INFORMATICA, 1999, Vol. 10, No. 4, 403–426

Agent Belief: Presentation, Propagation, and Optimization

Veska NONCHEVA
Faculty of Mathematics and Informatics, University of Plovdiv
24 Tsar Assen St., 4000 Plovdiv, Bulgaria
e-mail: nonchev@plovdiv.techno-link.com, wesnon@ulcc.uni-plovdiv.bg

Received: November 1999

Abstract. The aim of the article is to show a stochastic approach for both modelling and optimizing the statistical agent belief in a probability model.

Two networks are defined: a decision network $\mathfrak{D}$ of the agent belief state and a utility network $\mathfrak{U}$, presenting the utility structure of the agent belief problem.

The agent belief is presented via the following three items $(\mathfrak{B}, \mathfrak{D}, \mathfrak{U})$, where $\mathfrak{B}$ is a Bayesian network, presenting the probability structure of the agent belief problem.

Two propagation algorithms in $\mathfrak{D}$ and in $\mathfrak{U}$ are also presented.

Key words: agent belief, decision network, utility network.

1. Introduction

The present stage of statistical software development is creating a set of agents and their coming to life in multiagent systems for statistical investigations. An approach for building a multiagent system for time series forecasting is presented in (Prat *et al.*, 1994). Some desirable characteristics of multiagent systems for statistical investigations are discussed in (Noncheva and Stojanov, 1995).

The aim of our research work is to build an agent for statistical hypotheses testing. This agent embodies statistical expertise and possesses intelligent behaviour. It is conceptualized by using concepts which are usually applied to humans. Its mental state consists of components such as knowledge, belief, intention, and obligation.

The primary subject of this study is the agent belief in a probability model.

The necessity for defining the agent belief about the type of the probability distribution arises from the fact that the population distribution, which is background of many models of mathematical statistics, is unknown. As a result of this insufficient knowledge comes the indefiniteness of the appropriate method choice for statistical analysis, i.e., resulting in indefiniteness at the choice of the best behaviour of the agent, which agent is making the statistical analysis.

The agent belief is presented by a certainty threshold function. We say that the agent believes in the probability model if the agent certainty is not smaller than a threshold number that is known in advance.

However, the agent certainty is an unobservable random variable.

The agent belief about the probability model (i.e., about the probability distribution) must be based on the mechanism knowledge of the phenomenon under investigation. But if the phenomenon under investigation is unknown, the agent can make its own choice about the probability distribution after it has tested a Goodness-of-fit hypothesis. It is also possible that the agent may ask the user for his opinion.

Consequently, the agent will make use of a pre-test and a post-test when making a decision about the probability model. Usually, as a pre-test a statistical test about the form of the probability distribution is used, and after that the user is asked about his opinion. The results of those two tests are respectively $1-p$, where p is the p-value of test statistic and the user's degree of certainty, represented as numbers in the interval $[0,1]$.

It is natural to consider decision rules, which have a monotonous form. For example, the agent believes in normal population distribution, if the value of the variable, representing its certainty, is not smaller than the fixed in advance threshold value; otherwise the agent rejects the assumption for normality.

It is natural to expect also that the high values of the pre-test results will lead to lower requirements of the post-test results. That is, the received preliminary information influences the decision rules about the agent belief. Decision rules in which post-test decisions are functions of the concrete result from the pre-test, will be called weak rules.

Consequently, the problem for decision making about the agent belief is reduced to find threshold values for observable results, which are optimal with respect to Bayes' approach.

2. Modeling of the Agent Belief

Let $\boldsymbol{T}$ be a continuous random variable which has probability distribution in the set $\boldsymbol{\Omega}_t = [0,1]$. Let $\boldsymbol{t} \in \boldsymbol{\Omega}_t$ be a possible value of $\boldsymbol{T}$.

The random variable $\boldsymbol{T}$ is called agent *certainty*.

Let $\boldsymbol{\Omega}_t$ be divisible by $\{\boldsymbol{A}_k = [\boldsymbol{t}_k, \boldsymbol{t}_{k+1}],\ \boldsymbol{t}_k \in \boldsymbol{\Omega}_t,\ \boldsymbol{t}_{k+1} \in \boldsymbol{\Omega}_t,\ k = 0,1,\dots,n-1\}$, i.e., $\boldsymbol{A}_i \cap \boldsymbol{A}_j = \varnothing$ when $i \neq j$ and $\sum_{k=0}^{n-1} \boldsymbol{A}_k = \boldsymbol{\Omega}_t$.

