Virtanen, K., Karelahti, J., and Raivio, T., Modeling Air Combat by a Moving Horizon Influence Diagram Game, Journal of Guidance, Control, and Dynamics, Vol. 29, No. 5, 2006, pp. 1080-1091.

JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 29, No. 5, September–October 2006

Modeling Air Combat by a Moving Horizon Influence Diagram Game

Kai Virtanen,* Janne Karelahti,† and Tuomas Raivio‡
Helsinki University of Technology, FIN-02015 HUT, Espoo, Finland

DOI: 10.2514/1.17168

The paper describes a multistage influence diagram game for modeling the maneuvering decisions of pilots in one-on-one air combat. The game graphically describes the elements of the decision process, contains a model for the dynamics of the aircraft, and takes into account the pilots' preferences under conditions of uncertainty. The pilots' game optimal control sequences with respect to their preference models are obtained by solving the influence diagram game with a moving horizon control approach. In this approach, the time horizon of the original game is truncated, and a feedback Nash equilibrium of the dynamic game lasting only a limited planning horizon is determined and implemented at each decision stage. To demonstrate the influence diagram game and its aspects, examples with a point mass aircraft model are computed and analyzed. The game model presented in the paper offers a new way to analyze optimal air combat maneuvering and to develop an automated decision making system for selecting combat maneuvers in air combat simulators.

I. Introduction

THE paper introduces a nonzero-sum multistage influence diagram game that describes pilots' sequential control decisions in one-on-one air combat. The game takes into account the dynamics of the aircraft, the pilots' preferences, and the uncertainty related to the decision making in a structured and transparent way allowing, for example, the elicitation of preference information from the actual decision makers. As far as the authors know, the game model is the first application of influence diagrams in which the earlier ideas related to influence diagram game modeling and dynamic decision settings are taken into practice.

Because rigorous mathematical methods that result in practically feasible guidance solutions for realistic game models of one-on-one air combat are not available [1], it is imperative to consider some other approaches. An alternative is to consider how factors like preferences, perception, and beliefs reflect themselves in an air combat game setting. Overall, the modeling work should not bypass the fact that human influence is in some form always present in air combat. From a practical point of view, an important question is how to produce computer guided pilots that act in a controlled, explainable, and understandable manner? This issue is of particular interest to developers of decision making systems used, for example, in air combat simulators or unmanned aerial vehicles.

Influence diagrams [2,3] are directed acyclic graphs that describe a decision making situation and provide a methodological basis for the ranking of the available decision alternatives. They are closely related to decision trees (see, e.g., [4]) that originate from the theory of games in extensive form [5]. A diagram consists of decision, chance, and utility nodes and arcs connecting them. Influence diagrams separate the structural aspects of a specific decision problem from the aspects related to the opinions of the decision maker. The structured and well-laid foundations of the methodology make the modeling process transparent, traceable, and understandable also for the experts of the substance area. Considering the general usability of models, this is of utmost importance. For an introductory air combat related example of an influence diagram, see [4].

Originally, influence diagrams consider only one decision maker. An extension to multiple decision makers is mentioned in [6] and later implemented in [7,8] using concepts of noncooperative game theory [9]. On the other hand, in [10], an extension of the influence diagram concept into a dynamic multistage decision setting without game consideration is described. Although discussed on a general level in [11], this paper is the first elaboration where these ideas have been combined into a dynamic game setting. The setting is represented using a multistage influence diagram game containing an explicit model of the dynamic decision environment which is described with the equations of motion for an aircraft. Thus, here the term "dynamic game" refers to a game involving control and state variables that obey a given set of differential equations.

In the dynamic air combat influence diagram game, the preferences of the pilots are described by using utility functions [5]. The game model gives the cumulative utilities related to the discrete control alternatives of the players. The underlying uncertainties in the game combined with the cumulative utilities and the structure of the diagram result in a probability distribution of the utility for each control choice at each decision instant. The payoffs of the game associated with the cumulative expected utilities are calculated based on these distributions.

When solving a feedback Nash equilibrium [9] of the nonzero-sum influence diagram game, game optimal controls maximizing simultaneously both the payoffs are given as a function of the current state of the game. There exists a variety of solution techniques for ordinary influence diagrams [10,12]. For influence diagram games the situation is different. In [8], a divide and conquer-type solution approach where the game is divided into smaller subgames that are solved iteratively is presented. Unfortunately, the approach does not produce game optimal controls in a feedback form and thus it cannot be applied in the solution of the game at hand. If feedback solutions are preferred, dynamic programming [13] remains as the only alternative for solving the game.

A drawback attributed to dynamic programming is the combinatorial explosion of the computation. To circumvent this, we trade the solution of the complete game with computing time and apply a moving horizon control approach. The obtained solutions are suboptimal in the global sense, but they are calculated in a feedback form. Although originated from the field of control theory [14,15] where the approach is also known as receding horizon or model predictive control, it has been recently applied to dynamic games as

Received 14 April 2005; revision received 19 January 2006; accepted for publication 20 February 2006.

*D.Sc., Researcher, Systems Analysis Laboratory, P.O. Box 1100; kai.virtanen@hut.fi.

†Researcher, Systems Analysis Laboratory, P.O. Box 1100; janne.karelahti@hut.fi.

‡D.Sc., Docent, Systems Analysis Laboratory, P.O. Box 1100; tuomas.raivio@hut.fi.

well [16–18]. In the moving horizon control approach, the horizon of the original influence diagram game is truncated and a dynamic game with a short planning horizon is solved at each decision instant. The states of the players at the next decision instant are then obtained by applying the players' game optimal controls related to the truncated game.

II. Related Approaches

Quantitative analysis of air combat has its roots in optimal control theory [19] that allows the optimization of the controls of a single decision maker. For encounters with multiple decision makers, the theory of dynamic games [9] provides a suitable framework. Team problems involving M vs N players have received less attention in the open literature but situations involving two decision makers have been subject to a profound research effort.

In pursuit-evasion games [20], the roles of the two players are assumed fixed: one player, the pursuer, tries to reach the other player called the evader. The objective of the evader is to escape, but if this is not possible, maximize the cost of the capture that may be, for example, the elapsed time [21] or the distance to the border of a nation [22]. A capture is defined as the collection of states that belong to a specific target set of the game, that is, a capture means that the state of the game enters the target set.

The qualitative solution of a pursuit-evasion game divides the state space of the game into the sets of initial states that lead to a capture or evasion of the evader under the supposition that the players play optimally. Only an error of the adversary can change the outcome. For a capturability assessment of a missile and a fighter aircraft, see [23]. Quantitative solutions of the game announce the game optimal controls of the players for the states from which the pursuer can enforce a capture; see [24].

Pursuit-evasion games are suitable for describing situations where the roles of the players are fixed beforehand. Nevertheless, this is usually not the case for two aircraft committing to a combat. For such situations, two-target game models have been proposed [25,26]. In a two-target game, both players have a target set of their own, and the objectives of the players are to avoid the adversary's target set while attempting to drive the state of the game into his own target set. Four outcomes are possible: capture of one player by the other, joint capture, or draw.

Two-target games can be regarded as an extension to pursuit-evasion games. Similarly as in pursuit-evasion games, the state space can in principle be divided into regions with different outcomes by the qualitative solution of the game. For a comprehensive analysis with greatly simplified vehicle models, see [27].

Quantitative solutions for two-target games would give game optimal strategies for both players in different regions of the state space. However, there are severe modeling issues that hitherto have not been solved in a satisfactory way. First, even though the concept of game optimality is clear in the regions leading to a capture of either player, the fact that the game has two target sets has to be taken into account. For example, in the combat game approach [28], the condition that one must not enter the other player's target set leads to an ill-posed state variable inequality constraint involving both players but being on the responsibility of only the other player. In multicriteria games [29], the game has two objectives that represent, for example, the distances of the state of the game to the target sets of the players. Different solution concepts corresponding to concepts of multicriteria optimization are defined and outlined on a theoretical level [26], but none of these elaborations has, so far, led to practical applications.

The second unresolved aspect concerning quantitative solutions is that the concept of optimality in the regions not leading to a unique capture is not clear. In the region leading to a joint capture, the optimal behavior depends on the preference ordering of the players between joint capture and draw, and in the region leading to a draw, no unique way to determine the controls can be given. In [30], reprisal strategies in which the objective of the player is to reach the zone leading to his win before the adversary by utilizing nonoptimal behavior of the adversary are presented.

A further complication of the game approach is that in order to make sense, the quantitative solutions must be obtained in a feedback form. Complete feedback solutions of the games are computationally intractable. In practice, both the time and the control alternatives have to be discretized and the planning horizon must be truncated. In [18], a pursuit-evasion game is solved with a moving horizon control approach in which the evader is assumed to fly straight ahead for the duration of the planning horizon. In [31], one-stage matrix games with a heuristic payoff are used. In [32], an approximate solution where the time is discretized, and at each decision instant the players play a myopic matrix game with predetermined nonoptimal feedback guidance laws, is proposed. In [33,34], a helicopter duel as a myopic game tree where each node is associated with a score based on the players' kill and survival probabilities is considered.

Although the game approaches originate from profound mathematical analysis, the synthesis of control decisions is based on rule based reasoning or ranking approaches. The general principle is to first obtain the possible states of the combat after a given planning horizon by projecting each maneuver alternative into the future and by predicting the state of the adversary. Then, a score is assigned to each predicted combat state. Finally, the maneuver alternative leading to the highest score is executed.

In simple air combat expert systems, the combat states are evaluated by predetermined combat geometry rules [35]. More advanced systems [36] use a fixed set of questions representing different goals. The total value of each maneuver alternative is obtained by calculating the weighted sum of the normalized goal specific values. In a value-driven heuristic approach [37], the future combat states are described by a set of attributes, and different states are evaluated by using an additive value function (see, e.g., [38]) whose weights depend on the combat state and the goals of the pilots.

Common to these approaches is that unlike in game formulations, the behavior of the adversary is observed but nothing is assumed. This is inconsistent with the fact that also the adversary is a rational agent with similar goals as the decision maker. On the other hand, in game formulations, issues related to uncertainty and preference information have not been widely considered.

