

Metacognitive Behavior in Adaptive Agents¹

by Bryan Thompson, Dr. Marvin Cohen, and Dr. Jared Freeman

Cognitive Technologies, Inc.

4200 Lorcom Lane

Arlington, VA 22207

bbthomps@digex.net (Thompson)

cti@digex.net (Cohen & Freeman)

Abstract: This paper presents a novel multidisciplinary architecture which unites recent research in cognitive science and connectionism to develop an integrated solution, and computational model, for the acquisition and performance of recognitional and metacognitive skills by human subjects and intelligent agents. This work builds on three key components: (a) the Recognition / Metacognition model of human decision making; (b) the SHRUTI model of reflexive reasoning in a connectionist system; and (c) Adaptive Critics, a connectionist model of behavior learning. We are attempting to develop a hybrid computational realization of the Recognition/Metacognition model as a basis for the design of adaptive intelligent agents.

Introduction: Intelligent Agents (IA) seem to be getting all the press these days, but few of the issues posed by the agents paradigm are new, and fewer still are easy. These are the early days for intelligent agents, and even the term remains ill-defined. The most intriguing and challenging definitions carry the connotations of personality: information-gathering, decision-making, communication, and autonomous action. In this paper, we build on psychological evidence of the decision-making capabilities of humans to develop a computer architecture suitable for further exploring human decision making skills and building artificial, intelligent agents that model those skills. We draw, in the course of this paper, on a series of interviews with Naval tactical decision makers, and present an example of how an intelligent agent would implement behavior such as observed in a human decision maker.

Current Models of Decision-Making: Two quite different starting points have characterized recent discussions of how people learn and use decision-making skills. The *analytical model* (e.g., Raiffa, 1968, Keeney and Raiffa, 1976) specifies that decision makers: learn to generate options, outcomes, and goals; assess the probability of outcomes and the value of goals; aggregate the probabilities and values into a summary score for each option; and choose among the options. The observed effects of options serve as feedback with which the decision maker rationally updates her future assessments. The *recognition model* (Chase and Simon, 1973; Klein, 1993) supposes that experienced decision makers match perceptual cues and salient information to learned patterns, which in turn trigger specific responses. The success or failure of these responses alters the stored relations between cues, patterns and responses.

According to a large corpus of research, such models do not account for human behavior in critical decision-making tasks. Experienced decision makers do not engage in the exhaustive generation of outcomes or comparison of alternatives required by the analytical model. Examination of interviews with Naval officers conducted for the TADMUS project reveals few instances of analytical processes, and a wealth of instances in which officers attempted to weave coherent stories from events that described the intent and future actions of a radar contact. Where the analytical process simply aggregates concordant and conflicting data, we observed that officers treat the two types of data quite differently - conflicting

¹ To appear in the *CD-ROM Proceedings of the World Congress on Neural Networks*, 1995. Reprints are available from Thompson.

evidence (e.g., regarding intent) was used as a symptom of erroneous assumptions and led to efforts to find alternative interpretations of cues or alternative hypotheses, while concordant data were taken as confirmation of the current hypothesis (Cohen, 1986, 1989).

Psychologists have found systematic deviations from normative models at virtually every stage of the decision process (e.g., Kahneman, Slovic, & Tversky, 1982). Rather than label these as errors or biases, however, they should be examined as clues regarding highly adaptive decision processes that seem qualitatively different from the analytical approach (Cohen, 1993a).

The recognitional viewpoint, in turn, fails to account for the ability of officers to handle unfamiliar problems, update schemas in the face of novelty, handle conflicting and unreliable data, and to change their minds. In critical incident interviews with tactical Naval officers, we found numerous examples in which they: juggled competing hypotheses (e.g., one indicating hostile intent, the other indicating non-hostile intent), none of which perfectly matched the situation; generated alternative interpretations of cues in a remarkably controlled manner; and balanced the need to act with the need to think about the problem. None of these behaviors is well-explained under the recognitional model.

