SANDIA REPORT
SAND2001-3496
Unlimited Release
Printed November 2001

Foundations for Reasoning in Cognition-Based Computational Representations of Human Decision Making

Michael Senglaub, David Harris, and Elaine M. Raybourn

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of
Energy under Contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.doe.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/ordering.htm

SAND 2001-3496
Unlimited Release
Printed November 2001

Foundations for Reasoning in Cognition-Based Computational Representations of Human Decision Making

Michael Senglaub and David Harris
Network Systems Survivability & Assurance
Sandia National Laboratories
P.O. Box 5800, MS 0785
Albuquerque, NM USA 87185
mesengl@sandia.gov
dlharri@sandia.gov

Elaine M. Raybourn
Computational Initiatives
Sandia National Laboratories
P.O. Box 5800, MS 1188
Albuquerque, NM 87185-1188
emraybo@sandia.gov

Abstract

In exploring the question of how humans reason in ambiguous situations or in the absence of complete information, we stumbled onto a body of knowledge that addresses issues beyond the original scope of our effort. We have begun to understand the importance that philosophy, in particular the work of C. S. Peirce, plays in developing models of human cognition and of information theory in general. We have a foundation that can serve as a basis for further studies in cognition and decision making. Peircean philosophy provides a foundation for understanding human reasoning and capturing behavioral characteristics of decision makers due to cultural, physiological, and psychological effects. The present paper describes this philosophical approach to understanding the underpinnings of human reasoning. We present the work of C. S. Peirce, and define sets of fundamental reasoning behavior that would be captured in the mathematical constructs of these newer technologies and would be able to interact in an agent type framework. Further, we propose the adoption of a hybrid reasoning model based on his work for future computational representations or emulations of human cognition.

Acknowledgements

We thank Byron Dean, Ed Nozawa, Robert Burch, Stephen Littlejohn, and Kathy Domenici for their valuable insights and participation. This work was performed at Sandia National Laboratories. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed-Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Table of Contents

Figures

1. Introduction

Computational emulations of Human cognition and automate reasoning are technologies that seek to approximate the complex functions of human decision making. As these automated systems become more complex it is imperative that we, as developers and designers of these systems, fully understand the theoretical and philosophical underpinnings of human reasoning and incorporate this understanding into the computational representations of cognition or human emulations we develop. One approach toward instantiating human cognitive processes may involve the use of classic models and increased computational power in the divining of a solution. While a solution may exist using this approach, we believe the robustness of the solution and the applicability of that solution in divergent problem domains is likely to be minimized.

In the present Sandia National Laboratories (SNL) Laboratory Directed Research and Development (LDRD) project, we raised some fundamental questions concerning human reasoning and explored them within a context of issues including human reasoning in the absence of information and in the absence of expected cues, as well as the identification of new cues and the ability to connect information from diverse knowledge domains in order to generate hypotheses. Identification of a theoretical foundation for automated reasoning will enable us to develop automated reasoning agents and human emulators that can be used to populate complex high-consequence control systems.

Significant amounts of research have gone into describing human reasoning from a psychological and computational perspective; therefore we have prepared a bibliography of relevant papers and encourage the reader to use these references to familiarize her/himself with the concepts introduced in the present paper. In our literature search, we found that researchers from the human sciences and computer sciences seldom converge. For example, in the automated reasoning community there seems to be a bias toward a rule-based knowledge representation model. According to these models, human learning would reflect the processing of a specific knowledge construct at a time, while adding and refining rules. If human beings reasoned through the serial application and processing of rule sets, we would expect to see people locked into semi-comatose states while they processed all rules in their experience base prior to finding a solution. We do not see this, which implies we instead possess an ability to identify the information domain that is appropriate to the context of the situation. The secondary implication is that learning is not necessarily the sole process of rule creation and refinement but, as human learning theories would suggest-much more.

1.1 Objective

Many of the analytic technologies being developed seem to be technologies that add to the toolboxes of automated reasoning research. While these technologies are important, we still must identify the dynamic that leads to high-confidence solutions and that is unique to humans. A number of the analytic technologies that are being pursued to explore complex phenomena are based on some form of agent or Swarm type technology. The basic tenant of those technologies is that simple behavior of many interacting entities produce complex, unexpected behavior. The expected result is that we may be able to perform similar computations and analytical experiments, but in a reasoning domain. Modifications to classical Bayesian nets and fuzzy systems have been made that might lend themselves to this unique approach to reasoning. In particular, the developments of object-oriented Bayesian nets (OOBN) and stochastic programming, as well as sophisticated inference engines, are associated with these technologies. Assuming that human reasoning can be modeled as an interaction function of many sub-processing units, or clusters of neurons, this agent approach may be the correct environment for these reasoning studies.

In this report, we describe insights discovered through the literature search focusing on philosophical readings, and through workshop interactions with contemporary experts of the field of Peircean philosophy.

2. Background

2.1 Sources for Cognitive Models

The purpose of the research undertaken in this LDRD project was to better understand the fundamental underpinnings of human reasoning. To do so, we investigated the theoretical sources of cognitive models. After a significant amount of literature searches we found that in order to identify a foundation that crosscut human sciences and computer sciences, we had to look to philosophy for answers. Philosophers have looked inward and attempted to discover the mechanisms and functions that we refer to as reasoning or cognition. These studies started with Aristotle and continue today with the study of para-consistent logics. Para-consistent logic deals with information that is seemingly inconsistent.

There may be a growing trend in the United States for some brain-computer interface (BCI) developers to adopt B. F. Skinner's approach to the study of behavioral psychology, but we see a problem in this approach. A behavioral psychologist, Skinner, took an approach can that be crudely characterized as a stimulus-response dynamic in which the brain provides no function other than to identify a stimulus pattern and then react. While this is a component of cognition, it cannot explain the human aptitude for novel creation or scientific discovery. That is, while this approach may be useful to understanding cognitive functions under certain circumstances, there are huge gaps in our understanding of how human beings perform higher order cognitive functions. For example, if we were to apply the same research model to learning software behavior, we would expect to identify the functions and purposes of a large computer simulation by measuring the current into the computer and the changes in the adder stacks. While physiology (be it human brain physiology or a metaphor for computer hardware) is important for some problems, it alone does not provide the foundations needed to develop reasoning engines.

In our literature search, we identified a large body of information written by both computer scientists and by philosophers ruminating on the functions associated with human cognition. Each community offers a different perspective—the computer scientists favor an algorithmic perspective, and the philosophers prefer a more functional perspective. We do not address this literature here specifically, although we have provided our most relevant sources in the bibliography. In the present paper, we focus on the computer science research studies based on the work of C. S. Peirce, as they seem to explain a cognitive framework that appears to be extremely robust and is supported by mathematical tenets of logic.

This report describes a philosophical approach to understanding the underpinnings of human reasoning. We present the work of C. S. Peirce, and define sets of fundamental reasoning behavior that would be captured in the mathematical constructs of these newer technologies and could interact in an agent-type framework. Further, we propose the adoption of a reasoning model based on Peirce's work for future computational representations or emulations of human cognition.

2.2 The Promise of Peircean Thought

Charles Sanders Peirce was born in 1839 in Cambridge, Massachusetts. Harvard educated, Peirce ranks as one of our country's most prominent philosophers and the producer of the only major American philosophy: pragmaticism. During his career, Peirce was a scientist for the U. S. Coast Survey and a lecturer at Johns Hopkins. He was an engineering consultant and applied concepts of signs and logic to technical problems.

Primarily responsible for the advent of pragmaticism as a field of thought, Peirce aimed to create a human-centered philosophy and a fully developed scientific method. Pragmaticism concerns open, self-correcting systems in scientific and engineering terms. Based on logic and mathematical formulations and laboratory work, the theory is a laboratory philosophy, not a metaphysical or speculative philosophy.

Peirce is most well known for developing the field of semeiotics, or a study of signs based on the irreducible triad of three elements: object, sign, and interpretation. Meaning cannot be reduced to a simple dyadic relation between the sign and signified, as some semioticians have taught, but always includes a representation within a reasoning system.

As a normative and formal system, Peirce's theory focuses on a system of logic that can achieve the best possible conclusions based on available information. Complex adaptive systems use such logics not only to process information, but also to build a constantly evolving base of knowledge from which to hypothesize, predict, test, plan, and implement action. Because it aims to identify formal logical systems, Peirce's work opens the opportunity to articulate a variety of modes of reasoning that might be used by human beings or other complex systems.

The Peircean system does more than sense and respond. It constantly revises and expands a knowledge base by operating on information. The system plans responses based on a complex mix of operations such as comparison, prediction, and evaluation. Responses are not automatic, but consist of decisions informed by the knowledge base.

Heretofore, Western science has been dominated by nominalism, humanism, and logical positivism, generally neglecting the study of mental processes, form, and logic. Recognizing the importance of signs and representation, Peirce's ideas fill the gap too frequently ignored in the Western tradition.

The shift of awareness to information and logic has extremely valuable implications for system design and can have important national security implications. Scholars in other parts of the world, most notably the Russian Academy of Science, have been working with Peircean ideas for decades, and scientists and engineers in the U. S. are only beginning to see the importance of closing this knowledge gap. The potential effect of knowledge systems on the development of artificial intelligence would have fallout in many technical fields such as cognitive modeling, problem solving, robotics, knowledge representation, natural language, and learning.

2.2.1 A Peircean Model

Within Peirce's science of inquiry, the person (or other system) develops a knowledge base by hypothesis, deduction, and induction. Although deduction and induction were commonly known forms of inference for centuries, they cannot by themselves constitute a complete system of inquiry. Deduction is drawing logical consequences from premises, and induction is generalizing from a number of cases in which something is true and inferring that the same thing is true of a whole. Peirce therefore created abduction as a third class of reasoning. Abduction is the process by which a person develops a hypothesis, or reasonable conjecture, about what may be happening in a field of experience. More than a category, abduction is a set of logical operations that systems can use to create hypotheses.