Influence diagrams have been proven useful in modeling actions of a single pilot both in short-sighted [4] and long-sighted [10] decision settings. In the above-mentioned studies, the state of the adversary is given. Thus far, extensions to dynamic air combat settings where both players optimize their maneuvering decisions short sightedly are considered in [11]. The paper at hand extends the earlier influence diagram models to a multistage influence diagram game. It unifies the game analysis aspects with the simulation synthesis and provides a model that is capable of answering analysis questions and synthesis needs.

III. Air Combat Game

The players of the air combat game, hereafter referred to as blue and red, are assumed to maneuver in three dimensions and follow a set of differential equations (A17) described in the Appendix. The objective of the players is to reach their own target sets before the adversary. In practice, the target set could represent the firing envelope of a missile [39].

In the following, the air combat game is modeled by using the influence diagram that contains the dynamics of the aircraft, the pilot's preferences, and a probabilistic belief model of the pilots. The influence diagram representation of the air combat game, not to be confused with graph theoretic nets, is shown in Fig. 1. The upper and lower parts of the game model consist of blue's and red's variables at discrete time steps over the duration of the game, respectively. The influence diagram consists of decision, chance, deterministic, and value nodes depicted by squares, ovals, rounded squares, and diamonds, respectively. They represent the decisions to be made, uncertain variables, deterministic inputs, and quantities to be optimized. The directed arcs show the relationships between the various types of nodes in the diagram. Arcs directed toward a decision node imply the available information for the player at the moment of decision making. When directed toward a chance node,

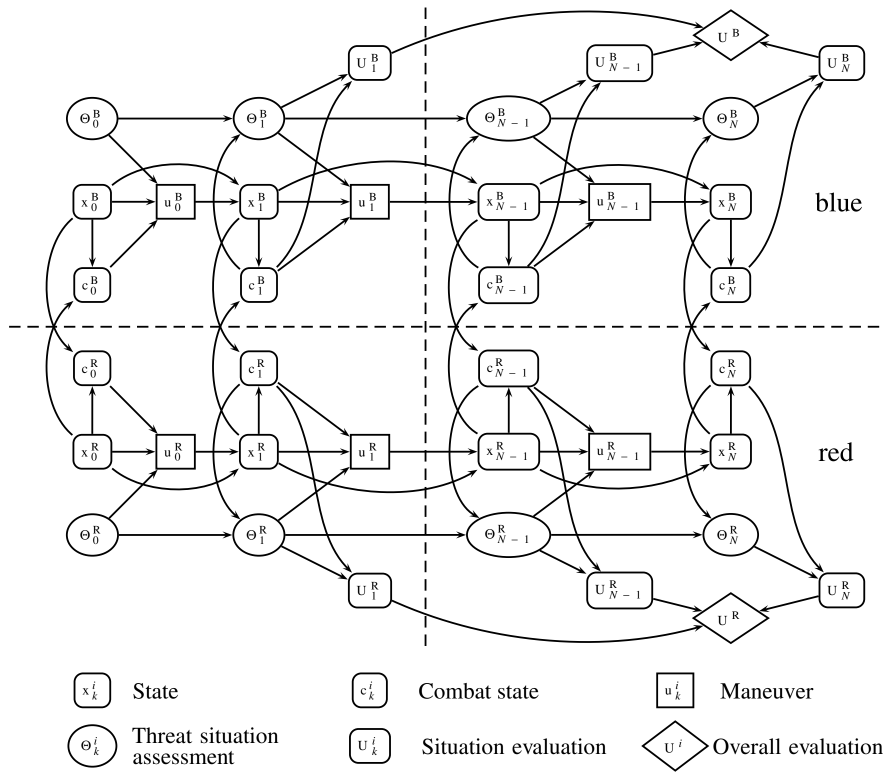


Fig. 1 Influence diagram representation of the air combat game.

an arc means that the node is conditionally dependent on the input node. An arc directed toward a deterministic or value node specifies that the value of the node is partially determined by the outcome of the input node.

In the air combat game, the control decisions of the players are made at discrete time instants $t_k = k\Delta t$, where $k = 0, \ldots, N-1$ is the stage counter and Δt is a fixed time interval between two successive decision stages. The terminal time of the game is given by $t_f = N\Delta t$ where N is the final stage of the game.

At stage k, the state of the player i is given in the deterministic state node by

$$\mathbf{x}_k^i = \left[x_k^i \quad y_k^i \quad h_k^i \quad \gamma_k^i \quad \chi_k^i \quad v_k^i\right]^T$$

$i = B, R$. The state variables x_k^i, y_k^i, and h_k^i are the horizontal coordinates and the altitude of the player, respectively, whereas the remaining state variables are the flight path angle γ_k^i, the heading angle χ_k^i, and the velocity v_k^i at stage k. The state is updated by integrating the state equations (A17) as

$$\mathbf{x}_{k+1}^i = \mathbf{x}_k^i + \int_{t_k}^{t_{k+1}} \mathbf{f}^i\left(\mathbf{x}^i, \mathbf{u}_k^i\right) \mathrm{d}t \tag{1}$$

where the control vector

$$\mathbf{u}_k^i = \left[\alpha_k^i \quad \eta_k^i \quad \mu_k^i\right]^T$$

containing the angle of attack α_k^i, the throttle setting η_k^i, and the bank angle μ_k^i is given in the maneuver decision node. The controls and the states of the players are constrained by (A10–A14) and (A15) and (A16) given in the Appendix. These constraints are written more concisely as

$$\mathbf{g}\left(\mathbf{x}_k^i, \mathbf{u}_k^i\right) \leq \mathbf{0} \tag{2}$$

$$\mathbf{h}\left(\mathbf{x}_k^i\right) \leq \mathbf{0} \tag{3}$$

The control and state constraints as well as the examination of the target set conditions are handled in the solution of the influence diagram game described in Sec. IV.A.

The state variables themselves are incomprehensible in the game analysis. We therefore define a function $\mathbf{q}^i$ describing the essential features as a combat state vector $\mathbf{c}_k^i$ that is computed in the deterministic combat state node as follows:

$$\mathbf{c}_k^i = \mathbf{q}^i\left(\mathbf{x}_k^B, \mathbf{x}_k^R\right), \quad i = B, R \tag{4}$$

The combat state vector can be defined in numerous ways [10,40]. Here, we select

$$\mathbf{c}_k^i = \left[\omega_k^i \quad \theta_k^i \quad d_k^i\right]^T \tag{5}$$

where the components of the combat state vector for blue are

$$\omega_k^B = \arccos\left[\mathbf{d}_k^B \cdot \mathbf{v}_k^B \Big/ \left(d_k^B v_k^B\right)\right] \tag{6}$$

$$\theta_k^B = \arccos\left[\mathbf{d}_k^B \cdot \mathbf{v}_k^R \Big/ \left(d_k^B v_k^R\right)\right] \tag{7}$$

$$d_k^B = \sqrt{\left(x_k^R - x_k^B\right)^2 + \left(y_k^R - y_k^B\right)^2 + \left(h_k^R - h_k^B\right)^2} \tag{8}$$

Above, $\mathbf{v}_k^i$ refers to the velocity vector of the player i given by

$$\mathbf{v}_k^i = \left[\dot{x}_k^i \quad \dot{y}_k^i \quad \dot{h}_k^i\right]^T \tag{9}$$

whereas the line-of-sight vector of blue equals

$$\mathbf{d}_k^B = \left[x_k^R - x_k^B \quad y_k^R - y_k^B \quad h_k^R - h_k^B\right]^T \tag{10}$$

Here the bearing $\omega_k^B \in [0, 180 \text{ deg}]$ and the angle-off $\theta_k^B \in [0, 180 \text{ deg}]$ are the angles between the line-of-sight vector of blue and the velocity vectors of blue and red, respectively, and $d_k^B \geq 0$ is the distance between the players at stage k; see Fig. 2. Red's combat

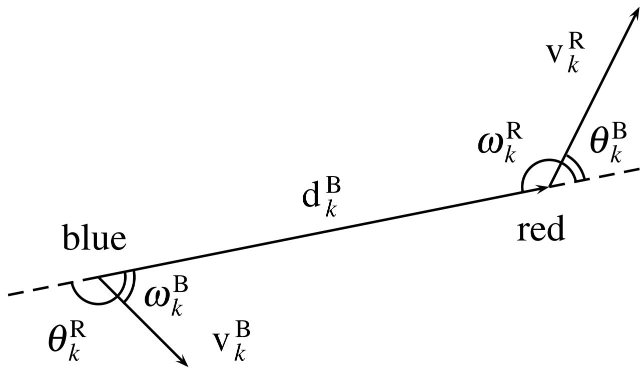


Fig. 2 Combat state variables.

state vector is obtained by swapping the indices B and R in Eqs. (6–8). The combat state vectors of the players describe the state of the duel from a viewpoint of a particular player.

During the game, each player tries to drive his combat state vector into his own target set and at the same time prevent the other player from reaching the other target set. The target set of player i is given by

$$T^i = \left\{\mathbf{c}_k^i \middle| \Psi\left(\mathbf{c}_k^i\right) \leq \mathbf{0}\right\}$$
$$\Psi\left(\mathbf{c}_k^i\right) = \left[\omega_k^i - \omega_f^i \quad \theta_k^i - \theta_f^i \quad d_k^i - d_f^i\right]^T \tag{11}$$

The threshold values of the combat state variables, denoted by ω_f^i, θ_f^i, and d_f^i, are fixed.

The game terminates when either or both players manage to drive their combat state vectors into the respective target sets. The game has four possible outcomes that are shown in Table 1. Furthermore, if neither combat state vector ends up in a target set within the maximum number of stages of the game, denoted by $N_{\max}$, the game terminates. The number of stages in the game is therefore given by

$$N = \min\left\{k \middle| \mathbf{c}_k^{\mathrm{B}} \in T^{\mathrm{B}} \quad \text{or} \quad \mathbf{c}_k^{\mathrm{R}} \in T^{\mathrm{R}} \quad \text{or} \quad k = N_{\max}\right\} \tag{12}$$

In addition to the state information, the players assess the threat situation of the combat at each stage. The threat assessment is modeled by a discrete random variable Θ_k^i given in the threat situation assessment node. The variable infers the threat situation from a viewpoint of the player at a particular stage. For blue, the four possible outcomes are shown in Table 2. For red, the outcomes are obtained from the same table by swapping the descriptions of the second and third rows.

