

1.1

Adversarial Models for Opponent Intent Inferencing

Eugene Santos Jr. and Qunhua Zhao

CONTENTS

1.1.1	Intent Inferencing.....	2
1.1.2	Representing and Reasoning under Uncertainty	5
1.1.3	Adversary Intent Inferencing Model (AII)	7
1.1.3.1	Architecture of AII Model.....	8
1.1.3.2	Model Construction	9
1.1.3.3	Reasoning over the Model.....	11
1.1.3.4	AII and Wargaming	12
1.1.4	Conclusions.....	19
	Acknowledgments	20
	References	20

Taking into account the characteristics and behaviors of one's adversary is essential for success in any competitive activity, such as in sports, business, or warfare. Obviously, if one's enemies are well understood, their actions can then be better anticipated and countered. To do so, the key is to capture the adversary's intentions. An intuitive approach that immediately comes to mind is to think what you would do if you were "in your opponent's shoes." However, "thinking as your enemy" is difficult because your perception of the world is quite likely to be very different from your opponents'. To address this problem and correctly infer adversary intent, the model of the adversary should capture critical aspects such as their history of movements and responses in different situations, their policies (e.g., military doctrines), capabilities, infrastructure, and human factors (e.g., social, political, cultural, and economic*). In this chapter, we focus primarily on adversary modeling (and, in particular,

* We only briefly touch upon human factors in this chapter. Chapter 1.2 provides a more detailed discussion of such factors.

adversary intent inferencing) for military planning and operations, but note that the concepts can be readily applied across a broad range of domains.

To successfully model the adversary, not only the opponent's capabilities but also the intent should be considered. The adversary's intent is composed of the adversary's desired end-states, reasons for pursuing such end-states, methods to achieve the goals, and the levels of commitment to achieving the goals. In this chapter, we present a comprehensive adversarial modeling approach that accounts for opponent intent inferencing in a dynamic and interactive environment. The model has been applied to military wargaming, resulting in an action-reaction-counteraction simulation environment where actions are initiative events (i.e., the offense from one side), reactions are the responses from the other side, and counteractions are the first side's responses to the reactions. In such a simulation environment, the sequence of action-reaction-counteraction is continued until the critical event is completed or the commander determines to use another *course of action* (COA) to accomplish the mission [20]. Thus, the model helped break the barrier of prescribed adversaries in wargaming, i.e., those that act in a predetermined fashion. Furthermore, we demonstrated that models of adversaries can be readily constructed and modified in real-time during a simulation to reflect the dynamic battlefield.

We begin with a discussion of intent inferencing and adversary intent inferencing, concentrating on providing a brief overview of concepts and definitions. We then present some background on representing and reasoning over noisy, incomplete, and uncertain information that is at the heart of modeling intent prediction, explanation, and understanding. Next, we describe our adversary intent inferencing framework, the semantics for building adversary models, and the inferencing process. With the framework in place, we describe its application to conflict analysis and wargaming. In particular, we provide details on some of the testbeds used in our experiments for inferring adversary intent. Finally, we present our thoughts and conclusions.

1.1.1 Intent Inferencing

In Bratman's belief-desire-intention (BDI) model [5,6], the *intentions* are viewed as partial plans committed by an intelligent entity to achieve certain goals (*desires*) based on the perception or knowledge of the world (*beliefs*). The *intentions* in the BDI model are also understood as a subset of desires upon which an entity chooses to act. In Geddes' view [11], intent inferencing involves deducing an entity's goals (desired end-states) based on observations of its actions. Such deduction typically involves the construction of one or more behavioral models that are optimized to represent the entity's behavior patterns. After data and knowledge representing observations of an entity, its actions, and its environment (collectively called *observables*) are

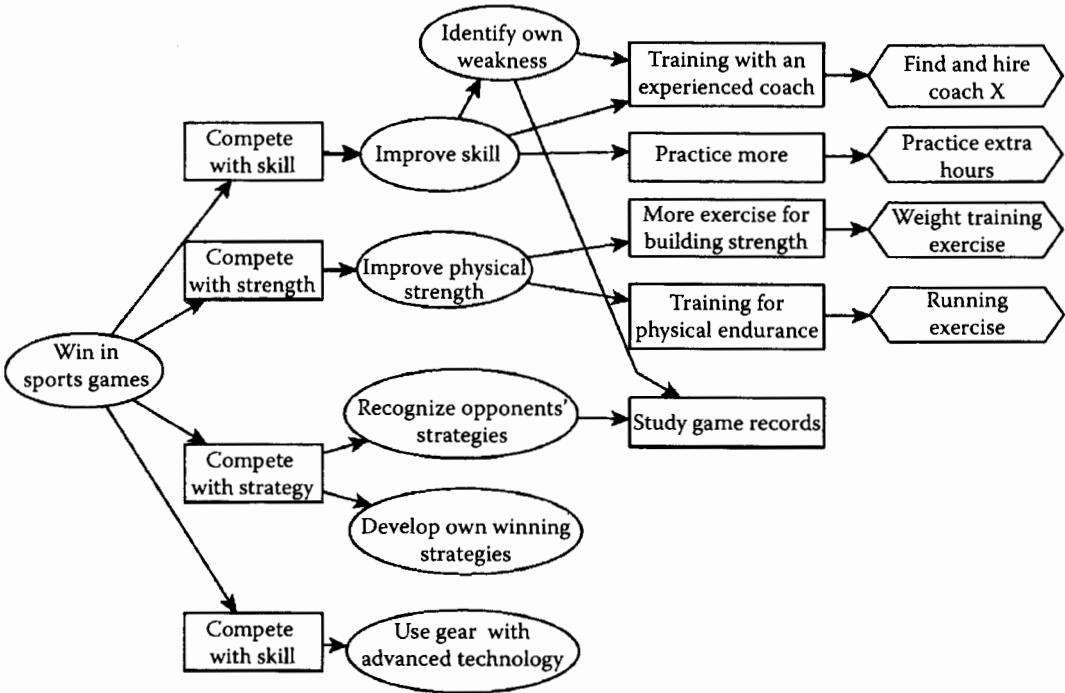


FIGURE 1.1.1

An example PGG for winning strategies in sports games.

gathered, these models attempt to match the observables against behavior patterns and derive *inferred intent* from those patterns [1,29].

Various methodologies have been developed for building computational models for intent inferencing. A plan-goal-graph (PGG) is a network of plans and goals where each high-level goal is decomposed into a set of plans for achieving it, and the plans are decomposed into subgoals that, in turn, are decomposed into lower-level plans [12]. Thus, the hierarchical structure of PGG is composed of alternating plans and goals. Intent inferencing is thus the process of searching for a path in the PGG from an observed action to a plan or goal. Figure 1.1.1 shows a PGG of an athlete's effort to win in sports games where ovals are goals and subgoals, rectangles are plans, and diamonds are the actions. As shown in the figure, the observation of a "running exercise" (action) can be traced back to the plan for training to improve physical endurance and further linked to the goal of improving physical condition, and so on.

