JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 27, No. 4, July–August 2004

Modeling Pilot's Sequential Maneuvering Decisions by a Multistage Influence Diagram

Kai Virtanen,* Tuomas Raivio,† and Raimo P. Hämäläinen‡
Helsinki University of Technology, FIN-02015 HUT Espoo, Finland

The paper presents an approach toward the off-line computation of preference optimal flight paths against given air combat maneuvers. The approach consists of a multistage influence diagram modeling the pilot's sequential maneuvering decisions and a solution procedure that utilizes nonlinear programming. The influence diagram graphically describes the elements of the decision process, contains a point-mass model for the dynamics of an aircraft, and takes into account the decision maker's preferences under conditions of uncertainty. Optimal trajectories with respect to the given preference model are obtained by converting the multistage influence diagram into a discrete-time dynamic optimization problem that is solved with nonlinear programming. The initial estimate for the decision variables of the problem is generated by solving a set of myopic single stage influence diagrams that anticipate the future state of the aircraft only a short planning horizon ahead. The presented solution procedure is illustrated by a numerical example.

I. Introduction

In this paper, an approach for modeling and solving an aircraft trajectory optimization problem by the methods of decision analysis (see Refs. 1–3) and nonlinear programming (see, e.g., Ref. 4) is introduced. The approach consists of a multistage influence diagram model[5] representing the sequential maneuvering decisions of a pilot against a hostile aircraft obeying a given air combat maneuver and a new off-line solution procedure for such a model. Hence, the approach provides a framework for a single decision maker to find an optimal flight path with respect to his or her preferences.

The proposed approach offers a new way to incorporate a realistic preference and uncertainty model into flight-path optimization problems and optimal trajectory planning in other application areas as well. It enables a pilot to be involved in the modeling process because the graphical representation of influence diagrams is easily understood by individuals with a little decision theoretic and mathematical background. As far as the authors know, this is the first application of influence diagrams that contains an explicit model of the dynamic decision environment, which is here represented by a set of difference equations.

Traditionally, in flight-path optimization problems, the decision dynamics is taken into account with the methods of optimization[4,6] and dynamic game theory.[7] Usually flight paths that maximize or minimize a given explicit objective or cost function are sought. For example, minimum time trajectories for a single aircraft (e.g., see Refs. 8 and 9) can be calculated by using optimal control theory[6] as well as nonlinear programming.[10] For an overview on aircraft trajectory optimization problems, see, for example, Ref. 9.

The literature on optimization theory seldom pays attention to the structure of performance criteria that model the preferences of human decision makers. However, this topic needs to be studied more carefully, for example, by developers of air combat simulators, because an essential part of a simulator is the model that imitates the decision-making process of a pilot. Approaches suggested earlier in the literature include knowledge-based expert systems,[11,12] heuristic value-driven systems,[13,14] or discrete dynamic games.[15,16] These models predict the future situation of the combat only a short planning horizon ahead. Thus, they do not produce optimal trajectories but rather myopic control commands.

The influence diagram, invented by Howard and Matheson,[5] is a tool from decision analysis for modeling and solving Bayesian decision problems. In such problems, the subjective probability interpretation (e.g., see Ref. 17) is applied, and the goodness of the decision alternatives' consequences is measured by the von Neumann–Morgenstern utility.[18] Influence diagrams represent decision problems as a graph of nodes and arcs. They allow quantitative analysis of uncertain and multiple attribute decision problems. Influence diagrams are closely linked to decision trees (e.g., see Ref. 3) that originate from the theory of game trees or games in extensive form, first defined by von Neumann and Morgenstern[18] (see also Ref. 7).

The multistage influence diagram used in this paper is based on the model developed in Ref. 19. In this model, a single maneuvering decision is not affected by the upcoming decisions because the future states of the aircraft are anticipated for a short planning horizon, one decision interval, ahead. Hence, this model is called a single-stage influence diagram. Its solution provides myopic maneuvering decisions that are optimal with respect to the given preference model and the available information.

To compute control sequences that are better with respect to the overall goals over the total flight time, the influence diagram must be able to predict the future states of the combat further than one decision stage ahead. In this paper, the interaction of several successive maneuvering decisions is taken into account by constructing a multistage influence diagram that offers a way to model and analyze sequential decision problems (see Refs. 1 and 20). The new extended model contains components describing the preferences of the pilot, uncertainty, as well as the decision and aircraft dynamics. It associates a probability and a utility measuring the overall preferences with each combat situation and allows a possibility to determine optimal foresighted flight paths against a given adversary trajectory with respect to the preferences of the pilot.

Traditional solution methods of influence diagrams[5,21,22] produce the best value for the decision variable as a function of the information available at the decision instant, that is, the solution is in a closed-loop form. However, these solution methods require an enormous computational effort when a multistage influence diagram with several decision alternatives and stages is to be solved. Hence, in practice, the length of the time horizon must be limited.

Received 16 March 2001; presented as Paper 2001-4267 at the AIAA Guidance, Navigation, and Control Conference, Montreal, Quebec, Canada, 6–9 August 2001; revision received 11 December 2003; accepted for publication 12 December 2003.

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

†Researcher, Systems Analysis Laboratory, P.O. Box 1100; tuomas.raivio@hut.fi.

‡Professor, Director of the Systems Analysis Laboratory, P.O. Box 1100; raimo@hut.fi.

Here, the computational difficulties are overcome by converting the multistage influence diagram into a discrete-time dynamic optimization problem that can be solved off-line using nonlinear programming methods (see Ref. 4). In the optimization over time, the sequence of the myopic single-stage closed-loop solutions is used as an initial estimate. In this way, preference optimal maneuvering decisions maximizing the overall utility are obtained. The best decision alternatives are now chosen without knowing the exact outcome of the different uncertainties at the decision stages, that is, the solution is in an open-loop form.

The paper is organized as follows. First, a short introduction to influence diagrams and their solution methods is given. In Sec. III, the sequential maneuvering problem is structured and modeled by a multistage influence diagram. In Sec. IV, first, the generation of myopic solutions using the single-stage influence diagram is introduced. Then, a dynamic discrete-time optimization problem representing the multistage model is formulated. In Sec. V, the solution procedure of the multistage influence diagram is demonstrated by a numerical example. In Sec. VI, the structure of the overall objective function, stability of control policies, and the utilization of the approach are discussed as well as improvements for refining the structure of the model are suggested. In addition, ideas related to the extension of the approach to game situations are given. Finally, concluding remarks appear in Sec. VII.

II. Influence Diagrams

An influence diagram[5] is a directed graph that consists of a set of nodes and arcs. The nodes include decision, deterministic, and random variables, and the arcs represent functional or probabilistic dependencies as well as the available information. A chance node contains a continuous or discrete random variable. An arc leading into a chance node implies that the probability distribution of the random variable depends on its predecessors. In addition, it can denote time precedence. A decision node contains a set of decision alternatives or a continuous decision variable. An arc into a decision node indicates that the values of the node's predecessors are known at the time the decision is made. A deterministic quantity or variable is modeled by a deterministic node whose value is either a constant or a function of its inputs. A utility node includes a utility function that expresses the preferences of the decision maker and evaluates the outcomes of decision alternatives.

A. Example

A simple one-stage decision situation is first considered. A single-stage influence diagram modeling this problem is shown in Fig. 1. It consists of decision, chance, and utility nodes. The decision alternatives are d_1 and d_2, and the chance node has the outcomes x_1 and x_2. The arcs imply that the chance node is independent of the other nodes. Thus, only the probabilities $P(x_1)$ and $P(x_2)$ have to be assessed. The utility function $u(d_i, x_j)$ depends on both the decision and chance nodes.

Next, a situation where decisions are made in several stages, and both decision variables and probability distributions are discrete, is considered. A multistage influence diagram that is essentially a chain of single-stage influence diagrams represents the decision problem of this type. An example of a multistage decision process is shown in Fig. 2.

There are two alternatives at both decision stages d_1^i and d_2^i, $i = 1, 2$. Here, the superscript refers to the decision stage. The direction of the arc between the decision nodes shows the chronological order

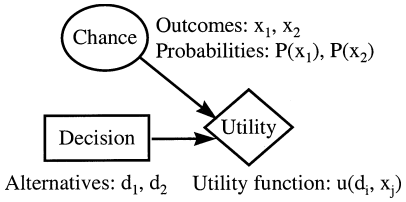

Fig. 1 Example of a single-stage influence diagram.

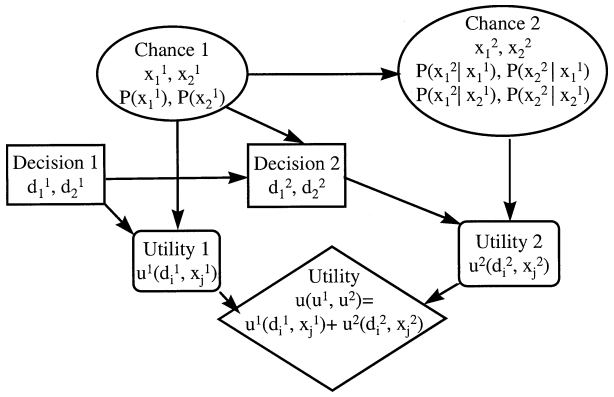

Fig. 2 Example of a multistage influence diagram.

of the decisions. The arc leading from the first chance node into the second decision node implies that the first chance node's outcome is known at the second decision stage. The first chance node is independent, but the second one depends on the outcome of the previous chance node. Therefore, the conditional probabilities $P(x_i^2|x_j^1)$, i, $j = 1, 2$, are needed. At each stage, the utility $u^i(\,)$, $i = 1, 2$, is a function of the decision and the outcome of the chance node. The aggregated utility $u(\,)$ is the sum of the individual utilities.

B. Solution Methods

Influence diagram analysis results in a probability distribution for the utility of each decision alternative. The most desirable decision alternative is selected on the basis of these distributions. One possible criterion is to maximize the overall expected utility (e.g., see Ref. 2). In general, the solution procedures provide the expected utility maximizing alternatives directly. If necessary, the probability distributions of utility can be computed during the solution procedure.

Howard and Matheson[5] introduce a way to solve an influence diagram by converting it into a decision tree (e.g., see Ref. 3) and solving this tree. In fact, any influence diagram can be converted into a symmetric decision tree and vice versa, although the transformation might require the use of the Bayes' theorem (e.g., see Ref. 17). This is illustrated by a simple example in Ref. 19. A more detailed description on the correspondence of influence diagrams and decision trees is given in Ref. 1. The modification procedure has been automated in the existing decision support software (e.g., see Ref. 23). The influence diagram and tree representations of decision problems have different advantages. The diagrams provide a compact representation, but they have shortcomings in asymmetric decision problems (see Ref. 24). The problems of this type can be presented in detail by decision trees, but they tend to become very large in complex problems.

When an influence diagram to be resolved contains only discrete decision variables and probability distributions, the most straightforward way to solve the corresponding decision tree is the rollback procedure (e.g., see Ref. 1). It is an application of dynamic programming[25] and proceeds in reverse chronological order from the leaf nodes toward the root node. The expected utility is calculated at each chance node, and at each decision node the decision alternative with the highest expected utility is selected. As a result, the branch of the tree leading to the highest expected utility is found. The use of the rollback procedure is illustrated by a simple example in Ref. 19. Shachter[21] presents an alternative solution method in which influence diagrams need not be converted into decision trees.

