

Influence Diagram Modeling of Decision Making in a Dynamic Game Setting^{*}

Kai Virtanen, Tuomas Raivio, and Raimo P. Hämäläinen

*Systems Analysis Laboratory
Helsinki University of Technology
P.O.Box 1100, FIN-02015 HUT, FINLAND*

Abstract

We introduce an influence diagram game that represents the control decisions of pilots in one-on-one air combat and produces the best myopic control strategies under the assumption that the adversary behaves in the worst possible way. Such a game model could be considered as the control system of an autonomous agent operating in an uncertain and hostile environment and could be utilized, e.g., in a guidance system that controls aircraft in an air combat simulator. We first present an influence diagram modeling the control problem from the viewpoint of a single player. On the basis of this model, a non-zero-sum influence diagram game containing explicit control variables for both players is structured. The game model graphically describes the elements of the decision process, contains the state equations for the dynamics of the players, and takes into account the preferences of the players under conditions of uncertainty. Depending on the information structure of the influence diagram game, it admits either a Nash or a Stackelberg equilibrium solution. These solutions are obtained by converting the influence diagram game into a game in strategic form or into nonlinear optimization problems and by solving them. To demonstrate the influence diagram game, a numerical example is presented.

1 Introduction

Air combat simulators contain computer-guided aircraft whose main component is a model that aims at imitating the decision process of a pilot and, on the other hand, producing as good combat decisions as possible. The decision model should represent and solve complex decision making problems, e.g., on maneuvering. The problems contain multiple conflicting objectives and must be solved based on uncertain and time varying information. In addition, the pilot's decision making process is complicated by the behavior of the adversaries. In this paper, these inherent features are taken into account in the modeling of the pilot's control decision in one-on-one air combat by utilizing the methodology of influence diagrams (IDs) (Howard and Matheson, 1984) and game theory (see, e.g., Basar and Olsder, 1995).

Traditionally, tools from noncooperative differential game theory (see, e.g., Basar and Olsder, 1995; Isaacs, 1965) have been utilized in the modeling of one-on-one air combat. Although game optimal maneuvers can be obtained by solving such games, pure

^{*}This is a revised version of a paper appeared in Proceedings of the 1st Bayesian Modeling Applications Workshop of the 19th Conference on Uncertainty in Artificial Intelligence, Acapulco, Mexico, 2003.

differential game models have not been used in the guidance models of simulators mainly because of the oversimplifications needed in differential game formulations. Existing guidance models are based, e.g., on the selection of predetermined control laws using game theory (Neumann, 1990). In addition, models that try to imitate the pilot's decision making have been developed. These models are knowledge based expert systems (McManus and Goodrich, 1989; Stehlin et al., 1994), heuristic value driven systems (Lazarus, 1997) or discrete dynamic games (Austin et al., 1990; Katz, 1994). Air combat models that combine artificial intelligence and game theory have also been introduced (see, e.g., Le Menec and Bernhard, 1995). The influence diagram game (IDG) approach presented in this paper offers an alternative way to include the model of the pilot's decision process into a guidance system.

The IDs representing the pilot's maneuvering decision in a duel between two aircraft have been developed earlier in Virtanen et al. (1999) and Virtanen et al. (2004a). In these models, the control decision process is represented by assuming that the pilot receives information about the state of the adversary via imprecise measurements on the momentary combat state. The rational behavior of the adversary is omitted. Here, we introduce an IDG that contains explicit control variables as well as other components needed for modeling the control problems of both the players concurrently. By solving the game, one can determine the best control for one player with respect to the worst action of the other player. The basis of the game model is an ID that represents the control decision from the viewpoint of a single player.

IDs have been used in the modeling of single decision maker problems. Although the possibility to use them in modeling games has been introduced in Shachter (1986) and there have been some attempts to utilize them in game situations (see, e.g., Smith, 1994), the first explicit ID representation, the multi-agent influence diagram (Koller and Milch, 2001; Koller and Milch, 2003) for describing games was presented only recently. The multi-agent ID represents a static game problem in a compact form and offers an efficient solution approach. There are also alternative approaches for applying decision theoretical principles in the modeling of behavior of rational agents in a game situation. For instance, in Gmytrasiewicz (2003), Markov decision processes are generalized for taking into account the actions of other players.

In this paper, the underlying dynamics of the control decision problem is taken into account by including differential equations representing the dynamics of aircraft in the IDG. Such a dynamic problem cannot be tackled with the formalism of multi-agent IDs because it does not treat games with a state and does not contain a model describing the dynamics of the players. The IDG representation under consideration is constructed by combining two separate IDs. One of these diagrams represents the control decision from the viewpoint of a single player whose belief about the decision process of the adversary is modeled by the other diagram. Utility functions (see, e.g., Keeney and Raiffa, 1976) are used to measure the overall preferences in different combat states. Because both players have utility functions of their own, the resulting game is non-zero-sum. The payoff functions of the game are associated with the expected utility. The IDG allows the use of both discrete and continuous control variables. In the former case, the most desirable control alternative is obtained by converting the diagram into a game in strategic form. The continuous control variables lead to optimization problems that can be solved using nonlinear programming.

The information structure of a game describes all the information available to players

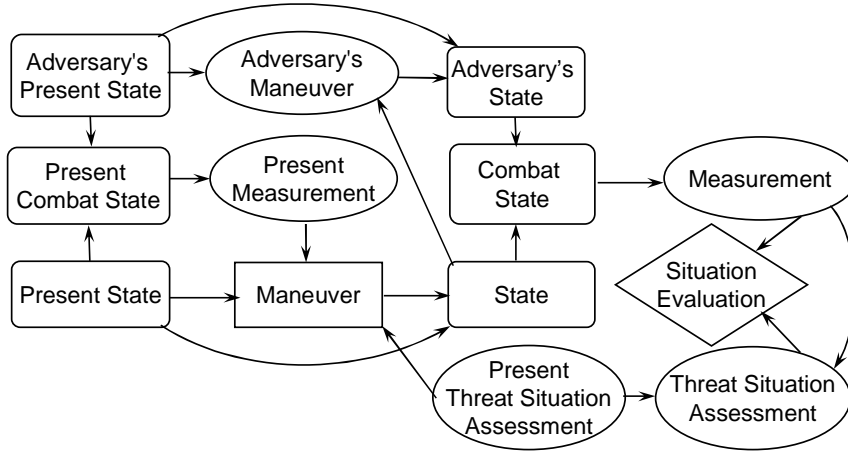


Figure 1: An ID representing the control decision from the viewpoint of the DM.

about the state of the game as well as about behavior of adversaries. In the IDG, it can be symmetric or asymmetric. The former structure refers to a situation in which both players are aware of all the elements of the game. Then, the solution of the game is a Nash equilibrium (e.g., Basar and Olsder, 1995). The equilibrium of this type does not always exist and is not necessarily unique (e.g., Basar and Olsder, 1995). This problem can be avoided by assuming that one player, the leader, has knowledge of all the elements of the game whereas the other one, the follower, makes the decision based on his or her probabilities and utilities as well as on the leader's action, i.e., the information structure is assumed asymmetric. Then, the IDG admits a solution called a Stackelberg equilibrium (see, e.g., Basar and Olsder, 1995).

