SAND2007-8130C

13th ICCRTS "C2 for Complex endeavors" Automated Decision Support in a Complex Information Space.

Topic 4: Cognitive and Social Issues

Topic 8: C2 Architectures

Topic 9: Collaborative Technologies for Network-Centric Operations

Michael Senglaub, PhD
Sandia National Laboratories
PO Box 5800 MS 1161
Albuquerque, NM 87185
505-844-9244
mesengl@sandia.gov

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company,for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

SAND2007-8130C

Automated Decision Support in a Complex Information Space

Michael Senglaub, PhD
Sandia National Laboratories
PO Box 5800 MS 1161
Albuquerque, NM 87185

Abstract

A decision support architecture and embedded functionality is described that supports a decision maker in very complex environments dealing with massive amounts of disparate data, information and knowledge. To demonstrate some of the existing capabilities a couple of application domains are discussed. The solution to this system is a hybrid solution employing a number of technologies that are based on Peircean reasoning, modal logic, and formal concept analysis. The primary requirement of all the supporting technologies is that they have a basis in mathematical theory which ensures a validation process of the results based on mathematical solutions as opposed to conjecture or some form of utility function lacking verifiability. While the capability is rather robust in its present configuration, areas for further development and some research areas are identified with the intent of defining a complete decision support solution that is adequate for multiple domains and process configurations.

1. Introduction

The trend in combat C2 as well as emergency management situations is to increase the flow of data into command centers. These data flows span dimensions of discipline, multi-national and government jurisdictions, with redundant sources, multiple protocol, and of varying levels of temporal importance. One problem facing the decision makers is a failure to transform data into information and knowledge prior to being integrated into the decision process. A result is an increase in the cognitive loads imposed on the decision maker, effectively overwhelming the process and leading to sub-optimal decisions. This paper defines the theoretical foundations and the engineering solutions used in the decision aid being developed at Sandia national Labs. This section articulates the paradigms used in the decision support solution. In section 2, the theoretical foundations for the knowledge representation technology is articulated. Section 3 covers the theoretical foundations of the Peircean based reasoning engines as well as the J.S. Mill's inspired knowledge operators.

Section 4 provides examples of applications in the domain of intelligence analysis and nuclear forensics. These examples begin to show the capability that currently exists, using the hybridized algorithms. Section 5 provides extensions to the model to add greater flexibility and robustness, followed by a section capturing some conclusion from the effort to date.

1.1. Heuristics vs Physics

A fundamental difference in representing human behaviors in information systems involves the source of those theories of representation. The source ranges from heuristics to physics and is similar to the problem in thermodynamics. In thermodynamics the 'laws' of heat transfer between bodies are a heuristic representation of a molecular phenomena. In order to understand the physics of this energy transfer process requires an understanding of statistical physics, Theories of human information processing can be found in the fields of cognitive psychology, neural-physiology or in philosophical theories. Modal logic and philosophy provide a heuristic assessment of the functions of human reasoning at what could be described as an engineering level. The neural-physiological models are the 'physics' of human cognitive functionality, it defines the electrical-chemical dynamics of the processes associated with this functionality. In the opinion of this author, cognitive psychology lies somewhere between these extremes. One problem seems to be sets of overly constrained, and under controlled experiments in which too much physics is read into the results of these experiments. At best you can define a heuristic at a functional level. A review of inductive reasoning (Ref. 7) highlights these problems, whether or not that was the intent.

1.2. Decision making Paradigm

The effort described focuses on the problem of decision support technologies in command systems. Decision support must be approached from a non-intrusive

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

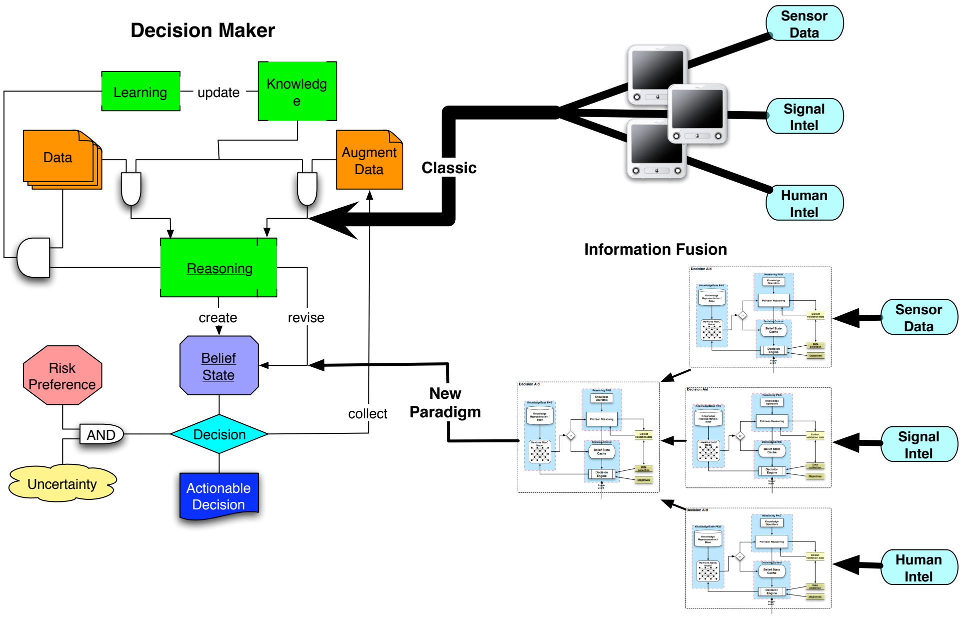

Figure 1. Model of the decision making process.

perspective and support a model of command. Our decision model is represented on the left in Figure 1.

While simplistic in design it captures a couple of key elements that seem to have been missed in the design of many decision aids. The first is the fact that data/information collected must be convolved with knowledge in order to construct situational awareness. The belief state representing situational awareness, provides the basis for decisions by a decision maker. Second, a belief state, can be in error due to errors in the data/information collected, or in the knowledge used to convolve that information. Any system design and its supporting logic systems must address both types of errors to deal with knowledge and belief modification.

Finally, in order to mitigate cognitive loads, the information systems need to be able to convert data to information so that they can interface to the decision makers belief state and not interface to their reasoning functionality. Having to reason about the information being collected adds to the cognitive load and ultimately leads to information overload. A well conceived design interface to the belief state would operate by postulating that a decision makers belief state may be in error. E.g., 'that pile of rocks on the side of the road may be an IED.'

2. Knowledge Representation

Development of reasoning based systems such as decision aids or data and information fusion engines requires attention to a knowledge representation technology that is flexible and enables one to transform the information into alternative representation domains to support the functional needs of a user community. In addition, the system must implement a robust suite of modal logics to ensure a theoretical rigor on all operations involving information and knowledge processed by the information system.

Our effort focused on the formal concept analysis technology with its strong mathematical foundations. Modifications or extensions permit us to treat continuous valued attributes and, with minor changes in our system state paradigm, to easily integrate temporal information into our information and knowledge space. Most application domains require an extensive integration of modal logic into the system to deal with knowledge, knowledge update, belief, belief revision and temporal logics. While not completely implemented, zero and first order solutions and a structure supporting these logics is in place.

This section lays the ground work for the robust knowledge representation domain used in our hybrid solution to information and decision support systems.

2.1. Ordered Set Theory

Interestingly order theory, with its roots in mathematics, seems to permeate all aspects of life. Even in pure chaos there exists an underlying order which can be seen in the phase trajectories representing the dynamics of the system. Order is not a property of an element but concerns the ability to compare objects. This aspect of order, is the ability to compare elements to determine which is greater, smaller or equal. Partially ordered set theory (Ref. 3) is based on a structure P = (P, ≤) consisting of a set P and a binary relation ≤. The resultant is a partial order if it is reflexive, antisymmetric and transitive. Therefore for all a, b, and c in P the following must hold:

$$
\begin{aligned}
&(i)\; a \leq a \\
&(ii)\; a \leq b \;and\; b \leq a \Rightarrow a = b \\
&(iii)\; a \leq b \;and\; b \leq c \Rightarrow a \leq c
\end{aligned}
\qquad \text{Eqn 2.1}
$$

these relations define reflectivity, antisymmetry, and transitivity respectively.

A mathematical construct from order theory that has proven useful involves the ideal and the filter of an ordered set. The definition for the ideal is the following; a subset L of $\boldsymbol{P}$ = (P, ≤)

$$
\begin{aligned}
&\forall x \in L,\, y \leq x \Rightarrow y \in L \\
&\forall x, y \in L, \exists z \in L : x \leq z \;and\; y \leq z
\end{aligned}
\qquad \text{Eqn 2.2}
$$

The definition for the filter is the following; a subset U of $\boldsymbol{P}$ = (P, ≤)

$$
\begin{aligned}
&\forall x \in U,\, x \leq y \Rightarrow y \in U \\
&\forall x, y \in U, \exists z \in U : z \leq x \;and\; z \leq y
\end{aligned}
\qquad \text{Eqn 2.3}
$$

The mathematics of order theory and/or ordered set theory sets the stage for the foundation of formal concept analysis. In formal concept analysis we operate on sets of sets which adds a degree of complexity to the concept of partially ordered sets and the comparator operator. The additional effort involves defining a binary relationship operator which compares a construct from formal concept analysis, the formal concept. In the remainder of this section we see the development of formal concept analysis and its use as a knowledge representation technology.