In another approach, the operator function model (OFM), an expert system uses a heterarchic-hierarchic network of finite-state automata for intent inferencing. In the network, nodes represent an entity's activities and arcs represent conditions that initiate/terminate certain activities [7,9,27]. From the top-most level to the lower-most levels, high-level activities (major operation functions) are hierarchically decomposed into subactivities, then tasks, and finally, actions. Each node has inputs and outputs: Inputs to a node are

outputs from higher-level nodes or exogenous events, in the case of the top network level (heterarchic level) [27]. For inferencing, hierarchical structures of goal-plan-task-action are derived from the OFM, and the inferencing process maps an observed operator action onto model-derived tasks that, in turn, support model-derived plans. As operator actions are performed, the system tries to connect each action to one or more appropriate activity trees [9,27]. A third approach is generalized plan recognition [8,17,22], which tries to recognize the entity's plan for carrying out the task based on observations, an exhaustive set of discrete actions (a plan library), and constraints. Generally when an action from a user is observed, the system tries to search and match the observation against the plan library and then choose the preferred plan or extend the currently identified plan according to given knowledge regarding typical user plans. Lastly, the Soar model [13] is based on operators (plans) and states (beliefs). Operators are selected based on preconditions (pattern matching) and executed based on the current state. Selecting high-level operators leads to certain related subgoals. In Soar, a selected operator constrains the option of new operators that the software agent intends to consider, which constrains the search space of the problem. This methodology is compatible with Bratman's insights in the BDI model [13].

Although intent inferencing has been applied to different fields, the usual goal is to infer the user's intention from monitoring the individual's interactions with the system and then anticipate the user's future actions to provide proactive support to them. An example is the application of a user model for predicting the intelligence analyst's intent during information retrieval and assisting him/her to achieve better results [32]. This intent inferencing model consists of three formative components: The first, *interests and foci*, captures the user's key interest in the process of retrieving the information. The second, *actions and preferences*, describes the activities that can be used to carry out the user's goals with an emphasis on how the user tends to carry them out. The third, *knowledge and reasoning*, provides insight into the deeper motivations behind the goals upon which the user is focused and illuminates connections among the goals. Assume that an intelligence analyst is seeking information on weapons of mass destruction (WMD) in a certain geographic region and that they issue a query about "facilities for WMDs in...". After receiving the query, the intent inferencing model tries to recall prior information about the user's searching behavior and discovers three items: (1) The user has been focusing in the biological warfare domain (which has been identified in the interests set); (2) that biological weapons are a type of WMD (from the knowledge of reasoning component); and (3) this analyst tends to quickly narrow down his query so as to search for detailed information (searching style, found in the actions and preferences component). The system then modifies the user query to reflect the analyst's interests and searching style and subsequently presents information, appropriately ranked, concerning biological weapon facilities. The analyst goes through the returned documents and indicates which are relevant. This relevancy information is then fed back into the system for the user intent

inferencing model to update itself [32,33]. Besides the intelligence domain, this user modeling approach has also been applied to the medical domain [4,23,31]. Many other applications of intent inferencing exist, including recommender systems [16,36], tutoring systems [3,9], and team intent identification [10].

We are interested in inferring the intent of a special entity, the adversary. As we shall see, many of the same underlying principles for intent inferencing hold for adversary intent inferencing, where observed adversarial actions are used to deduce the goals or plans that those actions try to achieve or carry out. Adversary intent inferencing starts by collecting information (observables) regarding the adversary from different sources, such as sensors and intelligence sources. The next step is to infer adversary intentions and goals with regards to given adversary perceptions and beliefs. Finally, the inferencing process can predict the adversary's COA. However, note that the adversary intent inferencing model should be constructed by taking into account the adversary's perception of the world, how they view the situation, what they believe about themselves and their opponents, what their desired end-states are, and what they can and intend to accomplish.

In this chapter, the discussion on adversarial modeling focuses primarily on intent inferencing. Alternative approaches exist not based on intent that have been applied to other domains and applications. For example, game-tree search approaches have been successful in games such as chess and cards. More discussion on these other approaches can be found in Chapter 3.1, Chapter 3.2, and Chapter 3.5.

1.1.2 Representing and Reasoning under Uncertainty

Capturing the uncertainties inherent in the adversarial model as well as those found in the observables is important. A wide variety of approaches to modeling uncertainty exist including fuzzy logic, possibility theory, Dempster-Shafer, and qualitative reasoning (see [19] for a brief survey of models for uncertainty). We focus here on probabilistic models, specifically discrete models. For probabilistic reasoning, random variables (abbreviated *r.v.s.*) are used to represent discrete events or objects in the world. By making assignments to these *r.v.s.*, the current state of the world can be modeled probabilistically. The reasoning process involves computing joint probabilities of the given *r.v.s.*

Bayesian networks (BNs) [25] are directed acyclic graphs in which the conditional dependency (such as a causal relationship) is represented through arcs between the *r.v.s.* When all parents of a given *r.v.* A are instantiated, that *r.v.* is said to be conditionally independent of the remaining *r.v.s.* that are not descendants of A given its parents.* This approach provides a structural and

* For more details on this, see *d*-separation in [25].

visual (graphical) organization of information and direct relationships among *r.v.s.*

While BNs have been successfully used to prototype intelligent systems, including a tool for adversarial intent inferencing [29] and a causal analysis tool for military planning and wargaming [21,26], limitations exist for constructing such networks due to BN requirements such as completeness of conditional probability tables. In this chapter, we recommend another uncertainty model called *Bayesian knowledge bases* (BKBs) [34].

BKBs are a generalization of BNs. BKBs have been extensively studied both theoretically [35] and empirically for use in knowledge engineering [28] in a wide variety of domains such as space shuttle engine diagnosis, medical information processing, and data mining [2]. BKBs provide a highly flexible and intuitive representation following a basic “if-then” structure in conjunction with probability theory. Furthermore, BKBs were designed with typical domain incompleteness in mind to retain semantic consistency as well as soundness of inference in the absence of complete knowledge. Conversely, BNs typically assume a complete probability distribution is available from the start. BKBs have also been shown to capture knowledge at a finer level of detail as well as knowledge that would be cyclical (hence, disallowed) in BNs.