Table 1 Outcomes of the air combat game

Outcome	Condition
Blue wins	$\mathbf{c}_N^B \in T^B$ and $\mathbf{c}_N^R \notin T^R$
Draw	$\mathbf{c}_N^B \notin T^B$ and $\mathbf{c}_N^R \notin T^R \Rightarrow N = N_{\max}$
Joint capture	$\mathbf{c}_N^B \in T^B$ and $\mathbf{c}_N^R \in T^R$
Red wins	$\mathbf{c}_N^R \in T^R$ and $\mathbf{c}_N^B \notin T^B$

Table 2 Outcomes of the threat situation assessment for blue

Outcome	Θ_k^B	Description
Neutral	1	Neither player pursues[a] the other one, or long distance.
Advantage	2	Blue pursues red who evades,[b] short distance.
Disadvantage	3	Red pursues blue who evades, short distance.
Mutual disadvantage	4	The players pursue each other, short distance.

[a]The player flies towards the adversary. [b]The player flies away from the adversary.

The probabilities of the outcomes given the momentary combat state are denoted by $P(\Theta_k^i = j \mid \mathbf{C} = \mathbf{c}_k^i)$, $j = 1, \ldots, 4$, where the elements of $\mathbf{C}$ are continuous random variables for the corresponding combat state variables. The probabilities sum up to 1, that is, $\sum_{j=1}^4 P(\Theta_k^i = j \mid \mathbf{C} = \mathbf{c}_k^i) = 1$. At each stage, these probabilities can be considered as prior beliefs of the succeeding stage's threat situation outcome probabilities. Based on these and the succeeding combat state, the posterior beliefs are computed using Bayes' theorem (see, e.g., [41]) as

$$\begin{aligned} P\left(\Theta_{k+1}^i = j \middle| \mathbf{C} = \mathbf{c}_{k+1}^i\right) &= \frac{P\left(\Theta_{k+1}^i = j\right) p^i\left(\mathbf{c}_{k+1}^i \middle| \Theta_{k+1}^i = j\right)}{\sum_{\ell=1}^4 P\left(\Theta_{k+1}^i = \ell\right) p^i\left(\mathbf{c}_{k+1}^i \middle| \Theta_{k+1}^i = \ell\right)} \\ &= \frac{P\left(\Theta_k^i = j \middle| \mathbf{C} = \mathbf{c}_k^i\right) p^i\left(\mathbf{c}_{k+1}^i \middle| \Theta_{k+1}^i = j\right)}{\sum_{\ell=1}^4 P\left(\Theta_k^i = \ell \middle| \mathbf{C} = \mathbf{c}_k^i\right) p^i\left(\mathbf{c}_{k+1}^i \middle| \Theta_{k+1}^i = \ell\right)} \end{aligned} \tag{13}$$

where the prior probabilities $P(\Theta_{k+1}^i = j)$ are equal to the preceding posterior probabilities $P(\Theta_k^i = j \mid \mathbf{C} = \mathbf{c}_k^i)$. It is assumed that the elements of $\mathbf{C}$ given a particular threat situation outcome are independent. Hence, the probability density function of the combat state given the threat situation outcome is calculated by

$$p^i\left(\mathbf{c}_k^i \middle| \Theta_k^i = j\right) = p^{\omega,i}\left(\omega_k^i \middle| \Theta_k^i = j\right) p^{\theta,i}\left(\theta_k^i \middle| \Theta_k^i = j\right) \times p^{d,i}\left(d_k^i \middle| \Theta_k^i = j\right) \tag{14}$$

Here, the time-invariant likelihood functions $p^{\omega,i}(\cdot)$, $p^{\theta,i}(\cdot)$, and $p^{d,i}(\cdot)$ represent the distributions of the combat state variables assuming that the player's threat situation outcome at stage k is j. Examples of likelihood functions are introduced in Sec. V.A. The probabilities of the threat situation outcomes at stage k can be written in shorthand as

$$\mathbf{p}_k^i\left(\mathbf{c}_k^i\right) = \begin{bmatrix} P\left(\Theta_k^i = 1 \middle| \mathbf{C} = \mathbf{c}_k^i\right) \\ P\left(\Theta_k^i = 2 \middle| \mathbf{C} = \mathbf{c}_k^i\right) \\ P\left(\Theta_k^i = 3 \middle| \mathbf{C} = \mathbf{c}_k^i\right) \\ P\left(\Theta_k^i = 4 \middle| \mathbf{C} = \mathbf{c}_k^i\right) \end{bmatrix} \tag{15}$$

where the P's are computed according to Eq. (13).

The combat situation at stage k is evaluated by a utility function that reflects the player's preferences in different combat states. The utilities associated with the different threat situation outcomes j are computed in the deterministic situation evaluation node as

$$\mathbf{U}_k^i\left(\mathbf{c}_k^i\right) = \left[U^i\left(1, \mathbf{c}_k^i\right) \quad U^i\left(2, \mathbf{c}_k^i\right) \quad U^i\left(3, \mathbf{c}_k^i\right) \quad U^i\left(4, \mathbf{c}_k^i\right)\right]^T \tag{16}$$

where

$$U^i\left(j, \mathbf{c}_k^i\right) = w_j^{\omega,i} u_j^{\omega,i}\left(\omega_k^i\right) + w_j^{\theta,i} u_j^{\theta,i}\left(\theta_k^i\right) + w_j^{d,i} u_j^{d,i}\left(d_k^i\right), \quad j = 1, \ldots, 4 \tag{17}$$

The utilities are linear combinations of single-attribute utility functions that map an attribute to a utility scale such that the worst and the best values of an attribute correspond to zero and one, respectively. Examples of suitable single-attribute utility functions are given in Sec. V.B. Each single-attribute utility function is multiplied by a positive weight representing the importance of the corresponding attribute. The weights are chosen so that they sum up to one.

The utility functions of the form (17) are called additive. Additive utility functions are appropriate if the attributes are mutually independent [17]. The attributes in (17) can be shown to satisfy this; see [10]. For an example on the use of more complicated preference models in air combat modeling, see [40].

The situation evaluation utility node contains the cumulative utility function over all the decision stages of the game. It is here computed by summing up the single stage utilities, that is,

$$\mathbf{U}^i = \sum_{k=1}^{N} \mathbf{U}_k^i\left(\mathbf{c}_k^i\right) \tag{18}$$

The definition of the game needs also the initial state vectors $\mathbf{x}_0^i$ as well as the initial threat probabilities $\mathbf{p}_0^i$ for both players $i = B, R$.

The players are assumed to be rational in the utility theoretical sense (see, e.g., [38]). The solution of the influence diagram game is then the sequences of the controls that provide the highest cumulative expected utilities for the players. On the other hand, the players are noncooperative, have perfect information concerning the current level of play, and act simultaneously. In the game setting of this type, the players end up playing a feedback Nash equilibrium which depends only on the current state of the system at each stage. It is unfavorable for either player to deviate from such an equilibrium alone.

To obtain a feedback Nash equilibrium of the air combat game, the payoffs of the players, that is, the cumulative expected utilities

$$\begin{aligned} &J^i\left(\mathbf{u}_0^B, \ldots, \mathbf{u}_{N-2}^B, \mathbf{u}_{N-1}^B, \mathbf{u}_0^R, \ldots, \mathbf{u}_{N-2}^R, \mathbf{u}_{N-1}^R\right) \\ &\quad = \sum_{k=1}^{N}\left[\mathbf{p}_k^i\left(\mathbf{c}_k^i\right)\right]^T \mathbf{U}_k^i\left(\mathbf{c}_k^i\right), \qquad i = B, R \end{aligned} \tag{19}$$

are simultaneously maximized subject to the constraints giving feasible controls and states.

IV. Moving Horizon Control Approach

A. Look-Ahead Strategy Definition

In the following, a K-step look-ahead strategy refers to a player's feedback Nash equilibrium strategy for a truncated game with a horizon of K stages. A moving horizon or K-step look-ahead solution of the air combat game is obtained by solving the players' K-step look-ahead strategies at each decision stage and implementing only the first components of the obtained control sequences. This is continued until termination of the game.

B. Moving Horizon Solution of the Air Combat Game

To make the computation of the look-ahead strategies easier, the control variables are now discretized. At each decision stage, the players can change their current controls in discrete steps. At stage k, available control alternatives are

$$\begin{aligned} &S_k^i = \left\{\mathbf{u}_{k-1}^{i*} + \Delta\mathbf{u}^i\left(s_1, s_2, s_3\right)\Delta t \mid s_1, s_2, s_3 \in \left\{-1, 0, 1\right\}\right\} \\ &\quad i = B, R \end{aligned} \tag{20}$$

where $\mathbf{u}_{k-1}^{i*}$ is the preceding stage's control vector and $\Delta\mathbf{u}^i$ is defined as follows:

$$\Delta\mathbf{u}^i(s_1, s_2, s_3) = \left[\, s_1\Delta\alpha^i \quad s_2\Delta\eta^i \quad s_3\Delta\mu^i \,\right]^T \tag{21}$$

Here $\Delta\alpha^i$, $\Delta\eta^i$, and $\Delta\mu^i$ are the maximum rates of change for the control variables, see Eq. (A11). In addition to (20), the controls are constrained by

$$\Omega_k^i = \left\{\mathbf{u}_k^i \middle| \mathbf{g}\left(\mathbf{x}_k^i, \mathbf{u}_k^i\right) \le \mathbf{0}, \mathbf{h}\left(\mathbf{x}_k^i\right) \le \mathbf{0}\right\} \tag{22}$$

Above, functions $\mathbf{g}(\cdot)$ and $\mathbf{h}(\cdot)$ refer to (2) and (3). In sum, the set of feasible controls at stage k is given by

$$U_k^i = S_k^i \cap \Omega_k^i \tag{23}$$

When solving one-step look-ahead strategies at stage k, the players maximize the payoffs

$$J_{k,k+1}^i\left(\mathbf{u}_k^B, \mathbf{u}_k^R\right) = \left[\mathbf{p}_{k+1}^i\left(\mathbf{c}_{k+1}^i\right)\right]^T \mathbf{U}_{k+1}^i\left(\mathbf{c}_{k+1}^i\right), \qquad i = B, R \tag{24}$$

The states of the aircraft $\mathbf{x}_k^i$, the combat state $\mathbf{c}_k^i$, and the probability vectors $\mathbf{p}_k^i$ at current stage k are known, and the states and the probabilities at the next decision stage $k+1$ are calculated according to Eqs. (1), (4), and (13).