2.2. Formal Concept Analysis

Formal concept analysis (FCA) is a knowledge representation development effort initiated by Ganter and Wille (Ref. 6) with foundations in ordered set and lattice theory. The mathematics of FCA lends itself to the rich representation capabilities of lattice theory. FCA is based on the idea of a formal context, $\mathcal{K}_{FC}$, defined by a 'triple' as the one in equation 2.4.

$$K_{FC} = (G, M, I) \qquad \text{Eqn 2.4}$$

In this expression, G and M are sets of objects and attributes respectively and I is a binary relation between the two sets. A formal context might be viewed as a block of information that is domain or topic consistent. E.g., it might represent the planets in our solar system, or the types of IED devices in common use.

The concept construct in formal concept analysis is a structure determined by its extent, a set of objects, and its intent, the corresponding set of attributes common to the set of objects. Using the solar systems planets as the context, the concept consisting of the extent, A={earth, mars} possesses an intent, B, consisting of the common attributes,{small-size, near-the-sun, has-moon}. This is an example of a concept (A,B). Ordering is based on the idea, a concept is less general if the extent of (A_1, B_1) is a subset of the extent of concept (A_2, B_2).

Following on the tails of the description of a concept, an operator is defined, $(\cdot)'$ which aids in the definition of formal concepts from the formal context.

$$
\begin{aligned}
(A)' &\equiv \{m \in M \,|\, (g,m) \in I, \forall g \in A\} \\
(B)' &\equiv \{g \in G \,|\, (g,m) \in I, \forall m \in B\}
\end{aligned}
\qquad \text{Eqn 2.5}
$$

In this expression, the operator action on the object set A, produces the set of attributes common to objects within the 'A' set. Likewise, application of the operator on a set of attributes B, produces the set of objects which posses those attributes in common. This operator permits us to construct concepts associated with a particular context, providing a basis for constructing lattices for use in visual interpretations of information and knowledge within the knowledge base.

The raw context must go through a formalization process in order to take advantage of the capabilities of lattice theory. These capabilities provide a basis for aiding the analyst in understanding the collected data/ information. This process requires the construction of a 'Begriff' which represents all concepts in a context. This 'Begriff' is used in the construction of the lattice. The Begriff, $\mathcal{B}$(G,M,I), is the ordered set of all concepts within a context. A concept, consisting of the set-of-sets (A,B), is defined by conditions in equation 2.6.

$$(A,B) \xrightarrow{fc} (G,M,I)$$
$$\Leftrightarrow$$
$$A \subseteq G,\ B \subseteq M \qquad \text{Eqn 2.6}$$
$$(A)' = B \,\&\, (B)' = A$$

The ordering of the concepts in $\mathcal{B}$(G,M,I) is defined in the next expression.

$$(A_1, B_1) \leq (A_2, B_2)$$
$$\Leftrightarrow \qquad \text{Eqn 2.7}$$
$$A_1 \subseteq A_2 \vee B_2 \subseteq B_1$$

An example of a lattice is provided from information developed by K. Wolff (Ref. 21) in his FCA tutorial. This example is a simple model capturing aspects of a knowledge base dealing with animals. The data consists of the following instances:

Lion={preying, mammal}, Finch={flying, bird}, Eagle={preying, flying, bird}, Hare={mammal}, Ostrich={bird}. The cross table representation of this information is provided in Table 1.

Animals	Preying	Flying	Bird	mam-
Lion	x			x
Finch		x	x	
Eagle	x	x	x	
Hare				x
Ostrich			x	
Bee		x		

Table 1. Cross-table of an animal context.

The lattice representation of this information produced by CONEXP (Ref. 25) is shown in Figure 2.

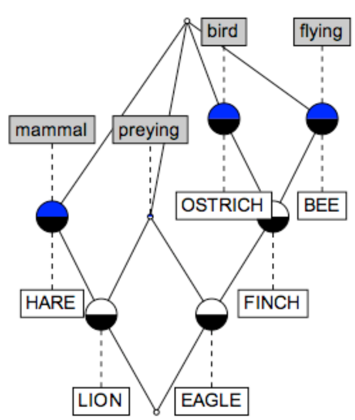

Figure 2. Lattice of animal domain.

The expansion capability of this technology is captured by the "Bee" entry in the matrix. The lattice prior to the addition of the information related to the bee consists of information in Figure 2 with the upper right node (BEE) removed. Expanding a knowledge base is a simple task in this technology. Likewise, the parsing of a lattice can be accomplished nearly as easily. What this does is give us the ability to structure the lattice at varying levels of knowledge abstraction and when additional detailed information is of interest we can "zoom" into an object node to see the additional structure of the knowledge base under the selected node. This mechanical process adds to the potential understanding of knowledge and data being worked with.

The situation we find in the real world is that attributes are often defined by continuous real, probabilistic temporal or even state variables. Working with continuous real variables in formal concept analysis is achieved by defining a special construct called a "many valued context"(Ref. 5 & 19). The structure of a many valued context is defined in the next expression.

$$\mathcal{K}_{mv} = (G, M, W, I) \qquad \text{Eqn 2.8}$$

G, M, and W are sets of objects, attributes, and attribute values. In this extension, the set of all values an attribute may assume is defined by the domain of that attribute.

$$D(m) \equiv g \in G \,|\, (g, m, w) \in I, w \in W \qquad \text{Eqn 2.9}$$

The conditional expression may be read as, the object g has the value w for attribute m. To use many-valued contexts in formal concept analyses these attributes must go through a scaling process in order to generate a formal context that identifies the presence or absence of an attribute. The scaling processes proposed in the literature, which is based on an inclusionary range, did not suite our needs. In order to handle real continuous valued attributes we employed fuzzy set theory with overlapping sets. The result of this approach resulted in our having to relax a condition imposed by Ganter(Ref. 5).

In his development he imposes a uniqueness on the attribute values as in the next expression.

$$(g, m, w_1), (g, m, w_2) \in I \Rightarrow w_1 = w_2 \qquad \text{Eqn 2.10}$$

Using overlapping fuzzy sets to define an attribute value requires us to relax this constraint which results in a more robust representation for real continuous attributes.

2.3. Mathematics of generating a Begriff.

Within the construct of this effort we have considered a context to represent a 'related block' of information, e.g. an explosives data set, or a sports car data set. The idea of a related block of information becomes important when we performed fuzzy variable transformations.

The Begriff is the set of all concepts of a context. The set of concepts can be defined by an application of the 'prime' operator discussed in the previous sections. The first step in defining the Begriff is to define an attribute set consisting of the intent of a context. The power set of this attribute set provides the basis set on which the prime operator is applied to produce the extent of all concepts in a context.

$$\begin{aligned} S &= \{Intent\}_{formal\ context} \\ P_s &= \{p\} = \{\mathcal{P}(S)\} \\ \mathcal{B} &= \{\textstyle\sum_k \{(p_k)'\}, \{p_k\}\} \end{aligned} \quad \text{Eqn 2.11}$$

P_s represents the power set of the contexts intent, and $\mathcal{B}$ is the Begriff which is a partially ordered set of concepts without duplications that may result from the process of applying the 'prime' operator on each member of the power set.

A procedural approach to defining a Begriff can be found in Davey & Priestley's (Ref. 3) book. This approach relies on a process that uses a series of set intersections as the context is processed. The effect is the same while the equations above are a rigorous interpretation of the process described.

2.4. FCA Variations

One of the first extensions implemented in the knowledge representation algorithms involves a modification of the interaction operator. In the general definition of a formal concept, $\mathcal{K}_{FC}$,

$$K_{FC} = (G, M, I) \quad \text{Eqn 2.12}$$

The interaction operator I is replaced by a set of interaction operators. Each member of this set corresponds to a specific 'predicate' in the information domain. This permits us to correlate information with a source or assign some descriptive property to the binary relationship between the objects and attributes. The structure enables us to apply specially designed operators to blocks of information to enhance user understanding.

The interaction operator is represented by a set of operators linked to a set of predicates.

$$I_s \equiv (P, I_p, X) \quad \text{Eqn 2.13}$$

In this expression P is the set of predicates, I_p is the set of interaction operators and X defines the linkage between predicates and interaction operators. There is a degenerate form of the operator I_s which represents all predicates and can be defined as a sum over all the predicate operators.

$$I = \sum_p I_p \quad \text{Eqn 2.14}$$

The resultant operator is the same as the operator found in all the FCA literature.

There are a number of implications or interpretations that may be imposed on the system using this construct. The use of predicates in the implementation can be viewed as a means of organizing orthogonal information while reducing the complexity of the information being processed by an analyst. Consider the problem of IEDs, we can define attributes associated with the predicate 'is-constructed-using' that provides insights into the construction, while a corresponding predicate "emits" provides information that is useful in the detection of the IED. Attributes associated with the two predicates define knowledge associated with the IED. The predicates enable very different functions to be addressed in pursuit of solutions to a problem.

The most generic variation involves the attribute descriptions themselves. We have seen the modifications employed to enable treatment of attributes associated with real variables which is only one type of attribute needed to handle a broader class of problems. We have included probabilistic variables which are treated in a manner identical to the real valued attributes. The architecture of the Peircean decision aid allows for the development of spectral attributes and a development effort has been proposed to permit state based attributes to be defined. In another situation a spectral attribute, is intended to permit efforts in a problem domain in which we want to include such things as the acoustic spectrums of vehicles or seismic signatures when operating on geophysical problems.

We intend to extend the class of attributes to include state dependent attributes as an alternative methodology for working temporal problems. The philosophy of the system design is to use the mathematics most relevant to the problem being pursued. In the case of temporal issues the trend towards temporal concept analysis seems to overload FCA and not produce a capability that simplifies the problems encountered by an analyst / decision maker. It is felt that the use of Markov type technologies at the attribute level would maximize the effectiveness of the FCA technology and augment it with a technology that is better suited to working on temporal issues.

