SAND2007-3483C

Knowledge Representation in Reasoning Systems

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

Abstract

An architecture was developed to support decision aids in which formal concept analysis(FCA) coupled to a Peircean reasoning engine was implemented to create a tailorable foundation for unique applications. The mathematical foundations associated with FCA permits transformations into alternative representations to take advantage of unique capabilities of information technologies maximizing the effectiveness of the information processing system. The reasoning engines and integrated modal logics provide a basis for verifiable information processing which is missing from many information processing systems.

1. Introduction

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 on information and knowledge processed by the information system. There are many examples of fragile systems that violate the most basic principles in logic and require process fixes to mitigate the potential problems. This class of systems also make the use of optimal technologies difficult if not impossible because of an inflexibility in representational space.

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. Our application domain requires 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 the structure supporting these logics is in place.

In this short article the knowledge representation approach and its extensions will be defined as well as the overarching reasoning system that supports decision aid design.

1.1. Architecture of Decision Aids

The problem addressed in the research and development of these technologies addresses the problem of decision support technologies in command systems. Decision support must be approached from a non-intrusive perspective and support a model of command. The command model is represented in the next figure.

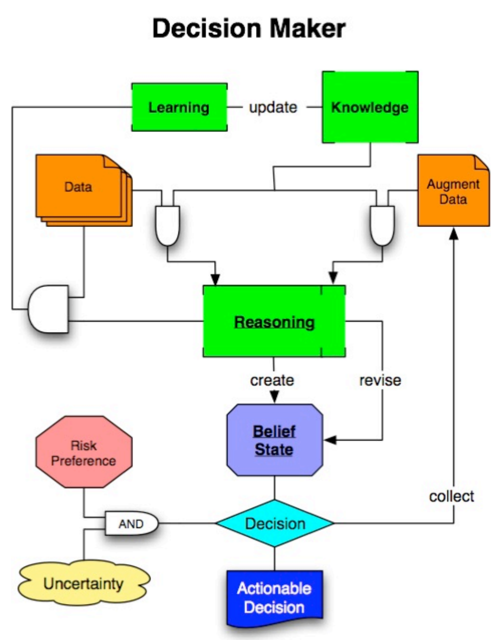

Figure 1. Model of the decision making process.

While simplistic in design it captures a couple key elements that seem to have 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.

2. Theoretical foundations of FCA

Formal concept analysis(FCA) is a knowledge representation development effort initiated by Ganter & Wille with foundations in ordered set theory. The mathematics of FCA lends itself to lattice theory and the rich representation capabilities of that domain. FCA is based on the idea of a formal context, $\mathcal{K}_{FC}$, defined by a "triple" as the one in equation 1.

$$K_{FC} = (G, M, I) \qquad \text{Eqn 1}$$

In this equation G and M are sets of objects and attributes respectively and I is a binary relation between the two sets. Within our problem domain we have refined the relation operator, I, to be a set of relations, each member corresponding 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.

There is an operator defined, $(\cdot)'$ which aids in the definition of formal concepts from the formal context.

$$\begin{aligned} (A)' &\equiv \{m \in M \mid (g,m) \in I, \forall g \in A\} \\ (B)' &\equiv \{g \in G \mid (g,m) \in I, \forall m \in B\} \end{aligned} \qquad \text{Eqn 2}$$

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 linkage to lattice theory provides avenues into a robust representation domain that can aid an analyst in developing an understanding of the collected information. The technologies use the "Begriff" of an identified context as the basis for the construction of that 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 3.

$$\begin{aligned} &(A,B) \xrightarrow{fc} (G,M,I) \\ &\Leftrightarrow \\ &A \subseteq G,\ B \subseteq M \\ &(A)' = B \,\&\, (B)' = A \end{aligned} \qquad \text{Eqn 3}$$

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

$$\begin{aligned} &(A_1, B_1) \le (A_2, B_2) \\ &\Leftrightarrow \\ &A_1 \subseteq A_2 \ \vee\ B_2 \subseteq B_1 \end{aligned} \qquad \text{Eqn 4}$$

An example of a lattice is provided from information developed by K. Wolff for his FCA tutorial. This example is a simple model capturing aspects of a knowledge base dealing with animals. The cross table of the information is the following.

Animals	Preying	Flying	Bird	mammal
Lion	x			x
Finch		x	x	
Eagle	x	x	x	
Hare				x
Ostrich			x	
Bee		x		

Table 1. Cross-table representation of an animal context.

The lattice representation of this information is shown in figure 2.