BKBs can also be depicted as directed graphs consisting of two types of nodes — *instantiation nodes* (I-nodes) and *support nodes* (S-nodes). In Figure 1.1.2, the I-nodes are labeled nodes that represent unique specific assignments to individual *r.v.s.* The I-nodes are related via edges to the S-nodes, which are depicted as dark nodes. Each *conditional probability rule* (CPR) is represented by an S-node where the parents of the S-node are the antecedents of the CPR and the child of the S-node denotes the

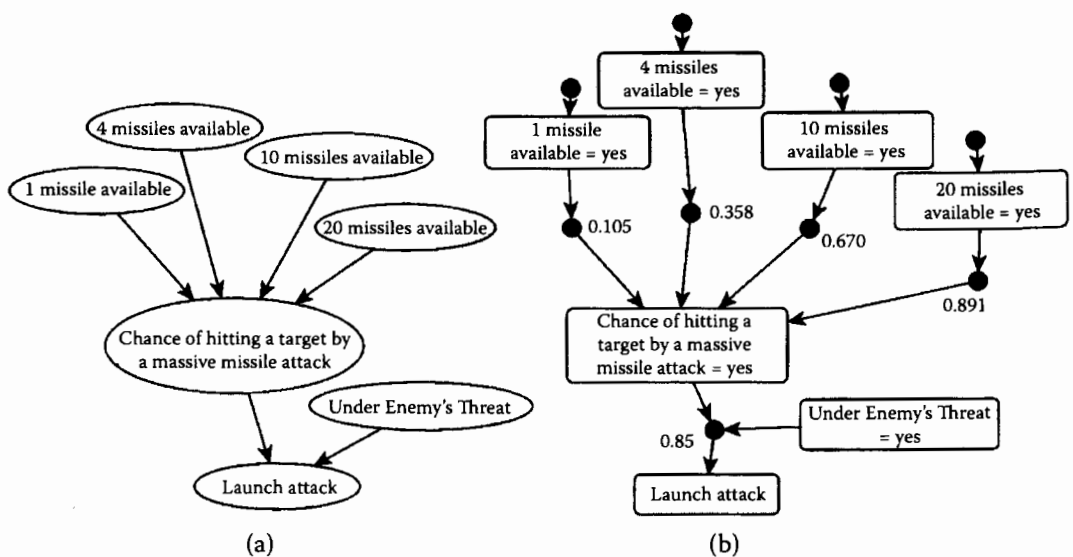


FIGURE 1.1.2

An example of probability networks for decision making process in (a) BN and (b) BKB.

consequence. Inferencing over BKBs can be conducted similarly to “if-then” rule chaining. As such, sets of CPRs collectively form inferences where the probability of the collection is a product of the probabilities associated with the CPRs.

As described in [35], probabilistic models exhibiting significant local structure are common. In such models, explicit representation of that structure — as done in BKBs — is advantageous as the resulting representation is much more compact than the full table representation of the *conditional probability tables* (CPT) in a BN. For example, as shown in Figure 1.1.2, the decision of whether to launch a missile attack against an enemy target is dependent upon two conditions: The chance of being able to hit and destroy/damage the target and whether a threat exists from that target. The chance of hitting the enemy target depends on the number of missiles that can be launched at the same time in the attack (consider the enemy having the ability to defend against those missiles). In this example, each random variable has only two states (*yes* or *no*). The structures of BNs and BKBs are very similar, except that BKBs have additional S-nodes (dark circles). However, in BNs, the CPT for the random variable “Chance of hitting the target by a massive missile attack” has 2^5 entries. The size of the representation of the conditional probabilities in terms of rules in the BKB is only 4.

Given the rule-based nature of BKBs, introducing new rules and modifying or removing existing rules is also relatively easy. This attribute allows the development of automatic algorithms for generating BKB fragments, i.e., small sets of CPRs, in real-time or stored in a BKB fragments library for use as needed. As we shall see, BKB fragments are successfully instantiated and used within our adversary intent inferencing framework for wargaming and mission planning and analysis.

1.1.3 Adversary Intent Inferencing Model (AII)

The *adversary intent inferencing* (AII) model strives to represent the perceptions of the adversary. It explicitly defines what the adversary believes about themselves, such as their capabilities and doctrines, and what the adversary believes about their opponents (see detailed discussion below). This is one major difference between the AII model and other approaches, such as the Soar system. Also, unlike the Soar and BDI models where the committed plans or chosen operators constrain the search space, the AII model’s reasoning space is defined by the current state of the world as seen through the eyes of the adversary. In addition to inferring the possible goals, intentions, and actions, the AII model also emphasizes the explanation of inferred results by relating them to the adversary’s beliefs. In general, intent inferencing should be able to provide three kinds of hypotheses: (1) Descriptive intent

inference provides insight into the motivations behind actions that have just occurred; (2) predictive intent inference can anticipate future actions given the individual's inferred goals; and (3) diagnostic intent inference detects deviations between predicted and observed actions to reveal possible errors [1]. Instead of just focusing on prediction, the AII model provides all three types of intent inferencing.

1.1.3.1 Architecture of AII Model

In the AII model, we decompose the adversary intent inferencing architecture into the what/how/why components to provide a natural and intuitive organization of the adversarial decision-making process. The three core components are as follows:

Goals/Foci (what): A prioritized, weighted short- and long-term goals list representing adversary objectives or foci, which evolves over time.

Rationale network (why): A probabilistic network representing the influences of the adversary's beliefs, both about themselves and about their opponents, on their goals and on high-level actions associated with those goals.

Action network (how): A probabilistic network representing the detailed relationships between adversary goals and possible actions to realize those goals.

Due to the uncertainty involved in adversary course of action prediction, we use BNs [25] or BKBs [34,35] as the knowledge representation for the rationale and action networks.* The *r.v.s* involved in the probabilistic networks are classified into four classes:

Axioms (X) represent the beliefs of the adversary about themselves. This can be an adversary's beliefs about their own true capabilities or even a fanatic belief of invulnerability. Axioms typically serve as inputs or explanations to the other *r.v.s*, such as adversary goals.

Beliefs (B) represent the beliefs regarding their opponent. For example, an adversary might believe that their opponent will conduct air strikes before moving ground troops but will not conduct carpet-bombing given their opponent's current political situation. Beliefs are further decomposed into tactical and strategic beliefs as we will describe in more detail later. (Note that for this chapter, we use *Red* to refer to the adversary and *Blue* for the friendly side as our models are used by the "friendly" side.)

* We initially employed BNs, but have moved to BKBs, as we mentioned earlier.

Goals (G) are the adversary's desired end-states. They are either short-term or long-term and are stored in a weighted, prioritized list. The goals can be further partitioned into two types: Abstract and concrete. Abstract goals are those that cannot be executed directly — for example, the abstract goal of damaging world opinion concerning Blue. Concrete goals could be something like destroying a Blue force checkpoint.

Actions (A) can be carried out to achieve adversarial goals. Actions typically can be observed by friendly forces — for example, launching a surface-to-air missile against Blue aircrafts.

These four *r.v.* types occur within the two networks. The rationale network contains all the Belief, Axiom, and Goal variables, as well as any Action variables that have Goals as inputs. This network is used to infer the adversary's short-term and long-term goals. Once the goals are determined, the action network is used to reason out what the most likely enemy actions will be. The action network contains the entire set of Action variables and any concrete Goal variables. These Action variables serve as the "connective" interface between the rationale and action networks. Figure 1.1.3 shows an example of a rationale network and an action network.

1.1.3.2 Model Construction

As described above, the AII model contains three major components. Besides the goal list, two probabilistic networks are also used: The action network and the rationale network. These two networks represent the knowledge of the adversarial decision-making process. Generally, the process of establishing any probabilistic network consists of three steps: (1) Identify the important random variables; (2) build the causal relationships among random variables and their assignments, which then gives a graphical structure; and (3) set the probability distribution values.