Multistage influence diagrams can also be converted into decision trees, but the size of the tree increases rapidly with the number of stages. Solving large models with the rollback procedure is practically impossible. Thus, alternative ways to analyze sequential decision models have been developed. In recursive methods[22,26] the structure of a decision tree is described by using the so-called next node functions. Then, decision alternatives with the highest expected utility can be calculated recursively. An approach that allows dynamic programming to be applied within the influence diagram

framework is suggested in Ref. 20. The traditional recursive equation of dynamic programming needs not be set up, but the curse of dimensionality (e.g., see Ref. 25) still persists. Warburton[27] moves away from recursive techniques and presents an alternative solution methodology. In this approach, a decision tree is converted into a set of linear and integer optimization problems. This technique offers a method to analyze tradeoffs among competing objectives, and thus the utility functions are not needed explicitly.

Influence diagrams can also be employed in problems with continuous decision variables and/or continuous probability distributions. The most straightforward way to solve such a diagram is to discretize the continuous variables and then apply the methods just discussed. However, if the model contains several continuous variables, they must be discretized coarsely enough to keep the problem size reasonable. Unfortunately coarse approximations can lead to inaccurate results.

If a diagram contains discrete decision variables and continuous probability distributions, it can be solved approximately by Monte Carlo simulation (see Ref. 28). The results give approximate distributions of the expected utility for each decision alternative. If the decision variables are continuous, approximate distributions can be calculated for discrete values of the variables. It is impossible to determine distributions related to all possible values of the continuous decision variables. Thus, the result of the analysis depends on the discretization.

Stonebraker and Kirkwood[29] present an approach for analyzing decision trees and influence diagrams with continuous variables using nonlinear programming. It is a generalization of the method presented in Ref. 22, which solves decision models containing discrete variables. In this approach, the formulation of a sequential decision model is based on the use of next node functions. Then, the recursive calculation of expected utilities is carried out with a nonlinear optimization method. Although the approach is appealing, it cannot be applied directly to an influence diagram that contains several decision stages and complicated interdependencies between the elements because the reformulation is a complex task and numerous optimizations are needed in the solution procedure.

C. Solution Example

Accurate results of small influence diagrams with continuous variables can be calculated by the rollback procedure. To demonstrate the connection of influence diagrams and decision trees as well as the use of nonlinear programming, the preceding two-stage diagram is considered. The discrete decision alternatives d_i^1 and d_i^2, $i = 1, 2$, are replaced with continuous decision variables that are denoted by d^1 and d^2. The decision-tree representation of this model is shown in Fig. 3.

When the two-stage decision tree is solved with the rollback procedure, the expected utilities of the second chance node $EU_i^2(d^1, d^2)$, $i = 1, 2$, where i refers to the outcome x_i^1 of the first chance node, are calculated as

$$EU_i^2(d^1, d^2) = P\left(x_1^2 | x_i^1\right)\left[u^1\left(d^1, x_i^1\right) + u^2\left(d^2, x_1^2\right)\right] + P\left(x_2^2 | x_i^1\right)\left[u^1\left(d^1, x_i^1\right) + u^2\left(d^2, x_2^2\right)\right], \qquad i = 1, 2 \quad (1)$$

The optimal value of the second decision variable is then determined. The optimal solutions $d^{2^*}(d^1, x_i^1)$, $i = 1, 2$, are obtained by solving

$$d^{2^*}\left(d^1, x_i^1\right) = \arg\max_{d^2} EU_i^2(d^1, d^2), \qquad i = 1, 2 \quad (2)$$

The maximization determines the optimal value of the second decision variable as a function of the first chance node's outcome and the first decision variable. When proceeding toward the root node, the next step is to calculate the expected utility at the first chance node:

$$EU^1(d^1) = P\left(x_1^1\right) EU_1^2\left[d^1, d^{2^*}\left(d^1, x_1^1\right)\right] + P\left(x_2^1\right) EU_2^2\left[d^1, d^{2^*}\left(d^1, x_2^1\right)\right] \quad (3)$$

where $EU_i^2[d^1, d^{2^*}(d^1, x_i^1)]$ is the expected utility with the given d^1 and the optimal value of the second decision variable determined by Eq. (2). Finally, the optimal value of the first decision variable is obtained by solving the optimization problem

$$d^{1^*} = \arg\max_{d^1} EU^1(d^1) \quad (4)$$

The solution of Eq. (2) must be obtained in a closed form in order to be able to express Eq. (3) explicitly and further to apply the rollback procedure. If the closed form does not exist, the solution of the influence diagram could be determined by solving a bilevel optimization problem (see Ref. 30) in which the objective function (4) is maximized subject to constraints consisting of the maximization operations given by Eq. (2).

The rollback procedure can also be applied in the analysis of models with continuous probability distributions. Then, at each chance node the expected utility is calculated by integrating the product of a probability density function and a utility function.

D. Open- and Closed-Loop Solutions

Let us point out the differences between closed-loop and open-loop solutions. Assume that one is dealing with an n-stage decision process and searches for the optimal decision sequence $d^1, d^2, \ldots, d^n$. Consider first an open-loop solution. Using the definition of Ref. 25, all of the decisions $d^1, d^2, \ldots, d^n$ are then made at stage 1, without knowing the outcomes of the chance nodes in the chain. In a closed-loop solution, the value of the decision variable d^k is selected at the latest possible moment. Hence, a decision at stage k is made only after the outcomes of the chances before that stage are known. The rollback procedure as well as the other aforementioned solution approaches produce closed-loop solutions.

The solution technique based on rollback and nonlinear optimization is applicable, if the number of decision stages as well as the number of discrete outcomes of chance nodes is not very high. However, the solution of large sequential decision models contains numerous optimization problems or a multilevel optimization problem. Furthermore, it might be impossible to find a closed-loop solution in an analytical form that is required, for example, in the recursive approach.[29]

A preference optimal open-loop solution is obtained by searching the sequence of decision variables such that the cumulative expected utility over all decision stages is maximized. In the example of the

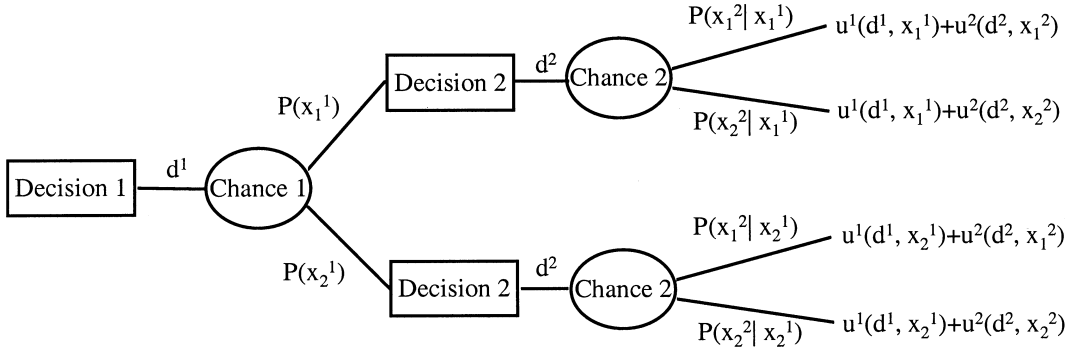

Fig. 3 Two-stage decision tree with continuous decision variables d^1 and d^2.

previous subsection, the optimization problem to be solved is

$$\max_{d^1,d^2} \sum_{i=1}^{2} \sum_{k=1}^{2} P\left(x_k^i\right) u^i\left(d^i, x_k^i\right) \tag{5}$$

Here the probabilities $P(x_i^2)$ can be solved using the formula of total probability. It can be shown (see the Appendix) that Eq. (5) can be expressed equivalently as

$$\max_{d^1,d^2} P\left(x_1^1\right) EU_1^2(d^1, d^2) + P\left(x_2^1\right) EU_2^2(d^1, d^2) \tag{6}$$

Note that Eq. (6) is similar to Eqs. (1–4), but now the optimal values of the decision variables are chosen concurrently. Therefore, in the determination of the optimal open-loop solution of multistage influence diagrams only one nonlinear optimization problem is solved.

III. Multistage Influence Diagram Model for Maneuvering Decisions

The basis for the modeling of the pilot's sequential maneuvering process is a single-stage influence diagram presented in Ref. 19. This model represents the single maneuvering decision of a pilot who is engaged in a duel between two aircraft using guns as their weapons. The maneuvering decision is modeled by assuming that the decision-making process of the pilot consists of the following steps:

1) Observe the state of the adversary.

2) Predict the future state of the adversary based on the available observations.

3) Predict the future states of own aircraft that can be reached by feasible maneuvering alternatives.

4) Evaluate possible future states of the combat.

5) Find an optimal maneuvering alternative in this state.

6) Implement the maneuver, return to step 1.

The maneuvering decision process involves different sources of uncertainty, like state measurements, adversary's goals and performance, and conflicting objectives. Such issues have been analyzed in detail in the model constructed in Ref. 19. In this paper, a simplified version of the model is used as the basic building block. Here, the state measurement is assumed exact. The impact of possible uncertainty could be modeled, for example, by using normal probability distributions (for details, see Ref. 19).

The flying actors are now called the decision maker and the opponent. The first term refers to the actor whose maneuvering is optimized. The other one is a nonreacting actor obeying a predetermined trajectory. The decision maker is assumed to have two objectives that are, in the order of importance, as follows: 1) avoid the opponent's weapons and 2) achieve a firing position.

Because of the gun duel, an aircraft is in a firing position whenever it is close behind the other. In other words, the decision maker tries to avoid situations where the opponent can acquire a tail position and, at the same time, aims at reaching the opponent's tail.

The decision maker's sequential maneuvering process is modeled by connecting several simplified single-stage models together. The resulting multistage influence diagram, shown in Fig. 4, represents n sequential maneuvering decisions that are made at time instants $t_i = i\,\Delta t, i = 0, \ldots, n-1$. Here, the time between two decisions Δt is called the decision interval.

The successive maneuvering decisions of the decision maker are represented by the nodes Maneuver at t_i, $i = 0, \ldots, n-1$. Each node has a continuous control vector $\boldsymbol{C}_i = [n_i, \mu_i, u_i]^T$, where the subscript denotes the decision stage. The variables of the control vector refer to the load factor, the bank angle, and the throttle setting. The values of these variables cannot be chosen freely because of the constraints set by the pilot and the aircraft itself. The load factor has lower and upper limits $n_i \in [n^{\min}, n^{\max}]$, and the throttle setting can vary between zero and one. In general, the limits of the load factor depend on the altitude and the velocity, but here they are assumed constant. The arcs leading into the decision nodes show that, at each decision instant, the decision maker knows the momentary state of his or her own aircraft, the momentary state of the combat, and the current threat assessment.

