SAMPLE: Situation Awareness Model For Pilot-In-The-Loop Evaluation

Sandeep S. Mulgund, Karen A. Harper and Greg L. Zacharias
Charles River Analytics Inc.
Cambridge, MA 02138

Timothy E. Menke
SIMAF ASC/HPMT
Wright-Patterson AFB, OH

ABSTRACT: *This paper describes the design and development of an intelligent agent model for human behavioral representation (HBR) in real-time air combat simulation. The essential feature of this model is its organization around the skilled human's situation assessment behavior in a complex multi-task environment. Its application is demonstrated for modeling tactical pilot behavior in a four-ship fighter sweep mission. A detailed knowledge engineering exercise was conducted with Air Force pilots to define relevant situation awareness (SA) and decision-making models. These models were inserted into a real-time tactical engagement simulator, along with a mechanism for inter-agent communication. Simulation results demonstrate the effectiveness of using SA-centered decision-making as a framework for modeling tactical piloting behavior. The agents demonstrate the ability to set up and execute intercepts, perform defensive reactions, make kill assessments, and re-engage when necessary. The inter-agent communications mechanism facilitates the development of shared situation awareness, and it provides a command mechanism for the flight leader to direct the behavior of other flight members.*

1. Introduction

The growing use of simulation for military training, systems analysis and acquisition, and command decision aiding has created a need to develop accurate computational models of individual and team behavior (Pew & Mavor, 1998). The goal of such models is to mimic the behavior of a single human or the collective behavior of a team of humans in some well-defined operational environment. For training purposes, they can provide a realistic threat environment when it is not feasible or desirable to fully populate a simulation with human players. To be effective, the models should demonstrate observable behaviors that are realistic to other participants in the simulation. Similarly, for systems analysis and evaluation, accurate human behavioral models can provide a realistic context for evaluating the performance and robustness of new design concepts.

One particular area where effective human behavioral models are needed is tactical air combat. Fighter pilots are facing increasing numbers of sophisticated threat weapons systems, and air superiority scenarios are becoming progressively more time-critical and complex. In response, significant effort is being placed on developing technologies that offer the potential to increase the pilot's *situation awareness* (SA). While many definitions of SA have been proposed (Pew & Mavor, 1988), a succinct one is offered by Endsley (1995): *situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status into the near future*. Few reliable tools exist at present for evaluating the utility of any given subsystem in terms of its direct potential for enhancing SA.

Many computer models have been developed for evaluating the impact of new technologies on overall mission success. A key U.S. Air Force (USAF) model is the Man-In-the-Loop Air-to-Air System Performance Evaluation Model (MIL-AASPEM), developed by The Boeing Company (Lawson & Butler, 1995). It is used extensively for subsystem effectiveness evaluation and air-to-air combat tactics development. MIL-AASPEM's strengths include capabilities for: 1) representing multiple types of aircraft and their associated avionics, sensor subsystems, displays, weapons and ground players; 2) conducting man-in-the-loop simulations using any combination of manned pilot stations and simulated pilots implemented via a rule-based pilot decision logic (PDL); and 3) simulating a many vs. many (MvN) engagement in real time. While MIL-AASPEM has been used effectively in several system studies, a number of key weaknesses have been identified. These include:

1) An inability to relate proposed display or subsystem enhancements to improved pilot situation assessment, decision-making, and overall performance.
2) The lack of an integrated architecture for representing advanced pilot behavior models needed for accurate predictions of engagement outcome.
3) The lack of a realistic representation of the pilot's capabilities and limitations in information processing, situation assessment, and decision-making.

Most of these problems can be traced directly to MIL-AASPEM's emphasis on a *rule-based* PDL, which does not accurately model situation assessment behavior. Since the rule-based PDL lacks capabilities for representing key SA concepts such as a tactical mental model, the process of evidence accumulation, and the quantitative representation of situation awareness, it cannot simply be expanded to properly account for SA-related pilot activities. The rule-based PDL model is thus considerably limited in its representation of the pilot's actual decision-making process. Consequently, no matter how well MIL-AASPEM represents the non-human components of the scenario (e.g., vehicle, subsystems, displays, etc.), its fidelity will always be limited by the validity of those components representing human decision-making behavior.

In light of these shortcomings an upgrade effort was initiated, focusing on the development of a high-fidelity agent model of SA-centered pilot decision-making, and the demonstration of its utility in MIL-AASPEM.

2. Background

The objective of this research was to develop an intelligent agent model that could be used to simulate skilled human behavior in a real-time tactical air combat simulation environment. This section provides background material that supported this modeling and development effort. Section 2.1 describes the Rasmussen Hierarchy of human information processing and skilled behavior, which is a conceptual framework for analyzing different types of human skills. Section 2.2 describes past research in the area of human decision modeling that supported the formulation of the intelligent agent model. Finally, section 2.3 describes some past efforts to develop HBR models that directly influenced this research. The reader is referred to Pew and Mavor (1998) for a broader overview of human behavioral modeling for simulation.

2.1 Rasmussen Hierarchy of Human Behavior

Rasmussen's three-tier model of human information processing and skilled behavior (Rasmussen, 1981, 1982), shown in Figure 1, provides a good unifying theoretical framework for analysis of different human skills that may be modeled within a real-time simulation.

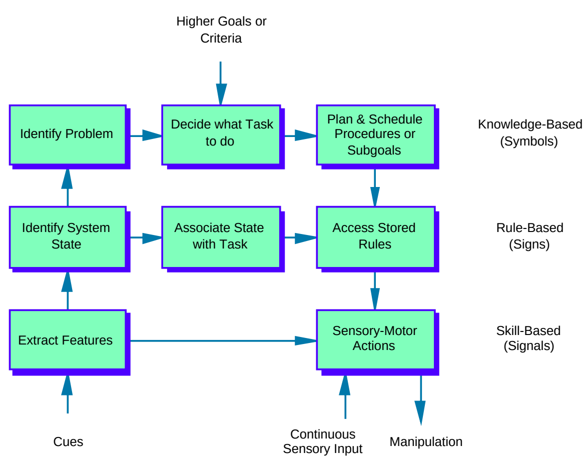

Figure 1: Rasmussen Hierarchy of Human Behavior

By dividing skilled behavior into categories based on the degree of automaticity, complexity, and level of cognitive processing, this framework supports systematic skill decomposition and measurement of individual aspects of the overall skill on a part-task basis. Each link in Figure 1 represents flow of information through the human information processing apparatus. Information flows through the system starting at the bottom left with environmental input, and flowing upwards along the left side of the diagram to the most complex level of knowledge-based processing. Decisions about behaviors propagate downward through the right side of the figure until they are executed by the motor organs. Aspects of the stimuli that can be handled at the lower levels are processed there (i.e., a cognitive "shortcut") and only situations that require more sophisticated processing reach the most complex knowledge-based level. Processing is divided into three broad categories, corresponding to activities at three different levels of complexity:

Knowledge-based behavior is the highest level of complexity, required for any complex problem solving that has not been fully automated and typically involves handling new or unusual situations where reasoning from first principles is required. These situations are often made more complex by the need to engage in several parallel tasks simultaneously.