The implication of this line of work in complex adaptive systems is tremendous. In order to reach conclusions and make decisions in a complex environment, a system must contain coded rules to allow it to create, test, refine, and optimize hypotheses through induction, abduction, and deduction, and thereby enlarge the knowledge base for prediction, control, planning, and future action.

The system first operates on the data by making an informed guess about the network of relations existing in the environment. This is a process of abduction. The system further reasons deductively that if the hypotheses are correct, then certain consequences will unfold.

Hypotheses are tested through observation and inductive reasoning, which helps the system further refine its hypotheses. The process of optimization, or developing conclusions quickly and efficiently within the constraints of available resources is important. It is not enough for the system to reason correctly, but it must do so in a timely, cost-efficient manner. For example, we may have a completely effective weapons system that is so slow or requires such complex resources as to be unrealistic in actual field conditions.

To summarize, a fundamental aspect of Peirce's belief is that knowledge is continuous and cumulative; that is, we build on previous foundations rather than throwing out "old knowledge." A seemingly counter example is the "flat worlds" model. Some would argue that this change in belief was based on the dismissal of a previous body of knowledge and the embracing of new knowledge. From a Peircean perspective, we add information to our model of a flat world until we encounter inconsistencies in the model at which time we seek modifications to the existing model to explain the new facts. This could be viewed as an example of Peirce's logic of abduction. Abduction assumes the role of idea generation while deduction evaluates the hypotheses and induction provides the mechanisms for justifying a hypothesis with data. A second component of Peircean cognition, or his science of inquiry, is deduction—founded on mathematical principles. This area has a significant foundation and is not discussed in detail in this report. The final component in his model involves induction. Peirce states that induction is the piece that is self-correcting; it provides the foundations for making changes in our accumulated knowledge base. If we were to represent this model of cognition functionally, it would resemble the model shown in Figure 1.

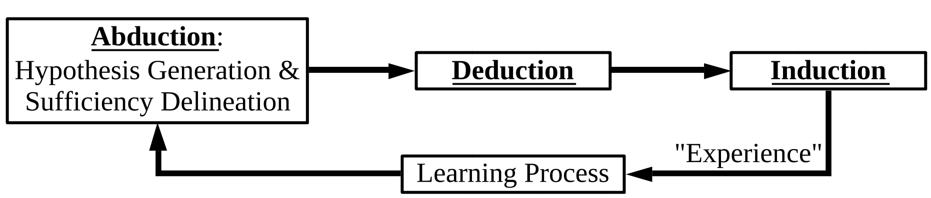

Figure 1. Peircean Model.

Reasoning individuals employ all three forms of logic in order to satisfy some inquiry. Abduction and deduction provide the foundations of a conceptual understanding, and induction provides the quantitative verification. Abduction creates, deduction explicates, and induction verifies.

2.3 Abduction Review

Abduction is the search for patterns in phenomena that suggest a hypothesis. (Peirce 1878)

The most formidable contribution to the study of the scientific method of inquiry or human cognition is Peirce's concept of abduction. Peirce in his later work explains abduction as a process for identifying hypotheses that can explain newly observed phenomena. It is similar to identifying disorders that might explain a set of medical symptoms, or identifying the best theory that might explain a set of experimental data. Abduction does not confirm a hypothesis; it is a process for arriving at plausible explanations of observable data. Abductive hypotheses must be plausible and likely. Plausibility is the condition that the hypothesis possesses the ability to explain the observed phenomena. Likelihood concerns the condition that the hypothesis has a good chance for explaining the data, i.e., the probability of this occurring is not low.

2.3.1 Models of Abduction

Abduction is perhaps the least well-known form of reasoning, and it still requires considerable development. All complex adaptive systems need the means for abduction, and artificial systems require coded algorithms to make this possible. John Stewart Mill provided a starting place for this work in his canons of causal reasoning. The Russian philosopher Finn has developed an intriguing logic based on what has come to be called the JSM (John Stewart Mill) method of direct agreement for small and large numbers of cases. Finn's work is strictly semeiotic, non-statistical, and qualitative.

The logics developed by Finn, based on Peirce and Mill, provide a set of relations that can be coded for the use of artificial systems. Faced with a complex array of potentially meaningful information, a system must go through a kind of inferential process, including abduction. In other words, it must speculate on what is happening in an efficient (optimal) manner so that, as an inquiring, reflexive system, it can act quickly to test, refine, and plan.

As reasoning agents, systems are far more complex than stimulus-response mechanisms. Previous work that concentrated on programming behavioral mechanisms to respond to environmental inputs failed because they did not allow for inquiry to occur. Moving from the behavioral-response model to that of reflexive systems opens avenues for complex perception, testing, planning, refinement, and adjustment in a system's response.

This is precisely the challenge facing engineers working on artificial, adaptive reasoning systems. Finn's models, employing JSM causal inference, provide the initial conceptual basis for this operational work. An advantage of this scheme is that it permits the use of a multi-valued logic for complex systems.

2.5 Analogical Reasoning Review

Peirce did not consider analogical reasoning to be a separate form of logical reasoning; he felt that it was a different weighting of the three basic forms defined earlier. He recognized the utility of analogical reasoning. A rather large amount of research can be found exploring the concept of analogical reasoning in computer and information sciences. Liang and Konsynski (1993) have worked on a tool called the Model Management System (MMS) which facilitates the use of mathematical models in deriving optimum solutions in various scenarios (see Figure 2). A key component of this technique is an analogical reasoning that aids in the construction of relationships between variables. Liang and Konsynski (1993) define analogy as “ a problem of the form A is to B as C is to D [A:B::C:D].” The last term D is the part that must be defined by some process based on the properties of A, B, and C. One example might be the attempt to model a Bohr atom and recognizing the similarity to planetary systems. One might conclude that the attractive forces in both situations obey an inverse square law.

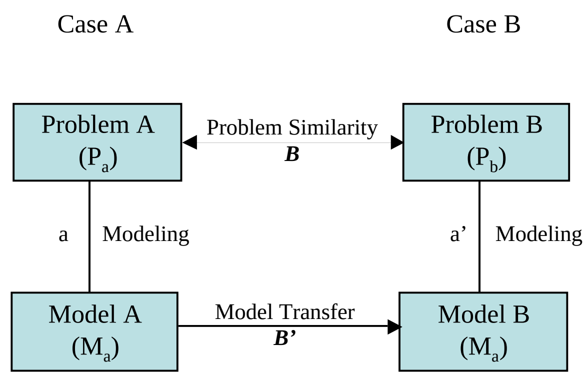

Figure 2. Model Management System (MMS).

2.6 Deduction Review

Deduction is drawing logical consequences from premises. The conclusion is true given the premises are also true (Peirce 1868).

The basis for deduction is found in mathematical reasoning. Deduction involves the link between premises and conclusions; if the premise is true, the conclusion must also be true. In Peirce’s terminology, deduction is a form of necessary reasoning. Necessary reasoning in this case does not concern itself with the truth of the premise; it concerns itself only with the relationship between the premise and the conclusion. The nature and structure of deduction has been studied since ancient times and has not been explored in any additional depth in this report. The work to be pursued in this area is the search for automated technologies that can perform the function of deduction in the broader context of a cognitive model. Examples of technologies that seem to perform deductive processes are technologies such as pattern recognition in targeting algorithms and aspects of Naturalistic Decision Making (NDM) in studies of cognition (Klein 1997). Greater detail can be found in mathematics texts on formal logic.

2.7 Induction Review

Induction is where we generalize from a number of cases of which something is true, and infer that the same thing is true of a whole class (Peirce 1878).

Induction is the reasoning function that draws rules from the results of sample cases. As such, it satisfies two functions in a cognitive model: it provides a validation mechanism for a hypothesis assertion, and it provides the foundations for learning or knowledge capture in the cognition process. Induction is a statistical construct that may not be fixed over an infinite time horizon, and is invalid when based on a single event. It can be likened to the problem of drawing a straight line through a single point; there is an infinite set of solutions. Another interesting characterization of induction is that it can be used to define empirical laws but in isolation cannot determine theoretical laws. Peirce, who has been characterized as a "statistical frequentist," believed that all empirical reasoning is a process of making inferences from samples about a population. Induction is justified under the "law of large numbers," we don't know the real probability due to our finite perspective, but because of the large number of cases, we are able to approximate the actual probability.

A large volume of work exists in the search for the ultimate technology to perform inductive learning. An excellent reference is the paper by M. A. Wani (2001) who compares a number of technologies that perform this inductive learning function. Basically, the effort involves the processing of large volumes of information in order to extract rules and/or concepts concerning the information. The objective is to then use the rules in decision algorithms. An interesting side study was that of Fogel, et. al. (1999) in which they used an induction model to explore the dynamic behavior of a number of agents in a postulated economic environment. Their study explored the impact of deterministic learning rules versus statistically based rules that might be an indicator of the effects of Lefebvre's work (discussed later).

2.7.1 Mayo's Theory of Induction

In researching aspects and functions of cognition, a number of references to D. Mayo's work (1996) showed up in conjunction with the work of Peirce. Her work centers on the concept of induction, but taken from a perspective of statistical concepts. Mayo, like Peirce, is a frequentist and does not support the ideas of Bayesian statistics. This position is somewhat reasonable when considering the perspectives of their work, which is an exploration of the foundations of scientific inquiry. The starting point for the comprehensive exploration of statistical induction follows the selection of a hypothesis that is believed to solve a specific problem. She does not explore the initial phases of identifying the likely hypothesis (abduction) or the final selection (deduction) of a hypothesis for test. Mayo breaks statistical or experimental inquiry into four steps, including primary modeling, experimental modeling, data modeling, and a fourth step that includes experimental design data generation and *ceteris paribus* conditions. Imposed on this construct is the concept of a "severe test" which constitutes conditions that are to be imposed on a statistical inquiry. Her objective is to define a construct of inquiry that permits a person to reason in the presence or under conditions of error.