The deterministic nodes State at t_i, $i = 0, \ldots, n$, contain the state vector of the decision maker's aircraft $\boldsymbol{X}_i = [x_i, y_i, h_i, v_i, \gamma_i, \chi_i]^T$, where the variables refer to the x range, the y range, the altitude,

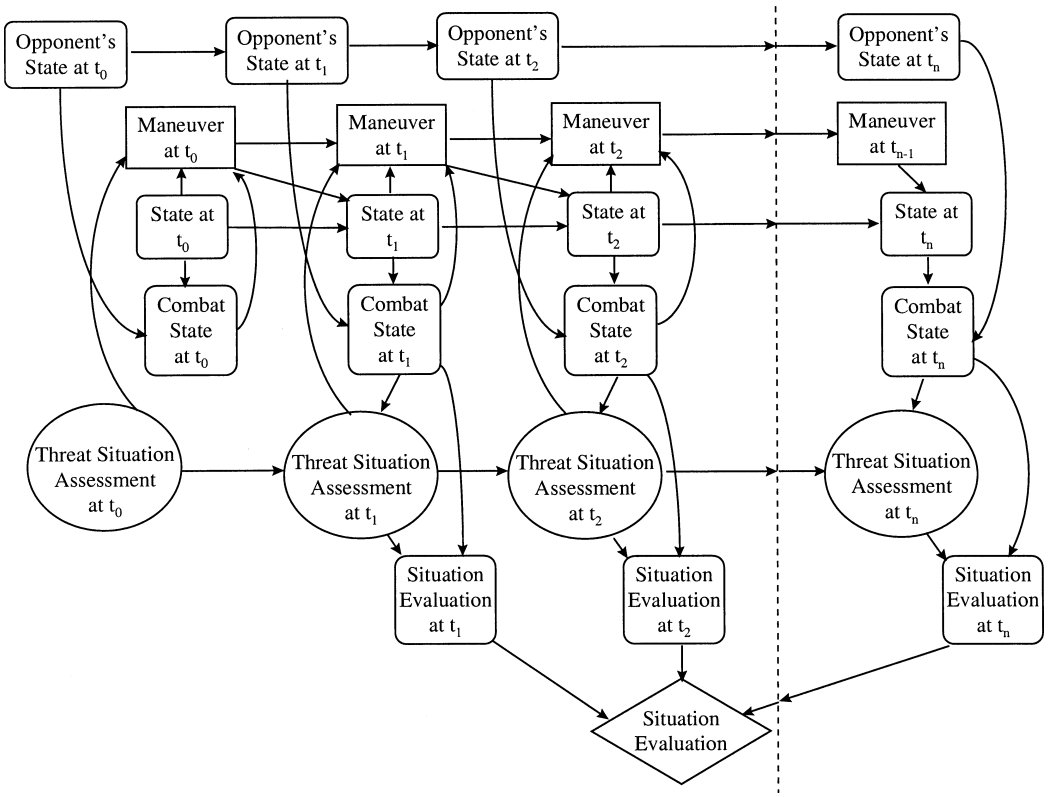

Fig. 4 Multistage influence diagram for a sequence of *n* maneuvering decisions.

the velocity, the flight-path angle, and the heading angle. The feasible region of stationary flight is defined by the minimum altitude constraint

$$H(h_i) = h^{\min} - h_i \le 0 \tag{7}$$

the minimum velocity constraint

$$V(v_i) = v^{\min} - v_i \le 0 \tag{8}$$

and the maximum dynamic pressure constraint

$$Q(h_i, v_i) = \tfrac{1}{2}\rho(h_i)v_i^2 - q^{\max} \le 0 \tag{9}$$

where $\rho(h_i)$ is the air density.

The arcs pointing to the state nodes imply that the current state depends on the control and the state at the previous decision instant. This relationship is taken into account with a three-degrees-of-freedom point-mass model that describes the motion of the aircraft. The equations of motion consist of nonlinear differential equations that are discretized using the Euler method. Hence, the evolution of the state is represented by the difference equations

$$\boldsymbol{X}_{i+1} = \boldsymbol{X}_i + \boldsymbol{f}(\boldsymbol{X}_i, \boldsymbol{C}_i)\Delta t, \qquad i = 0, \ldots, n-1 \tag{10}$$

where

$$\boldsymbol{f}(\boldsymbol{X}_i, \boldsymbol{C}_i) = \begin{cases} v_i \cos\gamma_i \cos\chi_i \\ v_i \cos\gamma_i \sin\chi_i \\ v_i \sin\gamma_i \\ \dfrac{1}{m}\{u_i T_{\max}[h_i, M(h_i, v_i)] - D[h_i, v_i, M(h_i, v_i), n_i]\} - g\sin\gamma_i \\ \dfrac{g}{v_i}(n_i\cos\mu_i - \cos\gamma_i) \\ \dfrac{g}{v_i}\dfrac{n_i \sin\mu_i}{\cos\gamma_i} \end{cases} \tag{11}$$

The aircraft mass m and the gravitational acceleration g are assumed constant. The Mach number $M(\,)$ as well as the density of the air are computed on the basis of the International Standard Atmosphere. $T_{\max}(\,)$ denotes the maximum available thrust force. The drag force $D(\,)$ depends on zero-lift and induced drag coefficients, and it is assumed to obey a quadratic polar. For the details of the model, see Ref. 9.

The predetermined trajectory of the opponent is given by the deterministic nodes Opponent's State at t_i, $i = 0, \ldots, n$. Each node contains the state vector $\boldsymbol{X}_i^{OP} = [x_i^{OP}, y_i^{OP}, h_i^{OP}, v_i^{OP}, \gamma_i^{OP}, \chi_i^{OP}]^T$ that is the true value of the opponent's state at time instant t_i. The meaning of the variables is same as in the decision maker's state vector. The arcs connecting the opponent's state nodes represent their chronological order.

The momentary state of the combat at time t_i is defined by the states of the aircraft. It is described by the combat state vector $\boldsymbol{CS}_i = [\alpha_i^1, \alpha_i^2, \beta_i, d_i, \Delta h_i, \Delta v_i^2, h_i, v_i^2]^T$ (see Fig. 5) computed in the deterministic Combat State at t_i node. The first elements of the vector are the angle between the line of sight (LOS) and the decision maker's velocity vector, called the deviation angle,

$$\alpha_i^1 = \arccos\left\{\left[\left(x_i^{OP} - x_i\right)\cos\gamma_i\cos\chi_i + \left(y_i^{OP} - y_i\right)\cos\gamma_i\sin\gamma_i + \left(h_i^{OP} - h_i\right)\sin\gamma_i\right]/d_i\right\} \tag{12}$$

the angle between the LOS and the opponent's velocity vector, called the angle off,

$$\alpha_i^2 = \arccos\left\{\left[\left(x_i^{OP} - x_i\right)\cos\gamma_i^{OP}\cos\chi_i^{OP} + \left(y_i^{OP} - y_i\right) \times \cos\gamma_i^{OP}\sin\gamma_i^{OP} + \left(h_i^{OP} - h\right)\sin\gamma_i^{OP}\right]/d_i\right\} \tag{13}$$

the angle between the velocity vectors of the aircraft

$$\beta_i = \arccos\left(\cos\gamma_i\cos\chi_i\cos\gamma_i^{OP}\cos\chi_i^{OP} + \cos\gamma_i\sin\gamma_i\cos\gamma_i^{OP}\sin\gamma_i^{OP} + \sin\gamma_i\sin\gamma_i^{OP}\right) \tag{14}$$

and the distance between the aircraft

$$d_i = \sqrt{\left(x_i^{OP} - x_i\right)^2 + \left(y_i^{OP} - y_i\right)^2 + \left(h_i^{OP} - h_i\right)^2} \tag{15}$$

The energy difference of the aircraft is taken into account by the difference between the altitudes

$$\Delta h_i = h_i^{OP} - h_i \tag{16}$$

and the difference between the squares of the velocities

$$\Delta v_i^2 = \left(v_i^{OP}\right)^2 - v_i^2 \tag{17}$$

In addition, the altitude h_i and the square of the velocity v_i^2 of the decision maker's aircraft are taken as combat state variables.

The chance nodes Threat Situation Assessment at t_i, $i = 0, \ldots, n$, infer the threat situation from the decision maker's point of view. Each node contains a discrete random variable Θ_i whose outcomes are associated with the given relative geometry of the combat. The possible outcomes are as follows: $\Theta_i =$ neutral, θ_1; $\Theta_i =$ advantage, θ_2; $\Theta_i =$ disadvantage, θ_3; and $\Theta_i =$ mutual disadvantage, θ_4.

At the decision instant t_i, the decision maker's belief on the threat situation is described by the probabilities $P(\Theta_i = \theta_1)$, $P(\Theta_i = \theta_2)$, $P(\Theta_i = \theta_3)$, and $P(\Theta_i = \theta_4)$ such that

$$\sum_{k=1}^{4} P(\Theta_i = \theta_k) = 1$$

They are calculated on the basis of the current combat state and the probabilities of the previous threat node. The probabilities $P(\Theta_i = \theta_k)$, $k = 1, \ldots, 4$, can be considered as the prior belief for the next decision moment t_{i+1}. The posterior belief of this stage is formed after the value of the combat state $\boldsymbol{CS}_{i+1}$ is observed, and it is given in the chance node Threat Situation Assessment at t_{i+1}. The posterior probabilities are determined by using the Bayes' theorem,

$$\begin{aligned} P(\Theta_{i+1} = \theta_k) &\hat{=} P\left(\Theta_{i+1} = \theta_k \middle| A^1 = \alpha_{i+1}^1, A^2 = \alpha_{i+1}^2, D = d_{i+1}\right) \\ &= \frac{\bar{P}(\Theta_{i+1} = \theta_k)\, p\left(\alpha_{i+1}^1, \alpha_{i+1}^2, d_{i+1} \middle| \Theta_{i+1} = \theta_k\right)}{\sum_{k=1}^{4} \bar{P}(\Theta_{i+1} = \theta_k)\, p\left(\alpha_{i+1}^1, \alpha_{i+1}^2, d_{i+1} \middle| \Theta_{i+1} = \theta_k\right)} \\ &= \frac{P(\Theta_i = \theta_k)\, p\left(\alpha_{i+1}^1, \alpha_{i+1}^2, d_{i+1} \middle| \Theta_{i+1} = \theta_k\right)}{\sum_{k=1}^{4} P(\Theta_i = \theta_k)\, p\left(\alpha_{i+1}^1, \alpha_{i+1}^2, d_{i+1} \middle| \Theta_{i+1} = \theta_k\right)} \\ &\qquad k = 1, \ldots, 4, \qquad i = 0, \ldots, n-1 \end{aligned} \tag{18}$$

where the prior probabilities $\bar{P}(\Theta_{i+1} = \theta_k)$ are equal to the previous posterior probabilities $P(\Theta_i = \theta_k)$, and A^1, A^2, and D are the continuous random variables for the deviation angle, the angle off, and the distance. Their conditional joint probability density is $p(\cdot|\cdot)$.

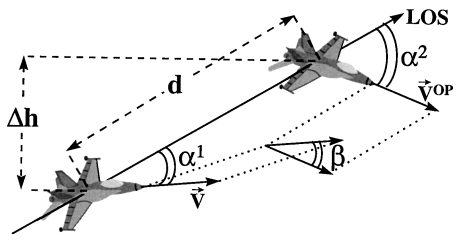

Fig. 5 Variables of the combat state vector.

A^1, A^2, and D are assumed to be conditionally independent given $\Theta_{i+1} = \theta_k$. Thus,

$$p\left(\alpha_{i+1}^1, \alpha_{i+1}^2, d_{i+1} | \Theta_{i+1} = \theta_k\right) = p_{A^1}\left(\alpha_{i+1}^1 \middle| \Theta_{i+1} = \theta_k\right) \times p_{A^2}\left(\alpha_{i+1}^2 \middle| \Theta_{i+1} = \theta_k\right) p_D(d_{i+1} | \Theta_{i+1} = \theta_k) \quad (19)$$

where the time-invariant likelihood functions $p_\cdot(\cdot|\Theta_{i+1} = \theta_k)$ represent the distributions of A^1, A^2, and D under the supposition that the outcome of the Threat Situation Assessment at t_{i+1} node is θ_k. The sketch of the likelihood functions used in this study is shown in Fig. 6.