2.5. Fuzzy set theory

Formal concept analysis is based on a binary relationship between objects and attributes, an attribute is associated with an object or it is not. The problem is that in real situations many of the attributes may be real or spectral in character among other types. In order to transform real world information into a form amenable to FCA we use a process based on fuzzy set theory(Ref. 8). Within a context, basically a block of information, we assume that a real attributes posseses common interpretation. Temperature in a materials context, might represent a melting temperature or a phase transition temperature. This temperature should not be associated with the temperature defined to represent an engine operating temperature. If all temperature interpretations were lumped and fuzzified over the combined range, significant biases could be introduced as well as introducing fidelity issues into the knowledge repository.

Fuzzy set theory is an extension of set theory in which the membership function associated with an element of the set can be represented by the next expression.

$$\begin{aligned} &\textit{crisp} \\ &\mu_A : X \rightarrow \{0,1\} \\ &\textit{fuzzy} \\ &\mu_A : X \rightarrow [0,1] \end{aligned} \qquad \text{Eqn 2.15}$$

In a crisp set the membership values are 0 or 1 while for a fuzzy set the membership function value ranges over the interval 0 to 1. The membership function used in this application is based on a Gaussian distribution as is defined in equation 2.16.

$$\mu_k = e^{[-(x-c_k)^2/2\sigma^2]} \qquad \text{Eqn 2.16}$$

Identifying and isolating a real variable is the first step of the process. The range of that variable is determined and "padding" of 10% is added to the maximum and minimum values to ensure a degree of robustness to the context classification, enabling a small degree of projection.

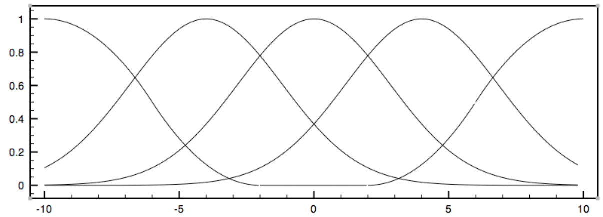

Figure 3. Fuzzification of a real variable over a range of -10 to 10.

Figure 3 above shows a real value fuzzification using 5 fuzzy levels. The membership functions are assumed to use the Gaussian membership function with sigmoid functions on either end of the range of values.

A variable value within the range covered by the fuzzy sets permits us to estimate the likelihood that the attribute belongs to each of the fuzzy intervals. In the implementation of the process we use an over lapping structure which permits a greater combinatory representation of a variable. For example a variable value of '5', has non-zero membership in 3 quantiles of the fuzzified variable. Effectively we have a 3-bit code representing the real attribute in some information domain.

A second feature of our implementation permits a user to define a threshold for membership. In this case the likelihood values must exceed the threshold in order for that quantile to be considered an attribute of an object. By defining the level of fuzzy set overlap and the threshold value we can change the degree of representation of real values in a particular context(information domain) . This gives us the ability to find a balance between uniqueness and computational effort. It also provides an analyst with a great deal of flexibility to discriminate information for use in a reasoning system.

3. Peircean Foundations

Charles Sanders Peirce was born on September 10, 1839 in Cambridge, Massachusetts, and he died on April 19, 1914 in Milford, Pennsylvania. His writings cover the period of about 1857 to 1914. His published works are in the neighborhood of 12,000 printed pages and his unpublished manuscripts number nearly 80,000 handwritten pages.

3.1. Definitions

The next few paragraphs are an attempt to identify elements that need to be better understood to assess wether we have achieved the objective of developing a truly Peircean decision aid. The definitions are from a number of sources (Ref. 14, 15, 16, 17, 18)

Category

The first dimension concerns the concept of category, of which Peirce identified three types/levels. This represents his effort to define a structure into which all phenomena can be grouped. These three categories or modes are 'Firstness', 'Secondness', and 'Thirdness' which give meaning to all phenomena and to all objects of thought. Basically, all phenomena may be regarded as manifestations of either 'Firstness' or quality, 'Secondness' or fact, and 'Thirdness' or laws that govern facts in the future. Peirce felt that these categories are evident/obvious to anyone who pays attention to what happens in the mind and that observation would produce these categories of knowledge.

To amplify slightly, firstness is a quality, such as the taste of banana, redness, or anxiety. Firstness is being "as is without regard for any other".

Secondness is the effect that firstness might have on us. One example cited involved pushing on a closed door, the resistance is an example of secondness, or the heat felt from a hot iron on our hand would be a form of secondness to the firstness of hot associated with the iron.

Thirdness begins to be more complicated, it mediates between secondness, fact, and firstness, possibility. Thirdness begins to address issues associated with rules or conditionals, enabling a predictive capability of information.

Peirce also talks about the degeneracy of secondness and thirdness but the most telling point is that all phenomena exhibit varying degrees of all categories. Which category is to be associated with a phenomena is the category that has the greatest extent or is most representative of that phenomena.

Sign

The second topic of importance to the total understanding involves the concept of 'sign'. In *The Essential Peirce*, (Ref. 15) he defines "a sign as a thing which serves to convey knowledge of some other thing, which it is said to stand for or represent." This thing is called the object of the sign; the idea in the mind that the sign excites, is called the interpretant of the sign. "*A sign stands for something to the idea which it produces or modifies... That for which it stands is called its object; that which it conveys, its meaning (the sign itself); and the idea to which it gives rise, its interpretant*.". Sign has a hypothetical 'if...then' status associated with it; firstness the potential (can be), secondness the factual (is), and thirdness the conditional (would be).

Peirce goes on to identify 3 forms of sign, an icon, an index, and a symbol. An icon is a sign that excites an idea naturally, they are a likeness of the object, for example the common symbol for man and women on a restroom door. "A*n index stands for its object by virture of a real connection with it, or because it forces the mind to attend to that object.*" Peirce uses the example of a weather vane as an indication/index of the direction of the wind. The symbol sign is the most complex of the signs. A symbol might be a word a phrase or even a treatise that references some phenomena, quantum mechanics could be viewed as a symbol. An icon can be linked to the category of firstness, an index to the category of secondness, and a symbol to the category of thirdness.

Logic

Logic generically is the theory and/or study of truth and the discovery of truth in signs (Ref. 9). Since all thought is conducted by the means of signs, Peirce describes logic as the science of the laws of signs, and divides logic into three areas of study: 1) Critic, which studies the relations of signs to their objects, by classifying arguments and assessing their validity. Critic is the logic that is normally understood by today's scientists, mathematicians, logicians, and engineers. From the Peircean perspective, critic also includes the logic of relations and the science of discovery or inquiry. 2) Speculative grammar, is the theory of the meaning of signs in all their forms. Speculative grammar, in simple terms, addresses the means by which logical thoughts may be represented. Peirce's triadic sign represents one approach. 3) Speculative rhetoric (methodeutic), studies the methods that ought to be pursued in the investigation and search for truth. It addresses methods of how signs maybe used to communicate from source to interpreter in an effort to maximize understanding. Peirce's speculative rhetoric provides a framework for a theory of communications that includes the utterer, the interpreter, and the sign

3.2. Reasoning

Reasoning is the process we as humans use to solve problems or make decisions. We all use reasoning, some use sophisticated philosophies, others use ad hoc reasoning, however,we all seem to be imbued with a basic inductive reasoning capability. The form taken is a function of our training and experience. Modal logic enters the equation in attempts to describe the flavors or nuances of the reasoning we employ. The ultimate form of reasoning is the method of scientific inquiry which was defined by C.S. Peirce (Re,f 15).

The reasoning engine implemented in this effort is based on C.S. Peirce's model of scientific inquiry. This philosophical construct provides the foundation for how we as humans reason about situations new to us. It consists of the three fundamental forms of reasoning ;deduction, induction and abduction. The logic associated with the abductive, deductive, and inductive forms of reasoning are captured in Figure 4 (Ref. 17) .

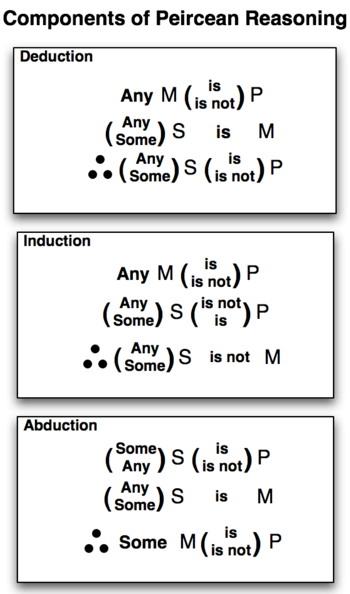

Figure 4. Formal representation of Peircean reasoning.

Peircean reasoning is a hybrid form that integrates these three foundational forms of reasoning into his method of scientific inquiry. Abduction is the more complex form of reasoning, it provides plausible hypotheses to explain an observation. Deduction provides a basis for selecting from that set of hypotheses. Deductive reasoning is based on a structure that concludes if the premise of an argument is true the resultant must be true, and provides the foundation for identifying what to expect with the selection of one of the abductively generated hypotheses.

Induction is the mechanism for validating the hypothesis selected. Induction can be viewed as a statistical collection of data used to confirm or support a hypothesis. Induction is used to support a reasoning process called inductive reasoning. Inductive reasoning operates on a principle that if '... I have thrown a ball in the air, and it fell to the ground every time...' I believe that the next time I throw the ball in the air it will fall to the ground.

The statistical validation must be tempered by maxims such as "severe" testing as defined by Mayo (Ref. 11). A second nuance of this problem is the frequentist perspective. Peirce and Mayo are frequentists and have developed theories from this perspective. The problem domain of decision support in command is really a Bayesian problem and these decision makers do not have the luxury of being frequentists, so the application must be tempered by Bayesian statistics.