In practice, the control and state constraints are handled by reducing the payoffs related to the control alternatives that violate constraints (2) or (3). This casts infeasible control alternatives as dominated strategies that cannot form an equilibrium solution.

Nash equilibrium strategies are obtained by utilizing optimal responses (see, e.g., [9]) of the players, giving the players' optimal controls against a given control of the adversary. At stage k, the optimal responses of the players as a function of the adversary's control are given by

$$\text{stage } k \begin{cases} r_k^B\left(\mathbf{u}_k^R\right) = \arg\max\limits_{u_k^B \in U_k^B} J_{k,k+1}^B\left(\mathbf{u}_k^B, \mathbf{u}_k^R\right) \\ r_k^R\left(\mathbf{u}_k^B\right) = \arg\max\limits_{u_k^R \in U_k^B} J_{k,k+1}^R\left(\mathbf{u}_k^B, \mathbf{u}_k^R\right) \end{cases} \tag{25}$$

If $\mathbf{u}_k^{B*} = r_k^B(\mathbf{u}_k^{R*})$ and $\mathbf{u}_k^{R*} = r_k^R(\mathbf{u}_k^{B*})$, then the pair of controls $\mathbf{u}_k^{i*}$, $i = B, R$ is a Nash equilibrium solution (see [9]) providing the players' one-step look-ahead strategies. In practice, the optimal responses are first solved for all feasible controls of the players, after which a strategy pair fulfilling the condition above is identified.

The solution of two-step look-ahead strategies requires application of dynamic programming in addition to optimal responses. The players then maximize the payoffs

$$\begin{aligned} &J_{k,k+2}^i\left(\mathbf{u}_k^B, \mathbf{u}_{k+1}^B, \mathbf{u}_k^R, \mathbf{u}_{k+1}^R\right) = \sum_{\ell=k+1}^{k+2}\left[\mathbf{p}_\ell^i\left(\mathbf{c}_\ell^i\right)\right]^T \mathbf{U}_\ell^i\left(\mathbf{c}_\ell^i\right) \\ &\quad i = B, R \end{aligned} \tag{26}$$

The optimal responses of the players at stages $k+1$ and k are given by

$$\begin{aligned} &\text{stage } k+1 \\ &\begin{cases} r_{k+1}^B\left(\mathbf{u}_k^B, \mathbf{u}_k^R, \mathbf{u}_{k+1}^R\right) = \arg\max\limits_{\mathbf{u}_{k+1}^B \in U_{k+1}^B} J_{k,k+2}^B\left(\mathbf{u}_k^B, \mathbf{u}_{k+1}^B, \mathbf{u}_k^R, \mathbf{u}_{k+1}^R\right) \\ r_{k+1}^R\left(\mathbf{u}_k^B, \mathbf{u}_k^R, \mathbf{u}_{k+1}^B\right) = \arg\max\limits_{\mathbf{u}_{k+1}^R \in U_{k+1}^B} J_{k,k+2}^R\left(\mathbf{u}_k^B, \mathbf{u}_{k+1}^B, \mathbf{u}_k^R, \mathbf{u}_{k+1}^R\right) \end{cases} \\ &\text{stage } k \begin{cases} r_k^B\left(\mathbf{u}_k^R\right) = \arg\max\limits_{\mathbf{u}_k^B \in U_k^B} J_{k,k+2}^B\left(\mathbf{u}_k^B, \mathbf{u}_{k+1}^{B*}, \mathbf{u}_k^R, \mathbf{u}_{k+1}^{R*}\right) \\ r_k^R\left(\mathbf{u}_k^B\right) = \arg\max\limits_{\mathbf{u}_k^R \in U_k^B} J_{k,k+2}^R\left(\mathbf{u}_k^B, \mathbf{u}_{k+1}^{B*}, \mathbf{u}_k^R, \mathbf{u}_{k+1}^{R*}\right) \end{cases} \end{aligned} \tag{27}$$

Two-step look-ahead strategies of the players are obtained by first solving the optimal responses of the players at stage $k+1$. By utilizing these responses, Nash equilibrium strategies $\mathbf{u}_{k+1}^{i*}$ can be identified for all feasible values of k-stage controls $\mathbf{u}_k^i$, $i = B, R$. Then, the optimal responses of the players at stage k and Nash equilibrium strategies $\mathbf{u}_k^{i*}$ are solved by applying the previously solved $\mathbf{u}_{k+1}^{i*}$ related to the particular k-stage controls of the players. The two-step look-ahead strategies of the players are then the game optimal control sequences $(\mathbf{u}_k^{i*}, \mathbf{u}_{k+1}^{i*})$, $i = B, R$.

The payoffs and optimal responses for K-step look-ahead strategies in which the players anticipate the evolution of the duel for K steps ahead from the current state can be formulated in a similar manner. The strategies can be solved with optimal responses and dynamic programming as described above.

The moving horizon control approach for obtaining the K-step look-ahead solution of the air combat game can be summarized as follows:

1) Set $k = 0$. Set the initial conditions $\mathbf{x}_0^i$, $\mathbf{u}_0^i$, and $\mathbf{p}_0^i$, $i = B, R$.

2) Determine the control sequences $(\mathbf{u}_k^{B^*}, \ldots, \mathbf{u}_{k+K-1}^{B^*})$ and $(\mathbf{u}_k^{R^*}, \ldots, \mathbf{u}_{k+K-1}^{R^*})$ by solving K-step look-ahead strategies for blue and red from states $\mathbf{x}_k^{i^*}$, $\mathbf{c}_k^{i^*}$ and probabilities $\mathbf{p}_k^{i^*}(\mathbf{c}_k^{i^*})$, $i = B, R$ by using optimal responses and dynamic programming.

3) Implement $\mathbf{u}_k^{B^*}$ and $\mathbf{u}_k^{R^*}$.

4) Solve $\mathbf{x}_{k+1}^{i^*}$ with Eq. (1), $\mathbf{c}_{k+1}^{i^*}$ with Eq. (4), and $\mathbf{p}_{k+1}^{i^*}(\mathbf{c}_{k+1}^{i^*})$ with Eq. (13) using $\mathbf{u}_k^{B^*}$ and $\mathbf{u}_k^{R^*}$.

5) If either or both players have reached their respective target sets (11) or if $k = N_{\max}$, stop. Otherwise, set $k = k + 1$ and go to step 2.

V. Numerical Examples

In this section, we demonstrate the air combat game and the solution approach with a set of numerical examples. The first example serves as a simple demonstration of the approach. In the second example, we solve several instances with different initial states of the players and compare the differences between the solutions produced by one- and two-step look-ahead strategies. The third example demonstrates how the variation of the blue player's likelihood functions affects the duel.

The aircraft model used in the examples corresponds to a generic fighter aircraft. The threat probabilities are initialized to

$$\mathbf{p}_0^i = \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 \end{bmatrix}^T$$

as well as the upper limits for the combat state variables in (11) are set to $\omega_f^i = 30$ deg, $\theta_f^i = 60$ deg, and $d_f^i = 1000$ m, $i = B, R$ in all the examples. The maximum rates of change of the players' control variables are set to $\Delta\alpha^B = 30$ deg s^{-1}, $\Delta\mu^B = 60$ deg s^{-1}, and $\Delta\eta^B = 1.0$ s^{-1} for blue as well as $\Delta\alpha^R = 10$ deg s^{-1}, $\Delta\mu^R = 20$ deg s^{-1}, and $\Delta\eta^R = 0.5$ s^{-1} for red. That is, blue is more agile than red. The initial controls of the players are set to $\mathbf{u}_0^i = [0\ 0\ 1]^T$, $i = B, R$. The angle of attack limit, the minimum altitude, the maximum dynamic pressure, and the maximum load factor are set for both players to $\alpha_{\max} = 30$ deg, $h_{\min} = 1000$ m, $q_{\max} = 90$ kPa, and $n_{\max} = 9$, respectively. Time between two successive decision instants is $\Delta t = 0.7$ s and the maximum number of stages of the game is $N_{\max} = 50$. The above parameters as well as the likelihood functions, the single-attribute utility functions, and the weights presented in the following subsections are chosen by the authors and are only demonstrative. In real life applications, they should be extracted from the experts. Numerical examples are calculated by using a personal computer equipped with a 1.40 GHz clock frequency CPU and 224 MB of random access memory.

A. Likelihood Functions

The likelihood functions for the threat situation outcomes are given in Table 3. The functions are characterized by parameters $a^i > 0$, $i = B$, R, and D which define the steepnesses of the likelihood functions and the maximum allowed distance between the aircraft, respectively. In the examples, the values $a^i = 0.1$, $i = B, R$ and $D = 10{,}000$ m are used. The distributions utilized in the examples are relatively simple and serve only demonstrative purposes. For convenience, linearly increasing and decreasing functions have been selected.

The likelihood functions are defined similarly for both players, but we next consider the threat situation from a viewpoint of blue. If the aircraft are flying away from each other, or the distance between the aircraft is large, the outcome is assumed to be "neutral," $j = 1$. In this case, the probability that the bearing is large and the angle-off is small is high. Hence, monotonously increasing and decreasing functions are potential choices for the bearing and the angle-off, respectively. A uniform density function is suitable for the distance, because when the distance between the aircraft is large, the uniform distribution gives larger probabilities than the distributions of the other outcomes.

The outcome "advantage," $j = 2$, refers here to a situation where blue is pursuing red. Therefore, it is highly probable that the angles as well as the distance between the aircraft are small. Now, a monotonously decreasing function is a possible choice for both angles because it gives higher likelihoods for small angles. For the distance, a monotonously decreasing function, giving higher probabilities for short distances, can be used.

The outcome "disadvantage," $j = 3$, is an opposite of the advantage outcome. The likelihoods should increase as the angles increase as well as the distance between the aircraft gets smaller. In this case, possible functions for the angles are monotonously increasing, and for the distance, a monotonously decreasing function. They give higher probabilities for large angles and short distances.