The outcome "advantage" of the threat assessment refers to a situation where the decision maker is reaching or has reached the opponent's tail. Then, there are high probabilities that the angles are small and the distance is short, and these probabilities increase when the angles and the distance get close to zero. The "disadvantage" outcome stands for an inverse situation: the likelihood functions increase when the angles approach to 180 deg and the distance approaches zero. If the aircraft are close to each other and, at the same time, fly toward each other, the outcome of the threat assessment is assumed to be "mutual disadvantage." High probabilities are now associated with a small value of the deviation angle and the distance as well as a large angle off. The "neutral" outcome implies that the aircraft move from away each other or are wide apart. The probability of a large deviation angle and a small angle off is high. The likelihood function of the distance is constant. Whenever the distance is long, its other likelihood functions give low probabilities. Thus, in the situation of this type, the probability of the neutral outcome, obtained from Eq. (18), becomes close to one regardless of the angles.

The combat situation is evaluated separately at each stage by a utility function $u(\,)$, calculated in the deterministic Situation Evaluation at t_i, $i = 1, \ldots, n$, nodes. The utility function reflects the pilot's preferences and, in practice, evaluates the possible consequences of the maneuvering alternatives. The preferable actions of the decision maker depend on the threat situation as well as on the state of the combat. Thus, the utility is a function of each component of the combat state vector and the outcome of the threat assessment. In practice, four different utility functions are used, and each of them is associated with the particular outcome of the threat assessment node. The utility functions are

$$U_i(\theta_k, \boldsymbol{CS}_i) \hat{=} u\left(\theta_k, \alpha_i^1, \alpha_i^2, \beta_i, d_i, \Delta h_i, \Delta v_i^2, h_i, v_i^2\right) = w_{\alpha_1}^k u_{\alpha_1}^k\left(\alpha_i^1\right) + w_{\alpha_2}^k u_{\alpha_2}^k\left(\alpha_i^2\right) + w_\beta^k u_\beta^k(\beta_i) + w_d^k u_d^k(d_i) + w_{\Delta h}^k u_{\Delta h}^k(\Delta h_i) + w_{\Delta v}^k u_{\Delta v}^k\left(\Delta v_i^2\right) + w_h^k u_h^k(h_i) + w_v^k u_v^k\left(v_i^2\right)$$
$$k = 1, \ldots, 4, \qquad i = 1, \ldots, n \quad (20)$$

Here $u_\cdot^k$ is a single-attribute utility function that maps an attribute onto a utility scale such that the best value of the attribute has a utility of one and the worst has a utility of zero. Positive weights $w_\cdot^k$ sum up to one and represent the importance of the attributes.

In Eq. (20), the overall utility is aggregated by calculating a linear combination of the single utilities. In the decision science literature, the form of aggregation of this type is called additive. It is a valid representation when the attributes are mutually utility independent (see Ref. 2). Otherwise, a multiplicative form can be used. It is composed by adding product terms of the single utilities to the

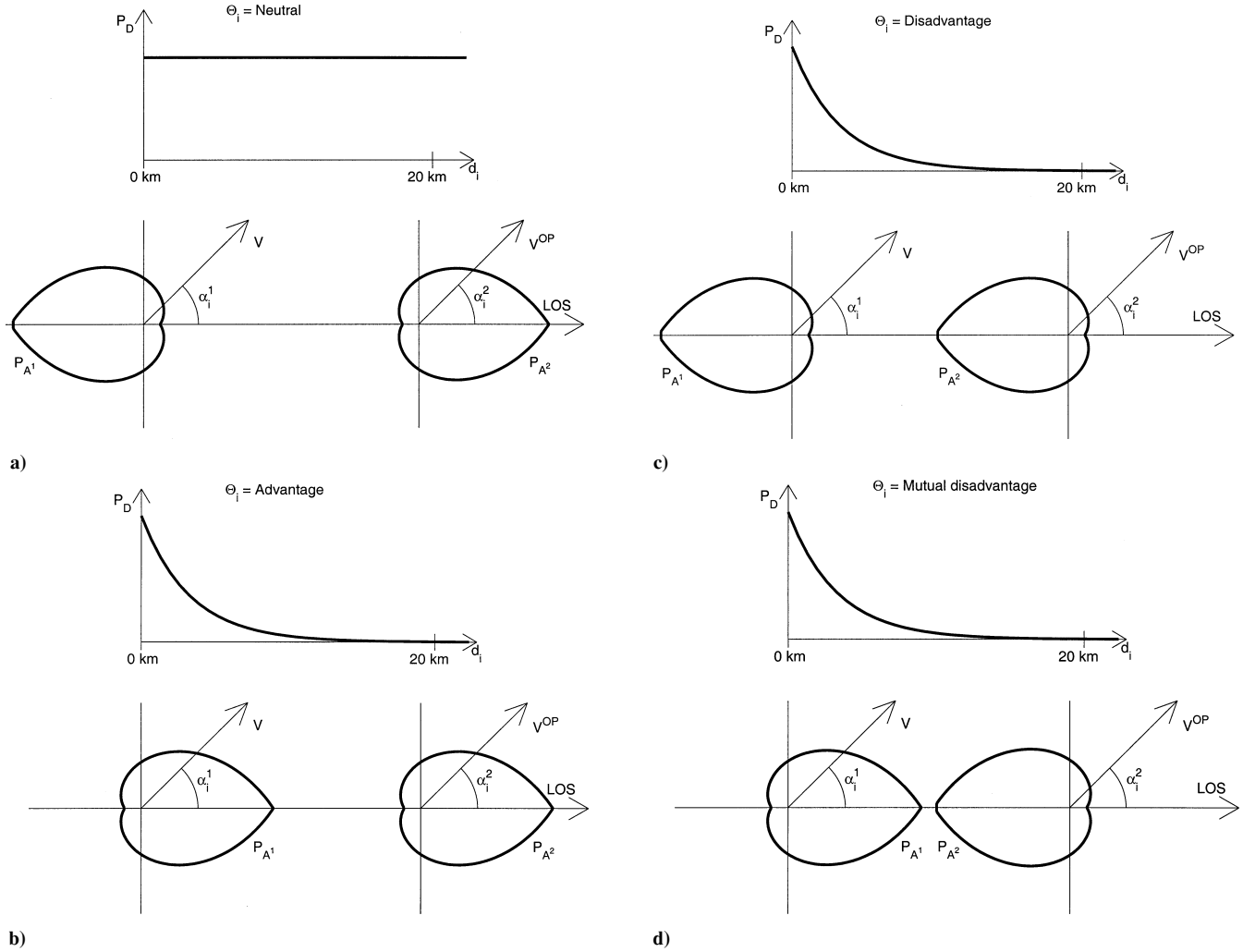

Fig. 6 Likelihood functions of the deviation angle, the angle off, and the distance.

Table 1 Weights and objectives of the utility functions[a]

Attribute	Weight	Objective
	Neutral	
d	0.3	$\infty \rightarrow 0$ m
α^1	0.21	$180 \rightarrow 0$ deg
α^2	0.19	$180 \rightarrow 0$ deg
h	0.15	$0 \rightarrow \infty$ m
v^2	0.15	$0 \rightarrow \infty$ m^2/s^2
	Disadvantage	
α^2	0.34	$180 \rightarrow 0$ deg
β	0.33	$0 \rightarrow 180$ deg
d	0.33	$0 \rightarrow \infty$ m
	Advantage	
d	0.31	$0, \infty \rightarrow 800$ m
α^1	0.27	$180 \rightarrow 0$ deg
α^2	0.27	$180 \rightarrow 0$ deg
Δh	0.09	$-\infty, \infty \rightarrow 0$ m
Δv^2	0.06	$-\infty, \infty \rightarrow 0$ m^2/s^2
	Mutual disadvantage	
α^2	0.25	$180 \rightarrow 0$ deg
α^1	0.21	$0 \rightarrow 180$ deg
β	0.21	$0 \rightarrow 180$ deg
d	0.18	$0 \rightarrow \infty$ m
h	0.08	$0 \rightarrow \infty$ m
v^2	0.07	$0 \rightarrow \infty$ m^2/s^2

[a]For example, the objective $\infty \rightarrow 0$ m related to the distance means that short distances are preferred, that is, the single-attribute utility function $u_d(d)$ is decreasing and $u_d(0) = 1$, $u_d(\infty) = 0$.

additive utility function. When the underlying uncertainties have been omitted, two attributes are independent, if the preference order for the level of one attribute does not depend on the level of the other attribute. Because each outcome of the threat assessment leads to the particular utility function, the attributes of Eq. (20) seem to satisfy this criterion. For example, if the threat situation is assumed to be advantage, short distances are preferred regardless of the levels of the other attributes. On the other hand, in a disadvantageous situation, the values of the deviation angle and the angle off do not affect the preference order for the levels of the distance because one prefers long distances at all of the cases. For an example on the use of multiplicative preference models, see Ref. 31.

The weights and the objectives in Eq. (20) are shown in Table 1. In an advantage situation, the decision maker tries to reach or stay at the opponent's tail. Hence, the deviation angle, the angle off, and the distance are assumed to be the most important attributes whose relative importance is almost equal. The less important attributes are the altitude and the velocity, which are aimed at matching with the opponent's aircraft. The values of the controls are selected such that the angles, the velocity, and altitude difference are minimized, and the distance is as close to 800 m as possible. If the outcome disadvantage of the threat assessment has a high probability, the aim is to avoid the front sector of the opponent's aircraft regardless of other factors, like the energy advantage. The angle off is minimized, and the angle between the velocity vectors of the aircraft as well as the distance between them is maximized. These attributes are assumed to be equally important.

When the situation of the combat is disadvantageous for both the actors, the main goal is to avoid the opponent's front sector, and the secondary goal is to maintain or even increase the energy that could be utilized in the future. Here, the angles are the most essential attributes. Their relative importance is slightly higher than the importance of the distance. The energy factor is taken account by maximizing the attributes h and v^2 whose weights are lower than the other weights. In a neutral situation, the decision maker aims at achieving the opponent's tail sector and increasing the energy. The importance of the distance is now the highest because the neutral outcome refers often a situation where the aircraft are far away each other. The second important attributes are the deviation angle and the angle off. The lowest weights are associated with h and v^2 that reflect the energy level. The attributes h and v^2 could also be replaced with Δh and Δv^2. Then, the importance of the energy advantage should be evaluated.

The utility node Situation Evaluation contains the aggregated utility function over all the decision stages of the diagram. It is the sum of the single-stage utilities

$$\bar{U}(U_1, \ldots, U_n) = \sum_{i=1}^{n} U_i(\theta_k, \boldsymbol{CS}_i) \tag{21}$$

Finally, the definition of the multistage influence diagram needs the initial state vector of the decision maker $\boldsymbol{X}_0$ as well as the initial probability distribution of the threat assessment $P(\Theta_0 = \theta_k)$, $k = 1, \ldots, 4$.

IV. Solution Procedure

Assume that a human expert whose preferences and opinions are captured into the influence diagram model is prepared to accept the utility theoretical definition of rationality, that is, the axioms of the utility theory (e.g., see Ref. 3). The optimal solution of the influence diagram is then the sequence of controls that provides the highest cumulative expected utility. The solution procedure for obtaining such controls is introduced by utilizing the decision-tree representation of the multistage influence diagram shown in Fig. 7. Note that there is a similar correspondence between the representations in Figs. 2 and 3, as well as in Figs. 4 and 7.