The paper is structured as follows. First, Section 2 describes an ID representing the control decision from the viewpoint of a single player. Section 3 formulates an IDG modeling the control decisions of both players involved in one-on-one air combat. Section 4 discusses the solution of the IDG. Section 5 gives a numerical example, and Section 6 concludes.

2 Qualitative influence diagram for the control decision

We first consider a control decision from the viewpoint of a single player called the decision maker (DM). Assume that the control decision is taken based on uncertain information that is received by observing the combat situation and anticipating the states of the players a short time interval ahead. Such a myopic decision process is represented qualitatively by the ID shown in Fig. 1.

At the decision instant, the momentary state of the DM is given in the *Present State* node and the state of the adversary (AD) in the *Adversary's Present State* node, respectively. The states define the current combat state that is calculated in the *Present Combat State* node. The *Maneuver* node represents the control of the DM. The future state of the DM after the given planning horizon is calculated in the *State* node that depends on the momentary state of the DM as well as on the employed control.

Arcs pointing to the decision node imply that the DM's information consists of his or her own state data and an uncertain observation about the combat state. The observation

process is modeled by the *Present Measurement* node. The threat assessment, utilized when selecting the best control, is represented by the *Present Threat Situation Assessment* node that infers the threat situation from the viewpoint of the DM.

The outcomes of the chance node *Adversary's Maneuver* describe the control alternatives of the AD. The probabilities of the outcomes reflect the DM's belief about the selection of the AD's control. The probabilities are conditioned on the momentary state of the AD and the future state of the DM. The anticipated states of the AD are included in the *Adversary's State* node. The future states of the players define the anticipated combat states that are represented by the *Combat State* node.

The *Situation Evaluation* node evaluates each combat state that can be achieved using the feasible control alternatives of the players. In practice, the evaluation can be carried out with a utility function (see, e.g., Keeney and Raiffa, 1976). The utilities associated with each combat state depend on the state measurement of the DM that is represented by the *Measurement* node. In addition, the utilities are affected by the *Threat Situation Assessment* node representing the upcoming threat situation. It is estimated with the help of the DM's threat assessment at the decision instant as well as of the measurement about the anticipated combat state.

3 Influence diagram game for the control decisions

We next introduce an IDG that contains decision nodes and other components needed for describing the control decisions of both the players. During the combat each player tries to

1. Avoid being captured by the other player
2. Capture the other player.

The aim of the IDG is to produce the best control for the DM with respect to the above goals under the assumption that the AD acts rationally according to his or her goals. In other words, the worst case control strategy of the DM is identified.

Assume that the DM's belief about the AD's representation of the control problem is the same as his or her own. Hence, the structure of the ID representing the control problem from the viewpoint of the AD is similar to the model shown in Fig. 1. When the control problems of both the players are included explicitly in the same model, a non-zero-sum IDG results, see Fig. 2. It is constructed by combining two IDs that both represent a one player control decision problem. Although the structures of the single player IDs are now similar, they could also differ from each other.

In the IDG, the underlying dynamics is taken into account by describing the evolution of the players' states with the differential equations

$$\dot{x}^k(t) = f^k(x^k(t), u^k(t)), \quad x^k(0) = x_0^k, \quad k = D, A, \quad (1)$$

where $x^k(t) \in R$ are the state variables, the control variables $u^k(t)$ belong to sets S^k , $k = D, A$, for all t , and x_0^k are the given initial states. The superscripts D and A refer to the DM and the AD, respectively. In practice, the state equation (1) represents the motion of an aircraft. The states of the players define the momentary state of the combat

$$c(t) = g(x^A(t), x^D(t)), \quad (2)$$

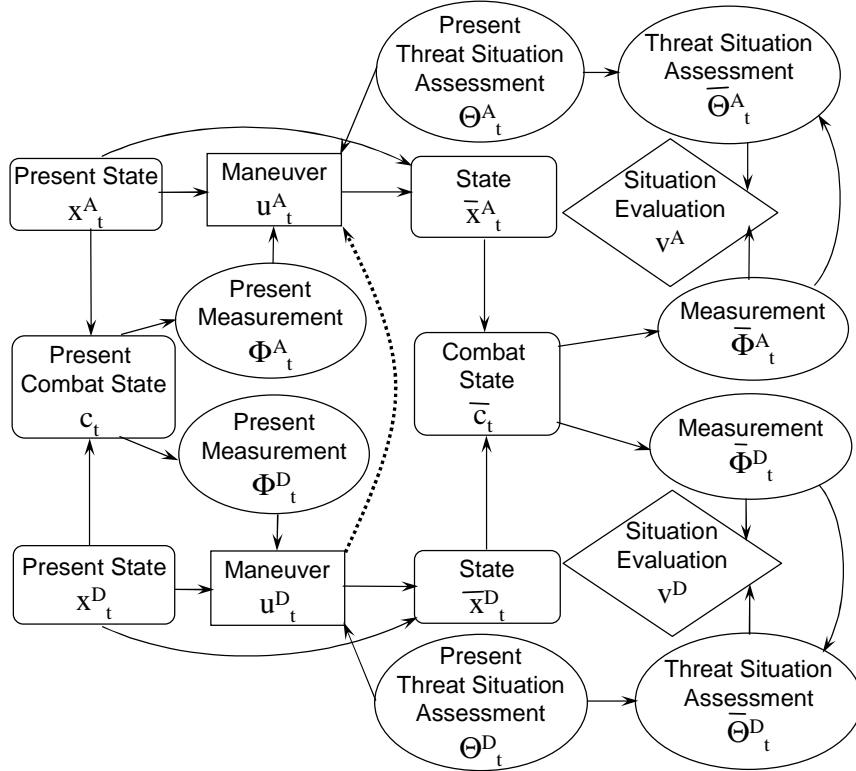


Figure 2: An IDG representing the control decisions of both the players. The dashed arc implies the information structure of the game.

where $c(t) \in R$ refers to the variable describing the combat state. For simplicity, we first assume that $x(t)$, $u(t)$, and $c(t)$ are scalars.