Let δ be a categorical variable which can take on values from the set

$$\boldsymbol{D} = \{\boldsymbol{a}_0, \boldsymbol{a}_1, \dots, \boldsymbol{a}_{n-1}\}.$$

We may define the random variable Bel as a function of the random variable $\boldsymbol{T}$ in the following manner:

$$Bel(\boldsymbol{T}) = \begin{cases} \boldsymbol{a}_0, & \text{if } \boldsymbol{t}_0 \leqslant \boldsymbol{T} < \boldsymbol{t}_1, \\ \boldsymbol{a}_1, & \text{if } \boldsymbol{t}_1 \leqslant \boldsymbol{T} < \boldsymbol{t}_2, \\ \dots & \\ \boldsymbol{a}_{n-1}, & \text{if } \boldsymbol{t}_{n-1} \leqslant \boldsymbol{T} < \boldsymbol{t}_n. \end{cases}$$

The random variable $Bel(\boldsymbol{T})$ is called the ***agent belief***. The values $\boldsymbol{a}_0, \boldsymbol{a}_1, \dots, \boldsymbol{a}_{n-1}$, which the random variable Bel can take on, are called ***agent belief states***. The set $\boldsymbol{D}$ is called a *set of the possible states of the agent belief.*

Our objective is to determine the value of the function $\delta = Bel(\boldsymbol{T})$.

However, the agent certainty is an unobservable random variable.

We shall emphasize once again that the choice of the appropriate probability model must be based above all on the understanding of the mechanism of the phenomenon under examination. However, when the phenomenon is unknown, the agent – after having tested statistical hypotheses – can make its own choice about the probability distribution. When the agent makes a decision about the form of the probability distribution, it can ask the user and make use of his expertise and intuition, too.

In accord with the results of the observations $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ we are to make a decision $\boldsymbol{\delta} = \boldsymbol{\delta}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) \in \boldsymbol{D}$. The function $\boldsymbol{\delta}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$, defined in the set of possible results from the observations and accepting values in set $\boldsymbol{D}$ of the possible states of agent belief is called a *decision-making rule for agent belief state* or – in short – a *decision rule*.

We need to define a *utility structure* for the agent belief problem. A *utility function* is called a function $\boldsymbol{u}_j(\boldsymbol{t})$ which describes the utility of the result from the appropriation of the state $\boldsymbol{a}_j$, $j = 0, 1, \ldots, n-1$ of the agent belief, when agent certainty is $\boldsymbol{t}$. A linear utility structure is used in (Noncheva, 1998). Gausian probability distribution functions are appropriate for determining optimal decision rules about agent belief, too.

We will consider four cases that are frequently met in the statistical analysis. For all of them we will define a decision rule for the agent belief in the probability model.

Case 1. *Normality at an assumption for symmetry*

Let us consider the agent belief about the normality of the population distribution.

Usually, as a pre-test is used Goodness-of-fit test and after that the user is asked about his opinion. The results of those two tests are respectively $1 - p$, where p is the p-value of test statistic and the user degree of certainty, represented as numbers in the interval $[0, 1]$.

Sometimes, e.g., in regression analysis, the user is first asked if the population distribution is normal, then the model is build and as a final step the residuals from the estimated regression model are tested for normality.

Let us mark the pre-test results with $\boldsymbol{X}_1$ and the post-test results with $\boldsymbol{X}_2$. Let us further assume that $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are continuous random variables with probability distribution in $\boldsymbol{\Omega}_1 = [0, 1]$ and $\boldsymbol{\Omega}_2 = [0, 1]$, respectively.

Let $\boldsymbol{T}$ be the agent certainty, i.e., a continuous random variable with a probability distribution in the set $\boldsymbol{\Omega}_t = [0, 1]$; and it cannot be observed.

Let us assume that the relation between $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ and $\boldsymbol{T}$ can be represented by their joint density $\boldsymbol{f}(x_1, x_2, t)$.

Then the weak monotonous rules for making a decision are defined as follows:

$$\boldsymbol{\delta}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \begin{cases} \boldsymbol{a}_0, & \text{if } \boldsymbol{X}_1 < x_1^c, \\ \boldsymbol{a}_1, & \text{if } \boldsymbol{X}_1 \geqslant x_1^c, \quad \boldsymbol{X}_2 < x_2^c(x_1), \\ \boldsymbol{a}_2, & \text{if } \boldsymbol{X}_1 \geqslant x_1^c, \quad \boldsymbol{X}_2 \geqslant x_2^c(x_1), \end{cases} \tag{1}$$

where $\boldsymbol{a}_j$, $j = 0, 1, 2$ are agent belief states, x_1^c and x_2^c are the threshold values for $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$.