Not addressed in this effort is analogical reasoning which is a form of abductive reasoning. The classic example of analogical reasoning is the Bohr atom example. Electron's were believed to revolve around the nucleus like planets revolve around the sun. Therefore, the forces in an atom should be able to be modeled using an inverse-square law. This form of hypothesis generation examines the detail of phenomena and looks for similarities at these levels of abstraction to draw higher level hypotheses.

3.3. Induction

A summery paper by Evan Heit (Ref. 7) provides a nice analysis of induction from a psychological perspective. The work focused on the evaluation of inductive arguments of a form given in equation 3.1.

$$\frac{\textit{Given a premise}}{\textit{The conclusion must follow}} \qquad \text{Eqn 3.1}$$

In deductive logic, if the premise is true the conclusion must be true, the inductive logic introduces an uncertainty. If the premise is true the conclusion is expected to be true. The psychological studies discussed in the review paper explored a number of factors associated with the argument and assessed the strength subjects placed on the conclusions.

The article describes conditions associated with good cases, sets of cases and properties for inductive reasoning. The examples they use, are presumed to be single premise cases. As an example they discuss examples such as:

$$\frac{\textit{Nearby houses were burgled}}{\textit{My house will be burgled}} \quad vs. \quad \frac{\textit{Houses } 50 \textit{ miles away were burgled}}{\textit{My house will be burgled}} \qquad \text{Eqn 3.2}$$

The first case leads to a stronger induction than the later in the study. They provide similar cases for problems involving animals described to exist on an isolated island in order to attempt some form of knowledge constraint. It seems that, what has been missed is an implicit multi-premise induction being performed. this 'implicit' set of premises was not captured. Subjects bring prior knowledge into the problem and evaluate the strength of the inductive arguments based on this knowledge. Effectively things such as; " ...50 miles from here is a large city with a high level of crime, my neighborhood lies in a very safe area so the likelihood of my house being burgled is low".

Similarly the animal examples fail to take into account the fact that we possess varying levels of biological knowledge and recognize that premise and conclusion involving species from a similar order may possess common anatomical characteristics or susceptibilities. Their discussion of 'property', reflects a similar issue. Certain classes of predicate carry varying levels of prior knowledge or meaning and impart varying levels of strength to the inductive arguments. One example uses the predicates 'thrives' versus 'secretes', the first is qualitative in nature while the second has a foundation in biology. The bottom line for me is, the experiments discussed could not be considered single premise experiments and results based on that supposition have to be flawed.

The discussion of sets of cases was of more interest in that it focused on the numbers and the diversity of the premises category. The greater the number of examples supporting a conclusion, adds strength as does a greater diversity of the examples (premises) in the inductive argument. It would seem that this is getting closer to the nature of induction and the inherent statistical character of the problem. Within the area of case sets no results dealing with counter examples was explored, This is likely due to the single premise constraint imposed on the experiments. In this situation, the set of attributes and corresponding set of objects can represent instances that would result in both true and false conclusions.

Induction Definitions

Within the domain of philosophy a significant body of research exists that addresses induction. The work used as the foundations for the Sandia decision aid is based on Peirce's modal of reasoning coupled to the work of V. Finn (Ref. 1, 4). Peirce defined induction '...as a form of reasoning from a sample to the whole sampled.' 'Induction is the mode of reasoning which adopts a conclusion as approximate.' Peirce indicated that there exists three kinds of induction all based on random samples. He seems to have used different terms for these types of induction but in general they consisted of a weak form, a strong form, and the 'gradual' form.

The strong form of induction consists of a sample or collection from a population in which it is possible to assess the proportion of the members of that population. The weak form deals with statements that could be disproved <u>if</u> a single counter example existed; e.g. 'liberals are intellectually bankrupt'. The third form of induction, gradual, is similar to the first form due to its quantitative nature. In this case an estimate of the population proportionality is made but each new sample acquired goes toward updating the proportions in the population.

Inductive Reasoning

Reasoning research in Russia has produced some very interesting results in which FCA is used as a knowledge representation technology and theories of abduction and induction have been developed. The work of V. Finn and V. Blinova(Ref. 1, 4) have used J.S. Mills (Ref. 12) canons as the guiding principles in developing a theoretical modal of inductive reasoning. Mills canons have been described by various researchers, as a set of inductive principles, a set of abductive principles and defined as principles describing causal reasoning. It seems that, in part due to ambiguities in definitions, that Mills canons have elements of each category.

Finn's (Ref. 1, 4, 10) work focused on the first of Mills five canons. His effort focused on an inductive 'learning' algorithm which may more accurately be characterized as an inductive reasoning algorithm.

Mill's (Ref. 12) canons consist of : the 'Method of agreement', 'Method of differences', 'Indirect method', 'Method of concomitant variation', and the 'Method of residues'. The descriptions of the canons that follow come directly from Mill's *System of Logic*.

The first canon: *If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree, is the cause (or effect) of the given phenomenon*.

Finn's implementation of the first canon involves considering positive and negative examples of an objective that draw from a single set of attributes, and one or more unknown examples. The objective is to classify the unknowns according to the reference sets. The model develops a Begriff for the positive and negative example sets, then uses the concepts in each Begriff to classify the unknown examples as either a positive, negative or as an indeterminate.

The example in Blinova's paper consists of the positive and negative sets in equation 3.3.

$$
\begin{array}{ll}
g_1^+ = \{a,b,c\} & \\
g_2^+ = \{a,b,d\} & g_5^- = \{a,c,d\} \\
g_3^+ = \{a,b,e\} & g_6^- = \{b,c,d\} \\
g_4^+ = \{a,c,e\} & g_7^- = \{a,d,e\}
\end{array}
\qquad \text{Eqn 3.3}
$$

The extra intent sets of the positive lattice consisted of {a,b}, {a,e}, {a,c}, and {a}. The extra intent sets of the negative lattice consisted of, {c,d}, {a,d}, and {d}.

The set of unknowns is defined in equation 3.4.

$$
\begin{array}{l}
g_8^\tau = \{a,b,c,e\} \\
g_9^\tau = \{c,d,e\} \\
g_{10}^\tau = \{a,b,c,d\}
\end{array}
\qquad \text{Eqn 3.4}
$$

Finn's method classified the first case as a positive example, the second as a negative example and the third as indeterminate due to both positive and negative intents being subsets of the example intent.

The one aspect of Finn's method that needs modification is the heuristic imposed on the positive and negative lattices. Finn requires that there be '2' or more examples in a concept before it can be classified as a positive or negative concept in the respective lattices. From a theoretical perspective, imposing a heuristic goes against the grain. This constraint is better suited for implementation in the engineered solutions where we might wish to establish a bias toward false positive or false negative classifications.

Inductive Learning

An excellent foundation for inductive reasoning has been established by Finn's work with Mill's first canon. In our implementation we have opted for an inductive learning system as opposed to the reasoning implementation. The algorithm builds on Finn's work with minor modifications and extensions. In Finn's work he defines a positive and negative lattice to use in the classification process. We needed to learn in situations in which the context possessed an arbitrary number of goals. E.g., the terrorist incident database used contained over 30 groups which represented the goal attribute. The algorithm developed is provided in the next paragraphs.

The algorithm begins by accepting the entire training set. In the example to follow a subset of the State Dept. terror incident database was used. The incidents covered a period from 1992 to 1998. The attributes describing the incident include; {date, location, target, tactic, result, organization}. Organization is the goal attribute. A Begriff is defined for each goal attribute in the training set. These Begriffs constitute the positive lattices defined in Blinova's paper.

$$
\begin{array}{l}
B_k \equiv (A_k, B_k) |\ A_k \subseteq A;\ B_k = (A_k)' \\
with \\
k = goal\ attribute
\end{array}
\qquad \text{Eqn 3.5}
$$

The first adjustment occurs in the construction of the goal oriented Begriffs. Engineering factors are introduced to bias the ultimate algorithms toward false positives or false negatives. This is done by setting likelihood thresholds to some fraction of '1.0'. This translates into the number of examples associated with a concept in the lattice. We can require that for a positive example we require two or more examples or in the case of counter examples two or more examples to reflect a counter instance. The next step of the algorithm is to define the resultant classification Begriff after any counter examples are removed from the goal Begriff.

$$
B_k^* \equiv B_k^\sim - \sum_{m \neq k} B_m^\sim \qquad \text{Eqn 3.6}
$$

In expression 3.6, B* represents the positive classification Begriff for a goal attribute, in our test problem this might be the group HAMAS. The Begriff class $B^\sim$, represents the post processed Begriff, with the engineering biases included. This product becomes the basis for the new classification context that will be constructed by the next steps.

Each remaining concept of the classification Begriff represent positive examples for identifying a specific goal. A concept, C = (A,B), represents a set of examples and the associated set of identifying attributes. Classification of 'k' begins by finding:

$$
(\{\ \}_k)' = \{a_k\} \qquad \text{Eqn 3.7}
$$

This is the set of all attributes associated with a goal. The next step is to define a likelihood for each attribute in the classification.

$$
lh_{k,m} = \sum_{j \in k} (A_{k,j})' / SizeOf(B_k^*) \qquad \text{Eqn 3.8}
$$

In expression 3.8, m is an attribute index in the classification Begriff k. j is an index for the concepts in the classification Begriff. SizeOf(.) is a function which determines the size of the Begriff or counts the remaining concepts in the classification Begriff. This process is repeated for each goal identified in the training set. The final classification context captures the essence of the training set in which each gaol attribute is an object in the constructed classification context, with each attribute assessed for likelihood as a discriminator for each object in the newly constructed context.