The control variables of the players at the discrete decision instant t , denoted by $u_t^k := u^k(t)$, $k = D, A$, are included in the decision nodes *Maneuver* of the IDG shown in Fig. 2. The variables can be continuous, $u_t^k \in S^k \subset R$, or discrete, $u_t^k \in S^k = \{d_1^k, \dots, d_{n_k}^k\}$. At the decision instant, each player knows the state of his or her own, $x_t^k := x^k(t)$, that is given in the deterministic *Present State* node. In addition, the player assesses the threat situation and receives an uncertain observation about the current combat state, $c_t := c(t)$, that is defined by the states of the players according to (2) and included in the deterministic node *Present Combat State*.

Recall that during the combat, both the players aim at capturing the opponent and at the same time avoiding being captured by the opponent. The target set of each player representing, e.g., the launching area of a missile is defined as

$$X^k = \{c(t) | \Psi^k(g(x^D(t), x^A(t))) \leq 0\}, \quad k = D, A, \quad (3)$$

where the function $\Psi^k(\cdot)$ measures the distance of the combat state from the target set of player k .

The threat assessment of player k is modeled by a discrete random variable Θ_t^k included in the chance node *Present Threat Situation Assessment*. It represents the prior belief of player k about the threat situation at time t . The outcomes of Θ_t^k reflect distance between the current combat state and the target sets. For the DM, the outcomes are

- Neutral, θ_1^D : c_t is far from X^D and X^A

- Advantage, θ_2^D : c_t is near X^D and far from X^A
- Disadvantage, θ_3^D : c_t is far from X^D and near X^A
- Mutual disadvantage, θ_4^D : c_t is near X^D and X^A

For the AD, the outcomes are obtained from the same list by swapping the indices D and A . For instance, the DM's probability of the "Advantage" outcome is high if the combat state is near the DM's target set and far from the AD's one. In general, the distance between the combat state and a target set can be considered small when the value of the function Ψ^k used in (3) gets close to zero. The probabilities of the outcomes are denoted by $P(\theta_i^k) := P(\Theta_t^k = \theta_i^k)$, $i = 1, \dots, 4$, $k = D, A$, and they sum up to one, i.e., $\sum_{i=1}^4 P(\theta_i^k) = 1$.

Often, the state observation process is continuous in its nature, and its outcome should be described by a random variable with a continuous probability distribution that depends on the true state and the device with which the state is being observed. For instance, the state observation can be assumed normally distributed with an expected value equal to the exact combat state c_t and a variance σ^2 describing the accuracy of the sensor device. In the forthcoming numerical solution process of the IDG, continuous distributions must finally be discretized. Since discrete distributions also simplify the presentation, we discretize the observation process already at this stage. Consequently, the observations of player k are represented by a discrete random variable Φ_t^k included in the chance node *Present Measurement*. The magnitudes of its outcomes depend on the combat state which can be expressed, e.g., as

$$\phi_j^k(c_t) = \Omega^k(c_t; j) = c_t + (j - \frac{n_\phi}{2} - \frac{1}{2})\sigma^2, \quad j = 1, \dots, n_\phi, \quad k = D, A. \quad (4)$$

Here, the outcomes are chosen symmetrically around the exact combat state. The dispersion of the outcomes can be affected by scaling the variance σ^2 with a suitable factor. n_ϕ is the number of the outcomes that can be interpreted as the coarsity of the discretization. The given probabilities of the outcomes are denoted by $P(\phi_j^k) := P(\Phi_t^k = \phi_j^k)$, $j = 1, \dots, n_\phi$.

The upcoming states of the players, \tilde{x}_t^k , $k = D, A$, after the planning horizon Δt are given in the *State* nodes. They are calculated by integrating (1) with u_t^k , i.e.,

$$\tilde{x}_t^k(u_t^k) := \tilde{x}_t^k(x_t^k, u_t^k) = x_t^k + \int_t^{t+\Delta t} f^k(x^k, u_t^k) dt. \quad (5)$$

The set of the predicted combat states, \tilde{c}_t , defined by the states of the players is represented by the deterministic node *Combat State*. Its outputs that can be obtained with the feasible control alternatives of the players are calculated according to (2), i.e.,

$$\tilde{c}_t(u_t^D, u_t^A) := \tilde{c}_t(\tilde{x}_t^D(u_t^D), \tilde{x}_t^A(u_t^A)) = g(\tilde{x}_t^D(u_t^D), \tilde{x}_t^A(u_t^A)). \quad (6)$$

The chance node *Measurement* contains a random variable $\tilde{\Phi}_t^k$, $k = D, A$, that models the player's observation about the future combat state. The outcomes of $\tilde{\Phi}_t^k$, denoted by $\tilde{\phi}_j^k$, $j = 1, \dots, n_\phi$, and their probabilities, $P(\tilde{\phi}_j^k) := P(\tilde{\Phi}_t^k = \tilde{\phi}_j^k)$, $j = 1, \dots, n_\phi$, are similar to those of the random variables included in the *Present Measurement* nodes

except that the magnitudes depend on the future combat state instead of the present one, i.e., $\tilde{\phi}_j^k(u_t^D, u_t^A) := \tilde{\phi}_j^k(\tilde{c}_t(u_t^D, u_t^A)) = \Omega^k(\tilde{c}_t(u_t^D, u_t^A); j)$ where $\Omega^k(\cdot)$ is given by (4).

The discrete random variable $\tilde{\Theta}_t^k$ representing the upcoming threat situation for player k at time $t + \Delta t$ is given in the chance node *Threat Situation Assessment*. The meaning of the outcomes, $\tilde{\theta}_i^k$, $i = 1, \dots, 4$, is the same as in the *Present Threat Situation Assessment* node. Their probabilities are conditioned on the future state observations $\tilde{\phi}_j^k$ and calculated by using Bayes' theorem as follows:

$$P(\tilde{\theta}_i^k | \tilde{\phi}_j^k(u_t^D, u_t^A)) := P(\tilde{\Theta}_t^k = \tilde{\theta}_i^k | C = \tilde{\phi}_j^k(\tilde{c}_t(u_t^D, u_t^A))) = \frac{P(\theta_i^k) p^k(\tilde{\phi}_j^k(\tilde{c}_t(u_t^D, u_t^A)) | \tilde{\Theta}_t^k = \tilde{\theta}_i^k)}{\sum_{l=1}^4 P(\theta_l^k) p^k(\tilde{\phi}_j^k(\tilde{c}_t(u_t^D, u_t^A)) | \tilde{\Theta}_t^k = \tilde{\theta}_l^k)}, \quad (7)$$

$$i = 1, \dots, 4, \quad j = 1, \dots, n_\phi, \quad k = D, A.$$

Here, the probability inference requires a continuous random variable, denoted by C , for the combat state variable. The likelihood function $p^k(\cdot | \tilde{\Theta}_t^k = \tilde{\theta}_i^k)$ represents the distribution of C under the supposition that the outcome of $\tilde{\Theta}_t^k$ is $\tilde{\theta}_i^k$. The forms of the likelihood functions used in Eq. (7) determine the rate of change of the threat probabilities. Note that the threat probabilities at time t , $P(\theta_i^k)$, $i = 1, \dots, 4$, are used as the prior probabilities in (7) when calculating the posterior probabilities $P(\tilde{\theta}_i^k | \cdot)$. For the details of the probability updating and the suitable likelihood functions, see Virtanen et al. (2004a).