In Mayo's efforts, primary modeling consists of a process of defining or identifying unique aspects of a hypothesis that can be statistically tested. This breakdown of a hypothesis is along the lines of questions that address standard errors. The types of errors she alludes to are those associated with chance effects vs. real effects: levels associated with a parameter; the inclusion of factors; and errors in experimental assumptions, such as the *ceteris paribus* conditions of the experimental setup. This tasking is in many ways outside the area of statistics and in the realm of the domain expert.

The second step or task involves experimental modeling. This modeling effort establishes a link between the primary model and the data. It identifies the number of trials, the kind of experiment that will be run, and any ancillary hypotheses that must be tested that would support H_0 if a hypothesis, H_0, is being tested. It does not involve a prediction of the expected outcomes of the experiment. Another aspect of experimental modeling involves the linkage between the experimental models and the data specifically. Hypotheses are not necessarily statistical in nature but can be framed statistically. The prediction can be framed statistically ($\mu \pm \sigma$) or considered statistically by the introduction of a test (t-test, etc.).

The third step involves data modeling. Data modeling is a transformation of raw data into information that correlates to the hypothesis being tested. An example described involved a validation of Einstein's theory that gravity could bend light. In the experiments conducted under an eclipse, star positions were measured. The raw data assembled involved detailed measurements of the star positions that were then transformed into deviations. The deviations directly related to the theory or hypothesis that was being studied. The final task of this inductive process simply involves identifying the number of replications to be executed and ensuring that other factors are the same under the series of experiments.

A unique aspect of Mayo's theory involves the idea of severe tests. She points out that for an experiment to be useful in the search for knowledge, the data generated must validate only one hypothesis. If the experimental results can validate multiple null hypotheses, the effort to gather this information is of little value. Mayo states the criteria as follows; "…a passing result is a severe test of hypothesis H just to the extent that it is very improbable for such a passing result to occur, were H false." A severe test distinguishes a hypothesis uniquely.

A final comment on Mayo's work involves her dislike (distrust) of the Bayesian approach to scientific inquiry. Her objectives are to identify unambiguous approaches to arrive at sound theories or hypotheses. A comment that seems to capture her position on Bayesian approaches occurred in her discussions of severity. She indicated that "…for the Bayesian, if two hypotheses entail evidence e, then in order for the two hypotheses to be differently confirmed there must be a difference in their prior probabilities." It is the subjective nature of Bayesian priors that is of concern. Her arguments seem to be that the biases of the scientist become innately integrated into the theory. It is her idea of test severity that appears to be the mechanism for ensuring the unbiased identification of the correct theory or hypothesis.

A question might be asked, "How does this relate to the problem of human reasoning in the absence of data, or to the process of scientific discovery?" A natural application of this theory of statistical induction involves the use of Taguchi techniques for system design or in performing sensitivity analyses. The process by which we perform variable and MOE analyses correlates with the experiment and data modeling tasks identified by Mayo. Additionally, the analysis of identifying the domain and the analytical objectives may be considered a form of primary analysis. We are theorizing settings or parameters that will lead to optimized performance of a system. If we consider the cognition process in total, this statistical induction methodology can be considered as either a validation mechanism for a postulated solution or as a theory that supports information gathered in a learning process.

3. Formal SNL Hybrid Solution

Significant effort in this LDRD project went into gaining a comprehensive understanding of the problem, the state of research, and on the possible solutions. We have come to the realization that our solution in the study of cognition lies in the domain of philosophy and formal logics, with neural science providing experimental or evidential clues to the mechanics of cognition. In this section, we present a formal functional model of cognition that identifies all components of reasoning as identified by Peirce. We also attempt to define and explore technologies that might be employed to instantiate the functions identified by the cognition model. Finally, we explore the impact of this reasoning construct on various activities that rely on some form of decision or cognition technology in subsequent sections.

3.1 Peircean Inspired Model of Reasoning

Through extensive literature review, we believe that the unique understanding of the philosophy of science as defined by C. S. Peirce provides the foundations needed to define a basic model of cognition. The model determined by Peirce is sufficiently robust to begin exploring extensions, which begin to address the problems we face in modern decision support systems. Our goal is to be able to capture and represent cognition in artificial systems that are needed in complex systems and systems control. The basic Peircean model is depicted in the Figure 3.

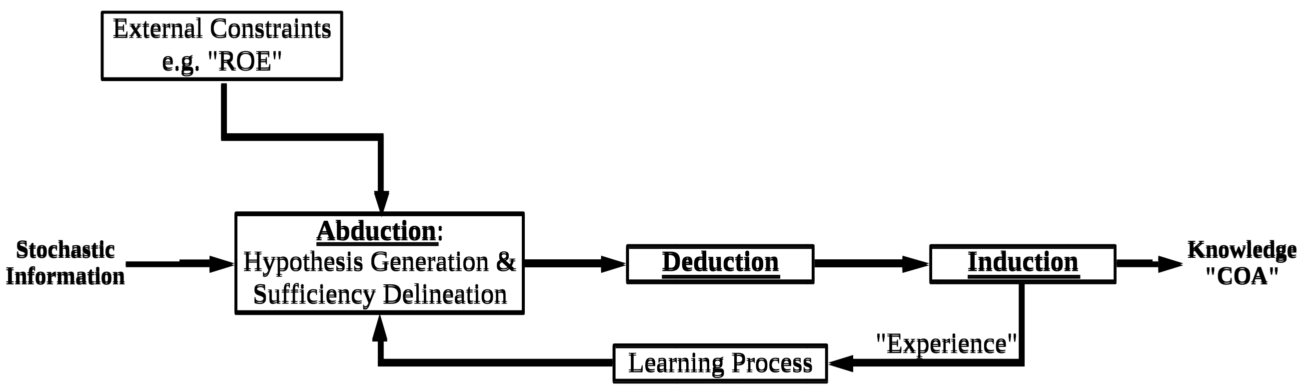

Figure 3. Basic Peircean Model with Nozawa's cognitive interpretation.

This model is a basic Peircean model of cognition with two minor modifications. The model captures the three principle components of cognition, abduction, deduction, and induction. The output of the model is knowledge or courses of action (COA). This is a result of the model becoming an integral part of a combat simulation system to augment command decision-making. As we have indicated, abduction provides the mechanism for generating potential theories or hypotheses to solve a problem. A refinement of the abduction process is to include a second-level cognitive analogical process. The analogical process involves the search for solutions through a process of analogy; e.g., early atomic models resembled planetary models; therefore, atomic forces may obey a similar inverse square law.

The deductive process is potentially the most understood part of the cognition model. Deduction is the process everyone schooled in mathematics understands. It is a process of demonstrating the connections between premises and conclusions. It can be considered the application of a specific rule to a single case. Significant levels of effort are expended in this area; e.g., pattern recognition is a form of deduction.

The final component of the model is the induction process. Induction draws a general conclusion or hypothesis from a sampling of cases. There are a number of interesting research efforts in this area, as well as very interesting work into the foundations of induction and its relationship to the scientific method. D. Mayo's work in this area is an excellent example. Unique to Peirce's model of cognition is the update or correction mechanism, which is represented by the lower feedback loop in Figure 3. Induction provides a mechanism for validating a hypothesis to a problem being considered, as well as providing the foundation for learning by the system.

The uniqueness of this model is the ability to naturally consider ill-defined problems. If we consider a model that is based solely on a deductive process, we are limiting the algorithms to regurgitate solutions that have already been used in situations defined by the parameters defining the problem. In some sense, it is the learned response of a ball player or a junior martial artist. A ball is thrown and the ball player automatically moves his glove to intercept the trajectory of the ball. A martial artist uses a middle block to prevent being struck by a blow to the ribs. It is the abductive process coupled to the deductive and inductive processes that provides a mechanism for solving problems that have not been pre-experienced. It is this functional model that solves the problem delineated in the research proposal. The problem is that the solution has opened the door to a large number of possibilities and technology questions that could ultimately impact information technologies and the decision sciences. We will try to address this in the final sections of the report.

The adaptations of Peirce's cognition model include an explicit mechanism that acts as a constraint on the abduction process and a mechanism that is a form of preliminary screening of potential solutions. The preliminary screening process could be represented as a high-level feedback loop that involves reasoning at a generic level. We know that drawing upon quantum concepts will not improve our billiard shooting ability even though aspects of quantum mechanics deal with elastic collisions. This aspect of the problem will be better understood if we have the ability to gain additional insights into Peirce's concept of "sign."

The second modification to the model is represented by the constraint mechanism feeding into the abduction process. The constraint mechanism might be to apply issues, such as rules of engagement (ROE). These are external constraints to the abduction process. Minimizing civilian collateral damage may rule out the use of certain types anti-personnel weapons. Potential solutions to a military problem have been effectively reduced by this constraint. It is the constraint mechanism that provides an avenue for the introducing qualitative factors into the decision-making process. By qualitative factors, we are referring to one's culture, to the psychology of an individual, and stress factors that may impact the decision process. Figure 4 includes a modification that could potentially address all these factors.