3.4. Abduction.

The basis for the abduction concept in this effort is again based on some of C.S. Peirce's (Ref. 15, 17) work. Trying to put this into some context has proven to be somewhat difficult and will have to evolve on a continuing basis as this author continues to learn more of Peirce's body of knowledge. This being no small undertaking and is unlikely to be accomplished by a single person.

Peircean abduction involves a broader understanding of Peirces ideas concerning categories, sign, and logic. After 4-5 years of building a system that attempts to follow his principles we have only scratched the surface and have not been able to put everything into a concise block of information that can be efficiently passed on to another engineer.

Abductive Reasoning

Peirce gives credit for the concept of abduction to Aristotle, however the literature seems to agree that Peircean abduction is unique to his body of work and should be credited to him. Peirce believes that abductive reasoning is the means for creating new knowledge. Abduction involves studying facts and information in order to devise some theory to explain these observations. Peircean abduction begins with the observation of some 'surprising' fact. A hypothesis or set of hypotheses are generated to explain that surprising fact. The hypothesis is of a form '..if theory A were correct then observation C would follow'. Therefore you are lead to suspect A to be the reason for the surprising observation. Surprise, in this context implies your belief state is now in doubt. Being a rational being we want to eliminate doubt and thus initiate the abductive process to resolve the surprising or disruptive fact.

An interesting paradox in this description is the point that should the law or theory A be known, C should not have been a surprising fact (Ref. 9). It is resolved by recognizing that A is a novel, creative solution to the doubt created by the surprising fact C. This creativity is the domain of research in trying to engineer an abductive engine. Peirce does provide a set of criteria for the degree of originality or quality of this creativity which could be used in the development process. The two relevant criteria involve a 'rearrangement' of old ideas to produce new insights while 'concept creation' which produces ideas new to the reasoner. This creative process is likely to produce multiple hypotheses which requires a selection process for identifying the best solution.

In order to identify the optimal hypothesis Peirce identified three criteria the most important being one of economy of effort. The criteria include explanation, verifiability, and economy. His economy of research involves assessing the cost to verify a hypothesis, the intrinsic value of the proposed hypothesis, and the effect of the hypothesis on future research. In terms of intrinsic value Peirce was attempting to find the simplest hypothesis that explains the facts.

These concepts form a foundation upon which abduction is based. In our limited effort we are attempting to adhere to these principles while still achieving an engineered solution to the problem of decision aids. It is for this reason and the criteria alluded to, that the process is an ongoing effort.

Mathematical Model of Abduction

The initial implementation of an abduction engine builds on the mathematical framework of formal concept analysis. The set structure of this knowledge representation technology permits us to use set operators to begin a process of defining hypotheses for observed data/information. In this first order system we are restricted to a very flat knowledge structure. Under these conditions blocks of knowledge within the knowledge base exhibit a degree of independence. When the knowledge base exhibits a hierarchical structure the hypotheses generated become more complex exhibiting an emergent characteristic that might be termed creativity. With the flat structure and knowledge independence some of Peirce's ideas concerned with surprise and the creativity of solutions is mitigated.

The process of abduction is initiated with the introduction of new externally derived information. this information is convolved with the resident knowledge incorporated in the knowledge base. The result is a set of potential hypotheses to the observed information. The operation is accomplished by performing a set intersection operation on the blocks of knowledge within the knowledge base and the observed information.

$$\{H\} \equiv \sum_j I \cap K_j \qquad \text{Eqn 3.9}$$

H represents the set of possible hypotheses, I, the set of unknown information and K_j a knowledge block. The summation is over the collected knowledge blocks in the knowledge base. The algorithms have been engineered to permit the user to set a threshold on the number of attributes that must be in common before a non-zero intersection is considered to produce an hypothesis. Because of the flat knowledge structure there is no automated integration of these hypotheses into higher level hypotheses which could conceivably lead to novel solutions to the observed information.

From a Peircean perspective the attributes associated with an object/instance in the knowledge block can be classified as a set of signs (of iconic and index form) which in total represent a sign (now a symbol) that must be interpreted as the object in the knowledge base. Expanding the capability of the system to handle hierarchical knowledge structures would add symbol to the type of attribute list defining an object in the knowledge base. This can be seen as objects at one level of knowledge being considered attributes at a higher level of knowledge abstraction. It might also be argued that even at the lowest levels, an attribute set may contain symbol signs as part of the argument list, which twe find agreement, but it is a bit easier to start at a level of abstraction in which the attributes can be classed as either icon or index.

Within the knowledge base are two types of knowledge, learned categorical knowledge and authoritative categorical knowledge. The difference is only in the means of construction, learned knowledge is generated by using the inductive learning engines in the Peircean decision aid and the authoritative knowledge is knowledge 'loaded' into the database from some outside source. The distinction comes in assessing the plausibility of an hypothesis. With authoritative knowledge all attributes are assumed to posses equal likelihood of viability or descriptiveness for the associated object. For learned knowledge we have varying degrees of attribute likelihood as a descriptor, and this can be taken into consideration for assessing the plausibility of a hypothesis generated by the set intersections. The result is a hypothesis may have greater or lessor plausibility for explaining some information.

Expanding the structure of the knowledge bases would modify the abductive process slightly permitting us to add depth and a degree of non-linearity to the process. The set operations would begin as in the earlier expression but a second stage would be added. Effectively another iterative loop would be added to the iterations. After the first pass in generating a set of hypotheses, the results would be added to the set of observed information.

$$\{I'\} \equiv \{I\} \cup \{H\} \qquad \text{Eqn 3.10}$$

The new information set I' is viewed as the observed information and then processed via equation 3.9. The result of this operation in a hierarchical knowledge base is to reproduce the hypotheses that resulted in the first expression and to add hypotheses that exist at higher levels of knowledge abstraction.

$$\{H'\} \equiv \sum_j I' \cap K_j \qquad \text{Eqn 3.11}$$

Performing a disjunction operation on the two sets of hypotheses produces the hypotheses at the next level of knowledge abstraction.

$$\{H_{lev2}\} = H \cap H' - H \qquad \text{Eqn 3.12}$$

The number of iterations to develop hypothesis depth should/could be limited to 8. This is the number layers in the neo-cortex of humans, where knowledge resides.

Construction of the Belief Cache

A belief cache is a mathematical representation of situational awareness. It is the product of observation and the knowledge brought to bear to understand the observations. The belief cache can be viewed as the tagged collection of validated hypotheses generated by the reasoning system. This cache contains the understanding to a point in time. The structure of this cache for our flat knowledge structure is defined in the next equation.

$$C_j \equiv \left\langle \begin{matrix} t_j, Active_j \\ \{h_{j,0}, \{d_n^s\}, \{d_m^v\}, \{d_r^{uc}\}\} \end{matrix} \right\rangle \qquad \text{Eqn 3.13}$$

These belief kernels consist of a time tag, t_j an activation flag, $Acitve_j$, a hypothesis, $h_{j,0}$, data collected that supports the hypothesis, $\{d^s{}_n\}$ and data that would validate the hypothesis, $\{d^v{}_m\}$. The last set of data consist of data collected that can not be resolved by selecting the indicated hypothesis.

For a hierarchical knowledge base, the structure of the belief cache would change to include hypotheses at different levels of knowledge abstraction. The form of this cache is represented in equation 5.6.

$$C_j \equiv \left\langle \begin{array}{c} t_j, Active_j \\ \{h_{j,0}, \{d_n^s\}, \{d_m^v\}, \{d_r^{uc}\}\} \\ \vdots \\ \{h_{j,r}, \{h_{j,r-x}\}\{d_n^s\}, \{d_m^v\}, \{d_r^{uc}\}\} \\ \vdots \end{array} \right\rangle \quad \text{Eqn 3.14}$$

What we have to do is add levels to the hypotheses that reflect the levels of knowledge abstraction. The second aspect of the change is hypotheses produced at lower levels of abstraction contribute to the hypotheses at a particular level. This construct for the belief cache, is needed to trace the impact of changes or updates to information at lower levels of abstraction.

3.5. Knowledge Operators

The operators being defined or designed for this reasoning construct are based on the five canons of John Stuart Mill (Ref. 12). Initial work by Burch and Finn (Ref. 2) have focused on the first canon and involved significant effort at validating these canons in a much larger philosophical and logic context. In this effort we are taking a more Peircean, pragmatic approach to selecting and implementing the operators. The five canons consist of those identified in the next list.

J.S. Mills (JSM) Canons

- Method of Agreement
- Method of Differences
- Indirect Method
- Method of Residues
- Method of Concomitant Variables

The first canon: *If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree, is the cause (or effect) of the given phenomenon*.

The second canon: *If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance save one in common, that one occurring only in the former; the circumstance in which alone the two instances differ, is the effect, or cause, or a necessary part of the cause, of the phenomenon.*

The third canon: *If two or more instances in which the phenomenon occurs have only one circumstance in common, while two or more instances in which it does not occur have nothing in common save the absence of the circumstance; the circumstance in which alone the two sets of instances differ, is the effect, or cause, or a necessary part of the cause, of the phenomenon.*

The fourth canon: *Subduct from any phenomenon such part as is known by previous inductions to be the effect of certain antecedents, and the residue of the phenomenon is the effect of the remaining antecedents*."

The fifth canon: *Whatever phenomenon varies in any manner whenever another phenomenon varies in some particular manner, is either a cause or an effect of that phenomenon, or is connected with it through some fact of causation.*

The descriptions of the canons come directly from Mill's *System of Logic*, (Ref. 12, pp224-233) and will form the basis for the knowledge operators in the system. Only the first 2 canons have been implemented in the coded algorithms.

Implementation of JSM-1.