The Recognition / Metacognition model: We have proposed an alternative model of the acquisition and performance of decision-making skills. It is a naturalistic approach, which begins with the way experienced, effective decision makers actually make decisions in real-world tasks (Klein, Orasanu, Calderwood, & Zsombok, 1993, Holyoak, 1991). The model is based on interviews with 14 tactical Naval officers (Cohen, Freeman, & Wolf, 1994) and 33 U.S. Army command staff (Cohen, Adelman, *et al.*, 1993). We have found (as have others in expert-novice research) that these officers use methods that combine pattern recognition with strategies for effectively facilitating recognition, verifying its results, and constructing more adequate models when recognition fails. From this foundation, we have developed an adaptive model of decision making that integrates recognition and metacognitive processes. We call this the *Recognition/Metacognition (R/M) model* (Cohen, 1993b).

The recognition function of the R/M model is responsible for transforming perceived and internally generated data into a cognitive model of the situation at hand, and for triggering an associated action or plan. In some cases, the situation model may reflect the activation of a familiar pattern, which in turn activates an associated response. For example, routine patrol or commercial air activities may be characterized by a consistent set of features corresponding to known routes, observed speed, altitude, and emissions, and associated activities, such as identification-friend-or-foe (IFF) transponder signals, warnings, and so on. In some very important cases, however, the available data do not satisfactorily match any previously experienced pattern. In these cases, metacognitive processes manipulate the recognition system, including memory, and information gathering processes, in order to construct coherent stories. According to Pennington and Hastie (1993) and others (e.g., Pearl, 1988), decision makers use causal representations when data are voluminous, complex, and interdependent. A story is a particular kind of causal structure that organizes information about human behavior, such as evidence presented at a trial or the indicators of the intent of an approaching surface or air contact. The main components of a story episode, or scenario, are initiating events (which elicit) goals (which motivate) actions (which result in) consequences.

Naval officers try to understand unfamiliar observations about a contact or set of contacts by constructing stories that involve assumptions of intent such as reconnaissance, harassment, search and rescue, or attack. In constructing stories, decision makers draw on case-specific data, general knowledge, and knowledge of similar cases. Based on general knowledge and experience, they develop an understanding of the elements required for a complete story.

The R/M model accounts for the ability of experienced decision-makers to recognize situations, to elaborate on and test those recognitions, and to determine when to cease elaboration in order to act. The model consists of several components: a situation recognition system, a gating function (called the Quick Test), and two types of metacognitive processes for editing the products of recognition: *critiquing* and *correcting*. Processing is cyclical. A situation may be recognized (albeit partially or imperfectly), the recognitional

product evaluated and modified, and then submitted for further evaluation and modification, before finally triggering an action associated with the recognized state.

A Hybrid Architecture for Agents with Metacognitive Behavior Learning: What kind of an architecture is suited to implementing such concepts? We suggest that the twin functions of recognition and metacognition, may be realized by a pair of cooperating adaptive critic architectures. (We use the term *Adaptive Critic* here to broadly refer to any of a variety of reinforcement learning methods that may be described as heuristic approximations of dynamic programming, including Heuristic Dynamic Programming (HDP), (Werbos, 1977), TD-Learning (Sutton, 1988), and Q-Learning (Watkins, 1990).) One adaptive critic architecture is focused on the recognitional modeling of reality (the domain), associated plans of action (policy), and the expected utility of the outcomes of such action. The other, the metacognitive critic architecture, is concerned with monitoring, critiquing, and correcting the recognitional products and with maintaining coherent stories that explain, and do not falsely aggregate, the available evidence. Backpropagated adaptive critics (BACs), (Werbos, 1990), are used for both recognition and metacognition. A SHRUTI network (Shastri & Ajjanagadde, 1993) is used in the recognitional model (**Figure 1**), as it is well-suited to the maintenance of structured representations of evidence and relationships. We find a role for traditional symbolic systems in managing the interface between the recognitional and metacognitive subsystems. We will now briefly review some aspects of the proposed architecture, including Adaptive Critics, SHRUTI, deictic representations and their roles and implications for this design.