If the aircraft are heading toward each other and the distance between the aircraft is small, the outcome is "mutual disadvantage," $j = 4$. In this case, it is likely that the bearing and the distance between the aircraft are small, but the angle-off is large. Potential likelihood functions for the bearing and distance are monotonously

Table 3 Likelihood functions for the different threat situation outcomes

j	Likelihood function	Range
1, 3	$p^{\omega,i}(\omega \mid \Theta^i = j) = [a^i\omega/180\ \text{deg} + 1 - a^i/2]/180\ \text{deg}$	$\omega \in [0, 180\ \text{deg}]$
2, 4	$p^{\omega,i}(\omega \mid \Theta^i = j) = [-a^i\omega/180\ \text{deg} + 1 + a^i/2]/180\ \text{deg}$	——
3, 4	$p^{\theta,i}(\theta \mid \Theta^i = j) = [a^i\theta/180\ \text{deg} + 1 - a^i/2]/180\ \text{deg}$	$\theta \in [0, 180\ \text{deg}]$
1, 2	$p^{\theta,i}(\theta \mid \Theta^i = j) = [-a^i\theta/180\ \text{deg} + 1 + a^i/2]/180\ \text{deg}$	——
1	$p^{d,i}(d \mid \Theta^i = j) = 1/D$	$d \in [0, D]$
2, 3, 4	$p^{d,i}(d \mid \Theta^i = j) = [-a^i d/D + 1 + a^i/2]/D$	——

Table 4 Single-attribute utility functions and weights

Outcome	j	$w_j^{\omega,i}$	$w_j^{\theta,i}$	$w_j^{d,i}$	$u_j^{\omega,i}(\omega)$	$u_j^{\theta,i}(\theta)$	$u_j^{d,i}(d)$
Neutral	1	0.2	0.1	0.7	$\frac{180\ \text{deg} - \omega}{180\ \text{deg}}$	$\frac{180\ \text{deg} - \theta}{180\ \text{deg}}$	$\frac{D-d}{D}$
Advantage	2	0.3	0.0	0.7	$\frac{180\ \text{deg} - \omega}{180\ \text{deg}}$	$\frac{180\ \text{deg} - \theta}{180\ \text{deg}}$	$\frac{D-d}{D}$
Disadvantage	3	0.0	0.7	0.3	$\frac{\omega}{180\ \text{deg}}$	$\frac{180\ \text{deg} - \theta}{180\ \text{deg}}$	$\frac{d}{D}$
Mutual disadvantage[a]	4	0.2	0.1	0.7	$\frac{180\ \text{deg} - \omega}{180\ \text{deg}}$	$\frac{180\ \text{deg} - \theta}{180\ \text{deg}}$	$\frac{D-d}{D}$
Mutual disadvantage[b]	4	0.2	0.1	0.7	$\frac{\omega}{180\ \text{deg}}$	$\frac{180\ \text{deg} - \theta}{180\ \text{deg}}$	$\frac{d}{D}$

[a]Player prefers joint capture to draw. [b]Player prefers draw to joint capture.

decreasing, and for the angle-off, monotonously increasing. They give higher probabilities for small bearing, short distance, and large angle-off.

B. Single-Attribute Utility Functions

The single-attribute utility functions used in (17) and the corresponding weights are shown in Table 4. In a neutral threat situation, the distance between the aircraft is large and it is hence presumed that blue should focus on pursuing red by minimizing the bearing and the distance between the aircraft. Blue should also try to avoid red from heading toward blue by minimizing the angle-off. Hence, small angles and short distances should be assigned high utility scores.

An advantageous threat situation is quite similar to the neutral one with respect to the player's preferences. However, because blue is behind red, it is assumed that red cannot head toward blue, and it is enough for blue to minimize only the bearing and the distance between the aircraft.

If the threat situation is disadvantageous, red is pursuing blue. Blue should now try to turn red's heading away from blue by minimizing the angle-off and maximizing the distance between the aircraft. Hence, small angle-offs and long distances should now be assigned high utility scores.

A mutually disadvantageous threat situation means here that both players are heading toward each other. If blue prefers joint capture to draw, he should head toward red by minimizing the angle-off, the bearing, and the distance. On the other hand, if blue prefers draw to joint capture, he should try to escape red by maximizing the angle-off and the distance as well as by minimizing the bearing.

In the following examples, blue prefers joint capture to draw, whereas red's preferences are opposite. That is, the blue and the red players' single-attribute utility functions for a mutually disadvantageous threat situation are given by the second to bottom and the bottom rows of Table 4, respectively. For the other threat situations, the players' utility functions and weights are the same.

C. Example 1

We first consider an example where the aircraft have already engaged in a dogfight and the less agile red is initially pursuing blue. The combat situation is initially advantageous for red and disadvantageous for blue. The initial states of the aircraft are shown in Table 5.

We study three cases where both players use one-, two-, and three-step look-ahead strategies, respectively. The three-dimensional trajectories of the aircraft for the three-step look-ahead solution are represented in Fig. 3. The histories of the control variables, the combat state variables, and the threat probability distributions of blue and red over the duel are shown in Figs. 4–6, respectively.

In the first case, the more agile blue, although being initially pursued, captures red in 21.0 s, or, in 30 stages. Figure 4 shows that both players apply the maximum throttle setting for most of the time. The players also use relatively large angles of attack and bank angles in magnitude, resulting in tortuous trajectories of Fig. 3. Figures 5a–5c reveal that although the bearing and the distance are within feasible intervals near the end, blue does not reach its target set until the angle-off of blue is driven below the upper limit of 60 deg.

Figure 6 shows that for blue, the probability of the disadvantage outcome is relatively high during the first 15 s, after which it declines rapidly. During the rest of the combat, the probability of the advantage outcome for blue rises toward one, whereas for red it decreases toward zero. The payoffs of blue and red, computed by

Table 5 Initial states of the aircraft in example 1

Player	x_0, m	y_0, m	h_0, m	γ_0, deg	χ_0, deg	v_0, m/s
Blue	3000.0	1000.0	6000.0	0.0	30.0	315.0
Red	1000.0	1000.0	6000.0	0.0	0.0	240.0

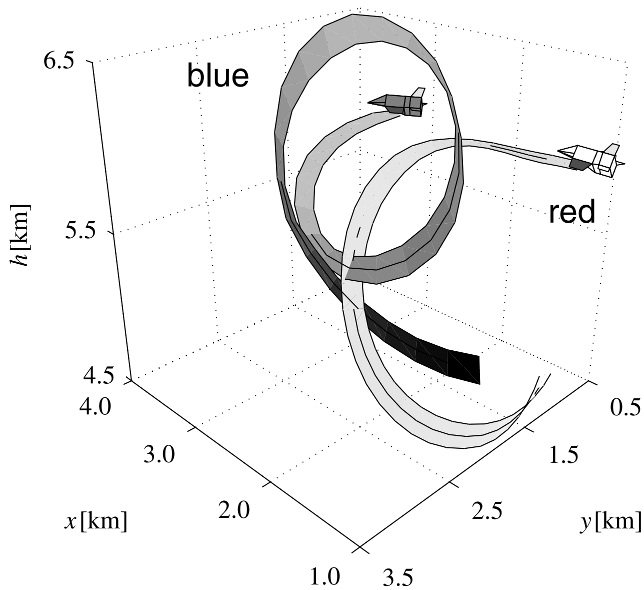


Fig. 3 3-D trajectories for the three-step look-ahead solution in example 1.

(19), are 17.81 and 15.56, respectively. The payoff ratio, that is, the ratio of blue's payoff to red's one, is 1.1449.

In the second and the third cases, blue again wins the duel by capturing red in 24.5 and 23.1 s, respectively. Comparison between the distributions of the first, second, and third cases in Fig. 6 shows that in the last two cases, the probability of the disadvantage outcome is slightly higher for blue than in the first one. On the other hand, the probability of the advantage outcome rises eventually higher in the second and the third case.

In the second case, the payoffs of blue and red are 21.43 and 17.21, respectively, and the payoff ratio is 1.2446, whereas in the third case the respective quantities are 19.59, 17.07, and 1.1472. The payoff ratios imply that with this particular initial setting, the more long-sighted strategies yield better payoffs for blue in relation to red than the first one.

According to Figs. 4–6, the controls, the combat states, and the threat probability distributions indicate some signs of convergence as the planning horizon is extended. Still, on the basis of the limited data, it is impossible to state whether the convergence implies that the moving horizon solution is close to the true feedback Nash equilibrium of the game. Computational burden prevents a direct comparison. The average computation times per decision instant in the one-, two-, and three-step look-ahead cases are 9.50×10^{-3} s, 3.62 s, and 1.89×10^3 s, respectively.

To summarize, the more agile blue wins the duel by rolling behind the red aircraft and capturing it. Red cannot exploit his advantageous initial state but is eventually forced to evade blue. Here, the longer planning horizons benefit the winner compared to the shortest one, but this is not always evident, as will be seen later.

D. Example 2

We next expand the first example by turning the initial heading angle of blue and applying one- and two-step look-ahead strategies for both players in each case. The initial states of the aircraft are shown in Table 5 with an exception of the heading angle of blue. It is shifted in steps of 30 deg for 180 deg resulting in seven different cases in total.

The payoff ratios for the different cases are represented in Table 6. Each case consists of two runs where the length of the planning horizon corresponds to number of the run. The average computation times per decision instant are similar to those of the previous example and are hence not mentioned here.

The three-dimensional trajectories produced by two-step look-ahead strategies are presented in Fig. 7 for the cases 2, 3, 5, and 6. In

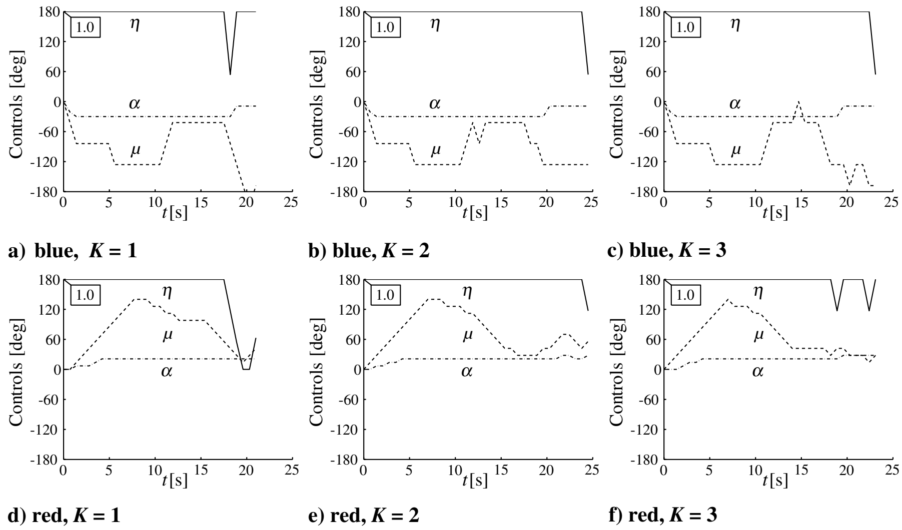


Fig. 4 Control histories of the players in example 1.