The first canon, the method of agreement, addresses issues of learning. The implementation of the first operator, which is based on Finn's method, recognizes two types of attribute; a goal attribute and a structural attribute. Structural attributes are those describing an instance. The goal attribute is an attribute which describes a common characteristic, e.g., a terrorist group, or a specific nuclear reactor. Two types of context are produced in the algorithm, the positive context captures examples which are representations of a goal attribute, the negative context provides counter examples.

Our implementation is different from Finn's method for a number of reasons, first, to eliminate a heuristic and second, to formulate a learning algorithm for creating knowledge as opposed to a reasoning algorithm. The knowledge base contains classification contexts and knowledge derived from learning contexts. The modification uses the first canon to construct a context in which conditions (attributes) are identified that are characteristic of the goal condition.

$$\begin{aligned} C_{lc} &\Rightarrow \mathcal{B}^+, \mathcal{B}^- \\ L_{goal} &= \mathcal{B}^+ - \mathcal{B}^+ \cap \mathcal{B}^- \end{aligned} \quad \text{Eqn 3.15}$$

A learning context is converted to a positive and negative Begriff which is subtracted from the positive

Begriff producing an incomplete goal lattice. The set nature of the Begriff requires that the subtraction operation be defined as in the expression above.

The engineering aspects come into play through a threshold based on a likelihood estimate the associates the likelihood of an attribute being a descriptor for an object (goal). Imposing a likelihood of greater than 0.2 on concepts in the positive Begriff means that a single example may not be sufficient to constitute a positive example. This process reduces the probability of predicting unknown instances but reduces false positives. Similarly, by imposing a threshold on the negative Begriff results in reducing false negatives in a prediction problem.

Implementation of JSM-2.

The second canon, the method of differences, can be characterized as a causal reasoning operator. In this case we again split evidential context into positive and negative Begriff's. The first step is to identify common attributes for the positive evidence. Once we have created that positive common lattice we subtract concepts of the negative Begriff from this common lattice. mathematically the operations are described in the next equation.

$$
\begin{aligned}
C_{ev} &\Rightarrow \mathcal{B}^{+}, \mathcal{B}^{-} \\
(\mathcal{L}^{+})_{common} &= \mathcal{B}^{+}_{1}\left(\sum_{k}\left(\cap\, \mathcal{B}^{+}_{k}\right)\right) \qquad \text{Eqn 3.16} \\
L_{cause} &= (\mathcal{L}^{+})_{common} - (\mathcal{L}^{+})_{common} \cap \mathcal{B}^{-}
\end{aligned}
$$

Implementation of JSM-3.

Interestingly, the third cannon is a very simple variation of the second. The difference is the intersection term in the 3rd expression of equation 3.16 is 'zero' based on the description above. The result is that the algorithm for JSM-2 will also support JSM-3.

4. Application

The focus of this application is to provide decision support capabilities and / or augment the efforts of an intelligence analyst. The architecture to support these functions is defined in Figure 5.

The focus is the construction or assembly of knowledge which provides the basis for evaluation information collected through sensor and intelligence sources. The system enables the inclusion of modal logics in support of the various functions of the system. Many of these linkages are 'zero order' at this point and can be tailored to support the application domain. For example, the requirements on the disjunctive logic required in an automated system are going to be more stringent than for an application supporting an intel analyst. Similar arguments apply to the modal logics associated with knowledge construction and revision.

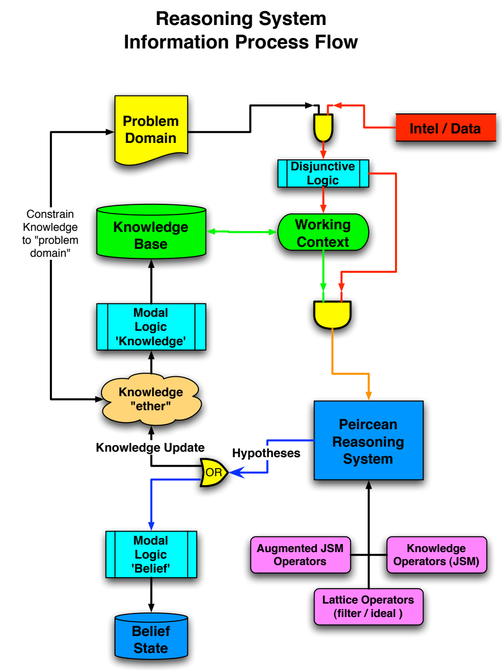

Figure 5. Decision support architecture.

4.1. Intel Analysis

One application domain we have been working supports an intel analyst. The rough scenario is an analyst is tasked with monitoring events for a possible terrorist attack in North America. Given this kind of problem there are many resources that may be utilized as preexisting knowledge bases, such as a terrorist incident database. In this case a database from state department reports from 1992-1998 was used in an effort to understand possible patterns of behavior and tactical preferences by various groups. The incidents were characterized by date, target, location, result, and the group responsible for the incident. The data was processed by the inductive learning engine in the Peicean Decision Aid (PDA) to construct the desired knowledge. The resultant knowledge base, in a lattice display, is depicted in Figure 6.

The highlighted section in the figure shows the result of a query concerning the area of operations. The knowledge base shows that 'HAMAS' and 'Islamic_Jihad' are the only groups operating in North America up to 1998. Similar queries show that HAMAS uses bombing as a tactic, impelling the analyst to key on information concerning the loss or theft of explosive material, as an example.

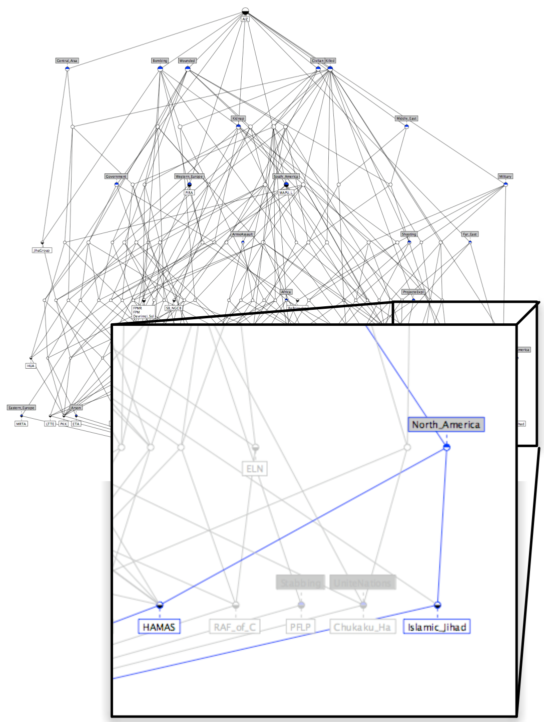

Figure 6. Terror incident database sample.

The raw data often conveys very little information to the analyst but in many cases it is possible to process the information, to convert it to knowledge and find interesting and useful knowledge in the transformed data. As in the case of the terror incident database, it is possible to process phone calls, bank transactions and other information bases to convert the information into knowledge permitting us to interpret data being collected by the analyst. Instead of seeing a series of transactions, we see linkages between banks known to launder money, the bank of a suspect, Confederate Bank, and a link to a new account / suspect at a Maryland bank, see Figure 7.

The zero order temporal logic in PDA, for example, provides a means to look for tactical trends by a group over time. This becomes critical, as behaviors can change over time and knowledge must be revised to reflect these shifting trends to ensure the hypothesis generation mechanism reflects current understanding.

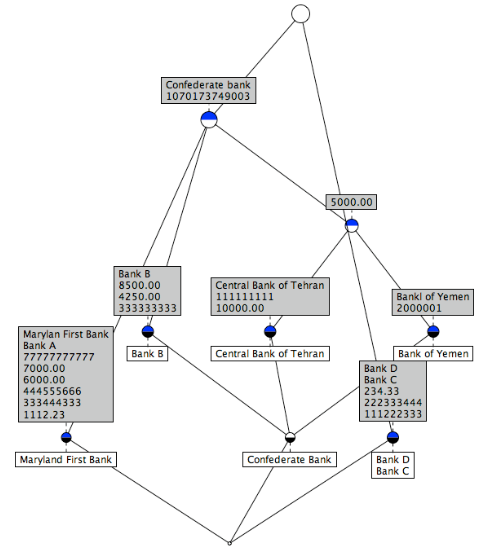

Figure 7. Knowledge associated with bank transactions.

Using the theories and technologies of Peircean reasoning provides the analyst with knowledge that can assist their daily activities as opposed to adding to their cognitive load. What we have is a capability that can process massive amounts of information that is likely to over whelm a decision maker faced with classical decision support technologies such as an air traffic control system. In these types of system raw data is presented to the decision maker who must reason or internally fuse the information provided.

4.2. Forensics Analysis

A second decision support area explored involved a problem of nuclear forensics. In this problem we have databases of assayed nuclear material as reference samples. Can we identify the source of black market interdicted material from these sample databases. The complexity of the problem can be rather daunting. The material carries its entire history in its chemical composition. The mining, enrichment, fabrication, operational history, and its reprocessing affects sample compositions. On top of this most materials are blended, i.e., new material is mixed with old material that may have gone through one or more life cycles.

This problem was appealing because it had the potential of breaking the algorithms as implemented within the PDA system. The sample data consisted of over 100 real valued attributes which were fuzzified into over 500 qualitative attributes for describing the materials in the database. What was discovered, we were able to break MS Excel but the PDA algorithms functioned without incident on the problem. In figure 8 is an example of a forensics lattice containing samples from various locations in eleven power reactors.