Adaptive critics: An integrated connectionist model of skill acquisition and performance. Adaptive critics implement incremental approximations to dynamic programming (DP). They compute, as their principal output, an approximation of a strategic utility function which may be used to adapt a plan of action, strategy, or behavior. Adaptive critics have been used widely in neurocontrol, where they guide the adaptation of a controller that learns to take actions to manipulate a process or control a plant. They may be used to solve problems in stochastic optimal control where the state space is so large, and the process dynamics so non-linear, that conventional control techniques, including dynamic programming, are not feasible. Adaptive critics may be viewed as implementations of a reinforcement learning strategy by which *behavior* is learned. It is with this insight that we propose to apply adaptive critics to the connectionist learning of *metacognitive* behavior, that is, to the problem of learning how and when to control the *recognitional* process. With reference to **Figure 1**, the larger boxes, labeled *Recognition* and *Metacognition* represent backpropagated adaptive critics (BACs). Each BAC is composed of three conceptually distinct, though integrated, multi-layer, differentiable networks: the *critic*, *model*, and *action* networks. While the recognition BAC is responsible for learning to interpret evidence and recommend actions, the metacognition BAC is responsible for learning to specify and apply metacognitive strategies that may improve recognitional results.

Briefly, in a BAC, the critic network is adapted by HDP, a reinforcement learning method; the action network is adapted by backpropagating from the critic *through* the model network to the action network, and the model network is adapted, e.g., by forecasting techniques such as backpropagation through time, to be predictive of future observed inputs and *not*, principally, in response to the backpropagated error signal from the critic. These issues are covered in detail by Werbos (1990).

The recognition subsystem is responsible for constructing a *situation model* given a sequence of cues, and for recommending actions given the current situation model. Intelligent agents typically live in highly structured information environments. To exploit the structured relationships present in such domains, the recognition model network utilizes SHRUTI, a connectionist solution for the representation of rules, variables, and dynamic bindings using temporal synchrony. Further, SHRUTI networks are differentiable and therefore suited for integration into BACs, whereas traditional symbolic inference engines are not. The context provided by metacognitive actions, including a dynamic attentional focus, are used to simplify the output of the recognition model (a SHRUTI network) as viewed by the metacognition model network and the recognition

critic and action networks (**Figure 1**). The recognition critic, which learns to estimate the expected value of outcomes associated with a story or scenario, and the recognition action network, which learns to suggest actions that lead to good outcomes (i.e. outcomes associated with a high expected value) are more traditional, distributed backpropagation networks, and their roles are typical of neuro-control.

The situation model encompasses all of the information actively instantiated in the recognition model network (**Figure 1**). That is: situation-specific observations and information activated in the long-term knowledge base that serve as evidence (in semantic frames), stories (causal networks of events, motives, and actions that include a key conclusion, such as *hostile intent*), and arguments (reasoning concerning the validity of the causal links that tie evidence to conclusions within stories). The situation model can have many active and supported stories. When these stories support different (mutually exclusive) conclusions, then conflict exists - which metacognition can try to resolve. Uncertainty can also exist with respect to a single story. There may be incomplete arguments (e.g., links in which the cause is not specified in sufficient detail to establish the effect, or the effect is not sufficiently specified to establish the cause), unreliable arguments (e.g., links in which unsupported assumptions are used to flesh out the cause or the effect), and conflicting arguments.

Reflexive reasoning is concerned with a limited set of inferences which people are able to make very quickly, within a few hundred milliseconds. Such inferences fill in causal links (i.e. The cat knocked over the vase. The vase broke. ergo The vase fell.) within selectional constraints (i.e. John hit the ball. vs. The book hit the ball., where book may not be an agent). We propose that: (1) the recognitional model performs reflexive reasoning when incorporating new evidence (i.e. track changed heading to own ship). (2) That this reasoning is done within story structures (e.g., The track will continue to converge with own ship. The contact intends to attack own ship. The contact has some method for localizing own ship. The contact has weapons capable of damaging own ship. The country of origin had some reason for selecting this platform for the attack. The country of origin has some motive for attacking a US ship. The country of origin or the contact had some reason for selecting own ship as the target). (3) That the propagation of learned, or chunked, arguments is reflexive (For example, if converging hostile, then hostile will attack, on account of Kadhafi's threats to defend Libyan territorial waters and own ship's being in same waters, unless hostile platform is inappropriate for attack on own ship class or own ship is not the actual target or contact is conducting a search and rescue mission, etc. And (4) that background knowledge for stories and pre-learned arguments are automatically retrieved and mutually constrain the flow of reflexive inference that elaborates stories from evidence.