We begin by examining the high-level variables belonging to Beliefs. In the AII model, Belief variables are independent and serve as inputs to Axioms, Goals, or Actions. Belief variables can be categorized into two basic types: Strategic Beliefs and tactical Beliefs. Strategic Beliefs include philosophy, strategic goals, and general characteristics or behaviors of Blue forces from the adversary's point of view. Tactical Beliefs represent actionable Blue events such as the physical repositioning of assets, specific kinetic attacks, and so on. While constructing a dependency structure among the Belief variables (especially, tactical beliefs) that represent sequences or hierarchies of Blue actions seems reasonable, this practice increases the complexity of the networks. The complexity is easily mitigated by the fact that when the Blue actions are known with certainty, the belief variables can be set as evidence, which has the effect of rendering the Beliefs independent [29,30].

The rest of the semantic structure is intuitive and is as follows: Axioms have strategic Beliefs as inputs and serve as inputs to Goals and other

Axioms. Goals have Axioms and Beliefs as inputs and serve as inputs to Actions or other Goals. Actions have Goals and tactical Beliefs as inputs and can only be inputs to other Actions. Basically, the structure follows an intuitive hierarchical pre- and post-condition organization. Hence, a natural dominance relationship exists between related variable types, for example, Axioms that are descendants of other Axioms. This dominance reflects the fact that one variable can be "more general," "more abstract," "aggregate," "precondition," and so on, with respect to a descendant variable.

To maintain the appropriate division of variables among the rationale and action networks, Goal variables are partitioned into abstract Goals and concrete Goals. Abstract Goals are Goals composed of additional Goal variables. Concrete Goals are Goals that are immediately actionable. As such, abstract Goals can only appear in the rationale network and are critical to providing the proper explanations for adversary rationale. Concrete Goals must appear in both networks and serve as the causal "glue" between networks.

Finally, two basic rules maintain the semantic integrity of the AII: (1) All Axioms, Beliefs, Goals, and Actions occur in at least one of the adversary rationale or action networks; and (2) given a concrete Goal G , if A is an Action node with input G , then G , A , and the inputs of A must occur in both networks with the same connection structure. The second rule makes sure that the propagation of reasoning between the two networks occurs properly. Such propagation actually happens in both directions, rationale to action to rationale, which reflects the predictive and explanatory process in the AII.

As we can see, the construction of the AII model is a process of identifying potential adversarial goals based on what the enemy believes about the Blue forces and the Red themselves, then generating possible actions that can be carried out to achieve these goals based on their capabilities and history movements. Human factors can be modeled by or show their impacts through the Axiom or Belief variables in the model as we shall see later.

1.1.3.3 Reasoning over the Model

Reasoning over the AII model is an iterative process that allows the adversary model to adapt to changes over time (Figure 1.1.4). It includes several steps: Take inputs (observables, currently inferred adversary goals, intelligence, and user feedback) as evidence and infer the new goals; use the new information as evidence to reason about the potential adversarial actions and present them to the user; update the model according to the environmental inputs, reasoning results, and user feedback; and get ready for the next cycle. More specifically, the intent inferencing and prediction process functions as follows:

- (1) Observables regarding the adversary are set as evidence in both rationale and action networks. Feedback from the analyst (if any) is set as evidence.
- (2) Current short- and long-term enemy foci from foci lists are also set as evidence in both networks.

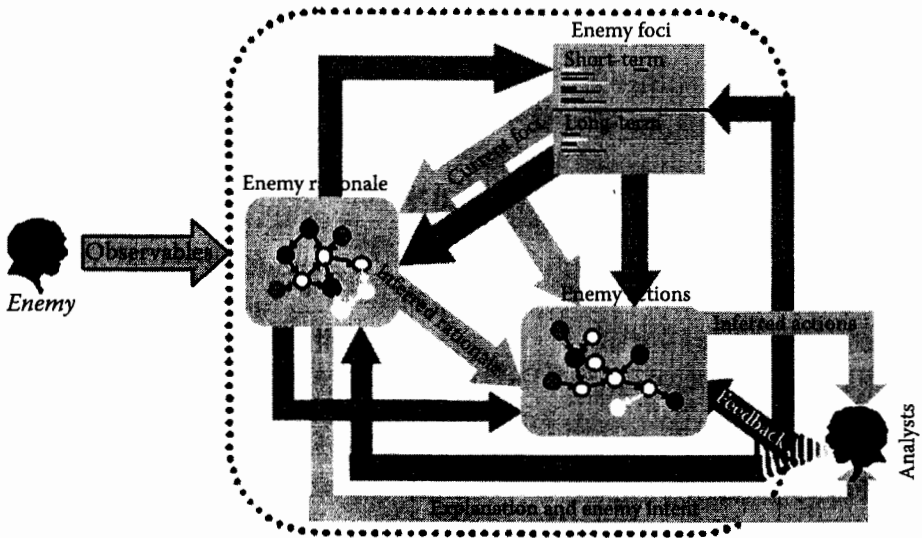


FIGURE 1.1.4
AII process (from [29]).

- (3) The rationale network infers new goals that are set as evidence for the action network.
- (4) The action network is now used to predict adversarial actions.
- (5) The analyst is presented with the inferred goals, predicted actions, and the explanation behind such predictions.
- (6) The analyst provides feedback in terms of corrected goals and actions, if desired.
- (7) The goals list is updated based on newly inferred goals and current strength of existing goals. If goals exceed a given threshold value, they are added to the list. If goals fall below a set threshold, they are removed. If goals in the short-term list persist beyond a given time threshold, they become long-term goals.
- (8) Networks are updated based on (8) analyst feedback (if any) and go to Step 1.

The inference process on both the rationale and action networks is based on Belief updating for probabilistic networks. In essence, given a target variable R and evidence set E , belief updating computes $P(R/E)$.

1.1.3.4 AII and Wargaming

Our AII model has been integrated into military wargaming environments to assist in simulating intelligent forces. This practice helps to reduce the overall level of human expert involvement while effectively assisting in the

rapid construction and execution of the simulations and “what-if” analysis in the after action review, where an assessment is conducted after the simulation to evaluate what has happened, why it has happened, what could happen, and what should be the response. In this section, we describe our experiments and results of prototyping and deploying AII.

The AII model was first prototyped and deployed in cooperation with Lockheed Martin’s Advanced Technology Laboratories. The goal of the prototype was to provide an initial proof-of-concept for AII based on a historical scenario, the Battle of Khafji during the Persian Gulf War in 1991. Khafji was a small abandoned town in Saudi Arabia near the Kuwait border, and this battle was the Iraqis’ only organized offensive during the war [18].