Case 2. *Symmetry*

In order to form its belief about the symmetry of the probability distribution of the population being investigated, the agent makes use of a statistical test for symmetry.

Let $\boldsymbol{X} = 1 - \boldsymbol{p}$, where $\boldsymbol{p}$ is the $\boldsymbol{p}$-value of the statistics of the test for symmetry and $\boldsymbol{T}$ be the agent certainty. Obviously $\boldsymbol{X}$ can be observed while $\boldsymbol{T}$ cannot be observed. We assume that $\boldsymbol{X}$ and $\boldsymbol{T}$ are continuous random variables and the relation between them can be represented by their joint probability distribution function $\boldsymbol{f}(x,t)$.

The agent belief in symmetry depends on the threshold value $\boldsymbol{t}_c$, which is fixed in advance and known to the agent. For example $\boldsymbol{t}_c = 0.95$. If the certainty value $\boldsymbol{T}$ exceeds the threshold value $\boldsymbol{t}_c$, then the agent is convinced in the symmetry of the population distribution under study. Otherwise the agent is convinced in the asymmetry of the distribution.

However, the random variable $\boldsymbol{T}$ is unobservable. Then the monotonous rules for making a decision about the agent belief, based on the observable random variable $\boldsymbol{X}$, will be defined as follows:

$$\delta(\boldsymbol{x}) = \begin{cases} \boldsymbol{a}_0, & \text{if } \boldsymbol{X} < x^c, \\ \boldsymbol{a}_1, & \text{if } \boldsymbol{X} \geqslant x^c, \end{cases}$$

where the states $\boldsymbol{a}_0$ and $\boldsymbol{a}_1$ represent the agent belief in the symmetry and the asymmetry of the distribution, and x^c is the threshold value for $\boldsymbol{X}$.

Case 3. *A combined model*

We are going to address the problem of formation of the agent belief in normality of the continuous probability distribution of the population under investigation. For this purpose a statistical test for symmetry, a statistical test for normality (pre-test for normality) and making an inquiry about the user opinion (post-test for normality) are used. The results from these three tests are respectively $1 - \boldsymbol{p}_1$, $1 - \boldsymbol{p}_2$, where $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are the $\boldsymbol{p}$-values of the statistics of the two tests, and the degree of the user's certainty, represented as a number in the interval $[0,1]$.

Let us designate with $\boldsymbol{X}_1$ the observed result from the symmetry test, with $\boldsymbol{X}_2$ – the observed result from the pre-test for normality, with $\boldsymbol{X}_3$ – the observed result from the post-test for normality, and with $\boldsymbol{T}$ – the agent certainty of normality, which cannot be observed.

Let us assume that $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$ and $\boldsymbol{T}$ are continuous random variables with a joint probability density function $f(x_1, x_2, x_3, t)$.

The rule for making a decision $\boldsymbol{\delta}(x_1, x_2, x_3)$ determines the state of the agent belief for each possible realization (x_1, x_2, x_3) of the random vector $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3)$.

The monotonous weak rule for making a decision about agent belief in this case has the form:

$$\boldsymbol{\delta}(x_1, x_2, x_3) = \begin{cases} \boldsymbol{a}_0, & \text{if } X_1 < x_1^c,\ X_2 \in [0,1],\ X_3 \in [0,1], \\ \boldsymbol{a}_1, & \text{if } X_1 \geqslant x_1^c,\ X_2 < x_2^c(x_1),\ X_3 \in [0,1], \\ \boldsymbol{a}_2, & \text{if } X_1 \geqslant x_1^c,\ X_2 \geqslant x_2^c(x_1),\ X_3 < x_3^c(x_1, x_2), \\ \boldsymbol{a}_3, & \text{if } X_1 \geqslant x_1^c,\ X_2 \geqslant x_2^c(x_1),\ X_3 \geqslant x_3^c(x_1, x_2), \end{cases}$$

where

- $\boldsymbol{a}_j$, $j = 0, 1, 2, 3$ are the following states of the agent belief:

  $\boldsymbol{a}_0$ – agent rejects the assumption for the symmetry of the distribution, describing the population under investigation;

  $\boldsymbol{a}_1$ – agent rejects the assumption for normality of the distribution;

  $\boldsymbol{a}_2$ – agent supposes (suspects) that the distribution of the population being investigated is normal;

  $\boldsymbol{a}_3$ – agent convinced that the distribution describing the population is normal.