The recognitional domain model will be implemented as a SHRUTI design. We are proposing a localist connectionist system to exploit the large body of structured domain knowledge. The SHRUTI system is one of a very few models of connectionist reflexive reasoning. In particular, it supports dynamic binding, via temporal synchrony, and very rapid inference for precisely the kinds of reflexive reasoning which the R/M model requires. Stories and argument structures will be implemented as n -ary predicates and as inferential links that support special limited classes of inference among those predicates. SHRUTI will also implement selectional restrictions required to match story components and argument components with new evidence and long-term background knowledge.

Arguments are structured sets of inferential links among story components, and other arguments, that are dynamically generated in response to a particular situation. They embody the evidence-conclusion-exception relations implicit in cause and effect reasoning on stories. Evidence-conclusion relations can point from cause to effect, from effect to cause, or from one effect to another effect of a common cause. Argument structures guide metacognition, which may implement searches in the knowledge base for rebuttals, i.e., alternative causes and effects with respect to arguments in the current story. These alternative causes and effects can inhibit propagation of inference along links in the current story. Metacognition can also implement searches in the knowledge base for information that will strengthen arguments in the current story, that is, it can support propagation along links in the current story. The inferential changes in the domain model which result from metacognition are implemented as structural modifications of the SHRUTI network. These modifications may be fleeting or they may become learned components of the domain model.

Dynamic binding, as provided by temporal synchrony of node firings, binds fillers to slots in story components. SHRUTI makes definite predictions concerning the limits of working memory based on the resolution available for dynamic binding before degradation begins. These predictions essentially limit the number of role fillers which can be active at any one time, but do not place limitations on the number of parallel inferences which can be drawn concerning those role fillers. For example, the SHRUTI model might observe *bearing(hostile-craft, own-ship) & speed-increasing(hostile-craft)* and rapidly draw the inferences required to build a hostile-intent story. Pre-learned arguments, stored as long-term facts, would automatically be activated by the propagation of activity along inferential links and constrain the generation of the story.

Deictic representations, focus, & selective attention. Our senses are a limited, though highly refined, resource through which we perceive and seek to understand the world. Given their limitations (e.g., finite line of sight, resolution, and field of view), we must allocate their resources carefully and dynamically. Many complex problems require the direction of sensory attention to objects in the external world. Learning for these problems includes learning which objects to selectively direct attention to, i.e. learning to *attend*.

The term ‘deictic’ is from the Greek ‘deiktikos’, meaning “to show directly”. Deictic representations (Agre & Chapman, 1987, Whitehead & Ballard, 1990) are used to address a special kind of dynamic binding problem in connectionist systems. They serve to dynamically bind key *representational patterns* that are the current subject of *attentional focus* into segments, known as *deictic markers*, of another representation. They may be used, as we propose, to interface the recognition model (with its potential for multiple instantiated schemas and a compositional, localist representation) to the metacognitive system (with its more traditional, vector style, distributed representation). The result is that metacognition is able to dynamically select (focus on) what information patterns it attends to and uses in developing and applying metacognitive strategies.

Considering a single story at a time. Evidence suggests that people focus on one story, and its associated arguments, at a time. The proposed design emulates this behavior by using metacognition to focus on a particular story, e.g., by generating assumptions (in the metacognitive action network) whose implications are reflexively propagated (in the recognition model network) and used to generate actions (by the recognition action network) and the expected value of outcomes (by the recognition critic network), which are conditional on the assumption that *this particular story* is correct.