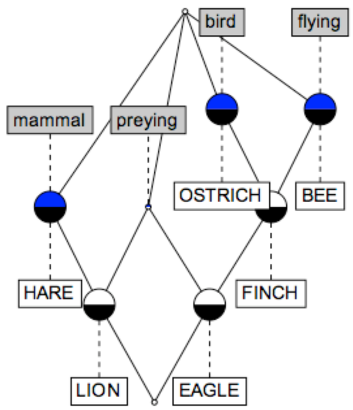

Figure 2. Lattice of animal domain.

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 or probabilistic variables. Formal concept analysis deals with attributes with continuous variables by defining a special construct called a "many valued context". This type of context is defined in the next expression.

$$\mathcal{K}_{mv} = (G, M, W, I) \qquad \text{Eqn 6}$$

As before, G is the set of objects, M is a set of attributes with values from the set W, defined by a ternary relational operator I. In this extension, the set of all values an attribute may assume is defined by the domain of that attribute.

$$D(m) \equiv g \in G \,|\, (g, m, w) \in I, w \in W \qquad \text{Eqn 7}$$

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 did not suite our needs. In order to handle real continuous valued attributes we employed fuzzy set theory with overlapping sets.

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 even spectral in character. In order to transform real world information into a form amenable to FCA we use a process based on fuzzy set theory. Within a context, basically a block of information, we assume that a real attributes posses common interpretation. Temperature in a materials context, might represent a melting temperature or a phase transition temperature. This temperature is not be associated with temperature defined to be an engine operating temperature. If all temperatures were lumped and fuzified 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.

$$\begin{aligned} &crisp \\ &\mu_A : X \rightarrow \{0,1\} \\ &fuzzy \\ &\mu_A : X \rightarrow [0,1] \end{aligned} \qquad \text{Eqn 8}$$

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 9.

$$\mu_k = e^{[-(x-c_k)^2/2\sigma^2]} \qquad \text{Eqn 9}$$

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.

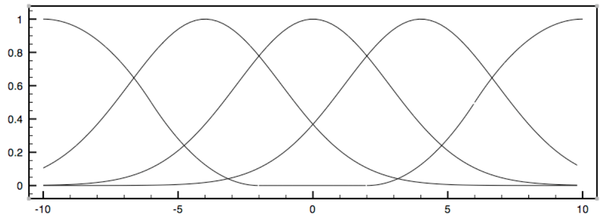

Figure 3. Fuzzification of a real variable over a range of -10 to 10.

The figure above shows a real 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 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, we all seem to be imbued with an 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 reasoning we employ. The ultimate form of reasoning is the method of scientific inquiry which was defined by C.S. Peirce.

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.

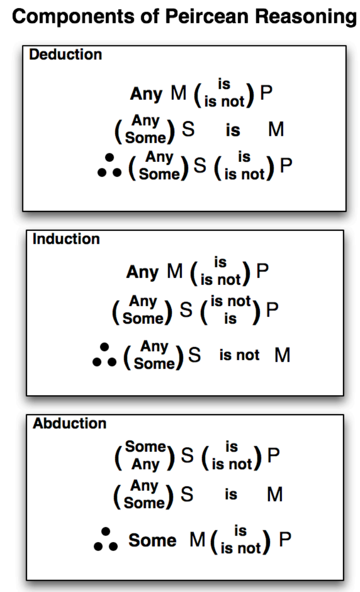

Figure 4. Formal representation of Peircean reasoning.

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. 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.

4. Application

The focus of the effort is to provide decision support capabilities and / or augment the efforts of an intel analyst. The architecture to support these functions is defined in the next figure.

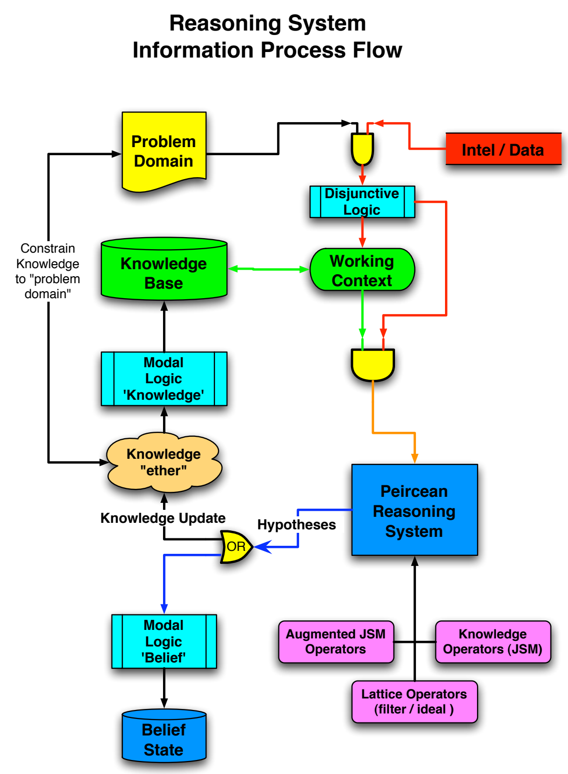

Figure 5. Decision support architecture.