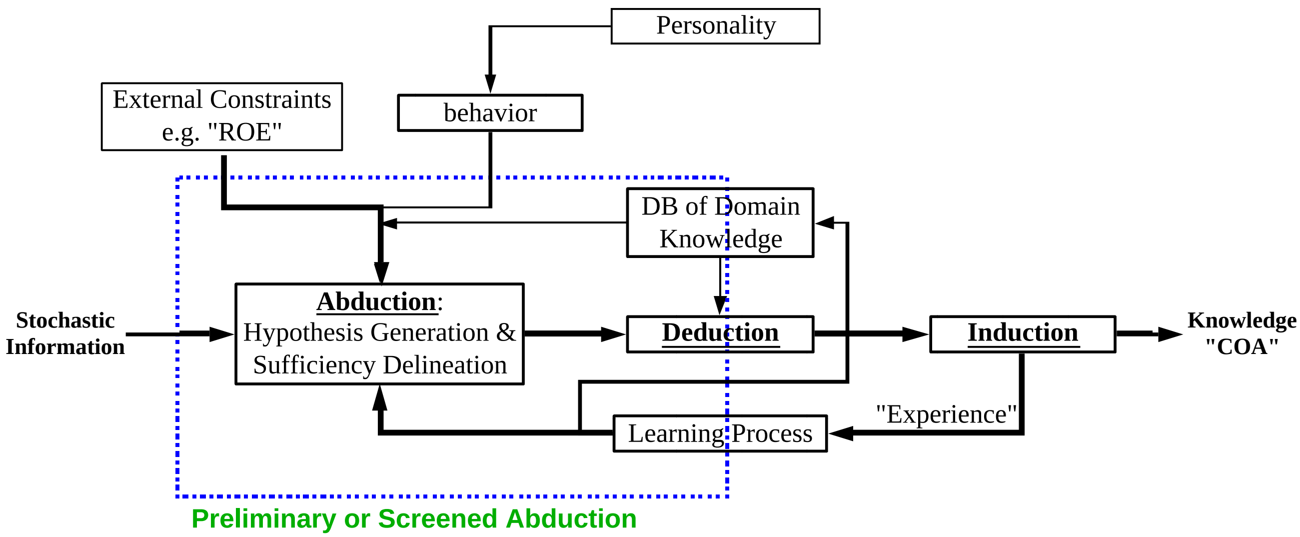

Figure 4. SNL Hybrid Solution.

While not appearing to be very different from the earlier figure, the natural interface of behaviors to the cognition process is obvious. We can view the process as a type of sliding window that filters potential solutions to a posed problem. In a purely logical process, a set of possibilities is determined and then used as the basis for the ultimate solution selection process. Environment, psychology, and culture affect where that window lies. In a military context, one culture may allow for the consideration civilian casualties while a second culture may not. A despot may consider the use of children as low-tech mine detectors while a democracy may not. The filtering of solutions changes depending on qualitative factors.

A second advantage to this construct is the potential for the inclusion of stress, fatigue, and other physiological factors into the decision process. Like the qualitative factors, we may consider these factors as a filtering of the abduction process. Stress and fatigue may result in a limiting of the potential solutions for consideration in solving a problem. In a pure decision support system, we do not need to consider these physiological or qualitative factors, since we are attempting to arrive at the best solution possible. When we are trying to capture the behavior of an adversary, we need to be able to represent the decision maker under various physiological and sociological conditions. In the technology section we will try to identify some promising theories that seem to complement the functional model identified in Figure 4 above.

3.1.1 Dynamics of Cognition

The dynamic aspects of cognition have not bee discussed at this point. Evidence supporting this consideration can be found in the neural sciences. Data suggests that the degree of neural activity increases with the complexity of the problem being solved. What is not answered by this observation is the process that is being executed. Another piece of information involves the human separation of long term and short-term memory. These pieces of data may provide a basis for exploring a Swarm approach to reasoning. Knowledge agents, representing domain expertise, interact with short-term agents constructed to represent the problem being prosecuted. These short-term agents self replicate when faced with a difficult problem so that they might interact with a greater number of domain-specific knowledge agents. Given a compact analytical representation of the cognition model, the next step would be to begin to explore the dynamics of reasoning and look for emergent properties that reflect experimental findings on the mechanics of human cognition.

3.2 Technologies Supporting Cognition Model

This section explores technologies that have been identified as being able to instantiate the functions defined in the functional model defined above. Figure 5 identifies potential analytic technologies that could provide a basis for representing the functions identified in the cognition model.

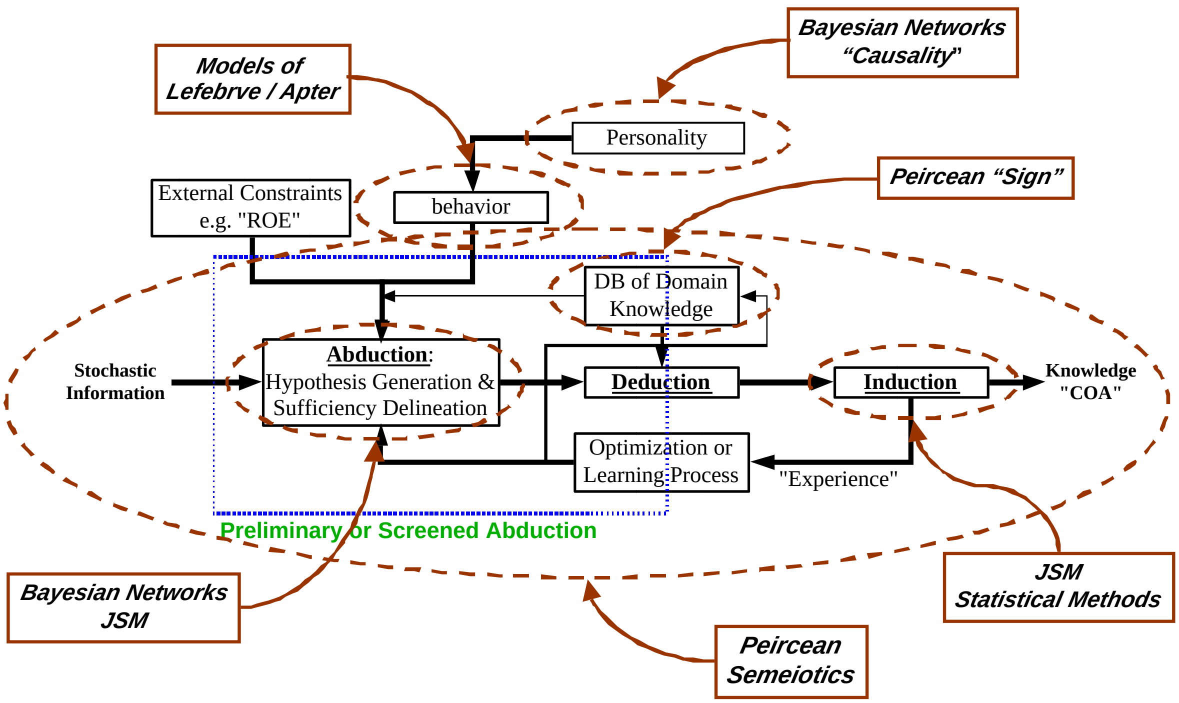

Figure 5. SNL Hybrid Solution, Continued.

A number of technologies could support the deduction process identified in this model including fuzzy logic, Bayesian nets, naturalistic decision-making, neural nets, and possibly other technologies. Figure 5 can be viewed as a research road map in which analytical experiments need to be performed to ascertain the suitability of the technologies as representations of the functional pieces of the cognition model. In the following sections, a subset of these technologies will be discussed because of experience with or application to the total concept of cognition delineated in the previous figures. It will not explore all the technologies delineated in Figure 5.

3.2.1 Reflexive Theory

Figure 5 indicates technologies that can be used to represent the behavior function in our cognition model. The technologies identified are those associated with Lefebvre (1998) and Apter. There are some similarities between these models, but Lefebvre's model seems to resonate because of its mathematical links to complex and chaotic dynamics. Both theories believe that there are "forces" or dimensions that affect our decision-making process. In Lefebvre's case these dimension are characterized as feelings, thoughts, and wishes. Apter has a somewhat larger set of dimensions. Both concepts believe that minor variations in our environment can force us to different ends of these axes, changing the resultant decisions we make. In the case of Apter, it appears to be a kind of binary set of axes, while Lefebvre uses a continuous axes. The interesting observation from both theories is the apparent bifurcation characteristic that emerges for the decision process. In the next couple of paragraphs, we will discuss in slightly greater detail the Lefebvre model, which appears to be slightly more developed.

The Lefebvre model was formulated by a very interesting and simple application of logic. In his paper, "Sketch of Reflexive Game Theory" (1998), he developed an algorithm that can be used to determine the dynamic of an individual confronted with a choice. The algorithm was based on a binary argument for clusters of neurons being in either a positive state or a negative state. The positive states are defined to be the socially acceptable solution, e.g., return money that a teller mistakenly gave to you. The argument goes that positive states are fixed, but negative states can change with time. Normalizing the number of states in a positive and negative state provides a structure for the probabilistic evolution of the neural states as they evolve to a positive state. The resultant differential equation was the logistic differential equation. This model is also what Lefebvre classifies as the realist model of human decision making. While this process shows the origins of the mathematical representations, the full model is defined by the equation below.

$$f(x_1, x_2, x_3) = x_1 + (1 - x_1)(1 - x_2)x_3$$
$$f_3(x, y) = 1 - y + x \cdot y$$
$$f(x_1, x_2, x_3) = f_3(x_1, f_3(x_2, x_3))$$

In this expression, the variables x_1, x_2, and x_3 represent feelings, thoughts, and wishes. The reflexive aspect of the theory comes in with the recursive nature of the model as seen in the last two equations. This characteristic can be defined as follows: We each have a world image that we use to make decisions, and that image may also include a representation of ourselves. In a two-player game, player 1 has a model that may include a representation of an adversary, player 2, who in turn has model which includes a representation of player 1, and so on. The depth of these models typically ends at three according to Lefebvre.

This model can be used to capture the behavior of a human component of a system, i.e., behavior characteristic of a pathological agent, of agents that are less than perfect, or of agents subjected to varying stress conditions. Conditions would determine the values of the three variables of the model that in turn determine the probability of which choice will be made in a decision process. An interesting aspect of this construct is the transition of cognition into the realm of nonlinear dynamics that we feel is a fundamental characteristic of cognition or any complex system. In our construct, we feel that this model would most likely impact the abduction process that is the preprocessor of potential solutions. The technology needs to be modeled and tested within a cognition setting. We think that a Bayesian net capturing the personality of the decision maker could be used to determine the values of these Lefebvre variables. In this case, by personality we mean the factors and causal relationships of these factors that would impact an individual's Lefebvre variables.