Rule-based behavior is at the core of high-performance skill as it involves well-practiced, often highly automatized behavioral sequences that comprise a set of *compiled situation-action pairs* (i.e., rules). It is demonstrated in situations requiring standardized procedures. The emphasis here is on accurate and timely situation assessment, followed by the appropriate procedural response.

Finally, **skill-based behavior** involves well-practiced sensorimotor skills that do not involve cognitive resources but are performed largely automatically in response to recognized stimuli. It is the most automated type of behavior demonstrated by almost unconscious performance of highly trained sensorimotor tasks.

The Rasmussen hierarchy is a good framework for decomposing a particular task-skill pair into constituent skill-types, each with different processing characteristics and requirements. This hierarchy of performance skills is a good tool for decomposing complex skills into their constituent elements. Although not posed directly as a computational representation, it does provide a good basis from which to approach the design of an intelligent agent for human behavioral modeling.

2.2 Human Decision-Making in Complex Task Environments

Human performance in decision-making has been studied extensively, primarily through empirical studies but increasingly with computational tools. These studies span the theoretical-to-applied spectrum and cover many domains.

Endsley (1993, 1995) and Adams, Tenney & Pew (1995) discuss a psychological model of decision-making, focusing in particular on situation awareness (SA), and the impact of particular system characteristics on the operator workload, attention and memory requirements, and the likelihood of errors. Klein (1989, 1994) has studied a particular type of decision-making predicated on the quick extraction of salient cues from a complex environment and a mapping of these cues to a set of procedures. Research indicates that such *Recognition-Primed Decision-making* (RPD) plays a major role in planning and it is therefore critical for decision-aiding systems to recognize and support this mode of human information processing. Situation-centered decision-making has been widely accepted as the most appropriate representation of actual human decision-making in high tempo, high value situations (Klein, 1989; Stiffler, 1988; Fracker, 1990).

While many definitions of situation awareness have been proposed (see Pew & Mavor, 1988 for an overview), a particularly succinct one is that offered by Endsley (1995), which divides SA into three hierarchical levels:

- **Level 1**: Perception of the key elements in the environment within a volume of time and space. This is the identification of all of the relevant elements that, when combined and interpreted, define the current situation.
- **Level 2**: Comprehension of the current situation. This consists of the mental "integration" of level 1 events into an overall, operationally relevant picture of the current situation.
- **Level 3**: Projection of the current situation into the near future. This is the prediction of the current situation into the future, to support short-term planning when possible.

These definitions are especially useful for the development of intelligent agent models of human behavior, because they can be used to as a starting point for knowledge engineering exercises with domain experts.

There is a significant difference between the SA-centered model and the conventional decision-centered model that views the decision maker as "faced with alternatives, and considering the consequences of each alternative in terms of analysis of future states (odds/probabilities) weighed against alternative goals" (Klein, 1989). In the SA-centered model, no utility or alternative is considered; instead, SA becomes the focus of all human actions. It defines a decision maker's view of the environment and characterizes the information needs that drive his/her experiential (*if-then*) decision-making. This view of human decision-making provides the basis for the intelligent agent representation developed under this research.

A number of studies have been conducted focusing on the differences between expert and non-expert performance. An experiment designed to determine differences in information usage by tactical planners indicated that 78% of critical facts identified by the experts were missed by the non-experts. The facts missed by non-experts included timing information, actions of adjacent units, changes in boundaries, enemy activities, terrain constraints, mobility, engineering capabilities, and logistical loads (Fallesen *et al*, 1992). Another critical difference between experts and non-experts is the use of uncertain information. Experts were more aware of uncertain assumptions and made explicit predictions of events that would confirm their expectations and thus confirm or disconfirm assumptions (Tolcott *et al*, 1989). A study of expert military tactical decision-making (Deckert *et al*, 1994) found that experts' performance differed along a number of dimensions, including awareness of enemy activities, learning from past mistakes, flexibility of planning, seeking of disconfirming evidence, deeper exploration of options, and better management of uncertain information. Although these characteristics are not captured within SAMPLE at present, they do provide a roadmap for modeling individual differences in human performance during future research.

2.3 Computational Models for Human Behavioral Representation

The modeling approach used here has its roots in development that began with the Optimal Control Model (OCM) (Kleinman & Baron, 1971). This is an information-processing model of the operator of a dynamic system, grounded in modern control and estimation theory, which accounts for closed-loop human/machine performance across a range of primarily continuous control tasks (e.g., flight-path control). Building on this model, the dynamic decision-making model (DDM) of Pattipati, Kleinman & Ephrath (1983) was developed to account for discrete decision-making behavior in a generic multi-task supervisory control environment.

The first attempt to integrate continuous and discrete information processing with the structured procedural activities of the flight crew was with PROCRU, a model developed to evaluate commercial approach procedures during landing (Baron, Zacharias, Muralidharan & Lancraft, 1980; Milgram, van der Wijngaart, Veerbeek, Bleeker & Fokker, 1984). A number of approach simulations were conducted under different procedural assumptions, to support crew information-processing requirements, performance, and workload. An upgraded version was also developed to model the anti-aircraft artillery (AAA) crew for military applications, and to evaluate performance against targets using a range of defensive countermeasures (Zacharias, Baron &

Muralidharan, 1982). Here, the task of situation assessment was first made explicit, and linked to procedurally-driven rules of engagement.

Finally, an enhanced model known as the crew/system integration model (CSIM) has been used for the analysis of anti-aircraft artillery crews (Zacharias & Baron, 1981), fighter attack missions (Zacharias & Baron, 1982), and in a precursor effort to the current research (Zacharias *et al*, 1996). CSIM provides a greater structural formalization of many of the functions included in previous versions, and makes explicit some of the modeling requirements needed for realistic human/systems analysis.

Figure 2 illustrates CSIM's block diagram. The block diagram is separated into two major components: one dealing with the aircraft, and one with the pilot. The display/control module acts as the interface between the two model components.

The **display/control interface** drives the pilot's **sensory channels**. The vehicle state, in conjunction with the terrain features generated by the external world, drive the pilot's extra-cockpit visual channel, and provide him with map-based navigation information. The aircraft displays, driven by both vehicle dynamics and subsystems (e.g., the altimeter driven by the vehicle state, or a situation display driven by radar), feed both visual and auditory channels. There is also a tactile/vestibular channel, to account for non-visual and non-auditory cues picked up by the pilot such as rotational accelerations.