The closed-loop solution of the decision tree cannot be obtained because it is impossible to express the optimal value of $\boldsymbol{C}_i$, $i = 1, \ldots, n-1$, as a function of the preceding control variables because of the nonlinearities in Eqs. (10) and (18) and, on the other hand, an n-level optimization problem is computationally intractable. An approximate closed-loop solution could be obtained by discretizing the continuous control variables. In general, n is so large in aircraft trajectory optimization problems that the solution is time consuming or even impossible by using available computers. Now, the initial estimate of a preference optimal flight path, the myopic closed-loop solution, is generated by shortening the time horizon of the model and solving the single-stage influence diagram containing discrete control variables at each decision stage.

In Sec. II.D, the analogy between the simple model shown in Fig. 3 and the optimization problem (5) is pointed out. Similarly, there is an equivalent nonlinear optimization problem corresponding to the decision-tree representation of the multistage influence diagram. The objective function of the problem, the cumulative expected utility, to be maximized is

$$\sum_{i=1}^{n} \sum_{k=1}^{4} P(\Theta_i = \theta_k) U_i(\theta_k, \boldsymbol{CS}_i) \tag{22}$$

It depends on the combat state and the threat probabilities. Hence, the interdependencies that map the states of the aircraft into the combat state as well as into the probabilities and the set of difference equations describing the evolution of the decision maker's state must be taken as the constraints of the optimization problem. By using the myopic solution as an initial estimate and solving the resulting discrete-time dynamic optimization problem, the open-loop solution

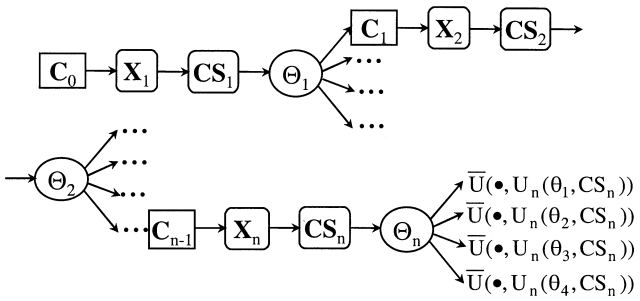

Fig. 7 Decision-tree representation corresponding to the multistage influence diagram.

that maximizes the cumulative expected utility over all the decision stages is obtained.

A. Myopic Solution

The solution of a myopic closed-loop maneuvering sequence is first described. In the beginning, the opponent's trajectory $\boldsymbol{X}_i^{OP}$, $i=0,\ldots,N$, is given. Here, N defines the maximum number of the decision stages as well as the maximum duration of the flight $T_{\max}=N\Delta t$. In addition, the decision maker's initial state $\boldsymbol{X}_0$ and initial probability distribution of the threat assessment $P(\Theta_0=\theta_k)$, $k=1,\ldots,4$, must be fixed.

The myopic controls are generated at time instants $t_i=i\Delta t$, $i=0,\ldots,$ by solving the influence diagram presented in Sec. III in which the future states are predicted only one decision interval ahead. During the solution, the distribution of the threat assessment is updated such that the prior probabilities at the current stage are associated with the posterior probabilities of the previous stage. The evolution of the decision maker's state is computed according to the equations of motion.

The continuous control variables are replaced with discrete control alternatives in order to be able to avoid the use of nonlinear programming. In this way, the initial estimate is found quickly because, in practice, the objective function is evaluated with all of the feasible decision alternatives, and then the alternative leading the highest expected utility is selected. In fact, the single-stage influence diagram is being resolved in real time, but the computation time increases linearly with the number of the single-stage diagrams to be solved. The control alternatives are

$$\boldsymbol{CA}_i(j,k,l)=\begin{bmatrix} n_i(j)\\ \mu_i(k)\\ u_i(l)\end{bmatrix}=\begin{bmatrix} n_{i-1}+jn_\Delta\Delta t\\ \mu_{i-1}+k\mu_\Delta\Delta t\\ u_{i-1}+lu_\Delta\Delta t\end{bmatrix}$$

$$j,k,l=\{-1,0,1\} \quad (23)$$

where $n_i(j)$, $\mu_i(k)$, and $u_i(l)$ refer to the values of the controls at time t_i. Similarly n_{i-1}, μ_{i-1}, and u_{i-1} refer to the values of the controls that were used during the preceding decision interval for the period Δt. The control rates of change n_Δ, μ_Δ, and u_Δ are fixed.

At time t_i, the expected utility maximizing control alternative $\boldsymbol{CA}_i^*$ from among Eq. (23) is determined by solving the problem

$$\boldsymbol{CA}_i^*=\arg\max_{\substack{\boldsymbol{CA}_i(j,k,l)\\ j,k,l=\{-1,0,1\}}}\sum_{k=1}^{4}P(\Theta_{i+1}=\theta_k)\times U_{i+1}\left[\theta_k,\boldsymbol{CS}_{i+1}\left(\boldsymbol{X}_{i+1},\boldsymbol{X}_{i+1}^{OP}\right)\right] \quad (24)$$

subject to

$$\boldsymbol{X}_{i+1}=\boldsymbol{g}\left(\boldsymbol{X}_i^*,\boldsymbol{CA}_i,\Delta t\right) \quad (25)$$

$$P(\Theta_{i+1}=\theta_k)=h\left[P(\Theta_i=\theta_k)^*,\boldsymbol{CS}_{i+1}\left(\boldsymbol{X}_{i+1},\boldsymbol{X}_{i+1}^{OP}\right)\right] \quad k=1,\ldots,4 \quad (26)$$

$$\boldsymbol{CA}_i\in[n^{\min},n^{\max}]\times[-\infty,\infty]\times[0,1] \quad (27)$$

$$H(\boldsymbol{X}_{i+1})\le 0 \quad (28)$$

$$V(\boldsymbol{X}_{i+1})\le 0 \quad (29)$$

$$Q(\boldsymbol{X}_{i+1})\le 0 \quad (30)$$

where the functions $\boldsymbol{g}$, h, H, V, and Q are given by Eqs. (10), (18), (7), (8), and (9), respectively, and the components of the vector $\boldsymbol{CS}_{i+1}$ are obtained from Eqs. (12–17). $\boldsymbol{X}_i^*$ refers to the optimal state and $P(\Theta_i=\theta_k)^*$, $k=1,\ldots,4$, the optimal probabilities that are obtained with the optimal control $\boldsymbol{CA}_{i-1}^*$ at the earlier time instant t_{i-1}, that is,

$$\boldsymbol{X}_i^*=\boldsymbol{g}\left(\boldsymbol{X}_{i-1}^*,\boldsymbol{CA}_{i-1}^*,\Delta t\right) \quad (31)$$

and

$$P(\Theta_i=\theta_k)^*=h\left[P(\Theta_{i-1}=\theta_k)^*,\boldsymbol{CS}_i\left(\boldsymbol{X}_i^*,\boldsymbol{X}_i^{OP}\right)\right] \quad k=1,\ldots,4 \quad (32)$$

The problem (24–30) is solved at times t_i until the terminal condition

$$\Psi\left[\boldsymbol{CS}_i\left(\boldsymbol{X}_i,\boldsymbol{X}_i^{OP}\right)\right]=\left[\left(\alpha_i^1-\alpha_f^1\right),(\beta_i-\beta_f),(d_i-d_f)\right]^T\le\bar{0} \quad (33)$$

with the fixed upper limits α_f^1, β_f, and d_f, is satisfied or the time t_i is equal to the maximum duration of the flight $T_{\max}$. The condition (33) refers to a situation where the decision maker has achieved the opponent's tail. This gives us the number of the multistage model's decision stages n as well as the total duration of the flight $n\Delta t$. If the decision maker cannot achieve the desirable terminal state within the maximum flight time, there is no a feasible initial estimate considering to the terminal condition for the preference optimal flight path. However, the resulting initial trajectory can be improved with the help of the discrete-time dynamic optimization problem in which the fixed flight time is equal to the maximum duration of the flight.

B. Optimal Solution

Let us next move to the optimization over time. Assume that an initial estimate satisfying Eq. (33) is generated. Then, the discrete-time dynamic optimization problem corresponding to the present multistage influence diagram is

$$\max_X\sum_{i=1}^{n}\sum_{k=1}^{4}P(\Theta_i=\theta_k)U_i\left[\theta_k,\boldsymbol{CS}_i\left(\boldsymbol{X}_i,\boldsymbol{X}_i^{OP}\right)\right] \quad (34)$$

subject to

$$\boldsymbol{X}_{i+1}=\boldsymbol{g}(\boldsymbol{X}_i,\boldsymbol{C}_i,\Delta t) \qquad \text{given} \qquad \boldsymbol{X}_0 \quad (35)$$

$$P(\Theta_{i+1}=\theta_k)=h\left[P(\Theta_i=\theta_k),\boldsymbol{CS}_{i+1}\left(\boldsymbol{X}_{i+1},\boldsymbol{X}_{i+1}^{OP}\right)\right] \quad \text{given} \quad P(\Theta_0=\theta_k), \quad k=1,\ldots,4 \quad (36)$$

$$\boldsymbol{C}_i\in[n^{\min},n^{\max}]\times[-\infty,\infty]\times[0,1] \quad i=0,\ldots,n-1 \quad (37)$$

$$H(\boldsymbol{X}_j)\le 0 \quad (38)$$

$$V(\boldsymbol{X}_j)\le 0 \quad (39)$$

$$Q(\boldsymbol{X}_j)\le 0, \qquad j=1,\ldots,n \quad (40)$$

$$\Psi\left[\boldsymbol{CS}_n\left(\boldsymbol{X}_n,\boldsymbol{X}_n^{OP}\right)\right]\le\bar{0} \quad (41)$$

$$\Delta t>0 \quad (42)$$

Here the vector $\boldsymbol{X}$ of the decision variables is defined by $\boldsymbol{X}=[\boldsymbol{C}_0^T,\ldots,\boldsymbol{C}_{n-1}^T,\boldsymbol{X}_1^T,\ldots,\boldsymbol{X}_n^T,P(\Theta_1=\theta_1),\ldots,P(\Theta_n=\theta_4),\Delta t]^T$. The opponent's trajectory $\boldsymbol{X}_i^{OP}$ is regarded as a given parameter. In the original influence diagram model, the decision interval Δt is assumed constant, but here it is assumed to be free to enable solutions that are optimal in their duration as well. The optimal decision interval is implicitly defined by the terminal constraint (41).

Even if the initial estimate does not satisfy Eq. (33), the optimization over time can be carried out. Then, the decision interval Δt is fixed, and the total flight time $n\Delta t$ is equal to $T_{\max}$. The corresponding optimization problem is similar to Eqs. (34–42) except Δt is not a decision variable and the constraints (41) and (42) are omitted.

Once the initial threat assessment distribution is fixed, the threat probabilities follow Eq. (36) and evolve deterministically. Therefore, the dynamic optimization problem (34–42) is in fact deterministic, and an open-loop solution for it can be computed by nonlinear programming techniques.

V. Numerical Example

In the following, an example preference optimal flight path is calculated. The discrete-time dynamic optimization problem is solved by the NPSOL subroutine,[32] which is a versatile implementation of the sequential quadratic programming (SQP).[4] The SQP method has proven to be an efficient and reliable approach for solving discretized dynamic optimization problems (see Refs. 8, 9, and 33).