Human decision makers typically act on the basis of a single scenario, e.g., a *worst case* or a *most likely* scenario, not on the basis of aggregate expectations for outcomes (i.e. an expected value situation analysis), which do not represent any potentially true scenario, or its outcomes. The evidence that people combine explicit structured reasoning about uncertainty with exploring single coherent scenarios (stories) has interesting implications for machine learning. All useful machine learning search techniques are biased, and each has its shortcomings. Arguably, it is both difficult and somewhat pointless to estimate expected values for aggregate situation models. By representing uncertainty explicitly, but only *exploring* a single story at a time (as mediated by metacognition), interesting hybrid search strategies can be developed which avoid making estimates of aggregate quantities which are never observed. This is likely to be far more valuable, and a better model of human decision making, than associating expected value estimates with the combined scenarios (stories) within the situation model. If desired, we can (via metacognition) compute the expected value associated with each plausible scenario by presenting them in turn to the recognition critic.

Presenting the recognition critic network with an estimated state representation based on a single story and its associated arguments *as if* we believed that to be the true scenario lets us constrain search in a very plausible way and, as alluded to above, dramatically reduces the complexity of the *output* representations generated by the recognitional model. By directing attention to, and elaborating arguments for multiple plausible scenarios, metacognition avoids biases inherent in exploring a single scenario and its associated strategies.

Example: Implementation of Metacognitive Strategies in the Hybrid Architecture. Figure 2 illustrates the cyclical nature of critiquing and correcting in one of the critical incidents reported in the TADMUS

interviews (Cohen, Freeman, & Wolf, 1994). In this incident, a slow, non-responsive aircraft was approaching the cruiser. Slow speed represented conflict with expectations generated by a hostile-intent story (step 1 in **Figure 2**). The Anti-Air Warfare Coordinator (AAWC) tried to explain this conflict, while retaining the belief that the aircraft was hostile, by dropping the assumption that the aircraft knew where its target was: The aircraft was moving slowly because it was looking for a target. This assumption, however, requires more detail. Since it does not specify the means used by the aircraft to look for a target, there is no clear argument supporting the prediction of slow speed. Thus, the need to fill a localization slot in the story model (step 2 in **Figure 2**): The AAWC must tell a plausible story about what the aircraft would be doing in order to locate a target.

The AAWC first tries to fill in the localization slot with the more specific assumption that the aircraft is searching visually. This assumption satisfactorily completes an argument for the prediction of slow speed. However, there is no particular support for this assumption, so it is unreliable (step 3). The AAWC chooses to test this unreliable assumption by mentally simulating the events that would be expected if it were true, and comparing them to actual observations. If the aircraft were searching visually, it would be flying erratically. Yet this aircraft is flying straight — a new conflict (step 4). The AAWC tries to resolve this new conflict by finding an assumption that rebuts the conflicting argument: he asks if the aircraft might have prior knowledge of where its target is: “I wonder what their intelligence capability is?... Do they know where I am?” If it knows precisely where its target is, the aircraft might be flying straight even though it is attempting to visually acquire the target. This assumption, however, is not directly supported, so it is unreliable (step 5). He tests it by mentally simulating the events to be expected if it were true: If it had precise intelligence, the aircraft would be heading toward the cruiser or some other ship in the area. But it is not. The assumption that it knows the location of the target must be dropped, leading back to conflict (step 6). This conflict is now resolved by discarding the assumption of visual search, leading back to incompleteness in the localization slot (step 7 in **Figure 2**). The AAWC now tries to fill this slot with a new assumption: that the aircraft is searching electronically (rather than visually) for a target. This assumption also is unreliable (step 8), so he decides to test it by mental simulation. If the aircraft was searching electronically, it would be emitting. But this aircraft is not emitting — again producing conflict (step 9). Electronic search has to be discarded, leading back to incompleteness (step 10). Finally, the AAWC considers filling the localization slot with a different assumption: “I can’t imagine him shooting blind. They could, though --- they did it once, shot something off and hoped it hit something... So you’re thinking, well, he probably won’t shoot, but he might.” This assumption, that the aircraft might shoot without knowledge of the location of its target, is a last-ditch effort to save the conclusion that the aircraft is hostile. The assumption cannot be directly tested, but is judged to be unreliable (step 11). Because of this failure to construct a plausible a hostile intent story, engagement of the approaching contact was delayed. It turned out to be a friendly aircraft.