Each combat state that can be achieved using the feasible control alternatives is evaluated by utility functions included in the *Situation Evaluation* nodes. The functions capture the preferences of the players and depend on the upcoming threat probabilities as well as on the combat state measurement. In general, the relative importance of the players' goals depends on the current combat state. Thus, each outcome of the *Threat Situation Assessment* node leads to a particular utility function.

Assume that the combat state is described by more than one variable, i.e., $c(t) \in R^p$. Then, the utility function associated with the outcome $\tilde{\theta}_i^k$ of $\tilde{\Theta}_t^k$ and the outcome $\tilde{\phi}_j^k = [\tilde{\phi}_{j,1}^k \dots \tilde{\phi}_{j,p}^k]^T$ of $\tilde{\Phi}_t^k$, can be expressed as

$$v^k(\tilde{\theta}_i^k, \tilde{\phi}_j^k(u_t^D, u_t^A)) = \sum_{l=1}^p w_l^{k,i} v_l^{k,i}(\tilde{\phi}_{j,l}^k(u_t^D, u_t^A)), \quad (8)$$

where the single attribute utility functions $v_l^{k,i}$ map each combat state variable into the interval $[0, 1]$. The weights $w_l^{k,i}$ describe the relative importance of objectives represented by the single attribute utility functions. The weights are limited to $[0, 1]$ and sum up to one. Because the single attribute functions and their weights are specific for each outcome of the threat assessment, the weighted sum is a valid representation of preferences (Virtanen et al., 2004a). In the IDG under consideration, the combat state is described by one variable. Therefore, there is no need for the weight vector and only one single attribute utility function is required in the aggregated function (8). A detailed description and discussion on the multiattribute utility model applied in IDs is given in Virtanen et al. (2004a).

The information structure of a game refers to all the information available to the players about the game situation. The IDG presented here contains imperfect information that

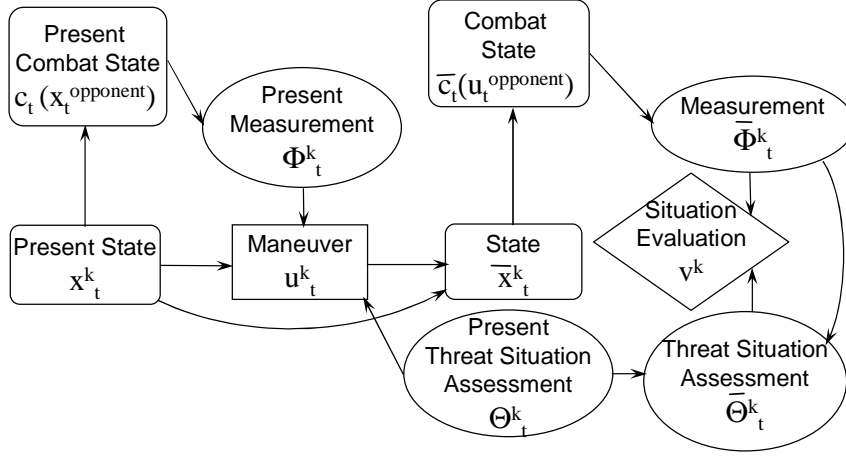


Figure 3: A parameterized ID representing the control decision of player k , $k = D, A$.

is either symmetric or asymmetric. The asymmetric information structure means that the players have different information either of the structure of the game or the current game situation (e.g., Gibbons, 1992). It is assumed that the players cannot observe the exact state of the combat, which causes the imperfect information (Basar and Olsder, 1995).

The information structure of the IDG cannot be defined by using the standard methodology of IDs because the time precedence of decision nodes should be fixed and defined by directed arcs. We solve this problem by introducing a new meaning for an arc in a two-player IDG. A dashed arc between the decision nodes of the players implies the information structure of the game. Asymmetric information is indicated by an arc that leads from one decision node to the other one. The IDG of this type represents a Stackelberg game that is a model for a leader-follower situation (e.g., Basar and Olsder, 1995). The initial node of the arc contains the controls of the leader who assumes or is aware of the opponent's preference model and the terminal node refers to the follower who makes the control decision based on state observations and utilizes information about the leader's decisions during the decision making process. A two-way dashed arc refers to a symmetric information structure, i.e., both the players have the knowledge of the opponent's probabilities, utilities, and information. Then, the IDG admits a Nash equilibrium (see, e.g., Basar and Olsder, 1995).

The dashed arc in the IDG shown in Fig. 2 implies that the DM knows or assumes the game situation of the AD, i.e., the DM acts as the leader of a Stackelberg game. By reversing the information structure arc, the AD would utilize knowledge of the DM's game situation in the decision making. A two-way dashed arc between the decision nodes of the IDG would imply the symmetric information, i.e., both the players act by knowing all the elements of the game.

4 Solution of the influence diagram game

When designing the best control strategy of the DM against the worst possible action of the AD, a Nash or a Stackelberg equilibrium of the IDG must be solved. In the first place, the information available to the players is assumed symmetric. Then, the Nash

equilibrium solution providing the best control of the DM under the assumption that the AD behaves in the worst possible way is determined. If there is no unique Nash equilibrium, the asymmetric information is assumed and the Stackelberg equilibrium solution is applied in the design of the worst case control. Such an equilibrium is always admitted by non-zero-sum games (see, e.g., Basar and Olsder, 1995).

A solution approach for the IDG introduced here is based on decomposition of the original game model as well as on decision tree representations. First, two parameterized IDs representing the control decision from the viewpoint of a single player are constructed. Then, the best control strategies providing the highest expected utility are determined by converting the IDs into decision trees and solving these trees. The use of the expected utility maximization criterion is justified by the utility theoretical definition of rationality (see, e.g., Keeney and Raiffa, 1976).

The decomposition of the IDG is carried out such that the control and state variables of the AD are considered as given parameters in the ID that models the control decision of the DM, and the control and the state of the AD are parameters in the DM's decision model, respectively. The resulting parameterized ID representation is shown in Fig. 3.