- $\boldsymbol{x}_1^c$, $\boldsymbol{x}_2^c$, $\boldsymbol{x}_3^c$ are the threshold points for $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ and $\boldsymbol{X}_3$.

Case 4. *A general model*

In order to form its belief about the type of population distribution, the agent can start with a statistical test for symmetry. In case when the hypothesis for symmetry is rejected, then a hypothesis for suspicious observations can be tested. If the latter is rejected, data transformation can be made.

If the hypothesis for suspicious observations cannot be rejected, the agent must ask the user, if there are grounds of non-statistical nature for the sharply deviating observations to be removed. The case in which the hypothesis for symmetry cannot be rejected, the agent can continue with an examination for normality.

Usually a Goodness-of-fit test is used as a pre-test and after that the user opinion is asked.

We shall represent the agent belief as a threshold function, which can take on one of the five possible states $\boldsymbol{a}_j$, $j = 0, 1, 2, 3, 4$, where:

$\boldsymbol{a}_0$ – agent is certain that the distribution is asymmetric and there are no suspicious observations. It will offer a suitable data transformation;

$\boldsymbol{a}_1$ – agent is certain that the distribution is asymmetric and that the asymmetry is due to suspicious observations. It will ask the user whether to remove them or not;

$\boldsymbol{a}_2$ – agent is certain that the distribution is asymmetric, but not normal. It will make use of the assumption for symmetry in the statistical analysis;

$\boldsymbol{a}_3$ – agent is certain that the distribution is symmetric and it assumes that it is normal, too. It will use only robust statistical tests;

$\boldsymbol{a}_4$ – agent is certain that the distribution is normal. It will use statistical tests, which are sensitive to the assumption of normality, too.

Let us designate with $\boldsymbol{X}_1$ the observed value $1 - \boldsymbol{p}_1$, where $\boldsymbol{p}_1$ is $\boldsymbol{p}$-value of the statistics of the test for symmetry. Let us designate with $\boldsymbol{X}_2$ the observed value $1 - \boldsymbol{p}_2$, where $\boldsymbol{p}_2$ is $\boldsymbol{p}$-value of the statistics of the test for suspicious observations. Let $\boldsymbol{X}_3$ be the observed result from the pre-test for normality. Let $\boldsymbol{X}_4$ be the observed result from the post-test for normality. Let $\boldsymbol{T}$ be the agent certainty of normality of the distribution, the value which cannot be observed.

Let us assume that $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$, $\boldsymbol{X}_4$ and $\boldsymbol{T}$ are continuous random variables.

Let us suppose that the connection between the measured results from the tests $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$, $\boldsymbol{X}_4$ and the variable $\boldsymbol{T}$ can be represented by the joint probability density $\boldsymbol{f}(x_1, x_2, x_3, x_4, t)$.

The decision-making rule $\boldsymbol{\delta}(x_1, x_2, x_3, x_4)$ determines for each possible realization (x_1, x_2, x_3, x_4) of the random vector $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3, \boldsymbol{X}_4)$ which state $\boldsymbol{a}_j, j = 0, 1, \dots, 4$, will be appropriated to the agent belief about the distribution of the population under investigation.

The weak monotonous rules for making a decision about the agent belief $\boldsymbol{\delta}(x_1, x_2, x_3, x_4)$ will be defined as follows:

$$\boldsymbol{\delta}(x_1, x_2, x_3, x_4) = \begin{cases} \boldsymbol{a}_0, & \text{if } \boldsymbol{X}_1 < x_1^c, \ \ \boldsymbol{X}_2 < x_2^c(x_1), \\ \boldsymbol{a}_1, & \text{if } \boldsymbol{X}_1 < x_1^c, \ \ \boldsymbol{X}_2 \geqslant x_2^c(x_1), \\ \boldsymbol{a}_2, & \text{if } \boldsymbol{X}_1 \geqslant x_1^c, \ \ \boldsymbol{X}_3 < x_3^c(x_1), \\ \boldsymbol{a}_3, & \text{if } \boldsymbol{X}_1 \geqslant x_1^c, \ \ \boldsymbol{X}_3 \geqslant x_3^c(x_1), \ \ \boldsymbol{X}_4 < x_4^c(x_1, x_3), \\ \boldsymbol{a}_4, & \text{if } \boldsymbol{X}_1 \geqslant x_1^c, \ \ \boldsymbol{X}_3 \geqslant x_3^c(x_1), \ \ \boldsymbol{X}_4 \geqslant x_4^c(x_1, x_3), \end{cases}$$

where $\boldsymbol{a}_j, j = 0, 1, 2, 3, 4$ are the states of the agent belief, x_1^c, x_2^c, x_3^c, x_4^c are the threshold values correspondingly for $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$ and $\boldsymbol{X}_4$.