The objective of this Iraqi attack is now accepted to be to attempt to engage the coalition forces in a ground battle while Iraqi mechanized forces could still maneuver in the Kuwait theater of operations, especially in light of the effectiveness of coalition air attacks. Several days before the attack, Iraqi forces massed behind the nearby battle line. Increased activities were detected including the digging of berms and reinforcement of artillery positions on January 26 and 27 by the Iraqis, and the movement of armored vehicles into position on January 28. On January 29, after nightfall, Iraqi tanks approached Khafji and made contact with U.S. Marine Corps outposts along the border. The outposts fell back to preplanned positions while coalition forces responded with air strikes. The Iraqis took control of Khafji and it was then the Coalition’s problem to determine Iraqi intentions, contain the offensive forces, and retake Khafji. The steady surveillance and constant availability of air power helped the Coalition stop the Iraqi attack in time to spoil their advantage of surprise.

In the simulation, working with the constructed AII model, the prototype successfully predicted as well as updated the predicted actions of the adversary over time as the events unfolded. At the beginning of the simulation, the AII model was set with the observation that an Iraqi offensive in the south was unlikely and Coalition attention and sensors were focused on the western reaches of Iraq in support of SCUD (a type of tactical ballistic missile) suppression, strikes on Iraqi Republican Guard divisions, and battle damage assessment [24]. As time passed and the simulation progressed with each new observation and intelligence report about Iraqi forces moving south, massing, trying to jam Coalition communications, and so on, the AII correctly reflected the intentions of the Iraqi forces by indicating that the probability of Iraqi forces crossing the border for a ground engagement was increasing. Because the AII was a model of Iraqi forces, the intentions of the Iraqis drove the appropriate selection of actions by the AII in the simulation and ultimately initiated the attack across the border. The selection of actions was based on choosing the highest probability actions supported by the inferred goals and intentions. The observations of events had been set as evidence into the AII, which strengthened the inferred goals and intentions over time, in this case, the Iraqi desire for a ground battle.

Later, the AII model was deployed with the Force Structure Simulations (FSS) wargaming system at the Air Force Research Laboratory, Information Directorate (AFRL/D) [14,37]. In this trial, a general adversary model was first created and then specialized to portray different adversaries. The general model contained random variables that accounted for variability in weapons capabilities, tactical/strategic strikes, and responses to the presence of Blue forces in any of the four directions. Two instances of the model with different belief systems were created with the AII, referenced as adversary A and B. Adversary A possessed a competent air force, a smaller ground force, and WMDs, whereas adversary B was lacking any WMDs, had much less air power, but had a powerful ground force.

The test was to see the effects on the wargaming simulation and how much these two adversaries would differ in countering Blue force actions. Two different sets of observations of Blue forces were created and fed into the system. In the first set of inputs for Blue, data indicated a strategic bombing campaign comprised of deploying sea forces and launching cruise missiles and air strikes at strategic targets. The second input set described a land invasion of military targets by Blue. Each step in the simulation covered a 30 minute time slice [37]. Table 1.1.1 lists some of the highly ranked actions and differences resulting from the input sets. Adversary A, while characterized by confidence in its higher technology capabilities, responded to Blue attacks with air counterstrikes and potentially employing their WMDs, whereas adversary B responded through ground actions. While the results might be obvious to a human, the goal of the simulation is to see how the automated adversary forces would act differently in different situations.

TABLE 1.1.1

Sample Results Matrix

COA Inputs	Adversary A	Adversary B
Set 1	Deliver Ultimatum Launch Air Attack Send Forces South Arm WMD Launch WMD	Launch Ground Assault Send Forces South Enemy Recon Probing Forces Cross Border Deploy Forces in Civilian Areas
Set 2	Deploy Forces in Civilian Areas Deliver Ultimatum Deploy Forces Along Border Arm WMD Conceal Assets Launch WMD	Deploy Forces in Civilian Areas Launch Ground Assault Send Forces West Send Forces North Enemy Recon Probing Forces Cross Border Deploy Forces along Border Conceal Assets

(Adapted from [37])

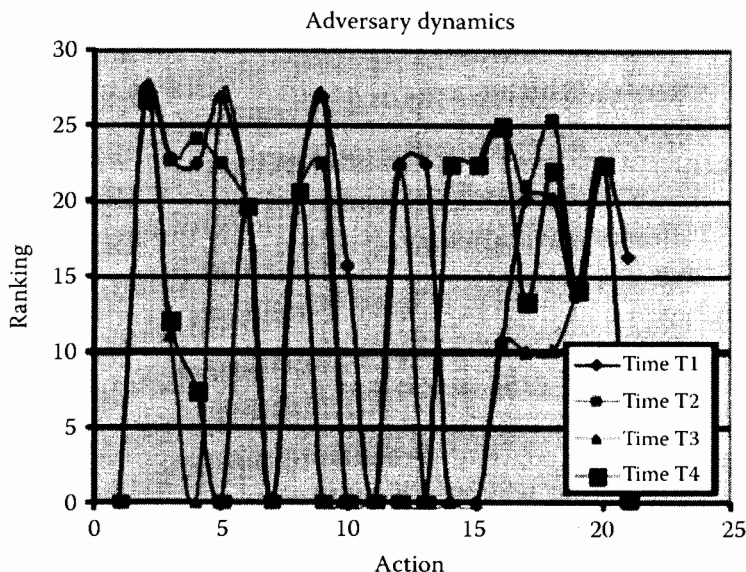


FIGURE 1.1.5
Dynamics of adversary’s action behaviors (from [37]).

Figure 1.1.5 and Figure 1.1.6 are plots of the adversary actions (referenced numerically along the x -axis) being pursued. More than 20 Red force actions were defined in the scenario, some of these shown in Table 1.1.1. Figure 1.1.5 shows the actions of adversary A over time as the first set of Blue force COA was applied. In each time step (T_1 to T_4), adversary A chose a different set of actions in response to the Blue force actions at that moment. For example,

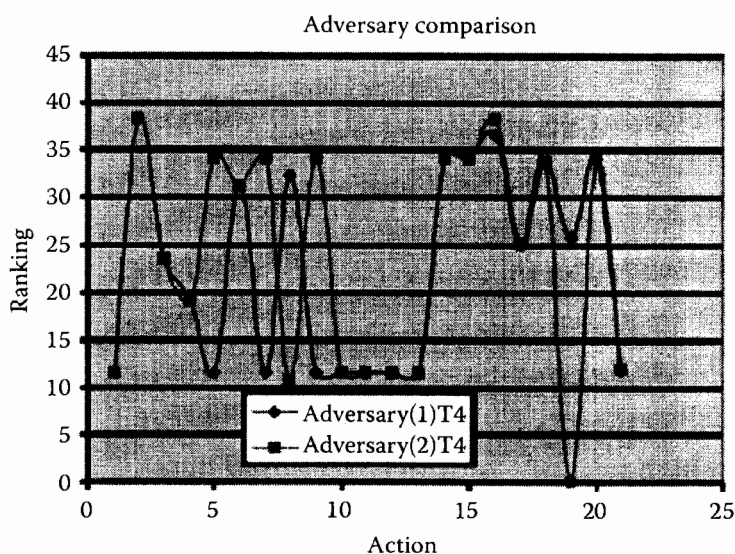


FIGURE 1.1.6
Action differences between adversaries (from [37]).

at T_2 , action 5 had a high probability, but at T_4 , action 5 would not be the choice of Red. The graph illustrates the dynamics in the anticipated behavior at each time interval, which shows that when the situation changes along the timeline, adversary A responded differently to Blue's attacks.