The parameterized IDs are solved by using their decision tree representations, see Fig. 4. Each path of the tree from the *Maneuver* node to the *Situation Evaluation* node gives a particular utility $V_{i,j}^k(u_t^D, u_t^A) := v^k(\tilde{\theta}_i^k, \tilde{\phi}_j^k(u_t^D, u_t^A))$ whose probability is $P_{i,j}^k(u_t^D, u_t^A) := P(\tilde{\phi}_j^k)P(\tilde{\theta}_i^k|\tilde{\phi}_j^k(u_t^D, u_t^A))$, $i = 1, \dots, 4$, $j = 1, \dots, n_\phi$, $k = D, A$. The expected utility obtained with the controls u_t^D and u_t^A is

$$J^k(u_t^D, u_t^A) = \sum_{j=1}^{n_\phi} \sum_{i=1}^4 P_{i,j}^k(u_t^D, u_t^A) V_{i,j}^k(u_t^D, u_t^A), \quad (9)$$

that is used as the game's payoff function for player k . The DM aims at maximizing J^D and the AD J^A , respectively.

Different solution types of the IDG are described by utilizing the concept of optimal reactions (see, e.g., Basar and Olsder, 1995). The optimal reaction of the DM, denoted by $u_t^{D^\circ}(u_t^A)$, is a function of the AD's control. It returns such a value for the DM's control that maximizes the payoff function J^D when the AD is employing the control u_t^A . The optimal reaction of the DM is expressed as

$$u_t^{D^\circ}(u_t^A) = \arg \max_{u_t^D \in S^D} J^D(u_t^D, u_t^A), \quad (10)$$

i.e., the payoff function of the DM is maximized with respect to the control of the DM subject to the feasible control set of the DM and the given control value of the AD. In the same way, one can define the optimal reaction for the AD that is denoted by $u_t^{A^\circ}(u_t^D)$.

The Nash equilibrium solution of the IDG, $(u_t^{D^N}, u_t^{A^N})$, must satisfy

$$\begin{aligned} J^D(u_t^D, u_t^{A^N}) &\leq J^D(u_t^{D^N}, u_t^{A^N}) \quad \forall u_t^D \in S^D, \\ J^A(u_t^{D^N}, u_t^A) &\leq J^A(u_t^{D^N}, u_t^{A^N}) \quad \forall u_t^A \in S^A. \end{aligned} \quad (11)$$

The Nash equilibrium is such that neither player wants to deviate from it. Note that the definition of the Nash equilibrium can also be expressed with the help of the optimal reactions by requiring $u_t^{D^\circ}(u_t^{A^N}) = u_t^{A^\circ}(u_t^{D^N})$.

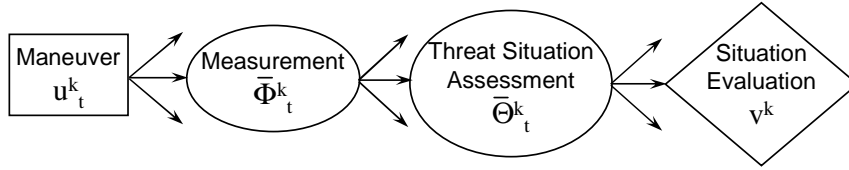


Figure 4: A simplified decision tree representation of the parameterized ID representing the control decision of player k , $k = D, A$.

If the design of the DM's best control strategy is based on a Stackelberg equilibrium, it is assumed that the DM, whose decision process and belief about the AD's decision process are represented in the IDG, acts as the leader. Then, the DM selects his or her best control by utilizing all the information about the game. The follower, i.e., the AD in the IDG shown in Fig. 2, chooses the best control of his or her own by knowing the decision of the DM. Because the DM is aware of the utilities and probabilities of the AD, he or she also knows the payoff function of the AD and thus the optimal reaction of the AD can be determined. The Stackelberg equilibrium solution of the IDG, (u_t^{DS}, u_t^{AS}) , must satisfy

$$\begin{aligned} J^D(u_t^D, u_t^{A^\circ}(u_t^D)) &\leq J^D(u_t^{DS}, u_t^{A^\circ}(u_t^{DS})) \quad \forall u_t^D \in S^D, \\ J^A(u_t^{DS}, u_t^A) &\leq J^A(u_t^{DS}, u_t^{AS}) \quad \forall u_t^A \in S^A. \end{aligned} \quad (12)$$

The definition of a Stackelberg equilibrium solution could also be given in the case in which the AD acts as the leader.

When the control variables of the players are discrete, i.e., $u_t^k \in S^k = \{d_1^k, \dots, d_{n_k}^k\}$, $k = D, A$, the IDG can be represented in strategic form that is an $n_D \times n_A$ bimatrix. The form displays the sets of feasible decision alternatives S^D and S^A available to each player and the values of the players' payoffs $J^k(d_i^D, d_j^A)$, $i = 1, \dots, n_D$, $j = 1, \dots, n_A$, $k = D, A$, defined by (9), related to each combination of the players' control alternatives. Nash and Stackelberg equilibria of the original IDG are obtained from this matrix game.

An ID containing continuous decision variables and discrete random variables can be converted into a nonlinear optimization problem. This correspondence is introduced in Virtanen et al. (2004a). Hence, when the control variables of the players are continuous, i.e., $u_t^k \in S^k \subset R$, $k = D, A$, the parameterized ID can be expressed as a nonlinear optimization problem in which the objective function is the expected utility and the constraints consist of the relations between the variables of the IDG. The parameterized ID of player k leads to the optimization problem of the form

$$\max_{u_t^k \in S^k} \sum_{j=1}^{n_\phi} \sum_{i=1}^4 P(\tilde{\phi}_j^k) P(\tilde{\theta}_i^k | \tilde{\phi}_j^k(u_t^D, u_t^A)) v^k(\tilde{\theta}_i^k, \tilde{\phi}_j^k(u_t^D, u_t^A)) \quad (13)$$

subject to

$$\begin{aligned} P(\tilde{\theta}_i^k | \tilde{\phi}_j^k(u_t^D, u_t^A)) &= H^k[P(\theta_i^k), \tilde{\phi}_j^k(u_t^D, u_t^A)], \\ i &= 1, \dots, 4, \quad j = 1, \dots, n_\phi, \end{aligned} \quad (14)$$

$$\tilde{\phi}_j^k(u_t^D, u_t^A) = \Omega^k(\tilde{c}_t(u_t^D, u_t^A); j), \quad j = 1, \dots, n_\phi, \quad (15)$$

$$\tilde{c}_t(u_t^D, u_t^A) = g(\tilde{x}_t^D(u_t^D), \tilde{x}_t^A(u_t^A)), \quad (16)$$

$$\tilde{x}_t^l(u_t^l) = x_t^l + f^l(x_t^l, u_t^l) \Delta t, \quad l = D, A. \quad (17)$$

Here, the functions g as well as H^k , Ω^k , f^k , and v^k are defined by (2), (7), (4), (1), and (8), respectively. For simplicity, the upcoming states of the players in (17) are determined by using the Euler method in the integration of Eq. (5). The given parameter set of the problem (13)-(17) contains the states of the players at the decision instant x_t^k , $k = D, A$, as well as the prior threat probabilities $P(\theta_i^k)$, $i = 1, \dots, 4$, the observation probabilities $P(\tilde{\phi}_j^k)$, $j = 1, \dots, n_\phi$, and the control of the opponent u_t^\bullet .