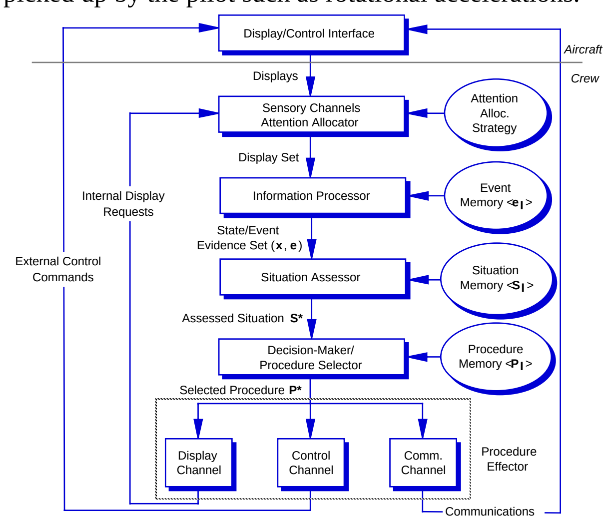

Figure 2: Crew/System Integration Model

The **attention allocator** accounts for the pilot's sensory limitations as well as for monitoring decisions on how to allocate attention among competing information sources. Visual sensory limitations are modeled in the same manner as in the OCM (Kleinman & Baron, 1971; Baron & Levison, 1977), except that the perceptual delay is neglected. An observation noise and a threshold are associated with each observed visual quantity. The thresholds are particularly important for external visual scene perception, for example, in visual target acquisition. Monitoring decisions reflect the fact that the pilot cannot process all sources of information simultaneously and must divide his attention (i.e., between the external world and within the cockpit; within the cockpit there is the question of which displays to fixate on, etc.).

The information processor submodel consists of two submodels, a continuous state estimator and a discrete event detector. The estimator can be identical to that used in the OCM, a Kalman filter designed to generate optimal estimates of the current vehicle/system state.

The outputs of the estimator are the estimate of the vehicle/system state, $\hat{\mathbf{x}}$ and the covariance of the estimation error, $\mathbf{\Sigma}$. Such states would include vehicle linear and angular velocities, position, and attitude, as well as the states of any targets or threats that might influence task execution. The estimator produces the state information needed for vehicle control, in addition to subjective probability estimates that can be used for event detection and situation assessment. The error covariance $\mathbf{\Sigma}$ also measures the pilot's uncertainty in the estimate $\hat{\mathbf{x}}$, and can be used to influence monitoring decisions.

The event detector generates occurrence probabilities of mission-relevant events, as perceived by the pilot on the basis of his dynamic information base. The event may be a failure, a request for action, a mission-related milestone, or some other discrete enunciated condition. The inputs to the event detector are visual and auditory discretes picked up by the display processor, and state estimator outputs.

The **situation assessor** block takes in the estimated states $\hat{\mathbf{x}}$ and the detected events **e**, and generates an assessed situation state **S***, which is a multi-dimensional vector defining the occurrence probabilities of the possible situations facing the pilot. For model tractability, a fixed and pre-defined set of candidate situations are assumed, determined solely by their task relevance. A situation defines an aggregated set of states, events, and possibly other situations which call for a given course of action or procedure execution. In simple terms, a situation is a predicate that must be satisfied to activate a procedure in a production rule system. Both the situation and production rule activities can be quite complex, depending on the modeled task.

The **decision maker** and **procedure selector** block takes in the assessed situation state **S***, and generates a selected task or procedure **P***, defined in the procedure memory shown. The definition of these procedures is an essential step in any task modeling effort, and it is important to note that the term procedure can apply to tasks in general; a procedure in these terms can have considerably more cognitive content than might normally be considered.

The selection and execution of a procedure results in an action or a sequence of actions. Three types of actions are modeled within CSIM: **control actions**, **display requests**, and **communications**. The control actions include continuous manual flight control inputs to the aircraft and discrete control settings. Display requests result from procedural requirements for specific information and, therefore, raise the attention allocated to the particular information source. Communications are verbal requests or responses as demanded by a procedure, and are modeled as the transfer of state, command, or event information.

3. Architecture and Implementation of Intelligent Agent Model

3.1 Overview

The SAMPLE architecture, shown in Figure 3, combines elements of the Rasmussen Hierarchy and the Crew/System Integration Model presented in the preceding section. It integrates the information processing/situation assessment/decision-making concepts of CSIM with the explicit delineation of skill based and rule-based behavior of the Rasmussen Hierarchy. Knowledge-based or problem-solving behavior is not currently supported explicitly, but it would be a straightforward extension of the existing architecture (although computational implementation of generic real-time planning or problem-solving algorithms would pose a considerable challenge).

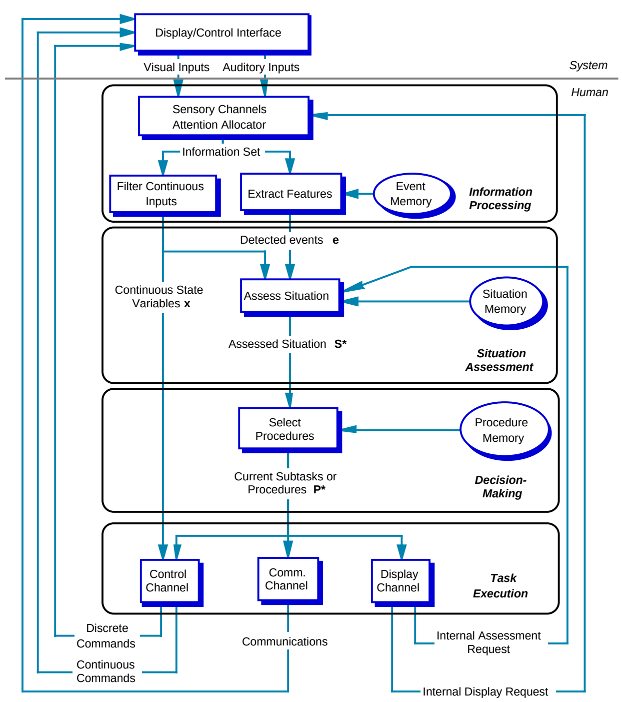

Figure 3: Agent Model Architecture

Two branches exist underneath the sensory channels and attention allocation block. The "shortcut" between continuous input filtering and the control channel serves the same function as the skill-based branch of the Rasmussen Hierarchy, while the path beginning with feature extraction and ending below the procedure selector is the equivalent of the rule-based branch. Like CSIM, a modeled behavior is that of directed attention allocation or situation assessment.

The integration of each of the key cognitive functions into a behavioral agent architecture called for the application of computational intelligence technologies. In this section the implementation approach used is described. Section 3.2 discusses the use of fuzzy logic for low-level event detection and information processing. Section 3.3 explains the application of belief networks for situation assessment modeling. Finally, section 3.4 describes the use of rule-based expert systems for modeling the final decision-making function.

3.2 Fuzzy Logic for Event Detection

The front end to the agent model's SA-based decision-making model is the information processing and event detection module, which transforms fused sensor data into situationally relevant semantic variables. These are the essential events that, as a group, define the overall situation facing the human operator.