With reference to **Figure 2**, the changes labeled on the arrows (e.g., drop assumption, collect new data; make assumption) are the direct result of primitive correcting *actions* taken by the metacognition system. Actions in the metacognition system are represented as a vector of actions, $\underline{u}(t)$, for each time t , and are *interpreted* by a symbolic subsystem (Inference Control) with respect to the current *focus* of the metacognition system and the active structures in the recognition model (i.e. the domain model, implemented as a SHRUTI network). Critiquing actions include those related to shifting the *focus* of the metacognition system along and between lines of reasoning, (e.g., tracing potential causes of an effect by following cause and effect links of arguments), mental simulation (e.g., what-if analysis), and the devil’s advocate strategy (falsifying one conclusion or argument to force the generation of another). Expectations of outcomes are generated by iterating the recognition model into a simulated future. Expectations of the valuation of outcomes are generated by the recognition critic for a given outcome. Explanations are generated by the recognition model via reflexive reasoning, which may include the instantiation of long-term memories that were not triggered when elaborating the original story. A particular critiquing strategy, e.g., identifying incompleteness, is realized by a sequences of actions in the metacognitive system. For

example, (a) shift the metacognitive focus along the story from conclusion to evidence and identify uncertain or missing story components. (b) Elaborate those story components, using facts in long-term memory to identify evidence that would reduce the uncertainty in this story. And (c), attempt to gather relevant new evidence if operators are available to do so, or explain how the existing evidence could be construed to reduce uncertainty.

Conclusions: We have proposed that metacognitive behavior operates over recognitional models of causal relations and associated measures of uncertainty and conflict. Further, metacognitive skill is not domain-independent, but is *enhanced* by knowledge of the domain. In the proposed architecture, information for the metacognitive system is filtered through a deictic representation, which is controlled by the metacognition subsystem via a dynamic attentional focus on structured subsets of the active assumptions, evidence, and relationships in the recognitional model.

Our approach has several broad implications. First, metacognitive behavior is organized about a small and regular subset of the instantiated structures in recognitional system model. (Pinker, 1989, discusses an interesting parallel in lexical semantic structures.) This suggests that metacognitive learning is neither intractable nor unapproachable with current methods. Second, metacognitive skills, once acquired, may be transferred to new domains. Metacognitive skill in the new domain will quickly improve as relevant domain information is incorporated by the metacognitive model to prioritize the already developed metacognitive strategies. In our research we are exploring the acquisition of metacognitive skills by an agent, such as described here, and comparing its skill acquisition and performance with that of human subjects in the same domain.

There must be boundaries on the recursive embedding of representations and computational structures in intelligent systems. This work describes a computational architecture that divides certain classes of cognitive behavior into recognitional processes (reflexive inference, action policies, and expected utility of outcomes) and metacognitive processes (monitoring, critiquing, and correcting) that we hope will further the understanding and construction of adaptive intelligent agents. Finally, we hope that this will provide a framework for the use of adaptive critic architectures as embedded components in complex system designs. In particular, it is interesting that some kinds of adaptive learning in the reflexive reasoning system simply are not possible without the parallel metacognitive system to focus attention on plausible and coherent interpretations of the available evidence.

References:

- Chase, W.G., & Simon, H.A. "The Mind's Eye in Chess," in W.G. Chase (ed.), *Visual Information Processing*. NY: Academic Press, 1973.
- Cohen, M.S. "Decision making 'biases' and support for assumption-based higher-order reasoning." In *Proceedings of the Fifth Workshop on Uncertainty and AI*. Windsor, Ontario, August 1989.
- Cohen, M.S., Freeman, J., Wolf, S.P., Militello, L., Klein, G.A., and Leddo, J. (1994, May 2). Recognizing, Thinking, and Acting: Training Anti-Air Warfare Decisions. Technical report prepared for the Naval Training Systems Center, Orlando, FL. Arlington, VA: Cognitive Technologies, Inc.
- Cohen, M.S. "The naturalistic basis of decision biases." In Klein, Orasanu, Calderwood & Zsombok (Eds.), *Decision Making in Action: Models and Methods*. (pp. 51-102). Norwood, NJ: Ablex Publishing Corp, 1993a.
- Cohen, M.S. "Metacognition strategies in support of recognition." To appear in *Proceedings of the 1993 Human Factors and Ergonomics Society 37th Annual Meeting*, Seattle, WA, 1993b.
- Cohen, M.S., Adleman, L., Tolcott, M.A., Bresnick, T.A., & Marvin, F.F. A cognitive framework for battlefield commander's situation assessment. Technical report 93-1. Arlington, VA: Cognitive Technologies, Inc. 1993.
- Holyoak, K.J. "Symbolic connectionism: Toward third-generation theories of expertise." In Ericsson, K.A. & Smith, J. (Eds.), *Towards a general theory of expertise*. Cambridge: Cambridge University Press, 1991.

- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). *Judgment under uncertainty: Heuristics and biases*. NY: Cambridge University Press, 1982.
- Keeney, R.L., & Raiffa, H. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. NY: Wiley and Sons, 1976.
- Klein, G.A. "A Recognition-primed Decision (RPD) Model of Rapid Decision Making." In Klein, Orasanu, Calderwood & Zsombok (Eds.), *Decision Making in Action: Models and Methods*. Norwood, NJ: Ablex Publishing Corp, 1993.
- Pennington, N., & Hastie, R. "A theory of explanation-based decision making." In Klein, Orasanu, Calderwood & Zsombok (Eds.), *Decision Making in Action: Models and Methods*. Norwood, NJ: Ablex Publishing Corp, 1993.
- Pearl, J. *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Mateo, CA: Morgan Kaufman Publishers, Inc., 1988.
- Pinker, S. (1989). "Learnability and Cognition: The Acquisition of Argument Structure." The MIT Press, Cambridge Massachusetts.
- Raiffa, H. *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Reading, MA: Addison-Wesley, 1968.
- Shastri & Ajjanagadde. "From simple associations to systematic reasoning: A connectionist representation of rules, variables, and dynamic bindings using temporal synchrony," in *Behavioral and Brain Sciences*, 1993, vol. 16.
- Sutton, R.S. (1988). "Learning to predict by the methods of temporal differences," in *Machine Learning* 3: 9-44.
- Watkins, C. (1989). Ph.D. diss., "Learning from delayed rewards." Cambridge University, Cambridge England.
- Werbos, P. (1977). "Advanced Forecasting Methods for Global Crisis Warning and Models of Intelligence," in *General Systems Yearbook*. (Appendix B).
- Werbos, P. (1990). "A Menu of Designs for Reinforcement Learning Over Time," in *Neural Networks for Control*. Miller, Sutton & Werbos, eds. Cambridge: MIT Press.
- Agre, P.E., & Chapman, D. "Pengi: an implementation of a theory of activity." In *AAAI*, pp 268-272, 1987.
- Whitehead & Ballard. Learning to perceive and act. Technical report TR 331, CS dept. Univ. of Rochester, 1989.