The maximum thrust force and the drag coefficients in Eq. (11) are approximated on the basis of tabular data that represent a generic modern fighter aircraft. The load factor is limited into the interval $[-3, 9]$. In the example, $h^{\min}$ is set to 500 m, $v^{\min}$ 50 m/s, and $q^{\max}$ 80 kPa, respectively. The terminal limits for the combat state in Eq. (33) are $\alpha_f^1 = 80$ deg, $\beta_f = 80$ deg, and $d_f = 1800$ m. The relatively large limits of the load factor as well as of the terminal condition are selected because in this way long flight times and difficulties in the visualization of trajectories are avoided. The control rates of change used in the myopic solution are $n_\Delta = 1$ 1/s, $\mu_\Delta = 1$ rad/s, and $u_\Delta = 0.5$ 1/s. The decision interval is 1 s.

The initial states of the aircraft are chosen such that the opponent's tail position is reached, that is, the terminal condition (33) becomes satisfied in both the myopic and optimal solutions. The state of the decision maker is initially

$$x_0 = 7000 \text{ m}, \qquad y_0 = 5100 \text{ m}, \qquad h_0 = 7000 \text{ m}$$

$$v_0 = 200 \text{ m/s}, \qquad \gamma_0 = 0 \text{ deg}, \qquad \text{and} \qquad \chi_0 = 0 \text{ deg}$$

and the given initial state of the opponent is

$$x_0^{OP} = 5000 \text{ m}, \qquad y_0^{OP} = 5000 \text{ m}, \qquad h_0^{OP} = 6000 \text{ m}$$

$$v_0^{OP} = 200 \text{ m/s}, \qquad \gamma_0^{OP} = 0 \text{ deg}, \qquad \text{and} \qquad \chi_0^{OP} = 0 \text{ deg}$$

On the predetermined trajectory, the opponent first increases and then decreases the altitude and turns to the right by increasing the velocity continuously. The initial states of the aircraft correspond to a disadvantageous situation where the opponent is flying behind the decision maker. The initial probability distribution of the threat assessment is uniform: $P(\Theta_0 = \theta_k) = 0.25$, $k = 1, \ldots, 4$. Computational experience has shown that the effect of the initial distribution decreases rapidly, and it does not affect strongly the optimal solutions.

The myopic closed-loop solution, the preference optimal open-loop solution, and the fixed trajectory of the opponent are shown in Fig. 8. The projections of the solutions in x, y; x, h; and y, h planes are presented in Fig. 9. The probability distribution of the threat assessment for the optimal flight path is shown in Fig. 10 and for the myopic one in Fig. 11. The myopic solution gives the total flight time of 66.0 s and the cumulative expected utility of 32.6 utility points. The maneuvering decision was made 66 times during the generation of the myopic solution. Thus, the corresponding multistage influence diagram consists of 66 decision stages. It is transformed into an optimization problem with 859 decision variables and 1060 constraints. The final time of the optimal solution is 65.4 s, and the optimal outcome is 37.9 utility points.

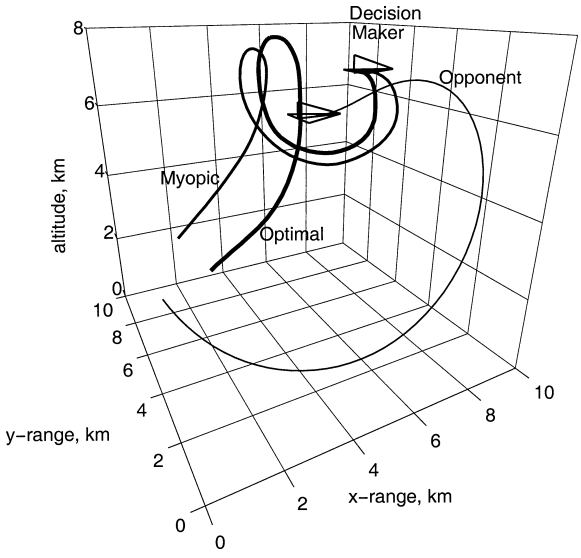

Fig. 8 Optimal and myopic solutions together with the opponent's trajectory.

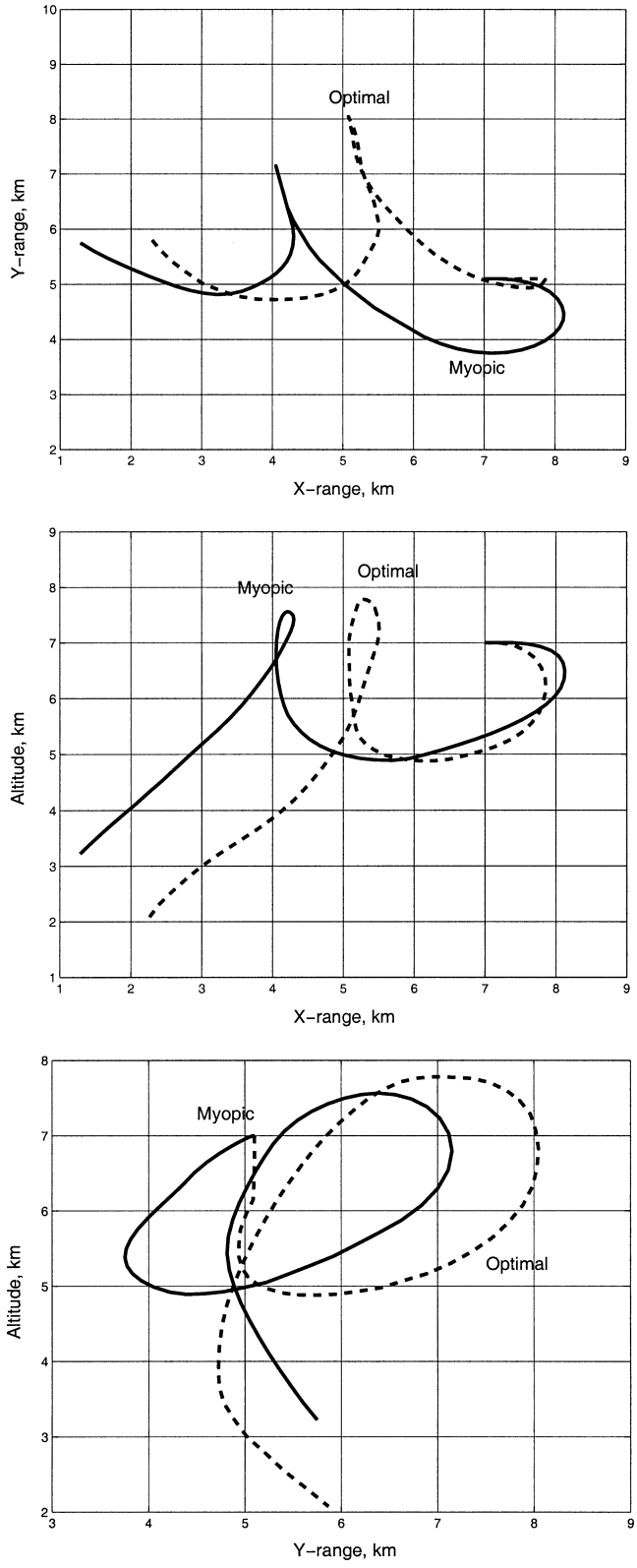

Fig. 9 Projections of the optimal (– – –) and myopic (——) solutions.

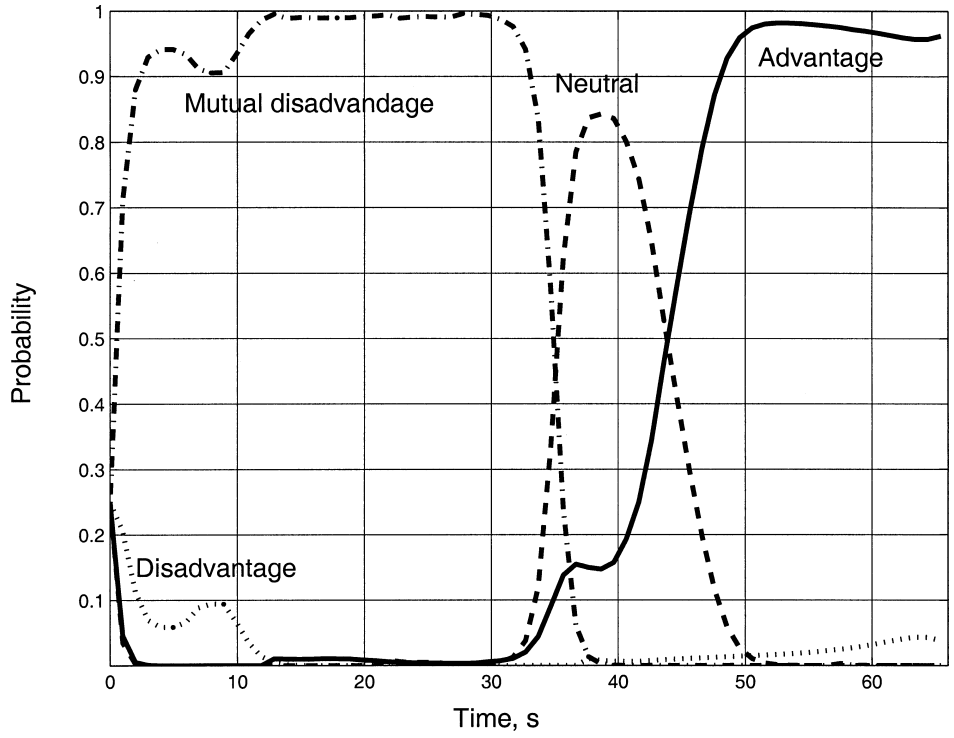

Fig. 10 Probability distribution of the threat situation assessment in the optimal solution.

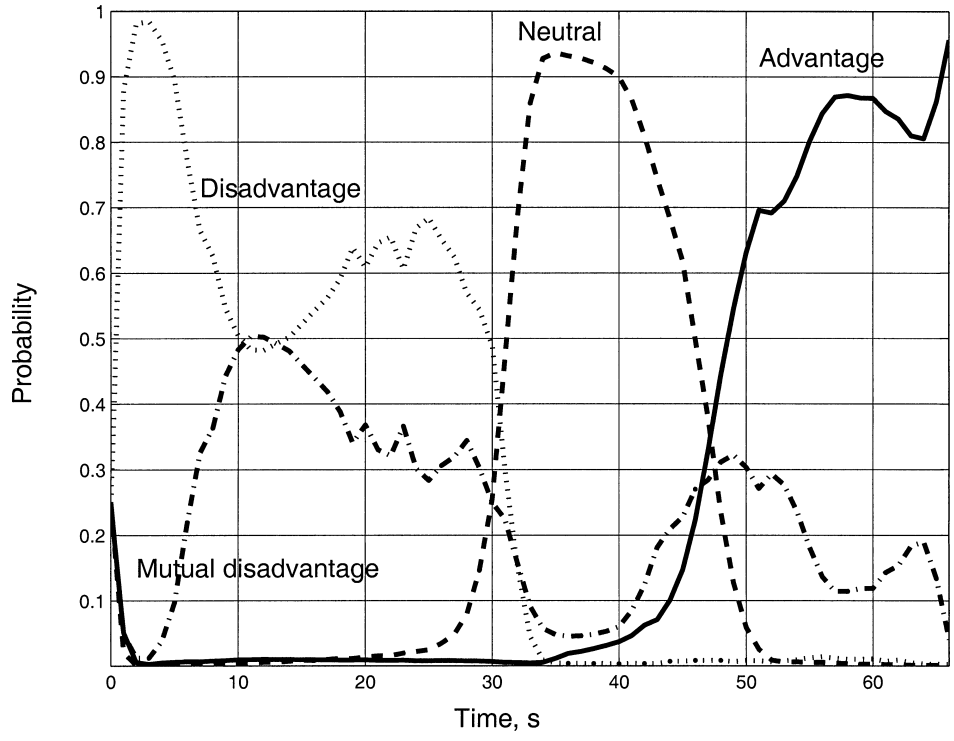

Fig. 11 Probability distribution of the threat situation assessment in the myopic solution.