SAMPLE uses fuzzy logic (FL) technology to implement event detection functionality. While some discrete event detection can be implemented using Boolean logic, most significant events required a more robust and flexible means of definition. This was achieved by the use of an FL-based event detector.

Fuzzy logic was proposed by Zadeh (1973) as a mathematical concept to deal with uncertainty in human decision-making. He was concerned with how humans can process imprecise non-numerical, or linguistic, information (i.e., *big*, *small*, *fast*, *heavy*, etc.) to perform a given task. He argued that if a human can perform complex tasks with this imprecise knowledge, then a machine would also benefit from such an approach.

Fuzzy reasoning employs the techniques of fuzzy logic in an IF-THEN rule base to perform a given task. A functional block diagram of fuzzy reasoning is shown in Figure 4. Fuzzy reasoning consists of three main blocks: a fuzzifier, a decision logic, and a defuzzifier.

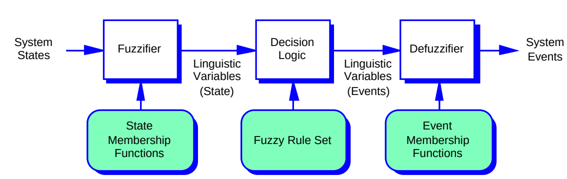

Figure 4: Functional Block Diagram for Event Detection using Fuzzy Logic

The fuzzifier acts on the system measurements and performs the mapping of deterministic numerical data (a crisp set) into fuzzy sets. Given a measurement value *x*, the fuzzifier interprets it as a fuzzy set A with membership function $\mu(x)$, where $\mu(x) \in [0,1]$. The fuzzification process involves the following steps: 1) determine the range of values of the measurement variables; 2) perform a transformation that maps the ranges of values of the measurements into the corresponding universes of discourse; and 3) perform the fuzzification function by transforming the measurement value into a suitable linguistic value which is viewed as labels of fuzzy sets.

After the fuzzification procedure, the fuzzy sets are processed via some decision logic consisting of linguistic rules. These rules are structured as follows:

IF X is A, THEN Y is B

where antecedent X is a measurement and consequence Y is the output. These rules thus relate the input measurements into the outputs. The mapping from the fuzzy set A into the fuzzy set B is called a fuzzy relation, and can be implemented via simple forward chaining algorithms.

The final element of the functional block diagram for fuzzy reasoning is the defuzzifier, which acts on the decision logic output variables or fuzzy control actions and performs the mapping to the corresponding deterministic numerical data (a crisp set). Three main techniques are employed to perform defuzzification: maximum membership, mean-of-maxima, and centroidal (Kosko, 1992).

3.3 Belief Networks for Situation Assessment

The second processing stage in SAMPLE is the situation assessor, which generates a high-level interpretation of the operational situation. The agent model design relies on *belief networks* (Pearl, 1988) for probabilistic reasoning in the presence of uncertainty, making use of computational models of situation assessment. These models emulate a skilled human's information fusion and reasoning process in a multi-task environment.

Any robust computational model of situation assessment requires a technology that has: 1) a capability to quantitatively represent the key SA concepts such as situations, events, and human mental models; 2) a mechanism to reflect both diagnostic and inferential reasoning; and 3) an ability to deal with various levels and types of uncertainties, since most real-world systems of any complexity involve uncertainty. Russell & Norvig (1995) cite three key reasons for this uncertainty:

1) **Theoretical ignorance**: All models of physical systems are necessarily approximations.
2) **Laziness**: Truly exceptionless rules require numerous antecedents and consequents and are therefore computationally intractable.
3) **Practical ignorance**: Even if all rules are known, there may not be enough time to measure all properties of the particular objects that must be reasoned over.

Belief networks (Pearl, 1988) offer considerable potential for meeting these requirements and modeling SA behavior. The principal advantages of belief networks over other uncertain reasoning methods are:

1) Its probability estimates are guaranteed to be consistent with probability theory. This stems from its Bayesian update procedure's strict derivation from the axioms of probability.
2) It is computationally tractable. Its efficiency stems principally from exploitation of conditional independence relationships over the domain.
3) The structure of a BN captures the cause-effect relationships that exist among the variables of the domain. The ease of causal interpretation in BN models typically makes them easier to construct, thus minimizing the knowledge engineering costs (Tversky & Kahneman, 1981) report that expert physicians prefer to assess disease-symptom probabilities in terms of causal reasoning rather than diagnostic reasoning), and easier to modify (Henrion, 1989).
4) The BN formalism supports many reasoning modes: causal reasoning from causes to effects, diagnostic reasoning from effects to causes, mixed causal and diagnostic reasoning, and intercausal reasoning. Intercausal reasoning refers to the situation in which a model contains two potential causes for a given effect. If we gain evidence that one of the possible causes is very likely, this reduces the likelihood of the other cause. Russell & Norvig (1995) assert that no other uncertain reasoning formalism supports this range of reasoning modes.

Belief networks provide the capability and flexibility of modeling situation awareness (SA) with its full richness without arbitrary restrictions. They provide a comprehensible picture of the SA problem by indicating the dependent relationships among the variables, both high-level (symbolic) and low-level (numeric), relevant to the SA problem. This provides a clearer view (than a low-level neural network-based approach would, for example) of how each individual piece of evidence affects the high-level situation characterization. They allow the incremental addition of evidence, concerning any of the domain variables, as it arrives, thus allowing for real-time SA update. The method is consistent with probability. Finally, the graphical representation in which the global CPT is decomposed into a set of smaller local CPTs each describing the interactions of a small subset of the domain variables renders probabilistic inferencing tractable, even for large domains such as tactical situation assessment.

3.4 Expert Systems for Decision-Making

Human decision-making is modeled using a cascade of two sub-models: a procedure selector and a procedure executor. In tandem, these emulate a human's rule-based decision-making behavior and psychomotor skills in executing a selected procedure. Decision-making behavior was implemented as a production rule system, with a general structure given by:

If (Situation = S_i) then (Procedure Set = P_i)

This supports selection of a procedure set that is pre-assigned to the assessed situation specified in the human operator's mental model. Details of the procedure set, and its linkages to the associated situation, are maintained in a procedural knowledge base.

A procedure set may contain one or more sub-procedures. After a procedure set is selected, each firing of a sub-procedure is determined using a forward chaining production system in response to event and state evolution.

3.5 Modeling of Inter-Agent Communications

An important component of any multi-agent system is the implementation of inter-agent communication. SAMPLE makes use of the concepts underlying the Command and Control Simulation Interface Language (CCSIL) as a framework for modeling inter-agent communications. CCSIL is a special language that has been developed by the DARPA Command Forces Project to model communication among entities in the Distributed Interactive Simulation (DIS) Environment (Salisbury, 1995). It is a component of the Command Forces (CFOR) Simulation (Calpin, 1998), which incorporates explicit modeling of battlefield command and control into virtual simulation. CCSIL includes a set of messages and a vocabulary of military terms to fill out those messages. CCSIL enables communications modeling via implementation of a message packet that comprises three fields:

- The sender, which is the alphanumeric representation of the sending unit.
- The destination, which is a list of one or more unit names of the intended message recipients.
- The message definition, which is a structured data packet that contains a well-defined set of data encoding the message's contents.