3.2.2 Bayesian Networks (Causality)

The technology has its roots in Bayesian probability theory (Pearl, 1988) in which the probability distributions are not known "a priori". The methodology is based on sets of prior distributions that may possess varying levels of knowledge and sophistication, and an update mechanism to improve these distributions as information is gathered through experiment, observation, analysis, or expert opinion. Bayesian network calculus begins with the idea of conditional probabilities. Conditional probabilities are simply stated as *given a state B, the probability of state A is x*. Mathematically this is written in Equation 1.

$$P(A \mid B) = x \tag{1}$$

The basic probability calculus rule is given in the next expression that leads to Bayes rule in Equation 3.

$$P(A \mid B)P(B) = P(A,B) \tag{2}$$

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)} \tag{3}$$

Bayesian networks (Jensen 1996; Pearl 1988) extend the fundamentals of Bayesian statistics to include a representation of information in formal directed graphs. The nodes in these graphs represent system variables with a finite number of mutually exclusive states. Therefore, a variable "A" with states a_i can be expresses as follows with associated constraints delineated in Equation 4.

$$\begin{aligned} &P(A) = (x_1, x_2, \cdots, x_n) \\ &x_i \geq 0 \\ &\sum_1^n x_i = 1.0 \end{aligned} \tag{4}$$

The arcs depicted in network diagrams represent the causal relationship between the system variables. Within the context of a network, a variable may have more than one parent. The resultant probability of a variable existing in a state is conditioned on the states of the parent nodes. Parent nodes are defined to be variable lying on the source side of the directed arcs in a diagram. The conditional probability for a variable "A" with parent variables, B_1,..., B_n, is represented in Equation 5.

$$P(A \mid B_1, B_2, \cdots, B_n) \tag{5}$$

The complexity of using this technology is defining the tables that correlate states and probabilities for the expressions represented by Equation 5. Our interest is in the joint probability distribution of the system. Given that conditional independencies hold for the network, the chain rule may be applied and the joint probability distributions may be defined as follows:

$$U = \{B_1, B_2, \cdots, B_m)$$
$$P(U) = \prod_i P(B_i \mid pa(B_i)) \tag{6}$$

pa(Bi) represents the parents of the variable Bi and U represents the set of variables comprising the Bayesian network. The advantage of the approach is the basic Bayesian nature of the problem in which information may be incorporated into the network as it becomes available. The distributions do not have to be known beforehand. Two papers by Spiegelhalter et. al. (Spiegelhalter, 1990; Spiegelhalter, 1993) provide a nice description of using data, sometimes sparse, to refine Bayesian network models.

Bayesian Modeling. Bayesian networks are the outgrowth of the failure of expert rule based systems to replicate the functions of domain experts. Expert systems are attempts to model the behavior of domain experts in computer algorithms. Early expert systems were rule based using classic *if-then* rules on discrete pieces of information. Rule based systems attempted to capture decisions in blocks of *if-then* rules and use an inference engine to determine actions given a set of evidence. Limitations of this technology involve the representation of uncertainty associated with the information, handling conflicting rule sets, and allowing for data refinement. Fuzzy logic mitigates problems with conflicting rules and some aspects of information uncertainty but does not provide effective data refinement methods.

Unlike expert systems, Bayesian networks are designed to model a domain. Execution of these models provides support function for the domain expert who must deal with complex issues and systems. Bayesian models are dynamic from a perspective that information may be appended to prior information allowing for continuous refinement of the information. Bayesian networks are not dynamic from a causality perspective. As new relationships between variables emerge, they will not be represented in the network. Bayesian networks have been used for medical diagnoses, for computer vision, meteorological prediction, and information processing.

Model Building (Variables, causality, data). One of the principle areas of effort associated with the construction of a Bayesian network involves the identification of the hypothesis variable(s) and the information variables. Hypothesis variables might represent a disease being diagnosed, or the state of a complex system. Information variables consist of the indirect observables that must be used to infer a hypothesis. Causality is the characteristic of one variable affecting a second variable. In a directed acyclic graph, the arrow connecting two nodes dictates the influence ordinality. The information delineating the degree of influence is defined in sets of conditional probability tables and represents information captured in expressions such as Equation 5.

The third type of variable is a mediating variable and is used for convenience. They are intended to ease the acquisition of conditional probabilities. These variables are used to collapse information into blocks were fidelity is unnecessary.

3.2.4 Echo with John Stuart Mill Enabled, Inductive Agents

The Echo model, described by Holland (1995), is an abstract, agent-based model used to simulate the behavior of complex adaptive systems and provides a cyber apparatus for investigations into these systems. The collection of agents within the model mimics Darwinian evolution via the use of biologically inspired operators of mutation and recombination with sexual reproduction. The individual agents compete to acquire, from a limited supply, the resources that are necessary for agent existence and reproduction. The agents within this model perceive their environment from the tags they expose and arbitrate their random encounters with other agents within the simulation by using internal control fields to reason about the tags of the agents that they meet. Agents that reason correctly about their environment are more able to acquire the resources necessary for reproduction and longer life since the outcomes of agent encounters are determined by their tags. Agents that are unable to cope with their environment will not survive, while agents better adapted to their environment will thrive and reproduce. Since the amount of environmental information that can be encoded into an agent's control field is severely limited, an individual agent is forced to make decisions with incomplete environmental knowledge. These Echo characteristics make it a useful experimental test bed for further investigations into the effectiveness of reasoning methods.

John Stuart Mill (1850) describes a system of logic (JSM) that includes methods for sound investigations into the determination of causal relationships between the observable features of a system and the outcomes or effects. His objective is to analyze the intellectual processes involved in reasoning or inference. Specific to experimental inquiry, he describes four methods of inference and five canons, the third canon being an improvement over and an extension of the first.

The first method, the method of agreement, begins to determine causation by investigating which of the observed properties is present in every instance a particular phenomenon is also observed and whose removal is not inconsistent with that phenomenon. This principle is captured in Mill's 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" (J.S. Mill 1850).

The second method, the method of disagreement, begins to determine which properties are present when the presence or absence of a phenomenon is observed. This principle is captured in Mill's 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" (J.S. Mill 1850).

The Echo test bed was used to evaluate an implementation of the first two JSM methods. A small set of Echo agents was randomly selected from the population. These agents were modified to include an ability to analyze a small sample of the agent collection and to select an internal control field determined by applying the JSM method of agreement and the method of difference. The agents inferred which tag elements to select by simulating the outcomes of encounters from this sample and determining which outcomes were favorable. Each of these simulations essentially represents an experiment that the agent performs. A particular Echo agent is successful if invariance in the patterns of tags of agents that it encounters match with its internal control field. Specifically, those tags or tag elements that describe agents that this particular agent can compete successfully against (the outcomes of the random encounters, which result in net resource gains) are encoded into the agent's internal control field. The JSM methods allow the agent to infer from its environmental sample which control field elements would cause more advantageous outcomes during its encounters. The measure of success in these agents is their ability to acquire resources and to reproduce more prolifically when compared to agents from a control set of unmodified agents. Results from initial Echo model executions indicate that agents that were enhanced with JSM inference methods consistently outperformed the unmodified control agents.

This page intentionally left blank.

4. Research Recommendations

In this section, we identify additional research and foundation research that should be pursued in order to acquire a complete characterization of robust reasoning algorithms. The areas cover a very broad range of skills and disciplines ranging from philosophy to non-linear dynamics and chaotic systems. In this section, we provide arguments for the inclusion of this area of research in a comprehensive cognition research program.

4.1 Formal Logics

Logic is dually housed in the academic worlds of philosophy and of mathematics. Its voice in either house is distinctly different from that of the other house. Yet the two share much that is the same. Consensus is that, although there is enough structure for application of logic to be meaningfully pursued, logic is still very much an open topic with even its basic foundations amenable to new structures, interpretations and applications. Logic is thought by many to be a foundation of mathematics itself, yet this "thought" can not be said to be demonstrated.

Formal logic has its foundations in the studies of the Greek philosopher Aristotle. His work was the first to explore the ideas associated with the theories of deduction. He also began the exploration of the idea that propositions can be arrived at through the structure or combinatorics of other propositions independent of their content. The ideas of formal logic reflect the fact that human reasoning in disparate branches of science can exhibit identical logical structures.

- A natural number is an integer, an integer is rational number, and therefore, a natural number is a rational number.
- An oak tree is a tree, a tree is a plant, and therefore an oak tree is a plant.

What can be seen from this simple example is a structure associated with the information that is identical in both cases. The validity of the propositions lies in the validity or truth of the structure and in the individual propositions comprising the premises. Formal logic provides a means of studying human reasoning since it goes beyond subject mater. Use of mathematical symbology provides a language that is unambiguous, unlike natural language. An example of this ambiguity is demonstrated in the next example.

- I have met Joe, Joe is an astronaut, and therefore I have met an astronaut.
- I have met someone, someone invented the radio, and therefore I have met the inventor of the radio.

While similar in structure like the proceeding example, the logic fails in this example because of the ambiguity associated with natural language.

A number of topics in the area of formal logics apply to the problems of cognition, including propositional logic, prepositional calculus, and predicate logic. Propositional logic involves the construction of formulas or functions dealing with propositions. It does not address the truth of the propositions themselves. While these basics rules can be defined using propositional functions that capture the rules of reasoning, how we manipulate and combine propositions or information to arrive at new information or propositions does not. Stolyar (1970) proceeds to develop and define a series of propositional algebras or formulas, what he calls tautologies that capture some law of logic. This law of logic represents a model of how we reason. A number of logics are being researched, including para-consistent logic and Peirce's tri-logic. Para-consistent logic is concerned with information that is seemingly inconsistent. For example:

- Penguins are birds, birds can fly, and penguins cannot fly.

If we used the logic structure of the first example, we would have come to the conclusion that penguins do fly. A fundamental uniqueness of Peirce's tri-logic is the idea of truth, falsehood, and indeterminacy. He has defined a logic constructed of propositional functions that reflect the tridicity of elemental propositions.