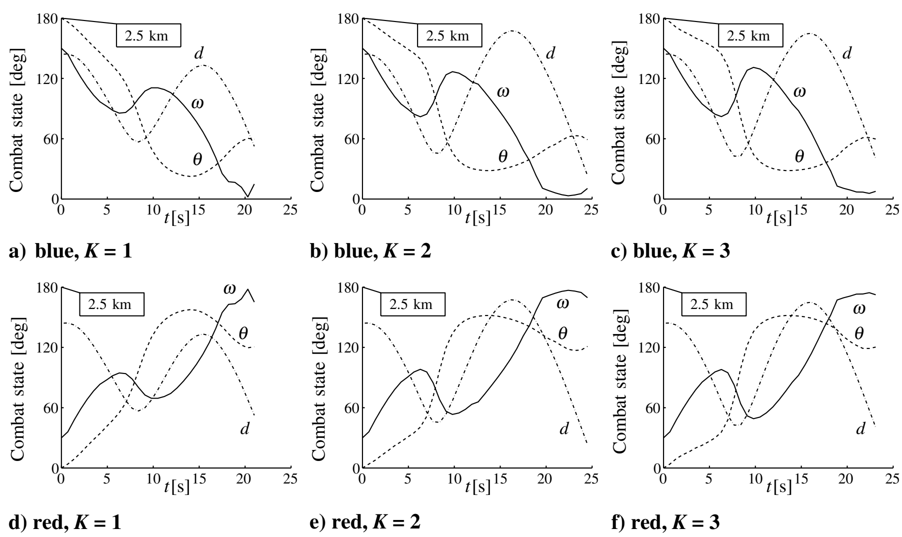


Fig. 5 Combat state histories of the players in example 1.

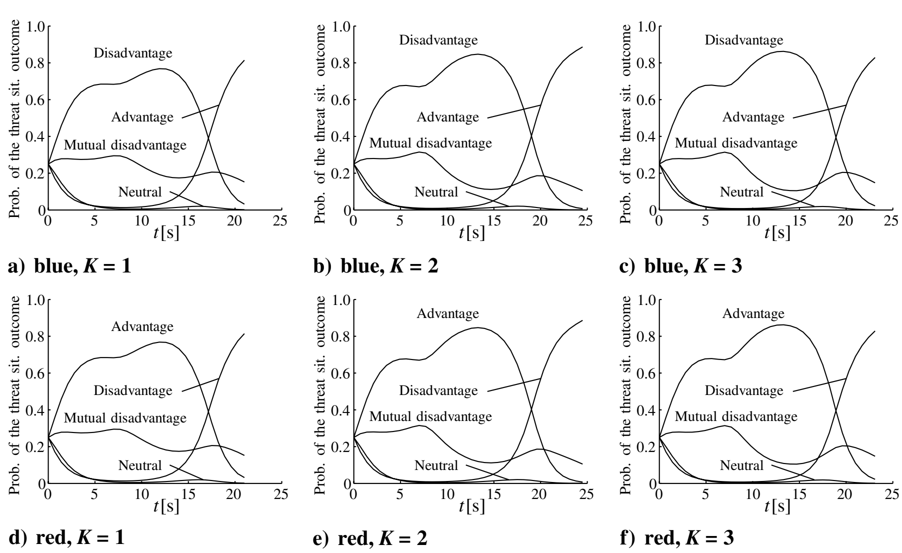


Fig. 6 Threat probability distributions of the players in example 1.

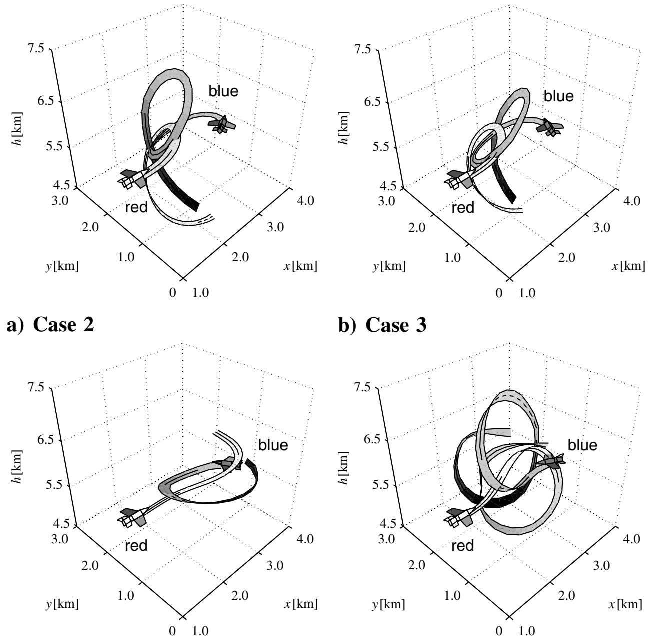
cases 2 and 3, blue and red consider being in a disadvantageous and advantageous threat situation, respectively. Therefore, blue tries to minimize the angle-off by turning toward red, whereas red tries to minimize the bearing by maintaining the heading toward blue. On the other hand, in cases 5 and 6, the players believe being in a mutually disadvantageous threat situation. In such a situation, blue and red prefer pursuit and evasion, respectively, resulting in the trajectories of Figs. 7c and 7d.

Table 6 Initial heading angles of blue, payoff ratios, and outcomes of the game in example 2

Case	χ^B, deg	$K = 1$	Outcome	$K = 2$	Outcome
1	0.0	0.4834	Draw	0.9125	Draw
2	30.0	1.1449	Blue wins	1.2446	Blue wins
3	60.0	1.4084	Blue wins	1.5271	Blue wins
4	90.0	1.9331	Blue wins	1.6752	Blue wins
5	120.0	2.6559	Blue wins	2.4501	Blue wins
6	150.0	3.0770	Blue wins	3.1093	Blue wins
7	180.0	1.7933	Blue wins	1.8407	Draw

Blue wins in all the cases except the first one and the second run of the last case where the outcome of the game is draw. The victories of blue are partly due to the different preferences of the players in a mutually disadvantageous threat situation. Such a situation is likely to turn advantageous for blue who prefers pursue and disadvantageous for red who prefers evasion, resulting eventually in a win of blue.

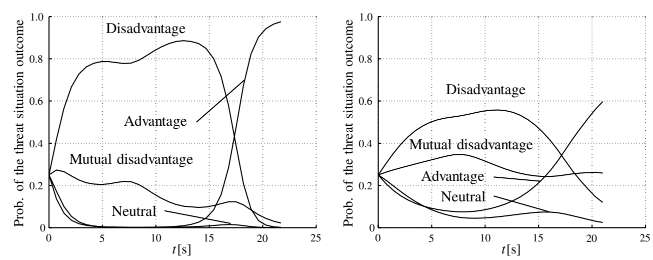
In most cases, the ratio of blue's payoff to red's one increases when the planning horizon is extended. However, in the fourth and the fifth cases, red gains better payoff in relation to blue due to the more long-sighted strategy. Further test runs not presented here validate that extending the planning horizon does not always benefit the more agile player or the winner. Note that since a player gains payoff also by maneuvering well in a disadvantageous situation, the optimal values of the payoffs do not necessarily correlate with the actual outcome of the game.



a) Case 2 b) Case 3

c) Case 5 d) Case 6

Fig. 7 3-D trajectories for the two-step look-ahead solutions in example 2.



a) $a^B = 0.15$, $a^R = 0.1$ **b) $a^B = 0.05$, $a^R = 0.1$**

Fig. 8 Threat probability distributions of blue in example 3.

E. Example 3

We finally consider how the payoffs are affected when blue chooses his likelihood functions differently. The likelihood functions shown in Table 7 are characterized by parameters a^i, $i = B, R$ that provide a straightforward way to change the steepnesses of the functions. The larger the parameter, the steeper the function. Steep likelihood functions should result in sharper threat situation assessments.

We study two cases in which blue's parameter a^B is set differently. In the first case, the value of the parameter is greater than in the second one. The initial states of the players are the same as in example 1 and the players use one-step look-ahead strategies. Blue wins the duel in both cases in 21.0 s. The values of the parameters, the players' payoffs, and the payoff ratios for both the cases are represented in Table 7.

The threat probability distributions of blue for both duels are presented in Fig. 8. Comparisons between the distributions and the payoffs suggest that with the particular initial setting, blue achieves a better payoff ratio by applying the nonsteep likelihood functions.

Steep likelihood functions produce alternating threat probability distributions in which only one threat situation outcome dominates at a time; see Fig. 8a. Nonsteep likelihood functions reflect a more conservative view about the threat situation related to a certain combat state. They produce distributions that are closer to the uniform distribution; see Fig. 8b. Recall that in Eq. (19), the utility associated with a particular threat situation outcome is weighted by the corresponding probability. Hence, with steep likelihood functions mainly a single, largely weighted utility contributes to the payoff, whereas with nonsteep likelihood functions the payoff is affected more evenly by all the utilities.

Table 7 Parameters of the players' likelihood functions, payoffs, and payoff ratios in example 3

Case	a^B	a^R	J^B	J^R	J^B/J^R
1	0.15	0.1	17.96	15.67	1.1461
2	0.05	0.1	18.85	15.56	1.2116

VI. Discussion

A. Comparison with Related Air Combat Modeling Approaches

In two-target games, the solution process consists of two phases. First, the qualitative solution attempts to classify the combat states into different sets each of them related to a certain outcome of the game. Then, the quantitative solution aims at providing game optimal controls for the players at each time instant. In the influence diagram game, this process has been unified. A priori knowledge, observations, and likelihood functions all contribute to the threat situation probabilities that classify the states, and the utilities and their weights associated with the combat state attributes guide the selection of the controls throughout the dynamic game setting.