Figure 8 captures some of the complexity of the information contained in this knowledge base. The lattice does show that we can uniquely define the source of interdicted material. The bottom row of the lattice has separate instances for each reactor which is why we can make that assertion. This leads to a question, could we go so far as identify where in a reactor the material was taken. PDA has a number of thresholds that can be utilized in the inductive learning engine, these were used to begin a study of refining the source of the material. At this point of the analysis it appears that with a normalized axial and radial position descriptor this additional capability may be possible.

The system can be modified to permit an analyst to propose hypotheses and then have the system define which attributes to measure in an effort to validate any of the pre-determined hypotheses. This could aid in a rapid identification of source as well as reducing the potential cost of assaying a material sample. This last possibility goes directly to Peirce's, economy of research idea (Ref. 14).

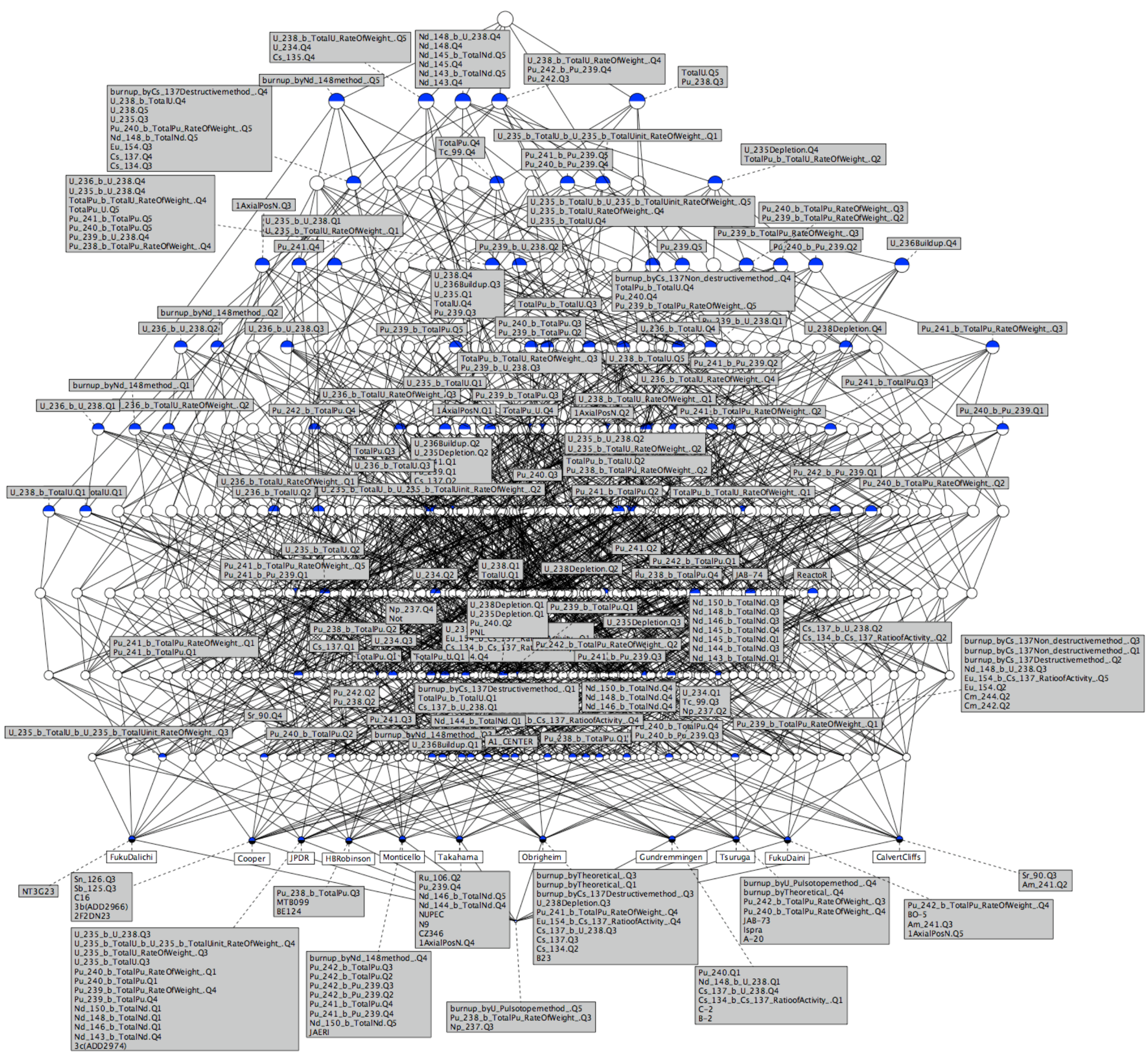

Figure 8. Example lattice of a forensics knowledge base (for purposes of demonstrating complexity).

5. Reasoning Extensions

The system has been developed with future extensions and enhancements in mind. The current version demonstrates the baseline potential and offers a 70-80% solution to a robust decision support capability. As with all development efforts there is never enough time or money to produce a complete solution. The following topics address approaches moving the capability towards the 90-95% solution.

5.1. Hierarchical Knowledge and Reasoning.

There are two approaches for extending the capabilities to include hierarchical reasoning. The first is a brute force approach which simply allows for objects to be used as attributes in higher level characterizations. This is a default structure and exists within the current system. If you use an ART neural net transformation of the FCA knowledge base, the knowledge structure could resemble the organizational structure seen in the Figure 9.

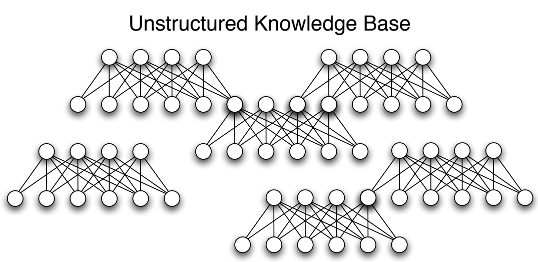

Figure 9. Unstructured knowledge base.

In this case you may find occurrences in which an object of one context is an attribute of another context as seen in the overlapping neural nets in Figure 10.

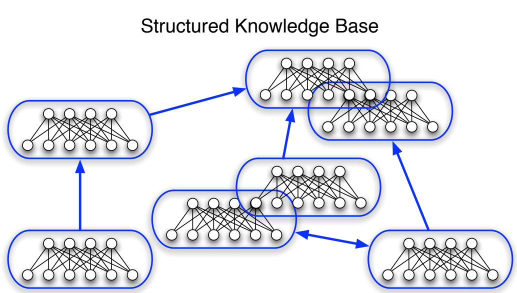

Figure 10. Structured knowledge base.

A better solution is to overlay the contexts within the knowledge base with a domain map, in which a single arrow indicates knowledge abstraction and double arrows signify a connection involving a similarity of the attributes within two contexts. 'Domain map' is used instead of 'concept map' because in formal concept analysis, concept has a specific mathematical meaning and over loading the term concept would lead to confusion.

The most important aspect of a structured knowledge base is the potential for defining an abductive reasoning engine that has the capability of employing analogical reasoning. It is this capability that would permit an automated decision aid to find truly novel solutions to new observations.

5.2. Temporal Logic and Reasoning

Andre' Trudel (Ref. 20) discusses a concept of temporal logic in which information collected or belief can affect not only future understanding but also past experience. Effectively we may re-interpret a past event based on new information. This perspective of time ties into our understanding of belief states and should be considered as belief is generated, or updated during the course of analysis.

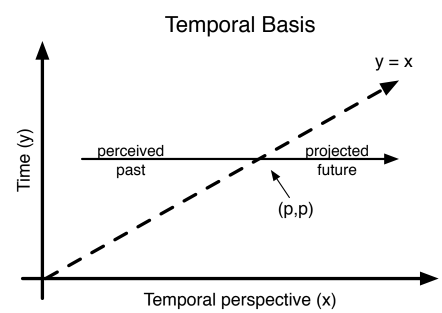

Figure 11. Model for temporal reasoning.

The premise is that 'the here & now' lies on the 45 deg. line defined by the x-y axis in Figure 11. This point is indicated by the point (p,p), the 'here and now'. The perceived past lies on a horizontal line emanating from (p,p) but for x < p. similarly the expected future lies on the line but for x > p. What the structure, or the paradigm brings to the table is a way to think about how new information can impact past belief which in turn can impact projections or predictions.

Temporal concept analysis is an extension of FCA in which the evolutions of the system or object are considered in conjunction with the conceptual aspects of the object. The principle researchers in the area, Wolff (Ref. 22, 23, 24) and Neouchi (Ref. 13), approach the problem by adding directed edges to the lattice to capture the evolutionary behaviors of the attributes.

Wolff's efforts have resulted in a very formal representation of the temporal extensions of FCA while Neouchi has focused on the development / definition of sets of operators that focus on issues associated with temporal concepts.

Wolff has approached temporal concept analysis by scaling the time and event space and adding directed edges to the concept lattice of the context. The potential difficulty of this approach can be seen in the simple example in Figure 12.

The blue vectors on the lattice in Figure 12 indicate the temporal evolution of the objects in the formal context. The red vectors show persistent states of objects in that context. What becomes clear, is the complexity of the display for even so simple an example. Complex information bases will rapidly overwhelm any advantages lattice representation brings to formal concept analyses.

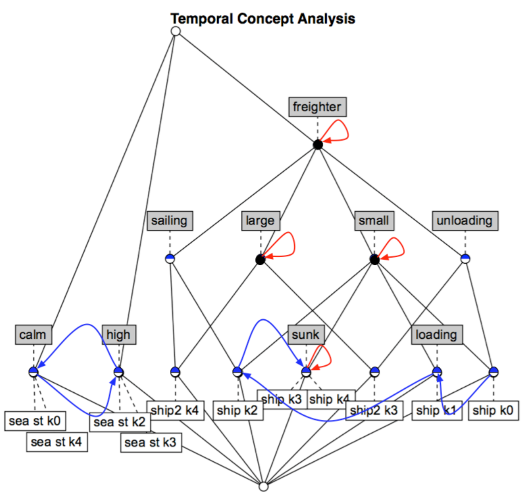

Figure 12. Example of lattice with directed edge overlay.