In the beginning of the flight, the decision maker correctly avoids the front sector of the opponent by decreasing the altitude on the optimal trajectory as well as by decreasing the altitude and turning to the right on the myopic trajectory. The turn rate used on the optimal trajectory is larger than on the myopic flight path. One can conclude that this arises from the additional order in the dynamics given by Eq. (23). In the beginning of the flight, the myopic controls, however, do not approach the limits of the control constraints, which means that the myopic solution could contain also tighter maneuvering. Hence, the change in the utility points is not solely caused by the control rate constraints.

Because of the tight turn on the optimal trajectory, the probability of the disadvantage outcome of the threat assessment does not increase above 0.1, whereas in the myopic solution this probability is almost one at $t = 3$ s (see Figs. 10 and 11). Thus, in the optimal solution the possibility that the opponent reaches the tail position first is reduced. Before 30 s, the probability of the mutual disadvantage outcome is at its highest in the optimal case, and the disadvantage

outcome dominates in the myopic case. After 30 s, the neutral outcome achieves the highest probability in both solutions, but in the optimal case the advantage outcome starts to dominate earlier.

On the preference optimal trajectory, the decision maker reaches the tail position earlier. A reason for this can be that in the optimal solution the decision maker flies a loop at a higher altitude than in the myopic solution. Overall, the optimal open-loop solution has the better value of the objective function, the firing position is achieved faster, and the cumulative time the decision maker is under the threat of the opponent's weapons is shorter than in the myopic closed-loop solution. The utility and likelihood functions sketched by the authors seem to reflect the reality to some extent because the front sector of the opponent is evaded and the advantageous terminal state is acquired.

VI. Discussion

A. Objective Function

The goals of the decision maker are to avoid the opponent's front sector and to reach its tail position. One could suggest an objective function that would measure the accomplishment of these goals at the end of the flight only. Then, the influence diagram could give solutions in which the tail position is reached at the terminal stage, but the decision maker could stay in the disadvantageous area for a long time during the flight. This is not desirable because in reality the decision maker does not necessarily survive through the time intervals where the disadvantage outcome of the threat assessment dominates. Problems of this type are now avoided because the overall objective function (22) evaluates the combat state at all of the decision stages. The decision maker aims at moving away from a disadvantageous position although this movement does not improve the chance to be on the opponent's tail at the end of the flight.

At each stage, the overall objective function (22) is of the form

$$\sum_{k=1}^{4} P(\Theta_i = \theta_k) U_i(\theta_k, \boldsymbol{CS}_i)$$

The single-stage objective function can be interpreted as a scalar cost function of a multiobjective optimization problem (e.g., see Ref. 34) that is constructed by forming a weighted sum of objectives and by requiring that the weights sum up to one. Here, the four objectives $U_i(\theta_k, \boldsymbol{CS}_i)$ to be maximized reflect the preferences in different threat situations of the combat and the probabilities $P(\Theta_i = \theta_k)$ have the role of the weights.

The weighting method is a widely used approach for solving multiobjective optimization[34] and control[35] problems. In this method, first the preceding weighted sum is formed, and a set of efficient solutions is obtained by solving the single objective problem with different weight combinations. Finally, the best solution in the expert's opinion is chosen from among the efficient ones.

In Eq. (22), the sum of the probabilities $P(\Theta_i = \theta_k), k = 1, \ldots, 4$, is one at each stage. Therefore, the optimal controls obtained by maximizing Eq. (22) belong to the set of efficient solutions. Equation (18) updates the probabilities according to the likelihood probabilities that reflect the opinions of a human expert. Thus, the resulting controls at each decision stage can be interpreted as the best efficient solution of the multiobjective problem subject to the given likelihood functions.

Kelley et al.[36] present the threat reciprocity concept for evaluating maneuvering alternatives in a matrix game that models one-on-one air combat. In such a game, the controls of the players are selected such that each player attempts to drive the combat state into his own target set without first being driven into the target set of the adversary. These goals are similar to those of the decision maker in the multistage influence diagram. In the threat reciprocity concept, the relative importance of the goals depends on the combat state. A similar approach is adopted in our model by calculating the weights in Eq. (22), the threat probabilities, based on the state-dependent likelihood functions. Hence, the relative importance of the decision maker's goals varies in a natural way as a function of the combat state.

B. Stability of Solutions

The moving one-stage planning horizon technique is a suitable way to solve a myopic control sequence, but it cannot be applied for generating an initial estimate for all discrete-time trajectory optimization problems. If the time horizon is infinite, the result can be an unstable and diverging initial estimate. Problems can also arise in finite horizon problems where the objective function depends only on the terminal state. Then, controls at early stages affect the value of the overall objective via the dynamics determined by the state equations, but a well-defined objective function cannot be formulated for the single-stage problems. A suitable objective function could be formed by approximating the cost-to-go function that measures the value of the overall objective function from a certain state and time to the terminal state. However, the construction of the approximation is not straightforward. Recently, however, a neurodynamic programming approach[37] has shed some light onto the generation of such approximations.

In the solution procedure of the multistage influence diagram, the moving planning horizon technique produces stable solutions. Because of the form of the objective function (24) and the short planning horizon of the single-stage model, the state of the aircraft cannot deviate excessively from the previous state. Divergence is also restricted by an additional order in the dynamics introduced by Eq. (23).

Existing nonlinear programming methods usually have a large convergence domain. Still, the initial estimate for decision variables of an optimization problem cannot be chosen completely arbitrarily. In the solution procedure, the myopic solution is used as the initial value of the decision maker's controls, states, and probabilities, which ensures reliable convergence when solving the discrete-time optimization problem (34–42).

C. Utilization of the Approach

Although a passive adversary aircraft flying a predetermined trajectory is only considered and the computation is not real time, the presented modeling and analysis approach gives information that is valuable in the planning and evaluation of air combat maneuvers as well as in the development of air combat simulators. For example, a defensive basic fighter maneuver could be improved by using it as the initial estimate and solving the preference optimal flight path against a fixed pursuing trajectory. In addition, it is possible to analyze how a change in the adversary's trajectory affects the optimal maneuvering sequence.

The impacts of preference assessments on optimal flight paths can be studied by solving the influence diagram repeatedly with different utility functions. By comparing the resulting trajectories, one can identify the functions that lead to the desired trajectory or basic fighter maneuver. These functions can be embedded in a single stage influence diagram that is a suitable real-time guidance model for air combat simulators. This analysis could be automated by adding a tool for modeling the preferences, like HIPRE,[38] to an aircraft trajectory optimization software, like VIATO.[9]

D. Model Improvements

The numerical example demonstrates that the influence diagram model produces reasonable results for the pilot's sequential maneuvering process. However, the accuracy and reality of the model can still be enhanced by describing the preferences of a pilot and the dynamics of an aircraft in more detail. The multistage influence diagram approach is flexible with respect to such generalizations.

To increase the quality of resulting preference optimal flight paths, the focus should be to improve the behavioral components. That is the preferences and the behavior of pilots should be captured into the multistage influence diagram by constructing the utility and likelihood functions in cooperation with pilots. The motion of the aircraft can be described more accurately by replacing the three-degrees-of-freedom point-mass model with, for example, a six-degrees-of-freedom model that also contains the equations of rotation. This extension would be rather straightforward, but it would also increase the number of variables in the influence diagram. The myopic moving horizon technique still remains a reliable and fast way to produce

an initial estimate for preference optimal flight paths. The more detailed model of the dynamics would complicate the solution of the dynamic optimization problem because of its size increases. For reasons of accuracy, the equations of motion should be discretized by using a higher-order scheme instead of Euler discretization. Naturally an improved discretization scheme could also be applied in the presented model to improve the accuracy of the solutions. The Jacobian of the constraints and the Hessian of the Lagrangian are almost block diagonal in dicsretized dynamic optimization problems. This allows the use of special algorithms for sparse matrices (e.g., see Refs. 33 and 39), which could decrease the computation times needed in the optimization.

E. Generalizing the Approach into Games

The presented modeling and analysis approach can be extended to a one-on-one air combat game in which both players aim simultaneously at capturing the adversary and avoiding being captured. Encounters of this type have been modeled using the concept of a two-target game (e.g., see Ref. 40). The extension of the multistage influence diagram could offer a way to incorporate uncertainty and preference models into a two-target game framework.

A game model can be formulated such that both the players are modeled and controlled with an influence diagram. One way to do this is to combine the multistage influence diagram modeling the maneuvering problem from the decision maker's viewpoint and the diagram representing the opponent's maneuvering process. The information structure of such an influence diagram game can be asymmetric or symmetric. In the former case, the decision maker knows or assumes the objective and preferences of the opponent. The latter information structure refers to a combat model where both the players receive the same information on the game situation.

Open-loop solutions of the game models could be obtained with the help of an extension of the introduced solution procedure. The symmetric information leads to the solution of two discrete-time dynamic optimization problems simultaneously. In the asymmetric case, a bilevel discrete-time dynamic optimization problem results (see Ref. 41).

A single-stage influence diagram game representing the maneuvering decisions in one-on-one air combat can also be resolved in a closed-loop form. This could offer a new way to produce so-called reprisal strategies[36] that utilize the nonoptimal behavior of an adversary.

VII. Conclusion

In an air combat, a pilot faces complicated multiobjective maneuvering decision problems. He or she has to evaluate maneuvering alternatives whose outcomes are known only under conditions of uncertainty. Multiple conflicting objectives as well as the uncertain nature of the decision-making environment should be taken into account in the planning of combat maneuvers. These features are included in the presented multistage influence diagram. It offers a new way to incorporate the structural model of the uncertainty and the pilot's preferences into trajectory optimization. Preference optimal open-loop maneuvering sequences against given trajectories of the opponent are obtained by converting the influence diagram into a discrete-time dynamic optimization problem and solving it with nonlinear programming. This new modeling and analysis approach also provides a good basis for considering one-on-one air combat models in which both players are active.