The other aspects of formal logic concern the assessments of the propositions themselves and predicate logic, and propositional calculus provides the language for manipulating formulas in order to create new formulas. Predicate logic decomposes propositions into subject and predicates and begins to explore the structure of propositions from this sub-level.

Formal logic provides a basis for validating information and knowledge bases. It provides an information or cognition scientist to unambiguously operate on information assembled for a particular problem domain. Formal logic also provides the basis for manipulating information in a consistent manner to generate new knowledge or solutions to problems present to the decision support system. There is speculation that it could also provide a basis for the development of a "language" for operating on information in a manner consistent with human reasoning models.

Logic had a major foundation placed for it by Boole who also opened the door to "technical applications" of logic as in the foundation of modern digital switching theory and computers. It was extended from propositional to predicate logic then to a predicate logic that culminated in the capstone version of most first order logic now used in applications. This was done by Peirce and Fregge who separately "invented" the existential and universal quantifiers, $\exists$ and $\forall$. Without these much of our more sophisticated logic-using applications could not be very well done at all.

The post-Boolean logic world can roughly be characterized by the three themes of: applications, theoretical results, and the search for a foundation for mathematics. The higher theoretical areas have been pursued most effectively, in modern times, by Göedle (spelling?) and Tarski. They showed us new truths about the limits of what can be decided and new structure in mathematical logic. Applications of note for logic can be said in the applied world to have included the Codd relational structures for databases. (Note that a relation is very nearly a predicate). Relational databases led to normal forms and efficient representations that are the foundation of most of what we loosely now call information systems. Codd's work is a logic application when looked at in the big picture of things.

In the late 50's, the application domain known as artificial intelligence was birthed by McCarthy and his invention of the Lisp programming language. This led to a 50-year effort to create a thinking software program, and apply it to real-world needs. When such software is applied it becomes an "inference engine" to reach conclusions from premises and "facts" within databases. Databases then became knowledge bases and knowledge engineering became a cousin to artificial intelligence in the arcana of pop science and futuristic hopes for better application software. However, the thinking machines and the robots have yet to be forthcoming.

The Lisp programming language itself actually followed upon an effort to create a new structure, the lambda calculus, that it was hoped would provide a foundation for mathematics and the logic of mathematics. Though this vision proved untenable within the lambda calculus, the calculus itself went on to be a major tool both for Lisp and for computer language studies and trial implementations.

After Lisp, the next big application step could be said to have been the invention of the Robinson resolution principle. It is a principle that greatly simplifies the search for proofs of statements made in first order predicate calculus. It is still being pursued today with success, just not the success that was hoped. Argonne National Labs is one of the leaders in Resolution Principle software for theorem provers today.

Similarly, although the Robinson Resolution principle proved to do less than was hoped, its use in applications have been notable. It can be said that this principle is the main enabler to the Prolog set of computer languages, which soon rivaled Lisp as the signature application language of artificial intelligence. We can now see these worlds as being a joining of symbolic computation and data bases, with efficient computer science and search algorithms, the latter having various and sundry more vibrant names for marketing use.

In all cases, a critical part of these advanced applications software constructs have always suffered under the threat of errors. Finding, fixing or "proving the absence of errors" and hence the correctness of software constructs became a new theme. It too has mostly born less fruit than was hoped. But proving the absence of error in databases is crucial to applications. Logic still is the hope of many for generating correct data base entries within applications programs.

All of these can be shown to have provided great insights and advances to technical applications and computer science in particular. Most of these, as good science will, have shown things that don't work and are not scientifically true. Many people are still trying to find these various holy grails. But there are few bright lights on the horizon of the future indicating that any grail might soon be found. So we can return to the past and pick up some of the dropped threads of past researchers.

We found at least three of these past threads to hold the possibility of contributing to the advance of these themes. Two of these are Peirce's Semeiotic, with its possibility for new representations/concepts such as triadicity, and his system of science. The third is Mill's methods of logic. These appear to be rich enough to be added to the suite of tools needed to make a better "inference engine" and better, more efficient, knowledge bases to contain the observations, facts and even rules of a master inference engine. But all-in-all reasoning under uncertainty remains yet an uncertainty. Including past threads into its pursuit is worth the effort to reduce this uncertainty, if not to totally eliminate it.

4.2 Semieotics, Inquiry, and Systems

Peirce's theory of signs is based on a triad of meaning that includes the object (O), representamen (R), and interpretant (I). Peirce wrote that sign is "anything (R) which is so determined by anything else called its object (O) and so determines an effect upon a person which I call its interpretant (I)."

The sign and signified cannot be separated from one another, nor can they be reduced to a simple dyadic relationship. Indeed, signs and objects are linked only by their mutual relationship with meaning in the person. The person (or other system) must form internal representations, or interpretations, in order for meaning to arise, an idea crucial to cognitive processing and mental models. The interpretant (I) is precisely this—the internal representation of the object made possible by the sign.

In human beings and, in theory, artificial systems, the internal representation is usually complex. It consists of an elaborate set of meanings derived from many signs linked to one another, as well as conclusions drawn through reasoning.

Metaphorically, the importance of the interpretant has been expressed as the "neglected 13 inches"—that space between the computer monitor and the human mind. Objects in the world are represented on the screen, but what happens between the screen and the person? Data per se is useless and must be transformed into meaningful information to become useful. Information, not data, is the basis of knowledge systems, prediction, and planning.

Sign theory is the most complex aspect of Peirce's theory. He identified three types of signs—(1) The icon, which resembles the object, (2) the index, which points to the object, and (3) the symbol, which arbitrarily stands for the object. There are perhaps 10 realistic subtypes.

Although sign theory is one of the most challenging aspects of Peirce's theory, it has important implications for deception detection. Deception and its detection are entirely a question of semeiotics. How do battlefield commanders deliberately mislead the enemy by attracting their attention to certain signs and manipulating how they might interpret these signs? Inversely, how do commanders read signs correctly and reason their way to valid conclusions in the battlefield situation?

Systems of representation can be complex, as signs are combined into larger signs within logical systems. This higher order of representation and interpretation gives rise to Peirce's development of logical systems of inquiry.

Within Peirce's science of inquiry, the person (or other system) develops a knowledge base by hypothesis, deduction, and induction. Although deduction and induction were commonly known forms of inference for centuries, they cannot by themselves constitute a complete system of inquiry. Peirce therefore created abduction as a third class of reasoning. Abduction is the process by which a person develops hypotheses, or reasonable conjecture, about what may be happening in a field of experience. More than a category, abduction is a set of logical operations that systems can use to create hypotheses.

The implication of this line of work in complex adaptive systems is tremendous. In order to do reach conclusions and make decisions in a complex environment, a system must contain coded rules to allow it to create, test, refine, and optimizing hypotheses through induction, abduction, and deduction, and thereby enlarge the knowledge base for prediction, control, planning, and future action.

The system first operates on the data by making an informed guess about the network of relations existing in the environment. This is a process of abduction. The system further reasons deductively that if the hypotheses are correct, then certain consequences will follow.

Hypotheses are tested through observation and inductive reasoning, which helps the system further refine its hypotheses. Important is the process of optimization, or developing conclusions quickly and efficiently within the constraints of available resources. It is not enough for the system to reason correctly, but it must do so in a timely, cost-efficient manner. We may have a completely effective weapons system that is so slow or requires such complex resources as to be unrealistic in actual field conditions.

4.3 Semeiotics ("Sign")

Knowledge representation, capture, and storage. Semeiotics is the study of truth, not what is true; rather it defines the conditions for truth. Within semeiotics or in another form, semeiotics is concerned with the concept of sign. Semeiotics addresses three generic sub-areas of study: grammar, logic, and rhetoric. Grammar is the study of the formal features of the sign; logic is concerned with the manner in which signs discern truth; and rhetoric is the study of how signs can be used to express ideas.

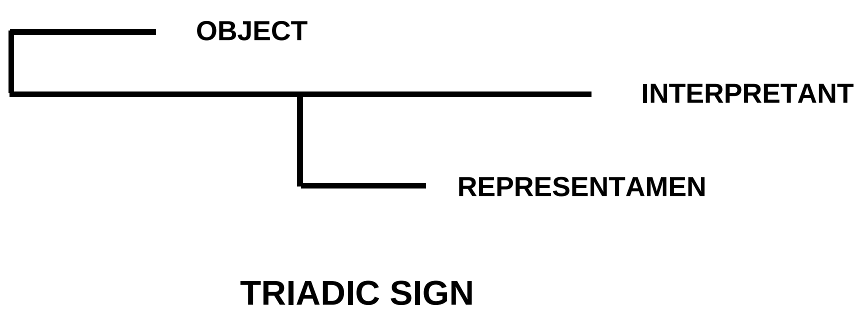

Figure 6. Peirce's Concept of Triadic Sign.

To Peirce a "…sign or representamen, is something which stands to somebody (the interpretant) for something (the object) in some respect of capacity." It is felt that using Peirce's ideas as a foundation, we may discover alternative information capture and representation technologies. The principle difference between this logic and other forms of formal logic is the inclusion of the interpretant into the model. A representation for an object might be some form of icon (one of three basics forms of Peircean representations) that must create a response in the person being communicated with. Considering the receiver in the equation involves recognition that an icon familiar to a physicist may not mean the same thing to a sheepherder.

4.4 Dynamics of Knowledge Agent Interactions

A final research potential involves the exploration of knowledge agents in a Swarm type environment. We speculate that knowledge in humans is likely clustered within groups of neurons. The question becomes, How are solutions generated via an abduction process in which groupings of these clusters are triggered? It would be interesting to explore the interactions of knowledge agents, each with specific expertise interacting with short-term agents tasked with finding solutions to a newly posed problem. The idea is to use newly created agents to represent the knowledge sought. The seekers are generated at a rate or possibly as a result of some threshold criteria that are functions of the complexity of the problem being solved. This might correspond to the phenomena of "spreading activation" observed in humans tasked with solving very difficult problems. The dynamics of this interactive behavior might provide interesting explorations into cognition and possible independent study related to Lefebvre's efforts into the dynamics of human cognition.