The threat probabilities can be interpreted as the weights of the payoff functions that are used for converting a multicriteria game into a single criterion zero-sum or nonzero-sum game (see [42,43], respectively). Instead of solutions with predetermined finite or infinite weights, the latter leading to state constrained quantitative pursuit-evasion games [28], the influence diagram game selects these weights, that is, the threat probabilities, in each combat situation differently. Besides overall threat probabilities, the weights of single-attribute utility functions can be chosen variously for the different threat situations. This means that the utility associated with the current combat state depends on the threat situation. For instance, if red is in the vicinity of blue's target set, that is, the threat situation is disadvantageous from a viewpoint of red, his objective is to avoid getting into the target set. In the influence diagram game, the control alternatives of red leading away from blue's target set become the most favorable ones, provided that the likelihood distributions are selected reasonably.

The influence diagram game shifts the focus of the classification of the combat states from rigorous mathematical analysis into a description enhancing the behavioral aspects of the situation. The crisp surfaces separating different states are substituted with probability distributions that describe the decision makers' beliefs. In this context, the approach also becomes more straightforward, as instead of very complex computations the classification is now based on expert opinions represented by the likelihood functions of the threat probabilities. An important consequence is that the influence diagram approach decouples the vehicle models from the analysis. Thus, it allows the use of arbitrarily complex aircraft dynamics models, although in the numerical examples, only a point mass model is applied for convenience. Naturally, application of more complex models increases the number of variables which in turn increases computational burden. Also, realistic modeling of rotational dynamics requires a relatively short integration step which limits the length of the planning horizon. However, the latter problem could possibly be overcome by applying denser discretization for the state variables and interpolating the control variables between the control discretization points.

The structured and widely accepted value-focused multicriteria decision making approach employed by influence diagrams offers a competitive alternative over many heuristic models representing the maneuvering decision of pilots described in Sec. II. Especially the possibility to treat uncertainty, human preferences, and attitudes toward risk in a well-established way make the approach an interesting choice over expert systems. The representation also makes it easier for persons of the application area to participate in the modeling process because the meanings of nodes and arcs in the graph can be understood with a little decision theoretic and mathematical background.

B. Maneuvering Decision Analysis

One of the aims of the model construction is to find how different functional parameters affect the behavior and perception of the players. As such, the detailed structure of the influence diagram model makes it possible to analyze the effect of each weight, utility function, and distribution separately. This is demonstrated for the likelihood functions in example 3. The results in the example are rather obvious, but the example is basically meant to illustrate the possibilities of the model.

It is interesting to note that even with completely similar preferences, the game is not zero sum in an arbitrary combat state. For example, when high probabilities are associated with the neutral outcome of both players' threat assessment, the game is nonzero sum. On the other end, when the threat situation is advantageous for one player and disadvantageous for the other one, the nature of the game is almost zero sum. Hence, the roles of the players change dynamically according to the combat state, and there is no need to fix them beforehand.

The combat state vector should be defined so that its elements would contain all the information on the players' states that is relevant regarding the combat. In the simplified examples of this paper, we use the distance between the aircraft and the angles between the velocity vectors and the line of sight. Other potential choices are, for example, features describing the energy difference of the players that is a crucial factor related to the success of the players in air combat [39].

One should note that in nonzero-sum games, a feedback Nash equilibrium in pure strategies does not necessarily exist and it is not necessarily unique. However, according to our computational experience, this is uncommon; less than 1% of the moving horizon subgames solved in this paper did not have a unique Nash equilibrium. If no unique Nash equilibrium exists, a Stackelberg equilibrium (see, e.g., [9]) can be used because a nonzero-sum game always admits this solution. In case of multiple equilibria, all the equilibrium solutions could be solved at first. Then, the worst one from a viewpoint of, for example, the blue player could be consistently implemented. In this paper, however, the first equilibrium that is found is chosen.

In the present game model, the control alternatives are always chosen under the supposition that also the adversary maneuvers optimally with respect to a given preference model. In most cases, this leads to decisions that guarantee a secure outcome for the player. It is obvious that if the opponent behaves nonoptimally, the assumption may be overly stringent. Nevertheless, a rationality assumption is often better than a mere prediction based on old observations from the state of the adversary. Considering the fact that the opponent is a rational agent with predetermined goals, the rationality assumption is likely to be correct at least for most of the time.

C. Combat Trajectory Synthesis

In this study, likelihood functions and single-attribute utility functions have been generated by the authors. For the sake of clarity, they have been kept reasonable but simple. In the field of decision analysis, there exist numerous methods for extracting probability distributions, single-attribute utility functions, and their weights from actual decision makers and experts. For basic approaches, see [44]. Application of such methods requires effort, but offers a possibility to transfer the perceptions, preferences, beliefs, and attitudes of an expert into the influence diagram game that can then be used as a decision making logic of an air combat simulator (see, e.g., [37]).

A simulator makes it possible to study air combat tactics and aircraft performance in a controlled and repeatable environment. With a planning horizon of the influence diagram game adapted to the available computational power, it is even possible to produce decisions in real time. We believe that a careful decision logic description combined with the modeling of the adversary's rational behavior offers an efficient way to produce game and preference optimal controls for computer guided aircraft.

An obvious shortcoming in all dynamic programming based methods is the computational explosion. Even the moving horizon approach allows only relatively short planning horizons. It is therefore obvious that increasing the computational power is hardly an option in deepening the search. In the future, the application of metaheuristics, like Nash genetic algorithms [45], or massively parallel computing approaches, might provide a way to improve the approach so that strategies for the players could be obtained in real time and also for longer planning horizons. However, even with one-, two-, and three-step look-ahead strategies, it is possible to study the effects of the planning horizon on the solutions.

The strategies of the players correspond to a Nash equilibrium solution that generally favors neither player. In [17], it is proposed that if the capabilities of the players are different, a longer planning horizon can be expected to benefit the more capable player. The results of example 2 indicate that for the game at hand, this holds true for some initial states, but not for all. To see why this happens, consider first a case where the maneuvering of the other player is completely constrained by the control rate constraints, and the problem is reduced to a single player optimal control problem. The global optimum can be guaranteed only by applying dynamic programming starting from the final stage. With limited planning horizons, it is possible that at a certain stage, a shorter planning horizon produces controls that coincide with the global optimum, whereas a longer one results in suboptimal controls for the particular stage. This applies obviously to a game situation as well. Nevertheless, it is reasonable to expect that extending the planning horizon benefits the more agile player on the average.

VII. Conclusions

In this paper, the successive control decisions of the pilots during one-on-one air combat are modeled using a multistage influence diagram game. A feedback Nash equilibrium solution is obtained by a moving horizon control approach that provides game optimal control sequences with respect to the preference models of the players. The computational complexity constrains the planning horizon of the moving horizon approach. However, the numerical examples presented in this paper indicate that the air combat game produces plausible control choices even with relatively short planning horizons for which the controls can be computed in real time. In addition to the synthesis of control decisions, the structure of the game model makes the backtracking of the solutions straightforward and offers several analysis possibilities.

The presented influence diagram game is the first elaboration in which the methodology of influence diagrams and the foundations of the theory of dynamic games are combined in air combat modeling. The game model and its aspects suggest that the combination of these disciplines offers a potential way for describing dynamic game situations, although a well-determined formalism related to dynamic influence diagram games is not yet defined. The results and the discussion of this paper indicate that the influence diagram representation of the air combat game overcomes several shortcomings that appear in traditional air combat modeling approaches.

Appendix

The dynamics of the players is described by a point mass model of an aircraft, that is, by the following system of differential equations [46]:

$$\dot{x} = v\cos\gamma\cos\chi \tag{A1}$$

$$\dot{y} = v\cos\gamma\sin\chi \tag{A2}$$

$$\dot{h} = v\sin\gamma \tag{A3}$$

$$\dot{\gamma} = \frac{1}{mv}\{\{L(\alpha,h,v,M(h,v)) + \eta T_{\max}(h,M(h,v))\sin\alpha\} \times \cos\mu - mg\cos\gamma\} \tag{A4}$$

$$\dot{\chi} = \frac{1}{mv\cos\gamma}\{L(\alpha,h,v,M(h,v)) + \eta T_{\max}(h,M(h,v))\sin\alpha\} \times \sin\mu \tag{A5}$$

$$\dot{v} = \frac{1}{m}\{\eta T_{\max}(h,M(h,v))\cos\alpha - D(\alpha,h,v,M(h,v))\} - g\sin\gamma \tag{A6}$$

where x and y refer to the horizontal coordinates and h to the altitude of the player. The remaining three state variables are the flight path angle γ, the heading angle χ, and the velocity v. The aircraft is guided with the angle of attack α, the throttle setting η, and the bank angle μ.

The acceleration due to the gravity as well as the mass of the aircraft, denoted by g and m, respectively, is assumed constant. $T_{\max}(\cdot)$ denotes the maximum available thrust force, $L(\cdot)$ the lift force, $D(\cdot)$ the drag force, and $M(\cdot)$ the Mach number. The lift force is given by

$$L(\alpha,h,v,M(h,v)) = C_L(\alpha,M(h,v))Sq(h,v) \tag{A7}$$

where $C_L(\cdot)$ is the lift coefficient and S the reference wing area. The dynamic pressure is

$$q(h,v) = \tfrac{1}{2}\varrho(h)v^2 \tag{A8}$$

where the air density $\varrho(h)$ is taken from the International Standard Atmosphere. The drag force is of the form

$$D(\alpha,h,v,M(h,v)) = C_D(\alpha,M(h,v))Sq(h,v) \tag{A9}$$

where $C_D(\cdot)$ denotes the total drag coefficient. This as well as the lift coefficient and the maximum thrust force of the aircraft is given as tabular data and is approximated with suitable continuously differentiable functions.

The control variables are constrained by lower and upper bounds

$$-\alpha_{\max} \le \alpha \le \alpha_{\max}, \qquad 0 \le \eta \le 1, \qquad -180 \le \mu \le 180 \text{ deg} \tag{A10}$$

and their rates of change are restricted as

$$-\Delta\alpha \le \dot{\alpha} \le \Delta\alpha, \qquad -\Delta\eta \le \dot{\eta} \le \Delta\eta, \qquad -\Delta\mu \le \dot{\mu} \le \Delta\mu \tag{A11}$$

The latter constraints are due to the flight control systems and the inertia of the aircraft.