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.

The first application has been in support of an analyst. One task of the system was to assimilate terror incidents documented in a State Department database over the years 1992 - 1998. The incidents were characterized by date, target,location, result, and the group responsible for the incident. This data was used by an inductive learning algorithm to provide a predictive capability for assessing which group might be involved in some future or past incident. The zero order temporal logic provided 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 a current understanding. The lattice representing a subset of the information is given in figure 6.

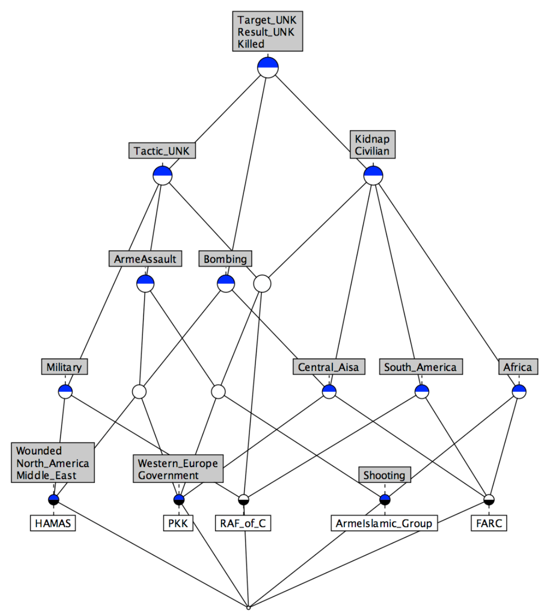

Figure 6. Lattice of terror incidents 1992-1998.

The structure of the information makes it practical to identify which groups might strike in North America, which groups prefer bombing tactics as opposed to kidnapping and what their main targets are. This knowledge becomes part of the basis for hypotheses generated by the system reflecting information collected.

5. Conclusion

What has been produced in this effort is a robust flexible decision support functionality. The system is a hybrid solution using a number of technologies in which function is matched to a technology strength. The knowledge representation methodology provides a mathematical basis permitting transformations to other representation technologies that support different technologies, ensuring an effective total solution. Currently , there are 3 one-to-one transformations that can be used to take advantage of matrix based technologies, or neural net technologies. This permits the development of fast real time engines as well as take advantage of technologies better suited for temporal analysis.

What we need to do is tailor the different modal logics to the needs and requirements of the application. Additionally, we need to expand the knowledge operator set to increase the flexibility of the hypothesis generation function. The system uses implementations of Mill's first two canons. The final extension would be to develop an analogical reasoning capability to give the system a more naturalistic ability to solve unique problems.

References.

(1) B. Ganter, G. Stumme, and R. Wille, "Applied Lattice Theory: Formal Concept Analysis," www.math.tu-dresden.de/~ganter/psfiles/concept.ps, 1997.

(1) B. Ganter, G. Stumme, and R. Wille, "Formal Concept Analysis: Theory and applications - J. UCS Special Issue," Journal of Universal Computer Science, vol. 10, pp. 926-926, 2004.

(2) E. Cox, The Fuzzy Systems Handbook. Cambridge, Ma: Academic Press, 1994.

(3) B. A. Davey and H. A. Priestley, Introduction to lattices and order, 2nd ed. Cambridge ; New York: Cambridge University Press, 2002.

(4) K.L. Ketner and H. Putnam. Reasoning and the Logic of Things: Charles Sanders Peirce. Cambridge, Ma: Harvard University Press, 1992.

(5) D,G, Mayo. Error and the Growth of Experimental Knowledge. Chicago, Il: University of Chicago Press, 1996.

(6) U. Priss, "Formal concept analysis in information science," Annual Review of Information Science and Technology, vol. 40, pp. 521-543, 2006.

(7) G. Stumme, "Formal concept analysis on its way from mathematics to computer science," Conceptual Structures: Integration and Interfaces, Proceedings, vol. 2393, pp. 2-19, 2002.

(8) R. Wille, "Concept Lattices and Conceptual Knowledge Systems," Computers & Mathematics with Applications, vol. 23, pp. 493-515, 1992.

(9) R. Wille, "Formal concept analysis as mathematical theory of concepts and concept hierarchies," Formal Concept Analysis: Formal Concept Analysis, vol. 3626, pp. 1-33, 2005.

(10) K.E. Wolff. "A First Course in Formal Concept Analysis" Adv in Statistical Software, 4, 429-438, 1993