Appendix: Equivalence of Eqs. (5) and (6)

This appendix shows the equivalence of Eqs. (5) and (6). First, the cumulative expected utility (CEU) is written according to Eq. (6):

$$
\begin{aligned}
CEU &= P\left(x_1^1\right) EU_1^2(d^1, d^2) + P\left(x_2^1\right) EU_2^2(d^1, d^2) \\
&= P\left(x_1^1\right)\left\{P\left(x_1^2 \middle| x_1^1\right)\left[u^1\left(d^1, x_1^1\right) + u^2\left(d^2, x_1^2\right)\right]\right. \\
&\quad \left. + P\left(x_2^2 \middle| x_1^1\right)\left[u^1\left(d^1, x_1^1\right) + u^2\left(d^2, x_2^2\right)\right]\right\} \\
&\quad + P\left(x_2^1\right)\left\{P\left(x_1^2 \middle| x_2^1\right)\left[u^1\left(d^1, x_2^1\right) + u^2\left(d^2, x_1^2\right)\right]\right. \\
&\quad \left. + P\left(x_2^2 \middle| x_2^1\right)\left[u^1\left(d^1, x_2^1\right) + u^2\left(d^2, x_2^2\right)\right]\right\} \\
&= P\left(x_1^1\right)\left\{u^1\left(d^1, x_1^1\right)\left[P\left(x_1^2 \middle| x_1^1\right) + P\left(x_2^2 \middle| x_1^1\right)\right]\right. \\
&\quad \left. + P\left(x_1^2 \middle| x_1^1\right) u^2\left(d^2, x_1^2\right) + P\left(x_2^2 \middle| x_1^1\right) u^2\left(d^2, x_2^2\right)\right\} \\
&\quad + P\left(x_2^1\right)\left\{u^1\left(d^1, x_2^1\right)\left[P\left(x_1^2 \middle| x_2^1\right) + P\left(x_2^2 \middle| x_2^1\right)\right]\right. \\
&\quad \left. + P\left(x_1^2 \middle| x_2^1\right) u^2\left(d^2, x_1^2\right) + P\left(x_2^2 \middle| x_2^1\right) u^2\left(d^2, x_2^2\right)\right\} \qquad \text{(A1)}
\end{aligned}
$$

Here $P(x_1^2|x_i^1) + P(x_2^2|x_i^1) = 1$, $i = 1, 2$, and thus Eq. (A1) yields

$$
\begin{aligned}
CEU &= P\left(x_1^1\right)\left[u^1\left(d^1, x_1^1\right) + P\left(x_1^2 \middle| x_1^1\right) u^2\left(d^2, x_1^2\right)\right. \\
&\quad \left. + P\left(x_2^2 \middle| x_1^1\right) u^2\left(d^2, x_2^2\right)\right] + P\left(x_2^1\right)\left[u^1\left(d^1, x_2^1\right)\right. \\
&\quad \left. + P\left(x_1^2 \middle| x_2^1\right) u^2\left(d^2, x_1^2\right) + P\left(x_2^2 \middle| x_2^1\right) u^2\left(d^2, x_2^2\right)\right] \\
&= P\left(x_1^1\right) u^1\left(d^1, x_1^1\right) + P\left(x_2^1\right) u^1\left(d^1, x_2^1\right) \\
&\quad + \left[P\left(x_1^2 \middle| x_1^1\right) P\left(x_1^1\right) + P\left(x_1^2 \middle| x_2^1\right) P\left(x_2^1\right)\right] u^2\left(d^2, x_1^2\right) \\
&\quad + \left[P\left(x_2^2 \middle| x_1^1\right) P\left(x_1^1\right) + P\left(x_2^2 \middle| x_2^1\right) P\left(x_2^1\right)\right] u^2\left(d^2, x_2^2\right) \qquad \text{(A2)}
\end{aligned}
$$

Using the probabilities of the decision tree shown in Fig. 3, the probabilities $P(x_1^2)$ and $P(x_2^2)$ can be expressed as

$$
P\left(x_i^2\right) = P\left(x_i^2 \middle| x_1^1\right) P\left(x_1^1\right) + P\left(x_i^2 \middle| x_2^1\right) P\left(x_2^1\right) \qquad i = 1, 2 \qquad \text{(A3)}
$$

Combining Eqs. (A2) and (A3) gives

$$
\begin{aligned}
CEU &= P\left(x_1^1\right) u^1\left(d^1, x_1^1\right) + P\left(x_2^1\right) u^1\left(d^1, x_2^1\right) + P\left(x_1^2\right) u^2\left(d^2, x_1^2\right) \\
&\quad + P\left(x_2^2\right) u^2\left(d^2, x_2^2\right) = \sum_{i=1}^{2} \sum_{k=1}^{2} P\left(x_k^i\right) u^i\left(d^i, x_k^i\right) \qquad \text{(A4)}
\end{aligned}
$$

that is the CEU according to Eq. (5).

Acknowledgment

This work has been carried out in collaboration with the Finnish Air Force.

References

[1]Clemen, R. T., *Making Hard Decisions, An Introduction to Decision Analysis*, 2nd ed., Duxbury Press, Belmont, CA, 1996, pp. 1–10, 50–60, 67–74, 105.

[2]Keeney, R. L., and Raiffa, H., *Decision with Multiple Objectives: Preferences and Value Tradeoffs*, Wiley, New York, 1976, pp. 1–30, 289, 131.

[3]von Winterfeldt, D., and Edwards, W., *Decision Analysis and Behavioral Research*, Cambridge Univ. Press, Cambridge, England, UK, 1986, pp. 63–89, 322–325.

[4]Bertsekas, D. P., *Nonlinear Programming*, Athena Scientific, Belmont, MA, 1995, pp. 361–382.

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

[6]Bryson, A. E., *Dynamic Optimization*, Addison-Wesley, Menlo Park, CA, 1999, Chap. 2, pp. 172–179.

[7]Basar, T., and Olsder, G., *Dynamic Noncooperative Game Theory*, 2nd ed., Academic Press, London, 1995, pp. 38–63.

[8]Raivio, T., Ehtamo, H., and Hämäläinen, R. P., "Aircraft Trajectory Optimization using Nonlinear Programming," *System Modeling and Optimization*, edited by J. Dolezal and J. Fidler, Chapman and Hall, London, 1996, pp. 435–441.

[9]Virtanen, K., Ehtamo, H., Raivio, T., and Hämäläinen, R. P., "VIATO—Visual Interactive Aircraft Trajectory Optimization," *IEEE Transaction on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, Vol. 29, No. 3, 1999, pp. 409–421.

[10]Betts, J., "Survey of Numerical Methods for Trajectory Optimization," *Journal of Guidance, Control, and Dynamics*, Vol. 21, No. 2, 1998, pp. 193–207.

[11]McManus, J. W., and Goodrich, K. H., "Application of Artificial Intelligence (AI) Programming Techniques to Tactical Guidance for Fighter Aircraft," AIAA Paper 89-3525, Aug. 1989.

[12]Stehlin, P., Hallkvist, I., and Dahlstrand, H., "Models for Air Combat Simulation," *Proceedings of the 19th Congress of the International Council of the Aeronautical Sciences (ICAS)*, Vol. 2, AIAA, Washington, DC, 1994, pp. 2190–2196.

[13]Bent, N. E., "The Helicopter Air-To-Air Value Driven Engagement Model (HAVDEM)," *Proceedings of the 19th Congress of the International Council of the Aeronautical Sciences (ICAS)*, Vol. 2, AIAA, Washington, DC, 1994, pp. 2181–2189.

[14]Lazarus, E., "The Application of Value-Driven Decision-Making in Air Combat Simulation," *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Vol. 3, Inst. of Electrical and Electronics Engineers, Piscataway, NJ, 1997, pp. 2302–2307.

[15]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.

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

[17]Lindley, D., "Foundations," *Subjective Probability*, edited by G. Wright and P. Ayton, Wiley, Chichester, England, UK, 1994, pp. 3–15.

[18]von Neumann, J., and Morgenstern, O., *Theory of Games and Economic Behavior*, 3rd ed., Princeton Univ. Press, Princeton, NJ, 1990, pp. 15–31, 77–79.

[19]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.

[20]Tatman, J. A., and Shachter, R. D., "Dynamic Programming and Influence Diagrams," *IEEE Transaction on Systems, Man, and Cybernetics*, Vol. 20, No. 2, 1990, pp. 365–379.

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

[22]Kirkwood, C. W., "Implementing an Algorithm to Solve Large Sequential Decision Analysis Models," *IEEE Transaction on Systems, Man, and Cybernetics*, Vol. 24, No. 10, 1994, pp. 1425–1432.

[23]Precision Tree, User's Guide, Palisade Corp., New York, 1996.

[24]Covaliu, Z., and Oliver, R. M., "Representation and Solution of Decision Problems Using Sequential Decision Diagrams," *Management Science*, Vol. 41, No. 12, 1995, pp. 1860–1881.

[25]Bertsekas, D. P., *Dynamic Programming and Optimal Control*, Vol. 1, Athena Scientific, Belmont, MA, 1995, pp. 4, 16–20, 246.

[26]Kirkwood, C. W., "Recursive Calculation of Probability Distributions for Sequential Decision Analysis Models," *IEEE Transaction on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, Vol. 28, No. 1, 1998, pp. 104–111.

[27]Warburton, A., "The Decision Tree Polytope and its Application to Sequential Decision Problems," *Journal of Multi-Criteria Decision Analysis*, Vol. 7, No. 6, 1998, pp. 330–339.

[28]@RISK, User's Guide, Palisade Corp., New York, 1996, pp. 256, 257.

[29]Stonebraker, J. S., and Kirkwood, C. W., "Formulating and Solving Sequential Decision Analysis Models with Continuous Variables," *IEEE Transactions on Engineering Management*, Vol. 44, No. 1, 1997, pp. 43–53.

[30]Bard, J. F., *Practical Bilevel Optimization, Algorithms and Applications*, Kluwer, Dordrecht, The Netherlands, 1998, pp. 5–8.

[31]Virtanen, K., Hämäläinen, R. P., and Mattila, V., "Team Optimal Signaling Strategies in Air Combat," *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Vol. 3, Inst. of Electrical and Electronics Engineers, Piscataway, NJ, 2001, pp. 2075–2080.

[32]Gill, P. E., Murray, W., and Wright, M. H., User's Guide for NPSOL, Ver. 4.0, Dept. Oper. Res., Stanford Univ., Rept. SOL 86-2, Stanford, CA, Jan. 1986.

[33]Gill, P. E., Murray, W., and Saunders, M. A., "Large-Scale SQP Methods and Their Applications in Trajectory Optimization," *Computational Optimal Control*, edited by R. Bulirsch and D. Kraft, Birkhauser, Basel, Switzerland, 1994, pp. 29–42.

[34]Miettinen, K., *Nonlinear Multiobjective Optimization*, Kluwer, Boston, MA, 1999, pp. 5, 78–85.

[35]Leitmann, G., *The Calculus of Variations and Optimal Control*, Plenum Press, New York, 1981, pp. 292–296.

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

[37]Bertsekas, D. P., and Tsitsiklis, J., *Neuro-Dynamic Programming*, Athena Scientific, Belmont, MA, 1996, Chap. 1.

[38]Hämäläinen, R. P., and Lauri, H., HIPRE 3+ User's Guide, Systems Analysis Lab., Helsinki Univ. of Technology, Finland, 1992, URL: http://www.hipre.hut.fi [cited 10 March 2003].

[39]Conn, A. R., Gould, N. I. M., and Toint, L., *LANCELOT: A Fortran Package for Large-Scale Nonlinear Optimization (Release A)*, Springer Series in Computational Mathematics, Vol. 17, Springer-Verlag, New York, 1992.

[40]Grimm, W., and Well, K. H., "Modelling Air Combat as Differential Game, Recent Approaches and Future Requirements," *Differential Games—Developments in Modelling and Computation, Lecture Notes in Control and Information Sciences*, Vol. 156, edited by R. P. Hämäläinen and H. Ehtamo, Springer-Verlag, New York, 1991, pp. 1–13.

[41]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.