5. Conclusion

5.1 Areas of Potential Application

The following section describes environments that could benefit from cognition-based decision support systems enhanced with the hybrid solution based on the work of C. S. Peirce. The common theme in these topic areas is the gathering and processing of information. Additionally, data fusion has been listed separately but is a foundation technology that applies to most, if not all, areas of decision support.

5.1.1 Data Fusion

Data fusion is a natural process in the human system in which data from a suite of sensors is combined to provide information about the current system state and environment. This suite of sensors includes the eyes, ears, nose and fingers. The variation in characteristics, spectrum, and spatial scope make the data fusion process a very complex analytic task. The process of data fusion is a natural extension or application of human cognition. Data fusion is the desire to derive more information about a situation through combining than could be gleaned from the data separately. A number of approaches have been applied to the problem of data fusion, including decision theory, estimation theory, association, and uncertainty management. Other approaches exist, but these appear to be the most significant approaches.

Decision or detection theory is based on the ideas of Bayes. In this approach, measurements are compared with alternative hypotheses in an effort to determine the best match. The process is effective because of the ability to update estimates as more information is gathered on the target. This technology is believed to play a significant role in the functionalization of the cognition model described earlier. Estimation theory is an extension of decision theory and finds its roots in statistical estimation methodologies. Basically, multiple measurements of a variable or parameter are executed, and estimates of the parameter are made based on these measurements.

Association is a method for dealing with multi-source information in a multi-sensor environment. In this situation, a resolution process must be undertaken in order to associate signal with source prior to making any kind of assessment. The uncertainty-management approach relies on Bayesian technologies as well as on fuzzy methodologies. In this case, information is gathered in an effort to reduce uncertainty in a decision, it is intended to operate in environments in which complete information is difficult or impossible to obtain.

This is an extremely short discourse into data fusion but it shows some of the problems being tackled by engineers attempting to develop data fusion algorithms. What we can see is that pieces of our cognition model can be used to perform the functions or techniques delineated in these constructs. The complete cognition model should be able to perform against all the problems characterized by data fusion objectives. Our initial objective was to develop a reasoning algorithm that can work in an absence of complete data-one of the more difficult data fusion problems.

5.1.2 Anti-Terrorism

Recent events have demonstrated a gap in our ability to assimilate information and predict the intentions of an adversary that does not reason on models we may be familiar with. We have seen a new dimension in unconventional tactics against targets that we have failed to recognize as targets. While efforts have been made to develop predictive models, we have not gone far enough to include a data fusion component with hypothesis generation that can augment that of our intelligence communities. The mass of information and misinformation makes prediction a very difficult process. The construct that is being proposed by our cognitive model may provide a foundation of technology that can begin to bridge that gap. The first component is a data fusion construct that can begin to use social dynamic networks to define one level of event linkage. The linkages can then begin to identify disparate events that are correlated. Once we begin to establish these correlations, the abduction process can begin to assemble hypotheses that might fit these correlated events. We would believe that the human should act as the deductive component of the intelligence process in which the deductive selection of a probable hypothesis is determined by the human in the loop. The inductive hypothesis test needs to be performed in a collaborative manner: man and machine. The need for a machine abduction mechanism becomes important because of the lack of cultural biases that a machine may possess. Humans have a continuous tendency to imbue an adversary with behavior and constraints that mirror their own. Letting the machine generate possibilities eliminates, or at least mitigates, that effect.

5.1.3 Cognition-Based Decision Making

The short discussion of data fusion provides a grounding for the broader context of cognition based decision making. Many of the problems we are facing are becoming far more complex and time critical than we as humans can deal with. The ability of machine aids that can perform some of the data fusion and information assembly could greatly enhance the effectiveness of the human decision maker. This is becoming more important as the volume of information available continues to increase. The effort required to sift through the volumes of information in order to ferret out the relevant pieces of information may render moot any decision made. The construct of our cognition model should be able to provide the construct for decision aids that can reduce and improve the timing of human decisions.

5.1.4 Autonomous System Control

Building on the last two topics, we see the potential for the development of very sophisticated controllers for use on autonomous systems. It is impossible for designers to anticipate all possible events and combinations of events that may be faced by systems in hazardous, hostile, or complex environments. Providing the system with a representation of human cognition and a knowledge base from which decisions may be developed will greatly expand the applicability of autonomous systems as solutions to complex problems. We also need to recognize that as systems continue to grow in complexity, we will reach a point where we will not only be unable to understand the behavior but will lack the ability to control that system with classical methodologies; human cognitive control will provide the means for establishing reliable control of these systems of systems.

5.1.5 Remote Medicine (Deep Space Medical Care)

A unique role for human cognitive decision making involves a situation in which the temporal considerations become important. In manned deep space missions, the communication times may become too long for effective treatment under emergency conditions. Development of a medical decision aid can greatly improve the survivability of explorers in these situations. The medical decision aid can begin diagnostics and certain triage type of activities while communication is established with medical teams at mission control. These decision aids need to be able to evaluate and make recommendations or perform tests without the direct interaction or intervention of human medics. Classical diagnostic models will be insufficient for use in these extreme conditions and with the temporal lags in communication with medics.

5.2 Summary

In the present paper, we described a philosophical approach that facilitates the understanding of the underpinnings of human reasoning. We presented the work of C. S. Peirce, and defined sets of fundamental reasoning behavior captured in the mathematical constructs of newer technologies that are also able to interact in an agent type framework. Further, we proposed the adoption of a reasoning model based on the work of C. S. Peirce for future computational representations or emulations of human cognition.

This page intentionally left blank.

6. Bibliography

Cognition (COG)

A. Garnham, 1993, Is Logicist Cognitive Science Possible? in *Mind and Language,* vol. 8(1), pp. 49-71.

G. Klein, 1997, An Overview of Naturalistic Decision Making Applications, in C. E. Zsambok and G. Klein *Naturalistic Decision Making*, Mowah, NJ: Lawrence Earlbaum, pp. 49-59.

B. N. Kokinov, 1994, A Hybrid Model of reasoning by Analogy, in J. A. Barnden (Series Ed.) and K. J. Holyoak and J. A. Barnden, (Vol. Eds.) *Advances in Connectionist and Neural Computation Theory: Vol. 2. Analogical Connections* (pp. 247-318). Norwood, New Jersey: Ablex Publishing.

J. J. Liszka, 1996, *A General Introduction to the Semeiotic of Charles Sanders Peirce*. Indiana University Press.

E.T. Nozawa, *Peircean Semeiotic A New Engineering Paradigm for Automatic and Adaptive Intelligent Systems Design*, LM Aero, Marietta Ga. 30063. www.d.kth.se/~tessy/Nozawa.pdf

E.T. Nozawa, 2000; *Reflexive Control Theory and the Scientific Philosophy of Charles Sanders Peirce*; LM Aero, Marietta Ga. International Reflexive Control Symposium 2000.

C. S. Peirce, 1868, Questions Concerning Certain Faculties Claimed for Man, in *Journal of Speculative Philosophy,* vol. 2, pp. 103-114.

C. S. Peirce, 1868, Some Consequences of Four Incapacities, in *Journal of Speculative Philosophy,* vol. 2, pp. 140-157.

C. S. Peirce, 1878, How to Make Our Ideas Clear, in *Popular Science Monthly,* vol. 12, pp. 286-302.

T. Steinert-Threlkeld, 2001, February 26, Moody Computers, in *ZDNet Interactive Week.* Retrieved from http://www.interactiveweek.com, Interactive button.

E. Stroulia, 1994, Reflective Reasoning and Learning [Abstract], in *Applied Artificial Intelligence,* vol. 2, p. 1493.

E. Stroulia and A. K. Goel, 1994, Learning Problem-Solving Concepts by Reflecting on Problem Solving, in F. Bergadano and L. D. Raedt (Eds.), *European Conference on Machine Learning, Vol. 784. Machine Learning ECML-94* (pp 287-306). Wein-New York: Springer.

E. Stroulia and A. K. Goel, 1995, Functional Representation and Reasoning for Reflective Systems, in *Applied Artificial Intelligence,* vol. 9, pp. 101-124.

C. H. Yu, 1994, April, *Abduction? Deduction? Induction? Is There a Logic of Exploratory Data Analysis?* Paper presented at the Annual Meeting of American Educational Research Association, New Orleans, Louisiana.

E. N. Zalta, (Ed.), 2001, *Stanford Encyclopedia of Philosophy*. Retrieved October 1, 2001, from http://plato.stanford.edu/entries/peirce-logic/

Analogics (AN)

B. N. Kokinov, 1994, A Hybrid Model of Reasoning by Analogy, in J. A. Barnden (Series Ed.) and K. J. Holyoak and J. A. Barnden, (Vol. Eds.) *Advances in Connectionist and Neural Computation Theory: Vol. 2. Analogical Connections* (pp. 247-318). Norwood, New Jersey: Ablex Publishing.

J. G. Carbonell and S. Minton, 1983, *Metaphor and Common-Sense Reasoning* (CMU-CS-83-110). Pittsburgh, Pennsylvania: Carnegie-Mellon University.

C. Dierbach and D. L. Chester, 1997, Abstractional Concept Mapping: A Foundational Model for Analogical Reasoning, in *Computational Intelligence*, vol. 13(1), pp. 32-86.

W. T. Hunt and M. McLeish, 1991, Semantics, in Analogical Reasoning, in *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 3, pp. 145-160.

M. T. Keane, T. Ledgeway, and S. Duff, 1994, Constraints on Analogical Mapping: A Comparison of Three Models, in *Cognitive Science*, vol. 18, pp. 387-438.

T. Liang and B. R. Konsynski, 1993, Modeling by Analogy: Use of Analogical Reasoning in Model Management Systems, in *Decision Support Systems*, vol. 9, pp. 113-125.

Abduction (AB)

V. Ajjanagadde, 1993, Incorporating Background Knowledge and structured Explanada in Abductive Reasoning: A Framework, in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23(3), pp. 650-664.