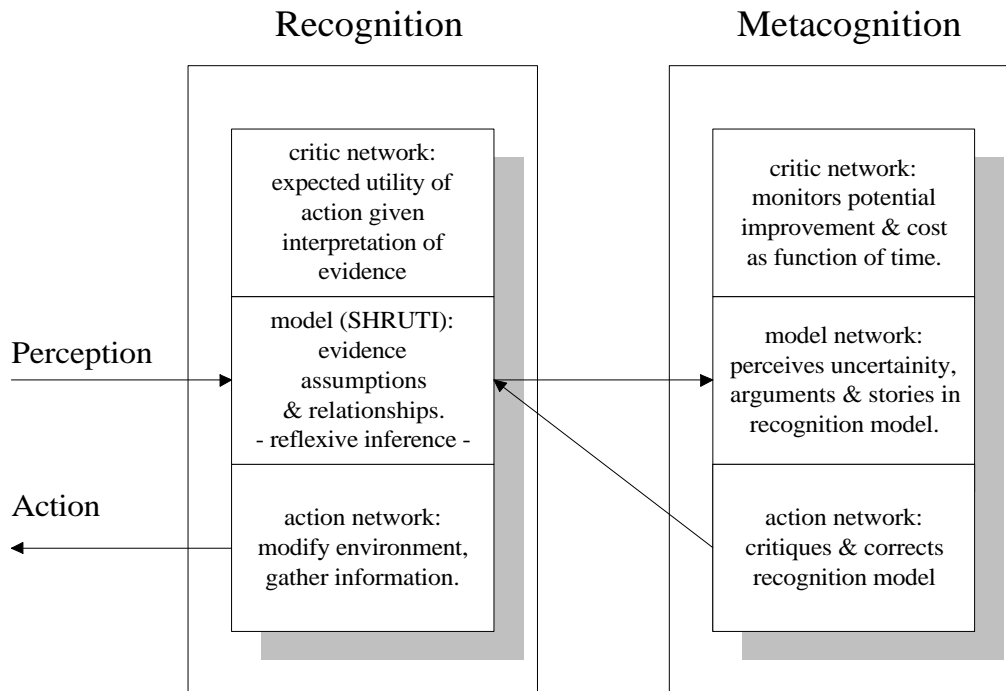


Figure 1. Proposed architecture for Agent with metacognitive behavior learning utilizing two backpropagated adaptive critics.

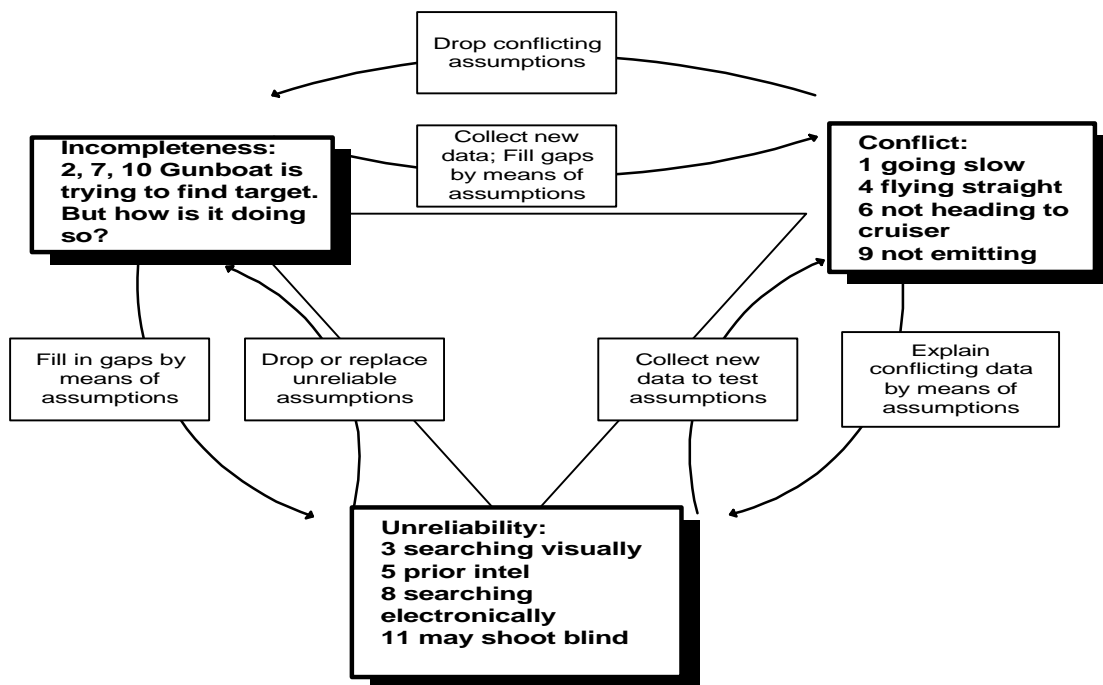


Figure 2. Examples of the cyclical character of critiquing and correcting. Problems discovered by critiquing are at corners of triangle. Correction steps are in boxes labeling arrows.