In the solution of a Nash equilibrium, the optimization problems representing both the DM's and the AD's control problem must be solved simultaneously. This leads to the solution of a set of nonlinear equations consisting of the necessary conditions of optimality for both the optimization problems. Due to the nonconvex objective function and constraints of the problems, the necessary conditions provide only a solution candidate for the Nash equilibrium. Its existence must be confirmed by examining that the solution satisfies the conditions (11).

From the optimization point of view, a two-player Stackelberg game is a bilevel optimization problem (see, e.g., Bard, 1998; Ehtamo and Raivio, 2001). The leader maximizes his or her payoff function subject to the follower's optimization problem that gives the optimal reaction of the follower. Recall that the DM acts as the leader and the AD is the follower. Then, a Stackelberg equilibrium for the IDG is obtained by solving the bilevel optimization problem that consists of an upper level problem referring to the DM's optimization task, i.e., k is set to D in (13)-(17), and a lower level problem referring to the AD's optimization task, i.e., k is set to A in (13)-(17). In the bilevel optimization, the latter problem is regarded as a constraint of the former one.

A way to solve the bilevel optimization problem is to take the necessary optimality conditions of the lower level problem as constraints to the upper level problem (see, e.g., Bard, 1998), and then to solve a standard optimization problem using nonlinear programming. The resulting solution is only a candidate for the Stackelberg equilibrium because of the nonconvexity of the objective function and the constraints. The validity of the solution candidate must be checked by studying whether it satisfies the conditions (12).

5 Numerical example

In this section, the presented IDG is utilized in a simulation procedure that produces a short-sighted qualitative solution for a two-target game (see, e.g., Grimm and Well, 1991) modeling one-on-one air combat. In such a game, both players have a target set defined by (3) and they attempt to drive the state of the game into own target set without first being driven into the target set of the adversary.

In the simulation procedure, the control decisions are taken at discrete decision stages and the players act simultaneously. In the beginning, the initial states and the initial threat probabilities of the players must be fixed. At each decision stage, the expected utility maximizing game optimal controls are obtained by solving the Nash, or if necessary, Stackelberg equilibrium of the IDG with a solution approach described in the previous section. The evolution of the players' states is computed by integrating the state equation (1) with the resulting optimal controls until the next decision stage is reached. During the simulation, the threat probability distributions are updated such that the prior probabilities at the current stage are associated with the posterior probabilities of the previous

stage. The simulation will be continued until one of the terminal conditions referring to the conditions $\Psi^k(\cdot) \leq 0$ in (3) is satisfied. This gives us the final time of the two-target game. If the states of the players do not satisfy the terminal conditions within a predetermined maximum duration, the game terminates in a draw. On the other hand, when the state of the game enters both target sets at the same time, the outcome is a joint capture.

In the following numerical example, a three degrees of freedom point-mass model representing the dynamics of the players' aircraft is used as the state equation (1). The control variables of the point-mass model are discretized and its parameters correspond a generic fighter aircraft. For illustrative purposes, the parameters are chosen such that the DM's aircraft is more agile than the AD's one. The utility functions and the probability distributions required in the IDG are identical for both the players. The terminal conditions refer to a combat situation where one player has achieved the other player's tail position. The aircraft and preference models are described in Virtanen et al. (2004b).