To avoid stall, the angle of attack must be chosen so that the lift coefficient does not exceed aircraft specific value $C_{L,\max}(\cdot)$ at a given altitude and velocity, that is,

$$C_L(\alpha,M(h,v)) - C_{L,\max}(M(h,v)) \le 0 \tag{A12}$$

The load factor that is given by

$$n(\alpha,h,v) = \frac{L(\alpha,h,v,M(h,v))}{mg} \tag{A13}$$

is limited by the structure of the aircraft. This imposes another constraint related to the angle of attack, altitude, and velocity:

$$n(\alpha,h,v) - n_{\max} \le 0 \tag{A14}$$

In addition, the altitude and the dynamic pressure are constrained by

$$h_{\min} - h \le 0 \tag{A15}$$

and

$$q(h,v) - q_{\max} \le 0 \tag{A16}$$

where $h_{\min}$ and $q_{\max}$ refer to the minimum altitude and the maximum dynamic pressure, respectively.

The set of differential equations (A1–A6) can be written in shorthand as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{u}), \qquad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{A17}$$

where

$$\mathbf{x} = \begin{bmatrix} x & y & h & \gamma & \chi & v \end{bmatrix}^T$$

is the state vector,

$$\mathbf{u} = \begin{bmatrix} \alpha & \eta & \mu \end{bmatrix}^T$$

is the control vector, and $\mathbf{x}_0$ is the initial state of the aircraft.

References

[1] Shinar, J., “Differential Games and Artificial Intelligence—A New Approach for Solving Complex Dynamic Conflicts,” *Lecture Notes in Control and Information Sciences, Differential Games—Developments in Modelling and Computation*, edited by R. P. Hämäläinen and H. K. Ehtamo, Vol. 156, Springer–Verlag, New York, 1991, pp. 100–110.

[2] Howard, R. A., and Matheson, J. E., “Influence Diagrams,” *The Principles and Applications of Decision Analysis*, edited by R. A. Howard and J. E. Matheson, Vol. 2, Strategic Decision Group, Palo Alto, CA, 1984, pp. 719–762.

[3] Howard, R. A., and Matheson, J. E., “Influence Diagrams,” *Decision Analysis*, Vol. 2, No. 3, 2005, pp. 127–147.

[4] Virtanen, K., Raivio, T., and Hämäläinen, R. P., “Decision Theoretical Approach to Pilot Simulation,” *Journal of Aircraft*, Vol. 36, No. 4, July 1999, pp. 632–641.

[5] von Neumann, J., and Morgenstern, O., *Theory of Games and Economic Behavior*, 3rd ed., Princeton Univ. Press, Princeton, NJ, 1990, pp. 15–31.

[6] Shachter, R. D., “Evaluating Influence Diagrams,” *Operations Research*, Vol. 34, No. 6, 1986, pp. 871–882.

[7] Koller, D., and Milch, B., “Multi-Agent Influence Diagrams for Representing and Solving Games,” *Proceedings of the 17th*

International Joint Conference of Artificial Intelligence, Morgan Kaufmann Publishers, Seattle, WA, Aug. 2001, pp. 319–328.

[8] Koller, D., and Milch, B., "Multi-Agent Influence Diagrams for Representing and Solving Games," *Games and Economic Behavior*, Vol. 45, No. 1, 2003, pp. 181–221.

[9] Başar, T., and Olsder, G. J., *Dynamic Noncooperative Game Theory*, 2nd ed., Academic Press, London, 1995, pp. 1–5, 138, 175, 284.

[10] 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, 2004, pp. 665–677.

[11] Virtanen, K., Raivio, T., and Hämäläinen, R. P., "An Influence Diagram Approach to One-on-One Air Combat," *Proceedings of the 10th International Symposium on Dynamic Games and Applications*, International Society of Dynamic Games, St. Petersburg, Russia, Vol. 2, July 2002, pp. 859–864.

[12] Jensen, F., Jensen, F. V., and Dittmer, S. L., "From Influence Diagrams to Junction Trees," *Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann Publishers, Seattle, WA, July 1994, pp. 367–373.

[13] Bertsekas, D. P., *Dynamic Programming and Optimal Control*, Vol. 2, Athena Scientific, Belmont, MA, 2001, pp. 16–20.

[14] Richalet, J., Rault, A., Testud, L., and Papon, J., "Model Predictive Heuristic Control: Applications to Industrial Processes," *Automatica*, Vol. 14, No. 5, 1978, pp. 431–428.

[15] Camacho, E. F., and Bordons, C., *Model Predictive Control*, Springer, London, 1999, pp. 1–11.

[16] Cruz, J. B., Jr., Simaan, M. A., Gacic, A., Jiang, H., Letellier, B., Li, M., and Liu, Y., "Game-Theoretic Modeling and Control of a Military Air Operation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 37, No. 4, 2001, pp. 1393–1403.

[17] Cruz, J. B., Jr., Simaan, M. A., Gacic, A., and Liu, Y., "Moving Horizon Nash Strategies for a Military Air Operation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 38, No. 3, 2002, pp. 978–999.

[18] Shinar, J., and Glizer, V. Y., "Application of Receding Horizon Control Strategy to Pursuit-Evasion Problems," *Optimal Control Applications & Methods*, Vol. 16, No. 2, 1995, pp. 127–141.

[19] Bryson, A. E., and Ho, Y., *Applied Optimal Control*, Hemisphere Publishing Corporation, New York, 1975, pp. 42–89.

[20] Isaacs, R., *Differential Games*, Krieger, New York, 1975, pp. 8–13.

[21] Lachner, R., Breitner, M., and Pesch, H., "Three-Dimensional Air Combat: Numerical Solution of Complex Differential Games," *Annals of the International Society of Dynamic Games*, edited by G. J. Olsder, Vol. 3, New Trends in Dynamic Games and Applications, Birkhäuser, Boston, 1995, pp. 165–190.

[22] Raivio, T., and Ehtamo, H., "Visual Aircraft Identification as a Pursuit-Evasion Game," *Journal of Guidance, Control, and Dynamics*, Vol. 23, No. 4, 2000, pp. 701–708.

[23] Raivio, T., "Capture Set Computation of an Optimally Guided Missile," *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 6, 2001, pp. 1167–1175.

[24] 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.

[25] Blaquière, A., Gérard, F., and Leitmann, G., *Qualitative and Quantitative Games*, Academic Press, New York, 1969, pp. 1–10.

[26] Grimm, W., and Well, K. H., "Modelling Air Combat as Differential Game, Recent Approaches and Future Requirements," *Differential Games—Developments in Modeling and Computation*, edited by R. P. Hämäläinen and H. Ehtamo, Springer, Berlin, 1991, pp. 1–13.

[27] Davidovitz, A., and Shinar, J., "Two-Target Game Model of an Air Combat With Fire-and-Forget All-Aspect Missiles," *Journal of Optimization Theory and Applications*, Vol. 63, No. 2, 1989, pp. 133–165.

[28] Ardema, M. D., Heymann, M., and Rajan, N., "Combat Games," *Journal of Optimization Theory and Applications*, Vol. 46, No. 4, 1985, pp. 391–398.

[29] Ghose, D., and Prasad, U. R., "Solution Concepts in Two-Person Multicriteria Games," *Journal of Optimization Theory and Applications*, Vol. 63, No. 2, 1989, pp. 167–189.

[30] Kelley, H. J., and Cliff, E. M., "Reprisal Strategies in Pursuit Games," *Journal of Guidance and Control*, Vol. 3, No. 3, 1980, pp. 257–260.

[31] Austin, F., Carbone, G., Falco, M., and Hinz, H., "Game Theory for Automated Maneuvering During Air-to-Air Combat," *Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 6, 1990, pp. 1143–1149.

[32] Neuman, F., "On the Approximate Solution of Complex Combat Games," *Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 1, 1990, pp. 128–136.

[33] Katz, A., and Ross, A., "One-on-One Helicopter Combat Simulated by Chess-Type Lookahead," *Journal of Aircraft*, Vol. 28, No. 2, 1991, pp. 158–160.

[34] Katz, A., "Tree Lookahead in Air Combat," *Journal of Aircraft*, Vol. 31, No. 4, 1994, pp. 970–973.

[35] McManus, J. W., and Goodrich, K. H., "Application of Artificial Intelligence (AI) Programming Techniques to Tactical Guidance for Fighter Aircraft," *AIAA Guidance, Navigation, and Control Conference*, AIAA, Boston, MA, Aug. 1989, pp. 851–858; also AIAA Paper 89-3525.

[36] Goodrich, K. H., and McManus, J. W., "Development of a Tactical Guidance Research and Evaluation System (TGRES)," *AIAA Flight Simulation Technologies Conference*, AIAA, Boston, MA, Aug. 1989, pp. 350–356; also AIAA Paper 89-3312.

[37] Lazarus, E., "The Application of Value-Driven Decision-Making in Air Combat Simulation," *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, Orlando, FL, 1997, pp. 2302–2307.

[38] Keeney, R. L., and Raiffa, H., *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, Cambridge Univ. Press, New York, 1993, pp. 90, 231.

[39] Shaw, R. L., *Fighter Combat: Tactics and Maneuvering*, Naval Institute Press, Annapolis, MD, 1985, pp. 104–111.

[40] Virtanen, K., Hämäläinen, R. P., and Mattila, V., "Team Optimal Signaling Strategies in Air Combat," *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, Vol. 26, No. 4, July 2006, pp. 643–660.

[41] Ross, S., *A First Course in Probability*, 7th ed., Pearson Prentice–Hall, Upper Saddle River, NJ, 2006, pp. 81–87, 293.

[42] Järmark, B., "A Missile Duel Between Two Aircraft," *Journal of Guidance, Control, and Dynamics*, Vol. 8, No. 4, 1985, pp. 508–513.

[43] Moritz, K., Polis, R., and Well, K. H., "Pursuit-Evasion in Medium-Range Air-Combat Scenarios," *Computers and Mathematics with Applications*, Vol. 13, Nos. 1–3, 1987, pp. 167–180.

[44] von Winterfeldt, D., and Edwards, W., *Decision Analysis and Behavioral Research*, Cambridge Univ. Press, Cambridge, U.K., 1986, pp. 322–325.

[45] Périaux, J., and Sefrioui, M., "Nash Genetic Algorithms: Examples and Applications," *Innovative Tools for Scientific Computation in Aeronautical Engineering*, edited by J. Périaux, P. Joly, O. Pironneau, and E. Onate, CIMNE, Barcelona, Spain, 2001, pp. 391–404.

[46] Miele, A., *Flight Mechanics, Vol. 1: Theory of Flight Paths*, Addison–Wesley, Reading, MA, 1962, pp. 48–50.