The divergence of selected actions between the two adversaries at a single time step is shown in Figure 1.1.6. Clearly at time T_4 , some actions were ranked quite differently (actions 5, 7, 9, and 19) by the two adversaries. This result shows that the differences in the Beliefs of two adversaries made a noticeable difference in their behaviors. Similar results have been obtained from other time slices.

More recently, the AII model has been deployed in the Emergent Adversarial Modeling Systems (EAMS) [15]. The AII engine was supported by the EAMS ontology that describes data and semantic relationships between the domain elements. The simulation was based upon the Deny Force Scenario used in the FSS [14], as documented by AFRL/IF, with a set of predefined Blue missions. The Blue force assets included the carrier *USS Roosevelt*, FA-18s on the carrier, Nellis Air Force Base, and the F-16s at the base. On the adversarial side, the Red force has surface-to-air missiles (SAM) located at the site named Twentynine Palms, 12 Seersuckers (anti-ship missiles) at Vandenberg, airports at Vandenberg and Twentynine Palms, command posts at Pendleton, Twentynine Palms and Meadows, and so on. The possible movements — which are commands that can be issued to FSS — include moving assets along known routes, operating assets, and engaging targets.

In the scenario, the Red force faced three main threats from the Blue force: Jamming of radar by EA-6 launched from *USS Roosevelt* at Meadows, Pendleton and Twentynine Palms; attack by FA-18s from *USS Roosevelt* at Meadows and Pendleton; and attack by F-16s from Nellis AFB at Twentynine Palms and Hesperia.

According to the Deny Force Scenario, the EAMS ontology was configured with a set of Actions that was possible for Red to perform, a set of Red Beliefs, a set of Red Axioms, and a set of Red Goals. An example where we identified critical information and generated a BKB (fragment) is shown as follows and can be seen in Figure 1.1.7. (Symbols *A*, *B*, *G*, and *X* represent random variables of Action, Belief, Goal and Axiom in the AII model.)

- The Red force collected observable information, where the Blue force has attacked with FA-18s launched from the carrier *USS Roosevelt*. The Red force was based at Vandenberg airport and has 12 Seersucker missiles in inventory.
- Based on this information, Red generated a Belief that the Blue force would strike the Red site with FA-18s from the *USS Roosevelt* (*B*, *Air Strike by FA-18 from USS Roosevelt*).
- Next, an Axiom was generated for each Red asset concerning the status and effectiveness of the asset. In this example, they were the status of the airport (*X*, *Vandenberg Airport Operational*) and Seersucker missiles (*X*, *Have Seersucker at Vandenberg Airport*), and the chance they

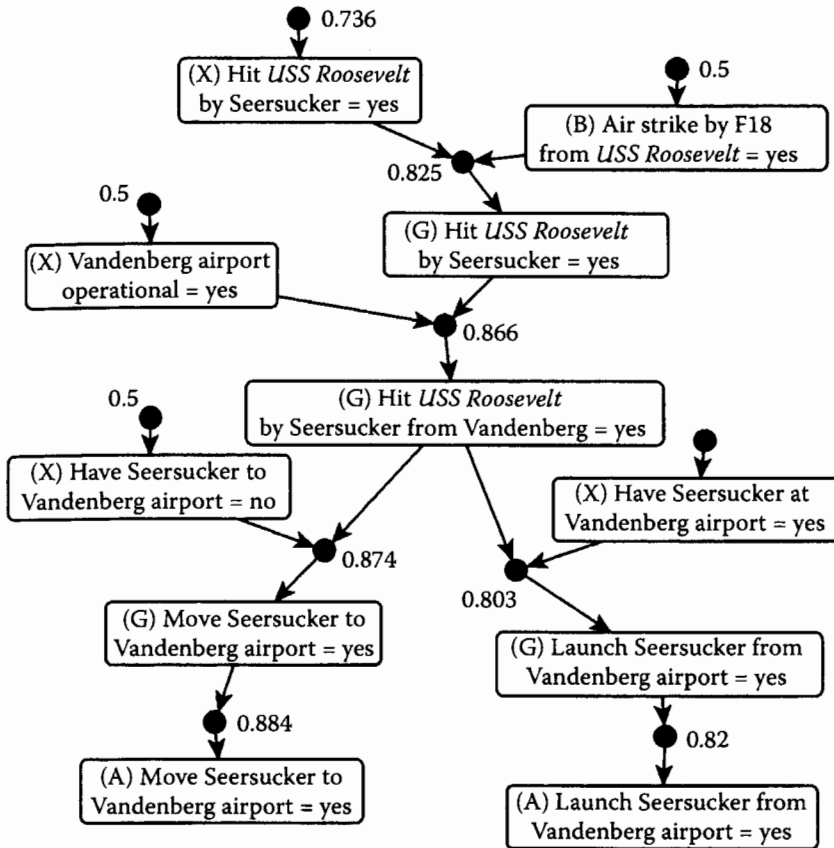


FIGURE 1.1.7

A partial rationale network (BKB) for simulation with Deny Force Scenario.

would be able use these missiles to damage the Blue carrier, which was the effectiveness of their weapon (*X, Hit USS Roosevelt by Seersucker*).

- According to the Belief that Blue's attacks were coming from the *USS Roosevelt*, Goals were then generated by pairing the Red and Blue capabilities that would be able to reach each other. One of the possible Goals was to attack the carrier with Seersuckers (*G, Hit USS Roosevelt by Seersucker*). This Goal was then detailed into a more concrete one, i.e., attack *USS Roosevelt* with Seersuckers from the Vandenberg airport (*G, Hit USS Roosevelt by Seersucker from Vandenberg*). Continuing this process, launching an attack against the *USS Roosevelt* from Vandenberg was further decomposed into two sub-Goals: To move missiles to the airport and to launch the missiles. Sub-Goals were provided by the adversary ontology, where if a weapon was to be used, it should be either already on-site or moved to the specific site.
- After the Goals have been generated, the Beliefs and Axioms were linked to Goals. Because generating Goals was a process of pairing

the Red and Blue capabilities for attacking each other or one defending against the other, proper Axioms that were about Red and Beliefs about Blue could be easily assigned to related Goals.

- Finally, Actions were added for the concrete Goals — in this case, corresponding to moving missiles to Vandenberg airport (*A, Move Seersucker to Vandenberg Airport*) and launching the missile attack from Vandenberg (*A, Launch Seersucker from Vandenberg Airport*). Adding Actions and sub-Actions were supported by the ontology, where knowledge was available for what Actions were necessary for carrying out certain concrete Goals.

The probability values for this fragment came from scenario settings and theoretical calculations. During the simulation, Red started with 12 missiles and could launch all of them in a massive attack. The probability of hitting the *USS Roosevelt* by such an attack (*X, Hit USS Roosevelt by Seersucker*) would be 0.736, which is calculated based on a single-shot hit probability of 0.7 and a 0.85 probability of each missile being blocked. Later in the simulation, Red force lost 8 of their missiles in action and had only 4 Seersuckers available to them. According to this change, the AII model recalculated the probability for (*X, Hit USS Roosevelt by Seersucker*) and dropped it to 0.358. The probabilities and method for computing them are retrieved from the ontology.