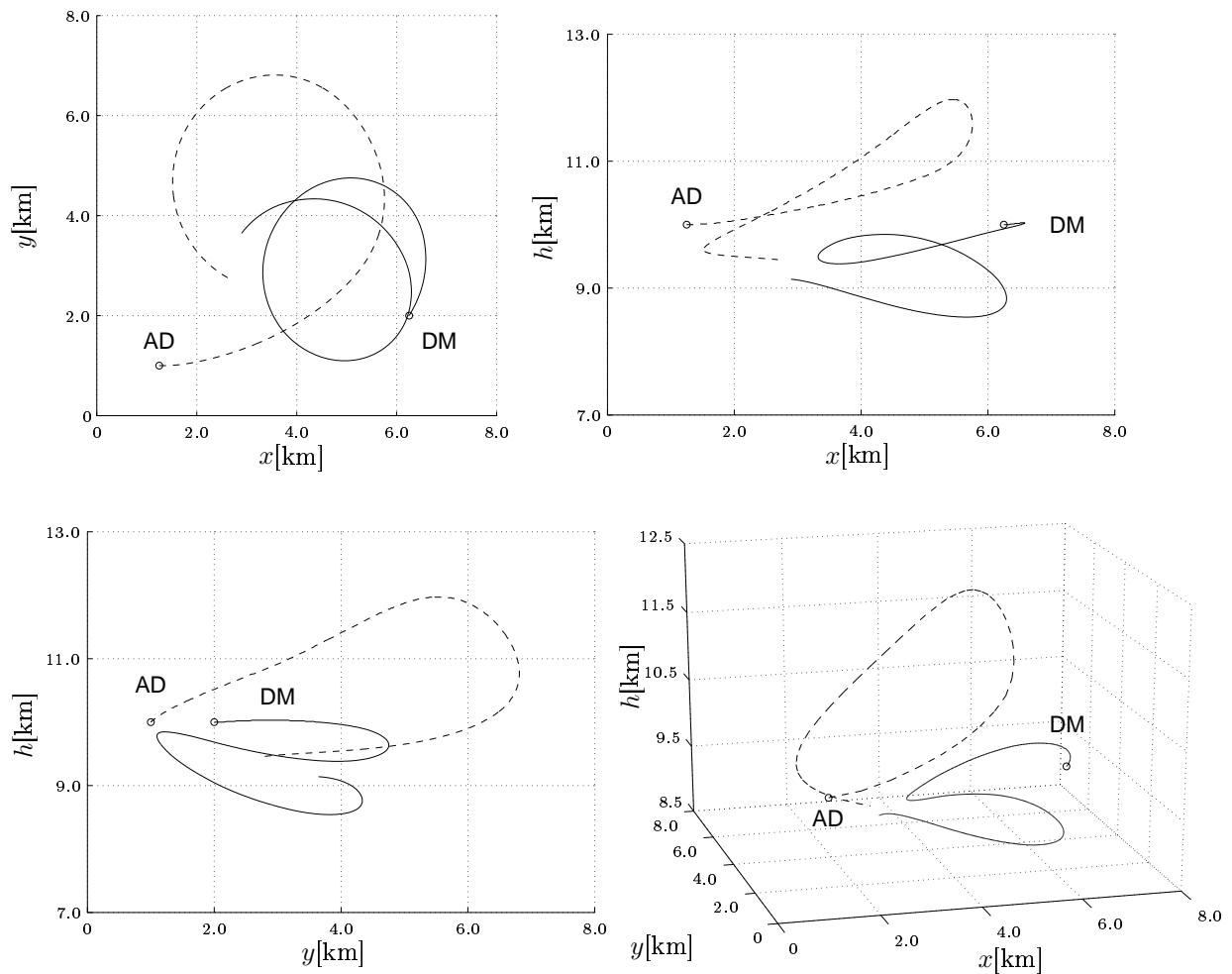


Figure 5: The expected utility maximizing game optimal trajectories of the players. The solid curves refer to the DM and the dashed curves to the AD, respectively. Circles denote the initial states of the players.

In the beginning of the combat game, the AD is heading toward the DM's tail whereas the DM is flying away from the AD. The combat situation is thus initially advantageous for

the AD and disadvantageous for the DM. The projections of the game optimal trajectories of the players on the x,y-, x,h-, and y,h-planes as well as the three-dimensional solution trajectories are shown in Fig. 5. The probability distributions of the players' threat assessments are presented in Fig. 6. The more agile DM, although being initially pursued, achieves a winning position in 98 s. The probability of the "Disadvantage" outcome of the DM's threat assessment is relatively high until 60 s, and after that, the probability declines rapidly. During the rest of the game, the probability of the "Advantage" outcome for the DM rises almost to unity whereas for the AD it descends practically to zero. In the end, the DM reaches the tail position of the AD and wins the combat game.

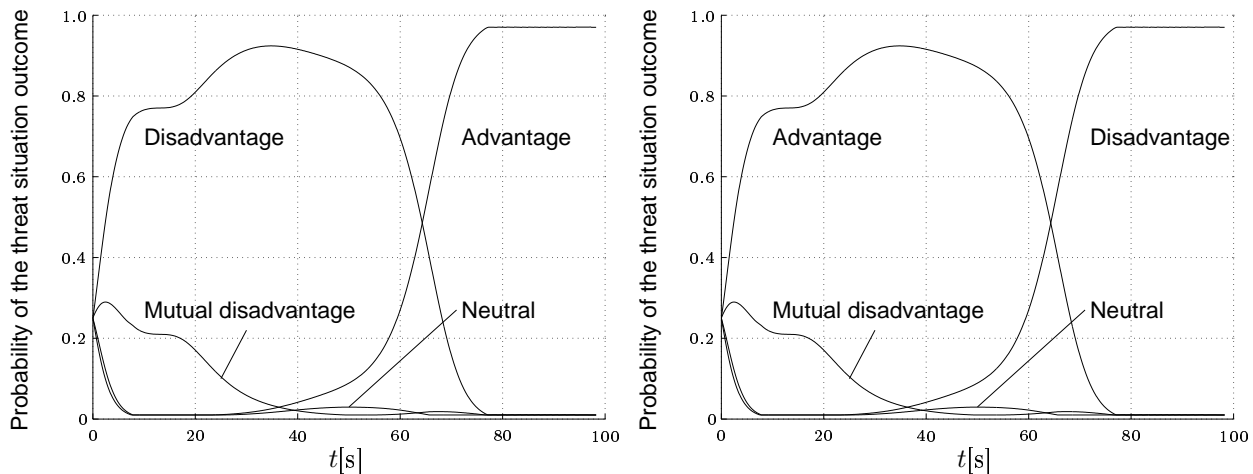


Figure 6: The threat probability distributions of the players. The left graph refers to the DM and the right to the AD, respectively.

Based on the solution of the example, the IDG seems to produce reasonable controls because the more agile DM wins the combat game by choosing the controls that eventually lead to an advantageous situation although the initial state is advantageous for the AD. Since the AD's aircraft is less agile, the AD cannot exploit his or her advantageous initial state but is eventually forced to evade the DM. It should be noted that due to the limited planning horizon of the IDG, the expected utility maximizing control sequences are only suboptimal solutions for the two-target game in the global sense.

6 Discussion

When determining control strategies in an air combat game, one has to take into account the preferences and goals of the players as well as the uncertainty associated, e.g., with state observations. In addition, the selection of controls is affected by the behavior of the adversary. In this paper, these features are incorporated into a pilot's control model by representing the multiobjective control decision problem with an influence diagram game (IDG).

Depending on the information structure of the IDG model, it admits a Nash or a Stackelberg equilibrium that can be considered as the expected utility maximizing control strategy for one player when the adversary behaves in the worst possible way. The equilibrium solutions are obtained in a feedback form, i.e., the best controls are given as a

function of all the available information at the particular decision instant but they are myopic due to a short planning horizon used in the IDG. If we aim at achieving better solutions considering the overall success of a pilot in the air combat, a control sequence maximizing the aggregated utility over the total duration of the combat should be found. To achieve this, the IDG must be able to predict the future states of the combat further than one decision interval ahead.

Virtanen et al. (2004a) introduce a multistage influence diagram that takes into account the interaction of several successive maneuvering decisions in a one-on-one air combat setting in which the trajectory of the adversary is predetermined. In the same way, one can extend the single stage IDG presented in this paper into a multistage IDG that represents the players' sequential decision making process (Virtanen et al., 2004b). Then, the combat situation is evaluated separately at each decision stage and the overall utility is calculated by summing the single utilities. The multistage IDG can be represented in the form of a discrete-time dynamic game and could be solved in an open-loop form with nonlinear programming. The open-loop form means that the best controls at all the decision stages are chosen concurrently without knowing the exact states of the game at the upcoming decision instants. On the other hand, suboptimal control sequences could be obtained in a feedback form by using a moving horizon control, or equivalently model predictive control (e.g., Camacho and Bordons, 1999), approach. In this approach, the time horizon of the original multistage game is truncated, and a feedback Nash or Stackelberg equilibrium of the dynamic game lasting only a limited planning horizon is determined and implemented at each decision stage. A similar moving horizon control approach for dynamic discrete-time games is presented in Cruz et al. (2002).