The developers of CCSIL have developed message content definitions for land, sea, and air operations for battlefield communications. The general philosophy underlying CCSIL's design was used for overall modeling of inter-agent communication, and its specific message packet definitions were used for the particular application described in the following section.

Although more generic approaches to inter-agent communication do exist (e.g., the Knowledge Query Manipulation Language, or KQML (Finn, Labrou, and Peng, 1998)), it was decided that the domain-specific needs of modeling air combat communications were met more readily by a customized approach such as that used here. This will also facilitate integration of SAMPLE into other military simulation models that use CCSIL for inter-agent communication.

4. MIL-AASPEM Integration and Demonstration

4.1 Development Scenario

The SAMPLE agents were developed to model pilot behavior in a *fighter sweep* mission, in which a four-ship formation of tactical aircraft deploys to clear the air of enemy fighters prior to other air missions entering the engagement area. Any airborne targets meeting a restrictive geometric commit criteria are engaged.

Figure 5 illustrates the general progression of the sweep mission. Five key phases were identified for this overall mission, as follows.

1) Ingress (Initial transit towards battle area)
2) Detection (Identification and assessment of detected threats)
3) Engagement (Tactical intercept leading to weapons employment)
4) Defensive Reaction (May occur at any time)
5) Egress (Exit back to friendly territory)

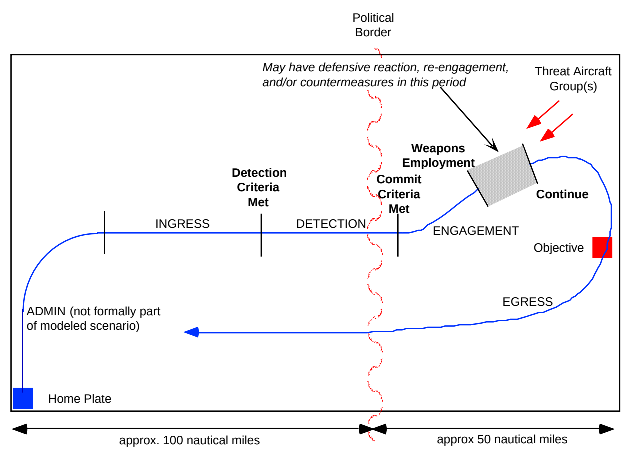

Figure 5: Sweep Mission Overview

The transitions between one mission phase and the next are event-driven rather than being constrained geographically or temporally, so they may occur at any time. The number of threat aircraft that may be encountered is unknown, and it is possible that some missions may not result in an engagement.

As the blue flight transits towards the engagement area, they build a shared picture of the tactical situation ahead through radio communications. This helps to ensure that all pilots have a common picture of the threat situation. The flight leader's decision matrix includes threat detection with onboard sensors, commit decision, positive identification, tactical intercept selection, threat group and within group sorting for flight members, hostile declaration and weapons employment. Note that only the flight leader can declare the decision to commit on a particular threat group; that decision is communicated to the other flight members via inter-ship communications. After weapons employment, all flight members monitor weapons success and re-engage as necessary. Post-engagement flight would continue running through decision matrices until the flight returns home.

4.2 Mental Model Development

The information processing, situation assessment, and decision-making models required to implement the SAMPLE agents were developed via an extensive knowledge engineering and cognitive task analysis (CTA) effort conducted with experienced USAF pilots. CTA determines the mental processes and skills required to perform a task at high proficiency levels (Redding 1992). The traditional method for task analysis is to decompose it into subtasks, skills, and knowledge. CTA improves the process by including the analysis of human thought processes during task performance. In this study, the objective was to identify the following:

- What activities, procedures, and tasks do the pilots perform during the sweep mission?
- What decisions support the performance of those tasks?
- What situation assessments support those decisions?
- What information is needed to develop these situation assessments?
- What is the nature of inter-ship communications over the course of the mission? How can they be modeled using the structured data packet representation described here?

A particular challenge in this effort was the development of a broad range of event detection and belief network models for tactical situation assessment. The approach taken here followed that prescribed by Russell & Norvig (1995):

1) *Decide what to talk about***.** Choose which factors will be modeled explicitly, and which will just be summarized by probability statements.
2) *Decide on a vocabulary of random variables*. Determine the variables to be used , and what values they can take on (e.g., target range may be *long*, *medium*, or *short*). For the purpose of fuzzification, it is useful to quantize continuous-valued variables into discrete ranges.
3) *Encode general knowledge about the dependence between variables*. This contains both a qualitative and quantitative part. During the former, general dependencies between variables are identified. During the latter, the specific probability values (i.e., CPTs) are identified. These values may come from the subject matter expert's experience, from measurements of frequencies in a database of past experiences, or some combination of the two.
4) *Encode a description of the specific problem instance, and pose queries to the inference procedure*. The most common query is to obtain the value of some hypothesis variable. For example, what is the risk posed by some hypothetical tactical situation? It is also common to use sensitivity analysis to determine the robustness of such answers with respect to perturbations in the CPT values.

The knowledge engineering effort yielded a network representation of the pilot's mental model for the sweep mission. This network consists of a large set of interconnected nodes, each representing a different event detection/situation assessment/decision-making function. Each SAMPLE agent instantiation is implemented as a feed-forward network, driven by a well-defined set of situational and mission context inputs.

Figure 6 illustrates one of the belief networks developed for situation assessment. Shown is the BN for computing mental posture, which quantifies the extent to which each pilot agent is in an *offensive* or *defensive* state of mind. Posture will drive decisions such as whether or not to abort an engagement, or the selection of an appropriate defensive reaction. This diagram shows two independent BNs linked via an algorithmic processing block that is essentially a lookup table followed by arithmetic computations. This assessment depends directly on three quantities:

- The current mission phase (ingress, detection, engagement, or egress)
- The level of confidence on whether or not the ownship has been targeted by enemy radar (high, medium, or low)
- The pilot's confidence that he has mutual support from his wingman (high, medium, or low)
- The pilot's offensive status (whether he has already taken a high p_k shot or a low p_k shot)

The pilot's belief as to whether he has been targeted depends on two pieces of evidence:

- Whether the radar warning receiver (RWR) detects any radar signals, and if so, their operating mode (none, search, or track)
- The ownship's location with respect to a potentially hostile aircraft's weapons employment zone, or WEZ (in or out)

