Adaptive mixtures of local experts. Neural Comput 3(1):79–87
231. Hopfield J (1982) Neural networks and physical systems with emergent collective computational abilities. In: Proceedings of the national academy of science, Washington, DC, vol 79(8), pp 2554–3558
232. Fishburn P (1989) Retrospective on the utility theory of Von Neumann and Morgenstern. J Risk Uncertain 2:127–157
233. Machina M (1982) Expected utility: analysis without the Independence Axiom. Econometrica 50:277–323
234. Akao Y (2004) Quality function deployment: integrating customer requirements into product design. Productivity PR
235. Suh N (2005) Complexity theory and applications. Oxford University Press, New York
236. Kandel E, Schwartz J, Jessel T The anatomical organization of the central nervous system. Principles of neural science (Ch. 17), McGraw Hill Professional, ISBN 978-0-8383-7701-1

Index

J.A. Crowder et al., *Artificial Cognition Architectures*,
DOI 10.1007/978-1-4614-8072-3, © Springer Science+Business Media New York 2014