A way around this complexity issue is to redefine how we think about systems / objects and the states of those systems. Traditionally, we view a system in a specific state as a unique object, so we are forced in a FCA paradigm to replicate an object as many times as we have states for it. If we instead view the system as being unique with sets of constant or time dependent attributes we can reduce the complexity of the lattice.

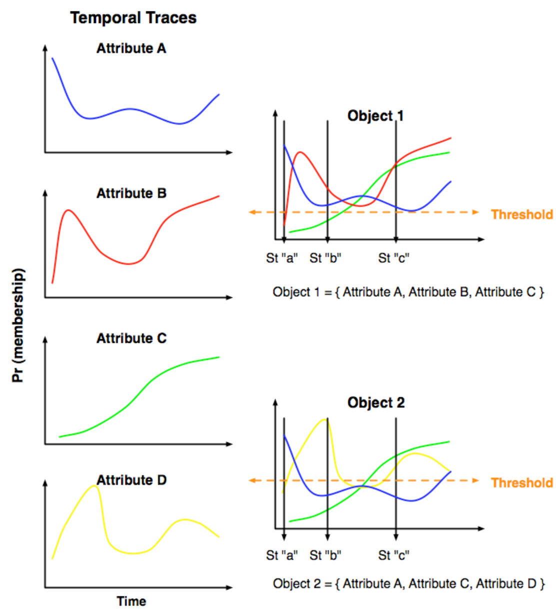

Figure 13. Temporal traces of four attributes and two objects with a mix of attributes.

We might be able to see these possibilities in more detail by considering the information in Figure 13. The notional example considers different temporal traces for the 4 attributes and a different set of attributes for two objects. We can see that taking a snapshot of these systems or objects at different points in time produces different collections of attributes describing the objects. This can also change with different threshold levels. At point 'a', object 1 is characterized by attributes A while object 2 by attributes A and D. If D was not in the data set the correct hypothesis could not be identified. Using a process of temporal matching could refine the hypothesis since A is present in object 1 at all three states while it is only present at state 'a' in object 2.

What we propose is to introduce a new attribute based on a Markovean type transition matrix. In this situation the Markovian attribute can define the state of the system which can be a simple temporal trace or a more complex material state that is the result of a complex life cycle process. For example, in the forensic problem discussed earlier, the material assay reflects the mining, enrichment, fabrication and operational history of a nuclear material. This information can be captured in the Markovian attribute.

5.3. Deductive-Inductive: Severe Testing

The final extension to be mentioned at this time is an automation of the hypothesis selection process. At present there is an assumption of an analysts intervention and interaction with the decision aid. In this case the analyst will exercise judgement in the selection and testing process for the validation and belief state generation process.

Mayo (Ref. 11) in conjunction with Peircean ideas provides a compelling theory for developing a deductive-inductive engine for evaluating hypotheses generated by our decision aid. She defines a framework in which to define the steps in a hypothesis testing function. Identified are three components consisting of a primary model, an experimental model and a data model.

The primary model involves assessing a hypothesis and identifying or hypothesizing signatures associated with the components of the model. The experimental model provides the linkage between the primary model and the data. The data model deals with what can be measured and how it relates to the hypothesis. The second major consideration talked about by Mayo is the idea of severe testing. In this case, the experiment that must be the result of the experimental design process is one that tests the validity of one hypothesis, with unambiguous results. A confirmation result can validate one and only one hypothesis.

While this appears easy enough we need to consider the economy of research idea that Peirce talks about. In some sense this may be considered an Ockham's razor type of consideration, the simplest answer is the best one. Any engine we design must adhere to these criteria. The mathematical structure of information and knowledge provides us with a foundation to build criteria meeting the requirements of Mayo and Peirce. The issue of Peircean cost may require the economic cost to be augmented with utility cost represented by some form of utility function. In this way a total cost may be assigned to a hypothesis. It is believed that a zero or first order solution should be very attainable.

6. Observations and Conclusions

What has been produced in this effort is a robust flexible decision support functionality that has its roots in reasoning, knowledge representation and logic theory. The system is a hybrid solution using these technologies in a manner in which the best technology is matched to function. We have applied the integrated solution to a couple problems utilizing differing process methodologies and have attempted to break the system by going beyond toy problems.

The development areas that could add benefit to the decision aid and strengthen solutions in weak areas have been identified. The architecture design is such that integrating more sophisticated solutions such as multiple modal logics will be readily implementable. We are also exploring a possibility of integrating the basic PDA library into a visual programming environment, such as Vipr, to enable a user to custom implement an analysis process that reflects the analyst's process used in performing their duties, as opposed to forcing them to adapt to a externally defined process paradigm.

The degree of hypothesis generation and selection can be automated to the degree of comfort desired by an analyst. The other aspect of the mathematical foundations being employed, gives us the ability to provide a justification or review capability of hypotheses generated by the system. There are no black boxes that obscure the logics of the underlying functionality.

Considering the linkage of the decision aid to the decision makers belief state and recognizing that the combined algorithms provide a reasoning based fusion capability, we have a solution that can have significant impact on the design of information systems. From figure 1 we see that integrating the Peircean algorithms into the information architecture close to the sensors permits us to minimize the flow of data to a central control facility. Instead of sending acoustic signatures to the decision maker, the system might simply send a message indicating a T-72 tank has been detected, significantly reducing bandwidth requirements.

6.1. Scalability

The knowledge representation methodology provides a mathematical basis permitting transformations to other representation technologies that support different technologies. One interesting application of this transformation capability is to structure the decision support algorithm into a real time and offline capability. The offline capability can work problems in back ground until unique knowledge bases are generated and transformed into neural net representations. Since neural nets are very fast running algorithms, and depending on the problem domain, very compact, we envision lightweight solutions that could in theory be implemented on small computational platforms, PDAs for example. Keeping in mind the decision support paradigm, in which we interface with the decision makers belief state, we could provide very compact, dynamic aids for use in dynamic environments such as the force protection environment.

References.

1. V. G. Blinova, D. A. Dobrynin, V. K. Finn, S. O. Kuznetsov, and e. al., "Toxicology analysis by means of the JSM-method," in Bioinformatics, 2003.

2. R. W. Burch, "Semeiotic data fusion," in Information Fusion, 2000.

3. B. A. Davey and H. A. Priestley, "Introduction to Lattices and Order," 2002, pp. 310.

4. V. K. Finn, T. Gergely, and S. O. Kuznetsov, "Plausible reasoning for open problem domains," in Intelligent Control, 1996.

5. B. Ganter, "Attribute exploration with background knowledge," in Theoretical Computer Science, 1999.

6. B. Ganter, R. Wille, T. H. Darmstadt, and F. Mathematik, "Applied Lattice Theory: Formal Concept Analysis," in cl.inf.tu-dresden.de, 1997.

7. E. Heit, "Properties of inductive reasoning," in Psychonomic Bulletin & Review, 2000.

8. G. J. Klir and T. A. Folger, "Fuzzy Sets, Uncertainty and Information," 1988.

9. G. J. M. Kruijff, "Peirce's late theory of abduction: A comprehensive account," in Semiotica, 2005.

10. S. O. Kuznetsov, "Machine Learning on the Basis of Formal Concept Analysis," in Automation and Remote Control, 2001.

11. R. D. G. Mayo, "Error and the Growth of Experimental Knowledge," 1996, pp. 509.

12. J. S Mill, "A system of logic, ratiocinative and inductive," 1856.

13. A. Neouchi, Y. Tawfik, and R. A. Frost, "Towards a Temporal Extension of Formal Concept Analysis," in Proceedings of the 14th Biennial Conference of the Canadian …, 2001.

14. E. T. Nozawa, "Peircean semeiotic: a new engineering paradigm for automatic and adaptive intelligent systems design," in Information Fusion, 2000.

15. C. S. Peirce, "The Essential Peirce: Selected Philosophical Writings. Volume 2 (1893-1913) - Page 361," 1998, pp. 640.

16. C. S Peirce, J. Hoopes (ed.), "Peirce on Signs: Writings on Semiotic - Page 189," 1991, pp. 284.

17. C. S Peirce, K.L. Ketner (ed.), "Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898," 1992, pp. 297.

18. C. S Peirce, P. Turrisi (ed.), "Pragmatism As a Principle and Method of Right Thinking: The 1903 Harvard ... - Page 6," 1997, pp. 305.

19. U. Priss, "Formal concept analysis in information science," in Annual Review of Information Science and Technology, 2006.

20. A. Trudel, "Atemporal structure that distinguishes between the past, ppresent, and future," in TIME-94 International Workshop on Temporal Representation and Reasoning, 1994, pp.146-152 .

21. K. E. Wolff, "A FIRST COURSE IN FORMAL CONCEPT ANALYSIS," in Soft-Stat, 1993.

22. K. E. Wolff, "Concepts, States, and Systems," in Computing Anticipatory Systems. CASYS99-Third International …, 1999.

23. K. E. Wolff, "Temporal Concept Analysis," in ICCS-2001 International Workshop on Concept Lattices-Based …, 2001.

24. K. E. Wolff and W. Yameogo, "Time Dimension, Objects, and Life Tracks-A Conceptual Analysis," in Analysis, UNK.

25. S. A. Yevtushenko. System of data analysis "Concept Explorer". (In Russian). Proceedings of the 7th national conference on Artificial Intelligence KII-2000, p. 127-134, Russia, 2000.