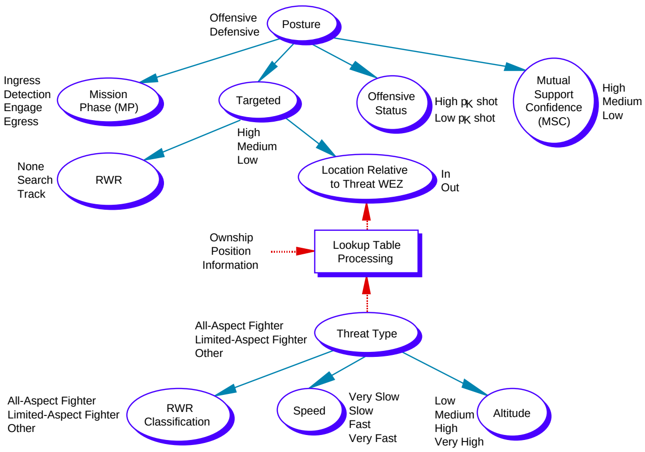

Figure 6: Example Belief Network: Mental Posture Assessment

Whether or not the pilot believes he is in a hostile aircraft's WEZ depends on two factors: the type of hostile aircraft (which can be used to estimate its WEZ), and the perceived relative location of the ownship with respect to the hostile aircraft. This assessment cannot be represented directly with a belief network, because it is not a probabilistic inferencing process. The block labeled Lookup Table Processing accepts two inputs: a hypothesis on the threat type (which is generated using a belief network), and information regarding ownship position relative to the threat aircraft. Lookup table operations are performed to determine the spatial extent of the WEZ based on the aircraft type, followed by geometric computations to determine whether the ownship is in or out of the WEZ.

The threat type hypothesis, which may be an all-aspect fighter, a limited-aspect fighter, or some other vehicle (e.g., a transport or bomber), depends on three pieces of evidence:

- The classification generated by the RWR (all-aspect fighter, limited-aspect fighter, or other)
- The vehicle's speed [very slow (less than 100 kt), slow (between 100 and 200 kt), fast (between 200 kt and Mach 1), or very fast (greater than Mach 1)]

The vehicle's altitude [low (less than 5,000 ft), medium (between 5,000 ft and 25,000 ft), high (between 25,000 ft and 50,000 ft), or very high (greater than 50,000 ft)]

These three pieces of evidence drive the threat type hypothesis, which in turn drives the (non-BN) computation of ownship's location relative to the WEZ.

Table 1 shows a representative example of SAMPLE's pilot decision matrices. Shown is the decision matrix for the type of tactical intercept to perform on a threat group. This decision is a function of the presence and arrangement of multiple threat groups, the number of threats, and constraints posed by the mission's rules of engagement (which are specified *a priori*).

Table 1: Decision Matrix for Tactical Intercept Selection

Threat Type	Multiple Group Arrangement	Number of Threats	Rules of Engagement Constraints	Intercept Tactic
All-Aspect Fighter		1 or 2 with certainty		Straight-in w/ pincer following missile shot
		≥3 or indeter-minate		Single-side offset
			Require Visual Identification of Target	Lead/trail
			Allow BVR Shot with Fox 3 ordnance	Lead/trail
Limited-Aspect Fighter or Other	Azimuth Presentation			Single-side offset
	Otherwise			Straight-in to pince, depending on spacing (i.e., larger bracket for greater spacings)

The main purpose of the intercept tactic is to maneuver to a point where the threat aircraft are within the WEZ, while staying outside the boundaries of the threat's WEZ. Figure 7 illustrates this concept, in the case where both aircraft have a WEZ of approximately the same size. The longest range at which either aircraft can launch its missiles such that its opponent is inside the WEZ is called the first launch opportunity (FLO). In the figure, the blue aircraft (on the left) has maneuvered such that the bandit is within its WEZ, but the blue aircraft is not within the bandit's WEZ. Blue therefore has FLO.

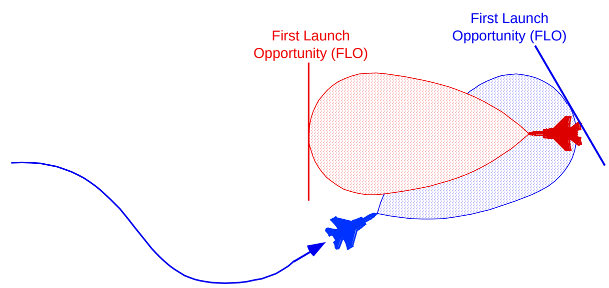

Figure 7: Weapons Employment Zones and First Launch Opportunity

Tables 2 and 3 show examples of the structured data packets developed to describe inter-agent communication, as discussed earlier in section 3.5. Table 2 shows the message definition packet for an *engagement advisory* message, which is used by a pilot agent to communicate some activity (firing missiles, dropping a target, etc.) in relation to a specific air entity that is being tracked or monitored. The *engagement keyword* field describes the activity, the *contact description* is an alphanumeric designator for the contact of interest, and the *contact location* is another structured data packet nested within this one. Table 3 shows its definition, which describes a contact's location in terms of its compass bearing, together with a group of optional variables that may be used to add additional details on the contact's location and motion.

Table 2: Message Packet for Engagement Advisory

Field Name	Description	Data Type	Allowed Values
Engagement Keyword	Type of Advisory	Enumerated	NOTCH ABORT SPIKED SPLIT FLOATING LOCKED FOX1 FOX2 FOX3 KILL DROP_BEAMER DROP_DRAGGER SORTED
Contact Description	Label used to identify the contact uniquely.	Alphanumeric	
Contact location	Location of the contact that is the subject of the message, or the direction in which a maneuver is being performed	Composite (Air Location Structure)	As defined in Table 3

The SAMPLE software was implemented using object-oriented C^{++} on a Unix workstation. Fuzzy logic reasoning is performed using software derived from the Matlab Fuzzy Logic Toolbox, while belief network modeling is implemented using a stand-alone C^{++} class. All expert system functions are performed using the CLIPS expert system shell (Giarratano, 1998).

Table 3: Message Packet for Air Location Structure

Field Name	Description	Data Type
Bearing	Bearing to contact	Bearing type
Range	Range to contact in nautical miles	Floating point
Altitude	Altitude of contact in ft	Floating point
Speed	Speed of contact in knots	Floating point
Heading	Heading of contact	Bearing
Aspect	Aspect angle of contact	Bearing
Bullseye	Location from which bearing, range, and heading or aspect are being calculated. If none specified, values are assumed to relative to message recipient.	Alphanumeric
Climb Rate	Rate of altitude change in ft/min	Floating point
Turn Rate	Rate of turn in deg/min	Floating point

4.3 Simulation Results

The results of a set of MIL-AASPEM-based simulation trials conducted to evaluate SAMPLE's performance in the sweep mission are now discussed. In all scenarios, the blue flight (which used the SAMPLE agents) was initialized at an altitude of approximately 25,000 ft, heading east towards the engagement area. Two 2v2 scenarios are presented. In both cases, threat team (red) behavior was driven by MIL-AASPEM's original PDL. The SAMPLE agents use the inter-agent communications mechanism described earlier in section 3.5 to support the development of shared situation awareness, and to allow the flight leader to direct the other flight members to engage the hostile threats.