N. L. Zhang and W. Liu, 1997, A Model Approximation Scheme for Planning in Partially Observable Stochastic Domains, in *Journal of Artificial Intelligence Research*, vol. 7, pp. 199-230.

Deduction (DED)

See Peirce's publications.

Induction (IN)

D. B. Fogel, K. Chellapilla, and P. J. Angeline, 1999, Inductive Reasoning and Bounded Rationality Reconsidered, in *IEEE Transactions on Evolutionary Computation*, vol. 3(2), pp. 142-146.

D.G. Mayo, 1996, *Error and the Growth of Experimental Knowledge*. The University of Chicago Press.

M. A. Wani, 2001, SAFARI: A Structured Approach for Automatic Rule Induction, in *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics,* vol. 31(4), pp. 650-657.

Formal Logic (FL)

A. Garnham, 1993, Is Logicist Cognitive Science Possible? in *Mind and Language,* vol. 8(1), pp. 49-71.

A. N. Kaplan and L. K. Schubert, 2000, A computational Model of Belief, in *Artificial Intelligence,* vol. 120, pp. 119-160.

J. S. Mill, 1850, *System of Logic, Ratiocinative and Inductive; Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation.* New York: Harper and Brothers.

A. A. Stolyar, 1970, *Introduction to Elementary Mathematical Logic*, Dover Publications, New York.

Technologies

J. H. Holland, 1995, *Hidden Order: How Adaptation Builds Complexity*. Addison Wesley, ISBN 0-201-44230-2

V. A. Lefebvre, *1998 Sketch of Reflexive Game Theory.* School of Social Sciences, U. of California, Irvine.

L. D. Miller and M. F. Sulcoski, Reflexive Models of Human Behavior: Variations on a Lefebvre Theme, in *Proceedings of the Workshop on Multi-Reflexive Models of Agent Behavior*.

L. D. Miller and M. F. Sulcoski, Application of Generalized Reflexive Behavior Models for Situation-Specific Decisions, in *Proceedings of the Workshop on Multi-Reflexive Models of Agent Behavior*.

J. Pearl, 1988, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Pub. Inc., San Francisco, Ca, ISBN 1-55860-479-0.

F. V. Jensen, 1996, *An Introduction to Bayesian Networks*, Springer-Verlag, NY, ISBN 0-387-91502-8.

J. Pearl, 2000, *Causality*, Cambridge University Press, ISBN 0-521-77362-8.

D. J. Spiegelhalter, A. P. David, S. L. Lauritzen, and R. G. Cowell, 1993, Bayesian Analysis in Expert Systems, in *Statistical Sciences,* vol. 8, No 3, pp. 219-283.

D. J. Spiegelhalter, and S. L. Lauritzen, 1990, Sequential Updating of Conditional Probabilities on Directed Graphical Structures, in *Networks*, vol. 20, pp. 579-605.

This page intentionally left blank.

Appendix. Summary of Peircean Workshop

A.1 The Seminar

A team of researchers at Sandia National Laboratories, consisting of Byron Dean, Dave Harris, Elaine Raybourn, and Michael Senglaub, explored the development of automated reasoning systems that approximate the complex functions of human decision making. To further their investigations, the group invited two internationally known specialists in Peircean Semeiotic to present a seminar on this subject and explore ways in which these concepts might be useful in advancing the automated reasoning project.

The seminar, held May 24 & 25, 2001, was presented by Dr. Robert Burch of Texas A&M University, and Ed Nozawa of Lockheed Martin. Communication consultants Kathy Domenici and Stephen Littlejohn facilitated the meetings.

The session was wide ranging and explored such topics as information processing, signs, logic, system control, philosophy of science, and national security. Moving from general ideas to very specific applications, the group talked in concrete terms about how to apply Peircean ideas, recent extensions, and applications in complex adaptive systems.

Some of the many ideas discussed are summarized in this report.

A.1.1 Discussion Themes

Some of the lines of discussion that emerged during the seminar are outlined below:

- Interacting agents pose an especially interesting and important problem for investigation. Each system must hypothesize about and act on the other. This means that each must have a view of self, other, and other's view of self. This interactive process is especially important in an adversarial situation, as when an air vehicle is effective in attacking the enemy.

- The interactivity problem does not minimize the importance of within-system problem-solving and reasoning. A fascinating question is how reasoning agents generate new forms of reasoning internally, as when the airplane itself must function effectively internally.

- A fruitful line of work would be to combine the interacting-agent problem with the within-system reasoning problem by viewing these as nested systems. If we view the interacting system as open and the within-system as closed, we can look at ways in which the closed system is nested within the open one. For example, in a battlefield situation, the system inevitably interacts with the adversary, but in order to do this, it must contain internal systems that function effectively within closed loops. In order to attack effectively, the airplane must have effective internal systems. This approach could be important to the artificial-reasoning team because it could allow them to integrate their separate interests and projects.

- Can an information base be "intelligent" in the sense that the information itself constrains the rules by which it can be processed? For example, in security systems, information may protect itself by not allowing an intruder to view it on an undesignated device.

- The adequacy of the internal algorithms of the system can be tested in a variety of ways. One criterion is success. We can test the adequacy of the inferential processes by running them over time and calculating the ratio of success to total runs. In the same way, the power of an inference rule can be tested by allowing it to run over time. Calculate the ratio of correct answers to total answers to get a measure of effectiveness. However, other criteria such as speed and cost may also be important.

A.2 Outcomes

The team was excited and encouraged by potential applications of seminar material to their project and made several decisions:

1. Peirce's theory of signs has much potential for offering a set of application models. The team will continue to learn more about the theory and its extensions through reading and additional consultation.
2. The JSM (John Stewart Mill) method of abduction, summarized in this report, has a great deal of potential for modeling causal hypothesis formation. The group will pursue the creation of a suite of algorithms for an actual computer application.
3. The team will explore the possibility of using JSM in a demonstration project to test ways in which reasoning processes might be coded and employed in a workable system. This will involve finding a manageable domain, building an inference engine, running it, refining it, and exploring ways in which it might modify itself. Such a prototype should show the feasibility of causal hypothesis formation in a larger system-control process. The project must be workable, affordable, relevant to the team's task, effective for showing potential, and useful for seeking additional funding. Various possibilities were discussed, including cyber-security, terrorism warnings and indicators, deception detection, and weapons-system and battlefield applications. SNL security itself may be an application worthy of consideration, especially to gain general Laboratory support.
4. Team members will explore ways to integrate their respective projects, including interacting systems and Echo.
5. Contracts with Dr. Burch and Mr. Nozawa will be pursued for further consultation on the team's project in general and a demonstration project in particular. This collaboration could also provide R&D support for Lockheed Martin's advanced mission management systems, C4ISR systems, and Information Operations research.
6. Funding avenues for a demonstration project will be explored.

Distribution

1 0129 Nigel Hey, 12600
1 0139 John Kelly, 9904
1 0157 Dorothy Stermer, 1322
1 0165 Chris Burroughs, 12640
1 0318 George Davidson, 9212
1 0449 Robert Hutchison, 6516
5 0785 David L. Harris, 6516
1 0451 Ron Trellue, 6501
1 0451 Tommy Woodall, 6502
1 0455 Reynold Tamashiro 6517
10 0785 Michael Senglaub, 6516
1 0747Allen L. Camp, 6410
1 0748 Robert Waters, 6413
1 0759 Mark Snell, 5845
1 0780 Sabina Jordan, 5838
1 0830 Caren Wenner, 12335
1 0830 Bobby Baca, 12335
1 0839 Gerry Yonas, 16000
5 0839 Advanced Concepts, 16000
1 1004 Ray Harrigan, 15221
1 1010 Dan Small, 15222
1 1109 Rich Pryor, 9212
1 1110 William Hart, 9211
1 1125 Keith Miller, 15252
1 1137 John Ganter, 6535
1 1137 John Mitchiner, 6534
1 1140 Larry Ellis, 6502
1 1153 Stewart Cameron, 15336
1 1170 Russ Skocypec, 15310
1 1176 Robert Cranwell, 15312
1 1188 John Wagner, 15311
10 1188 Elaine M. Raybourn, 15311
1 1188 Eric Parker, 15311
1 1188 Chris Forsythe, 15311
1 1188 Steve Tucker, 15311
1 1165 William Guyton, 15300
1 1204 Craig Dean, 5902
1 9003 Ken Washington, 8900
1 9011 Nina Berry, 8920
1 9012 Paul Nielan, 8920
1 9012 Carmen Pancerella, 8920
1 9019 Brian Maxwell, 8945
1 9012 Ernest Friedman-Hill, 8920

1 9037 Jim Berry, 8935
1 9201 Howard Hirano, 16000
1 9018 Central Technical Files, 8945-1
2 0899 Technical Library, 9616
1 0612 Review & Approval Desk, 9612
for DOE/OSTI
1 0188 Donna Chavez, 1030, LDRD
Office

1 Gary Klein
Klein & Associates, Inc.
1750 Commerce Center Blvd. North
Fairborn, Ohio 45324-3987

1 Army Research Lab
AMSRL-SL-EA
ATTN: Tom Reader
White Sands Missile Range,
NM 88002

1 Dr. Raymond Bernstein
New Mexico State University
Physical Sciences Lab
Las Cruses, NM 88003-0001

1 Dr. Donald Birx
New Mexico State University
Physical Sciences Lab
Las Cruses, NM 88003-0001

1 Byron Dean
Cedar Crest, NM 87008

1 M.F. Sulcoski
National Ground Intelligence Center
220 7th St. N. E.
Charlottesville, VA 22902

1 US Army Criminal Investigation Com
ATTN: LTC C.W. Hunt
6010 6th St.
Fort Belvoir, VA 22060-5506

1 Ed Nozawa
Lockheed Martin Aeronautics Co.
86 south Cobb Drive
Marietta, GA 30063-0264

1 Robert Burch
Dept. of Philosophy
Texas A&M University
College Station, TX 77843-4237