In the simulation, Red commander actions were generated real-time by the AII model in response to Blue movements and thus became the executed Red missions. Two test scenarios were developed and simulated to highlight the notion that given the execution of a set of Blue missions, two adversarial commanders — differing only by their intent — will perform different sets of actions in response to the same set of Blue missions. In the experiment, one Axiom variable defined simply as “Behavior” was added into the rationale network in the AII model. It has three states: Aggressive, neutral, and passive. The three states of the variable “Behavior” simply reflect possible high, medium, and low levels of aggressiveness of Red. It is introduced as a parent node for each high-level Goal variable and the state of the “Behavior” variable influenced the probability of each Goal. Assume that the Red Goal of attacking the Blue force has a probability value of p_n that reflects a neutral commander. To automatically reflect the other commander types, we have probabilities for an aggressive commander of $p_n + 0.33 \times (1.0 - p_n)$ and, for a passive commander, $(1 - 0.33) \times p_n$.^{*} This allows us to easily introduce different human factors into the AII model by encapsulating the behaviors in a way that permits us to automatically perturb probabilistic behaviors from a single baseline model. For the experiment, one commander was

^{*} Note that in the case where $p_n = 1$, the aggressive and passive probabilities are 1.0 and 0.33, respectively. Because these are conditional probabilities representing $P(G = \text{true} \mid \text{Behavior} = \{\text{aggressive, neutral, passive}\})$, the respective values are probabilistically valid. Intuitively, the probability numbers will still be in a range from 0 to 1. The numerical influence of behavior is derived from the EAMS ontology reflecting the players in the battlespace.

TABLE 1.1.2

Red Commander's Actions in Simulation

Scenario 1: Commander Intent — Aggressive Attack	Scenario 2: Commander Intent — Passive Defense
Defend Initial Attack	Defend Initial Attack
Move SAMs into Meadows from Pendleton	Move SAMs to Meadows from Pendleton
React to Destruction	Continue to Defend
Launch Seersuckers to <i>USS Roosevelt</i> from Vandenberg	Move SAMs to Twentynine Palms from Pendleton
Continue to Defend	Defend with Authority
Move SAMs to Twentynine Palms from Pendleton	Operate all SAMs

selected to be “aggressive” while the second was “passive.” In the two scenarios, Scenario 1 represents the aggressive Red commander and Scenario 2 represents the passive commander.

The significant events chosen for the simulation were: *Meadows Detects FA-18s*, *Meadows Experiences Destruction*, and *Twentynine Palms Detects F-16s*. Upon the occurrence of a significant event, the AII model was triggered to generate the list of potential Red actions that correspond to the event. Once the candidate actions were generated, EAMS generated the specific instance of a mission, which is then used by FSS to execute the mission.

Not surprisingly, the simulation demonstrated that the aggressive commander was likely to actively respond to Blue threats whereas the passive commander merely defends (Table 1.1.2). What was surprising is that the passive commander caused more serious damage to the Blue force than the aggressive commander. This outcome resulted from the fact that the passive Red commander chose to preserve assets by shutting down all equipment (SAMs, etc.), which caused targeting problems for Blue force bombers. Finally, when Red did decide to attack, Blue aircraft were already turning around and disengaging and were thus caught in a disadvantageous firing position.

The simulations above demonstrate that the AII model can support existing systems and simulation applications; descriptive elements of adversary composition can be properly classified allowing for the rapid assembly and modification of an adversary force (exploiting an ontology); and intent has great influence on the actions of an adversarial force, where soft factors such as the aggressive stance of an adversary force commander can alter adversary response and mission results.

1.1.4 Conclusions

Correctly predicting the adversary's intentions, actions, and reactions can lead to effectively responding to those actions, as well as planning ones own operations. The AII model provides a systematic approach to modeling the

adversary, and predicting adversary's intentions and potential future actions in a dynamic environment, while providing explanations of adversary's goals and actions. It employs explicit representation of adversarial actions, associated goals, and their beliefs behind those goals. The inferencing results can be used as a primary driver of adversary behaviors and responses as demonstrated in our experiments. Furthermore, it provides a tool that evaluates both the adversarial capabilities and potential future movement in light of their intentions. Putting all this together, the AII model has achieved capabilities beyond the traditional goal, plan, task, or action searching and matching approaches to adversarial behavior modeling.

Acknowledgments

This work was supported in part by Air Force Research Labs, Information Directorate, Grant No. F30602-01-1-0595, Air Force Office of Scientific Research Grant No. F49620-03-1-0014, Department of Defense Grant No. FA8750-04-C-0118, and Office of the Secretary of Defense Grant No. FA8650-04-M-6435.

Thanks to Bob Hillman, Jim Hanna, Duane Gilmour, Dan Fayette, Joe Carozzoni, Al Sisti, Dawn Trevisani, and many other folks at AFRL/IF, AFRL/HE, and AFOSR; Lee Krause, Lynn Lehman, Bruce McQueary, and Tony Stirtzinger at Securboratorion, Inc.; Axel Anurak, Sergio Gigli, and Frank Vetesi at LM ATL; and Scott Brown, Ben Bell, Joshua Surman, Hua Wang, and Alex Negri for all their invaluable assistance in moving the ideas in this research forward. Finally, special thanks to John Graniero, Nort Fowler, and Barry McKinney for their efforts in getting all of this off the ground.

References

1. Bell, B., Santos, E. Jr., and Brown, S.M., Making adversary decision modeling tractable with intent inference and information fusion, in *Proc. of 11th Conf. Comput. Generated Forces Behav. Represent.*, Orlando, FL, 2002, 535–542.
2. Ben-Eliyahu-Zohary, R., Domshak, C., Gudes, E., Liusternik, N., Meisels, A., Rosen, T., and Shimony, S.E., FlexiMine—a flexible platform for kdd research and application development, *Ann. Math. Artif. Intell.*, 39(1-2), 175–204, 2003.
3. Benyon, D. and Murray, D., Adaptive systems: from intelligent tutoring to autonomous agents, *Knowl.-Based Syst.* 6(4), 197–219, 1993.
4. Brown, S.M., Santos, E. Jr., and Banks, S.B., Active user interface for building decision-theoretic systems, in *Proc. 1st Asia-Pacific Conf. Intell. Agent Tech.*, Hong Kong, 1999, 244–253.
5. Bratman, M.E., *Intention, Plans, and Practical Reason*, Cambridge and London: Harvard University Press, 1987.