4.3.1 2v2 Scenario, Threat Group Initially in CAP

Figure 8 illustrates the evolution of a scenario in which a 2-ship blue flight (coming from the left) engages a 2-ship red group. The red group is initially in a combat air patrol (CAP) orbit, which it breaks to engage the incoming blue flight. In this scenario and those that follow, only missile launches resulting in a kill are shown, unless a re-engagement is required.

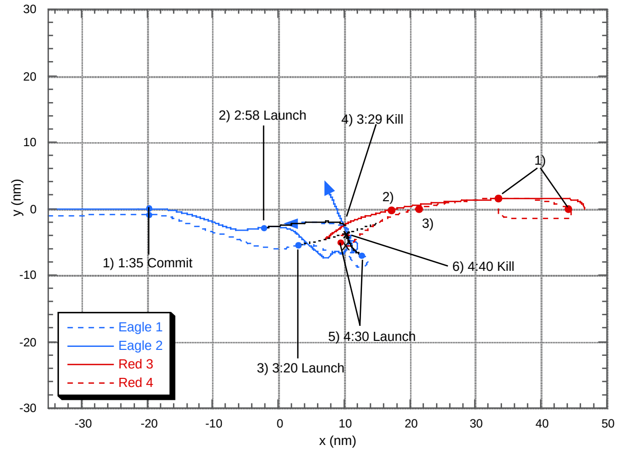

Figure 8: 2v2 CAP Scenario Evolution

In this scenario, the blue flight commits on the threat group at t = 1:35 (time index 1 in the figure), at which point the centroid of the threat group is 60 nm away. The commit means that the flight leader decides to engage on the incoming threats, and communicates this decision to his wingman. The red circles shown along the threat trajectories at time index 1 indicate their location at the instant that the blue flight entered the engagement phase. As the red CAP breaks its orbit, it assumes a range-split formation (i.e., one aircraft is ahead of the other, relative to their direction of flight). As such, the blue flight performs a *range sort* during the engagement: Eagle 2 (i.e., blue flight member #2) sorts on the frontmost threat, while Eagle 1 sorts on the trailer. Since the player ratio is 1 and the threats are all-aspect fighters, the blue flight performs a single-side offset intercept (SSO) (see Table 1). The goal of an SSO intercept is to maneuver around a threat's WEZ while seeking to place the threat in the ownship's WEZ, as shown earlier in Figure 7. Because Eagle 1 is engaging on the trailer, it makes a wider turn during the offset to avoid the red leader's weapons employment zone (WEZ).

Eagle 2 launches its missiles at t = 2:58 (time index 2), and then performs an F-pole (the standard post-launch maneuver following the release of radar-guided missiles). The objective of an F-pole is to steer the aircraft in such a way as to keep the threat just on the edge of radar coverage. The solid black line emanating from Eagle 2's location at 2:38 shows the trajectory of the first missile that hit the target (at t = 3:29). It then continues F-poling to avoid missiles launched by Red 3 (not shown in the figure). At t = 3:20, Eagle 1 launches its missiles on Red 4, but must re-engage because the threat evades the missiles (the dotted lines emanating from Eagle 1 at 3:20 show the trajectory of the evaded missile). Both Eagle 1 and Eagle 2 fire two missiles each at their target, in accordance with the relevant decision matrix for weapons quantity selection. At *t* = 4:30, Eagle 1 fires a second round of missiles at Red 4, which hit the target at *t* = 4:40. Both blue players then disengage and head home base.

4.3.2 2v2 Scenario, Threat Group in Azimuth-Split Formation

Figure 9 shows the evolution of a 2v2 scenario in which the threat group is initially in an azimuth-split formation (abreast with respect to each other, relative to the direction of flight), heading west to engage the blue flight. The blue flight commits on the threat group at *t* = 0:34, at which point the threat group centroid is just under 60 nm away. The blue flight performs an *azimuth sort* on the threats, using a single-side offset intercept. Eagle 1 chooses the sort logic and intercept type, and communicates that to his wingman. Like the previous scenario, a single-side offset is used because the player ratio is 1.0 and the threats are all-aspect fighters. Both blue players launch their missiles between *t* = 2:05 and *t* = 2:08, and then perform an f-pole maneuver to the right (defensive reactions against incoming missiles take place during this time). Both Eagle 1 and Eagle 2 hit their target on the first attempt, with Eagle 1 making a kill at *t* = 2:37 and Eagle 2 at *t* = 2:40. The blue flight then disengages and egresses towards home base.

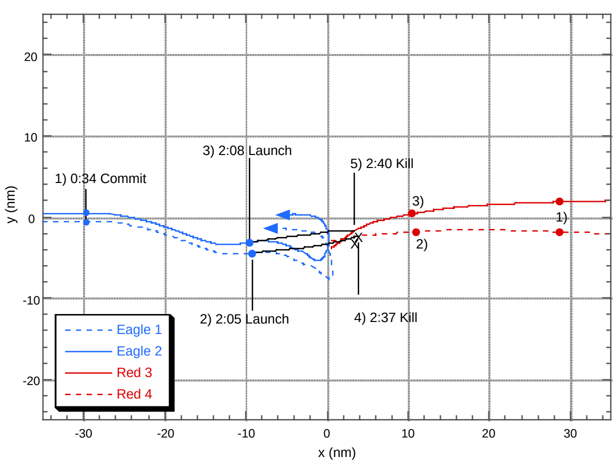

Figure 9: 2v2 Scenario Evolution

5. Summary

The design and development of an intelligent agent model for human behavioral representation (HBR) in real-time simulation has been described. This agent model uses a probabilistic representation of human information processing and situation assessment as the foundation for modeling complex decision-making behavior in multi-task environments. The simulation results demonstrate the feasibility and effectiveness of using such SA-centered decision-making for modeling tactical piloting behavior. The tactics employed by the SAMPLE players were consistent with the results of the knowledge engineering exercise that supported this study. SAMPLE demonstrated the ability to set up and execute intercepts, perform defensive reactions, make kill assessments, and re-engage when necessary. The inter-agent communications mechanism facilitated the development of shared situation awareness, as well as providing a command mechanism for the flight leader to direct the behavior of other flight members.

These results lay the foundation for the development of a full-scope SA-based model of tactical piloting behavior across a range of mission scenarios, and the deployment of the agent model into other operational domains. Work is underway to apply the generic agent architecture for modeling dismounted infantry decision-making in hostile urban threat environments, and commercial pilot decision-making under distributed air/ground traffic management (Harper *et al*, 1998).