The single stage IDG in which the control variables are discrete can be solved in real-time and could be applied, e.g., in decision making systems of air combat simulators. It could also provide a way to produce so-called reprisal strategies (Kelley et al., 1980) that utilize the nonoptimal behavior of an adversary in a two-target game. In addition, combining the IDG with a simulation procedure presented in the paper, one can obtain short-sighted game optimal controls over the total duration of two-target games. Thus far, such games have proven to be intractable and solutions are obtained only for small problems containing simple dynamics models.

Overall, the presented IDG offers a structured and transparent way to model and simulate the pilot's control decision in a dynamic and uncertain one-on-one air combat setting. The graphical representation of the IDG enables air combat experts to be involved in the modeling and structuring process of the game because it can be easily understood by individuals with a little decision theoretic and mathematical background. It should be noted that the graphical representation of the IDG coincides with the multi-agent influence diagram framework introduced by Koller and Milch (2001, 2003). However, the game formulation described in the paper contains an explicit model for the underlying dynamics of the decision environment that is omitted in the formalism of multi-agent influence diagrams. Although the paper at hand does not introduce a well-defined formalism related to dynamic influence diagram games, a similar IDG representation may also be utilized in other application areas in which a dynamic and uncertain system is controlled by several agents whose goals are conflicting.

References

- Austin, F., Carbone, G., Falco, M., Hinz, H., and Lewis, M., "Game Theory for Automated Maneuvering During Air-to-Air Combat," *Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 6, 1990, pp. 1143-1149.
- Bard, J.F., *Practical Bilevel Optimization, Algorithms and Applications*, Kluwer, Dordrecht, Netherlands, 1998, pp. 5-8.
- Basar, T., and Olsder, G., *Dynamic Noncooperative Game Theory*, 2nd ed., Academic Press, London, 1995.
- Camacho, E.F., and Bordons, C., *Model Predictive Control*, Springer, London, England, 1999.
- Cruz, J.B., Simaan, M.A., Gacic, A., and Liu, Y., "Moving Horizon Nash Strategies for a Military Air Operation," *IEEE Transaction on Aerospace and Electronic Systems*, Vol. 38, No. 3, 2002, pp. 989-999.
- Ehtamo, H., and Raivio, T., "On Applied Nonlinear and Bilevel Programming for Pursuit-Evasion Games," *Journal of Optimization Theory and Applications*, Vol. 108, No. 1, 2001, pp. 65-96.
- Gibbons, R., *Game Theory for Applied Economists*, Princeton University Press, Princeton, NJ, 1992.
- Gmytrasiewicz, P.J., "Issues in Rational Planning in Multi-Agent Settings," *Proc. of the 36th Annual Hawaii International Conference on System Sciences*, Hawaii, 2003, pp. 84-91.
- Grimm, W., and Well, K.H., "Modelling Air Combat as Differential Game, Recent Approaches and Future Requirements," In Hämäläinen, R.P., and Ehtamo, H., eds., *Differential Games - Developments in Modeling and Computation*, Springer, Berlin, 1991, pp. 1-13.
- Howard, R.A., and Matheson, J.E., "Influence Diagrams," In Howard, R.A., and Matheson, J.E., eds., *The Principles and Applications of Decision Analysis*, Vol. 2, Palo Alto, CA, Strategic Decision Group, 1984, pp. 719-762.
- Isaacs, R., *Differential Games*, Wiley, New York, 1965.
- Katz, A., "Tree Lookahead in Air Combat," *Journal of Aircraft*, Vol. 31, No. 4, 1994, pp. 970-973.
- Keeney, R., and Raiffa, H., *Decision with Multiple Objectives*, Wiley, New York, 1976.
- Kelley, H.J., Eugene, M.C., and Lefton, L., "Reprisal Strategies in Pursuit Games," *Journal of Guidance and Control*, Vol. 3, No. 3, 1980, pp. 257-260.
- Koller, D., and Milch, B., "Multi-Agent Influence Diagram for Representing and Solving Games," *Proc. of 17th International Joint Conference on Artificial Intelligence*, Seattle, Washington, 2001, pp. 1027-1034.
- Koller, D., and Milch, B., "Multi-Agent Influence Diagram for Representing and Solving Games," *Games and Economic Behavior*, Vol. 45, No. 1, 2003, pp. 181-221.
- Lazarus, E., "The Application of Value-Driven Decision-Making in Air Combat Simulation," *Proc. of the IEEE International Conference on Systems, Man, and Cybernetics*, Orlando, FL, 1997, pp. 2302-2307.

- Le Menec, S., and Bernhard, P., "Decision Support System for Medium Range Aerial Duels Combining Elements of Pursuit-Evasion Game Solutions with AI Techniques," In Olsder, G.J., ed., *Annals of the International Society of Dynamic Games, Vol. 3: New Trends in Dynamic Games and Applications*, Birkhäuser, Boston, 1995, pp. 207-226.
- McManus, J.W., and Goodrich, K.H., "Application of Artificial Intelligence (AI) Programming Techniques to Tactical Guidance For Fighter Aircraft," *Proc. of the AIAA Guidance, Navigation, and Control Conference*, Boston, Massachusetts, 1989.
- Neumann, F., "On the Approximate Solution of Complex Combat Games," *Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 1, 1990, pp.128-136.
- Shachter, R.D., "Evaluating Influence Diagrams," *Operations Research*, Vol. 34, No. 6, 1986, pp. 871-882.
- Smith, J.Q., "Decision Influence Diagrams and Their Use," In Rios, S., ed., *Decision Theory and Decision Analysis: Trends and Challenges*, Kluwer, Boston, 1994, pp. 33-51.
- Stehlin, P., Hallkvist, I., and Dahlstrand, H., "Models for Air Combat Simulation," *Proc. of the 19th ICAS Congress, ICAS-94-8.6.2*, Anaheim, CA, 1994, pp. 2190-2196.
- Virtanen, K., Raivio, T., and Hämäläinen, R.P., "Decision Theoretical Approach to Pilot Simulation," *Journal of Aircraft*, Vol. 36, No. 4, 1999, pp. 632-641.
- Virtanen, K., Raivio, T., and Hämäläinen, R.P., "Modeling Pilot's Sequential Maneuvering Decisions by a Multistage Influence Diagram," *Journal of Guidance, Control, and Dynamics*, Vol. 27, No. 4, 2004a.
- Virtanen, K., Karelahti, J., Raivio, T., and Hämäläinen, R.P., "Modeling Air Combat by a Moving Horizon Influence Diagram Game," manuscript, submitted for publication, 2004b.