6. Bratman, M.E., Israel, D.J., and Pollack, M.E., Plans and resource-bounded practical reasoning, *Computational Intelligence*, 4, 349–355, 1988.
7. Bushman, J.B., Mitchell, C.M., Jones, P.M., and Rubin, K.S., ALLY: an operator's associate for cooperative supervisory control systems, *IEEE Trans. Syst., Man, and Cyber.*, 23(1), 111–128, 1993.
8. Carberry, S., Modeling the user's plans and goals, *Comp. Ling.*, 14(3), 23–27, 1988.
9. Chu, R.W., Mitchell, C.M., and Jones, P.M., Using the operator function model and OFMspert as the basis for an intelligent tutoring system: towards a tutor/aid paradigm for operators of supervisory control systems, *IEEE Trans. Syst., Man, and Cyber.*, 25(7), 1054–1075, 1995.
10. Franke, J., Bell, B., Mendenhall, H., and Brown, S., Enhancing teamwork through team-level intent inference, in *Proc. Int. Conf. Artif. Intell.*, Las Vegas, NV, 2000.
11. Geddes, N.D., The use of individual differences in inferring human operator intentions, in *Proc. 2nd Ann. Aerosp. Appl. Intelli. Conf.*, Dayton, OH, 1986, 31–41.
12. Geddes, N.D., A model for intent interpretation for multiple agents with conflicts, in *Proc. of IEEE Int. Conf. Syst., Man, and Cyber.*, San Antonio, TX, 1994.
13. Georgeff, M., Pell, B., Pollack, M., Tample, M., and Wooldridge, M., The belief-desire-intention model of agency, in Muller, J. P., Singh, M., and Rao, A., Eds., *Intelligent Agents V, Lec. Notes in AI, Vol. 1365*, Berlin: Springer-Verlag, 1999.
14. Gilmour, D.A., Hanna, J., Koziarz, W., McKeever, W., and Walter, M., High-performance computing for command and control real-time decision support, *AFRL Technol. Horizons*, 2005, <http://www.afrlhorizons.com/Briefs/Feb05/IF0407.html>
15. Gilmour, D.A., Krause, L.S., Lehman, L.A., Santos, E., Jr., and Zhao, Q., Intent driven adversarial modeling, *10th Intl. Command and Control Res. and Tech. Symp., The Future of C2*. McLean, VA, 2005.
16. Good, N., Schafer, J.B., Konstan, J.A., Borchers, A., Sarwar, B., Herlocker, J., and Riedl, J., Combining collaborative filtering with personal agents for better recommendations, in *Proc. 1999 Conf. Am. Assoc. Artif. Intell. (AAAI-99)*, 1999, 439–446.
17. Goodman, B.A. and Litman, D.J., Plan recognition for intelligent interfaces, in *Proc. of 6th IEEE Conf. on Artif. Intell. Appl.*, Santa Barbara, CA, 1990, 297–303.
18. Grant, R., The epic little battle of Khafji, *Air Force Mag. Online*, 81(2), 1998, <http://www.afa.org/magazine/Feb1998/0298epic.asp>
19. Halpern, J., *Reasoning About Uncertainty*, Cambridge, MA: MIT Press, 2005.
20. Headquarters, Department of the Army, Field Manual, No. 101-5. Staff Organization and Operations, 1997, http://www.dtic.mil/doctrine/jel/service_pubs/101_5.pdf
21. Lemmer, J.F. and Gossink, D. E., Recursive noisy OR—a rule for estimating complex probabilistic interactions, *IEEE Trans. on Syst., Man, and Cyber., Part B*, 34(6), 2252–2261, 2004.
22. Lesh, N., Rich, C., and Sidner, C.L., Using plan recognition in human-computer collaboration, in *Proc. 7th Int. Conf. User Modelling*, Banff, Canada, 1999, 23–32.
23. Nguyen, H., Saba, M.G., Santos, E., Jr. and Brown, S.M., Active user interface in a knowledge discovery and retrieval system, in *Proc. 2000 Int. Conf. Artif. Intell. (IC-AI 2000)*, Las Vegas, NV, 2000, 339–344.
24. Palmer, P.S., Scott, D.J., and Toolan, J.A., The battle of Khafji: an assessment of air power, *Research Report AU/AWC/192/1998-04*, Air War College, Air University, Maxwell Air Force Base, AL, 1998.

25. Pearl, J., *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, San Mateo, CA: Morgan Kaufman, 1988.
26. Phister, P.W., Jr. and Plonisch, I.G., Military applications of information technologies, *Air Space Power J.*, 18(1), 77–90, 2004.
27. Rubin, K.S., Jones, P.M., and Mitchell, C.M., OFMspert: inference of operator intentions in supervisory control using a blackboard architecture, *IEEE Trans. Syst., Man, and Cybern.*, 18(4), 618–637, 1988.
28. Santos, E., Jr., Verification and validation of knowledge-bases under uncertainty, *Data Knowl. Eng.*, 37, 307–329, 2001.
29. Santos, E., Jr., A cognitive architecture for adversary intent inferencing: structure of knowledge and computation, in *Proc. SPIE 17th Ann. Int. Symp. on Aerospace/Defense Sensing and Controls: AeroSense 2003*, Orlando, FL, 2003, 182–193.
30. Santos, E., Jr. and Negri, A., Constructing adversarial models for threat intent prediction and inferencing, in *Proc. SPIE Defense and Security Symp.*, 5423, Orlando, FL, 2004.
31. Santos, E., Jr., Nguyen, H., Zhao, Q. and Pukinskis, E., Empirical evaluation of adaptive user modeling in a medical information retrieval application, in Brusilovsky, P., Corbett, A., and de Rosis, F., Eds., *Lecture Notes in Artif. Intelligence 2702: User Modeling 2003*, Johnstown, PA: Springer, 2003, 292–296.
32. Santos, E., Jr., Nguyen, H., Zhao, Q. and Wang, H., User modelling for intent prediction in information analysis, in *Proc. 47th Annu. Meet. for Hum. Factors and Ergonomics Soc. (HFES-03)*, Denver, CO, 2003, 1034–1038.
33. Santos E., Jr., Zhao, Q., Nguyen, H. and Wang, H., Impacts of user modeling on personalization of information retrieval: an evaluation with human intelligence analysts, in *4th Workshop Eval. of Adaptive Syst.*, in conjunction with UM'05, 2005, 27–36.
34. Santos, E., Jr. and Santos, E.S., A framework for building knowledge-bases under uncertainty, *J. Exp. Theor. Artif. Intell.*, 11, 265–286, 1999.
35. Santos, E., Jr., Santos, E.S., and Shimony, S.E., Implicitly preserving semantics during incremental knowledge-base acquisition under uncertainty, *Int. J. Approx. Reas.*, 33(1), 71–94, 2003.
36. Schmitt, C., Dengler, D., and Bauer, M., The MAUT machine: an adaptive recommender system, in *Online Proc. ABIS Workshop, 2002*, http://www.kbs.uni-hannover.de/henze/lla02/abis_proceedings.html
37. Surman, J., Hillman, R., and Santos, E., Jr., Adversarial inferencing for generating dynamic adversary behavior, in *Proc. SPIE 17th Ann. Int. Symp. Aerosp./Defense Sensing and Controls: AeroSense 2003*, Orlando, FL, 2003, 194–201.