6. Acknowledgments

This research was funded under Contract F33657-96-C-2066 from the Aeronautical Systems Center at Wright-Patterson Air Force Base. The authors are grateful to Col. Steve Rapp and Major Art Crain of USAF for providing their subject matter expertise on air combat to support this effort.

7. References

Adams, M., Tenney, Y., & Pew, R. (1995). “Situation Awareness and the Cognitive Management of Complex Systems.” Human Factors, 37(1), 85-104.

Baron, S., Zacharias, G., Muralidharan, R., & Lancraft, R. (1980). “PROCRU: A model for analyzing flight crew procedures in approach to landing,” *Proceedings of the Eigth IFAC World Congress*, Tokyo, Japan.

Calpin, J. (1998). “Performance of the DARPA Command Forces Program in the STOW97 ACTD,” *Proceedings of the Eighth Conference on Computer Generated Forces and Behavioral Representation*, Orlando, FL.

Deckert, J., Entin, E.B., Entin, E.E., MacMillan, J. & Serfaty, D. (1994). “Military Command Decisionmaking Expertise,” Report 631. Alphatech Inc., Burlington, MA.

Druzdzel, M. (1996). “Five Useful properties of Probablistic Knowledge Representations from the Point of View of Intelligent Systems.” *Fundaments Informaticae*, (Special issue on Knowledge Representation and Machine Learning).

Endsley, M. (1993). “A Survey of Situation Awareness Requirements in Air-to-Air Combat Fighters.” *Int’l Journal of Aviation Psychology*, Vol. 3: 157-168.

Endsley, M. (1995). “Toward a Theory of Situation Awareness in Dynamic Systems.” *Human Factors*, 35(1): 32-45.

Fallesen, J., Carter, C., Perkins, M., Micehl, R., Flannagan, J., & McKeown, P. (1992). *The Effects of Procedural Structure and Computer Support upon*

Selecting a Tactical Course of Action (Tech. Report 960): US Army Research Institute for Behavioral and Social Sciences.

Finin, T., Labrou, Y., & Peng, Y. (1998). "Mobile Agents Can Benefit from Standards Efforts on Interagent Communication." *IEEE Communications Magazine* (July).

Fracker, M. (1990). "Attention Gradients in Situation Awareness." *Situational Awareness in Aerospace Operations AGARD-CP-478*. Neuilly-Sur-Siene, France. Advisory Group for Aerospace Research and Development.

Giarratano, J.C. (1998). *CLIPS User's Guide Version 6.10*, NASA Johnson Space Center.

Harper, K., Mulgund, S., Zacharias, G., and Kuchar, J. (1998). "Agent-Based Performance Assessment Tool for General Aviation Operations under Free Flight." *Proceedings of the 1998 Guidance, Navigation, and Control Conference*, Boston, MA (August).

Henrion, M. (1989). "Some Practical Issues in Constructing Belief Networks." In Kanal, Levitt, & Lemmer (Eds.), *Uncertainty in Artificial Intelligence 3*, (pp. 161-173). North Holland: Elsevier Science Publishers.

Klein, G. (1989). "Recognition-Primed Decisions." *Advances in Man-Machine Systems Research*, Vol. 5: 47-92.

Klein, G. (1994). "Naturalistic Decision Making: Implications for Design." *Gateway*, 5(1): 6-7.

Kleinman, D., Baron, S., & Levison, W. (1971). "A control theoretic approach to manned-vehicle systems analysis." *IEEE Transactions on Automatic Control*. No. 61, pp. 824-832.

Lawson, K.E. & Butler, C.T. (1995). *Overview of Man-in-the-Loop Air-to-Air System Performance Evaluation Model (MIL-AASPEM) II*. Boeing Company. D658-10485-1 (September).

Kosko, B. (1992). *Neural Networks and Fuzzy Systems*. Englewood Cliffs, NJ: Prentice Hall.

Milgram, P., van der Winngaart, H., Verbeek, O., & Fokker, B. (1984). "Multi-crew model analytic assessment of landing performance and decision making demands," *Proceedings of the 20th Annual Conference on Manual Control*, Moffett Field, CA: NASA Ames Research Center.

Pattipati, K., Ephrath, A. & Kleinman, D. (1979). "Analysis of Human Decision-Making in Multitask Environments." Department of Electrical Engineering, University of Connecticut. UCONN EECS-TR-79-15.

Pattipati, K., Kleinman, D. and Ephrath, A. (1983). "A Dynamic Decision Model of Human Task Selection Performance." *IEEE Trans. on Systems, Man & Cybernetics*, 13(3): 145-156.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann.

Pew, R. & Mavor, A. (Eds) (1998). *Modeling Human and Organizational Behavior: Application to Military Simulations*. Washington, D.C.: Nat'l Academy Press.

Rasmussen, J. (1981). "Models of Mental Strategies in Process Plant Diagnosis." *Human Detection and Diagnosis of System Failures*. J. Rasmussen and W.B. Rouse, ed. Plenum.

Rasmussen, J. (1982). "Human Errors: A Taxonomy for Describing Human Malfunctions in Industrial Installations." *Journal of Occupational Accidents*, Vol. 4: 311-335.

Redding, R. (1992). *A standard procedure for conducting cognitive task analysis*. McLean, VA: Human Technology Inc.

Russell, P. & Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Englewood Cliffs, NJ: Prentice Hall.

Salisbury, M. (1995). "Command and Control Simulation Interface Language (CCSIL): Status Update," *Proceedings of the Twelfth Workshop on Standards for the Interoperability of Defense Simulations*, pp. 639-649, Orlando, FL (March).

Shaw, R. (1988). *Fighter Combat: Tactics and Maneuvering*, Annapolis, MD: Naval Institute Press.

Stiffler, D. (1988). "Graduate Level Situation Awareness." *USAF Fighter Weapons Review*.

Tolcott, M., Marvin, F., & Lehner, P. (1989). "Expert decisionmaking in evolving situations," *IEEE Transactions on Systems, Man, and Cybernetics*. 19(3).

Tversky A., & Kahneman (1981). "The Framing of decisions and the psychology of choice," *Science* (211), pp. 453-458.

Zacharias, G., & Baron, S. (1982). *A Proposed Crew-Centered Analysis Methodology for Fighter/Attack Missions* (4866): Bolt Beranek and Newman Inc.

Zacharias, G., Baron, S. & Muralidharan, R. (1981). "A Supervisory Control Model of the AAA Crew," *Proceedings of the 17th Conference on Manual Control*, Los Angeles, CA.

Zacharias, G., Miao, A., Illgen, C., Yara, J. & Siouris, G. (1996). "SAMPLE: Situation Awareness Model for Pilot in-the-Loop Evaluation," *Proceedings of the First Annual Symposium on Situational Awareness in the Tactical Air Environment*, Patuxent River, MD (June).

Zadeh, L. (1973). "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes." *IEEE Transactions on Systems, Man and Cybernetics*, 3(1): 28-44.