3. A Decision Network

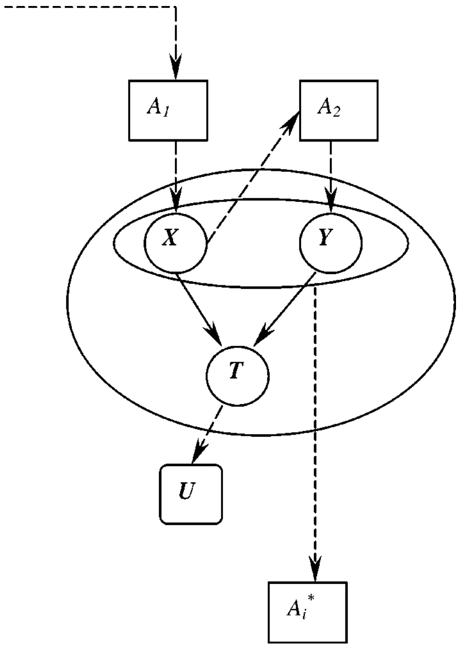
Decision networks are graphic structures, that represent probability relations and information flows (Shachter, 1986; Shachter, 1988).

A decision network consists of the following items:

- a set of basic operations;
- a probability structure;
- a utility structure;
- a decision rule;
- an optimal action, through which the expected utility is maximized.

DEFINITION 1. *A decision network* comprises of the following set of items: $((\boldsymbol{A}, I)$, (X, P), $\boldsymbol{A}^*, u, \delta)$. It is within this particular set where they are as shown below:

- (A, I) is recognized as a directed graph of basic operations, where $\boldsymbol{A} = \{A_i,\ i = 1, \dots, k\}$ is the set of these basic operations, and $I = \{I_{l,m} = (A_l, A_m), l \neq m,\ l = 1, \dots, k;\ m = 1, \dots, k\}$ is the set of directed information arcs;
- (X, P) is a Bayesian network;
- $\boldsymbol{A}^*$ is the decision set;
- $u: X \to R$ is the utility function, where R represents the real numbers set;
- $\boldsymbol{\delta}$: $\boldsymbol{A} \to \boldsymbol{A}^*$ is the decision rule.



Nodes:
Random variables
$\boldsymbol{X}$ – pre-test result
$\boldsymbol{Y}$ – post-test result
$\boldsymbol{T}$ – agent certainty
Operations:
A_1 – statistical test
A_2 – in formation required from the user
A_i^* – optimal aagent belief state, maximizing the excepted utility
U – utility structure

Fig. 1. Decision network for agent belief problem presented in Case 1.

Let us now postulate a definition of a decision network for the agent belief state.

DEFINITION 2. A *decision network for the agent belief state* is a decision network within which we recognize the following parameters:

- $\boldsymbol{A}$ is a set of (pre- and post-) tests;
- $\boldsymbol{A}^*$ is a set of the possible states of the agent belief;
- $u: X \to [01]$ is a utility function;
- $\boldsymbol{\delta}$: $\boldsymbol{A} \to \boldsymbol{A}^*$ is a weak monotonous decision rule.

Fig. 1 is a graphic representation of the decision network, presenting the agent belief problem in Case 1. The basic operation nodes $A_i \in \boldsymbol{A}$, $i = 1, 2$, completing the tests, are represented in the figure by squares, whereas the circles represent the random value nodes. The dotted lines represent the information arcs, which arcs, in fact visualize the information flow; whereas the continuous lines show the relations into the probability model. The decision rule $\boldsymbol{\delta}$: $\boldsymbol{A} \to \boldsymbol{A}^*$ is defined through (1).

Finally we are ready to postulate a propagation algorithm for decision network for agent belief state. It is as shown below.

Algorithm 1. An algorithm determining agent belief state.
Input: A decision network for agent belief state.
Output: The optimal agent belief state.

Steps:
1. Do the basic operations in the order specified by the decision network information arcs.
2. Apply the decision rule to the basic operation results in order to get the optimal agent belief state.
3. End.

4. A Utility Network

4.1. *Event Tree*

Let us assume that $\boldsymbol{X} = \{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n\}$ is a set of continuous random variables.

In the present discussion we will examine a binary treelike structure with the following properties:

- the nodes and the leaves are mapped events;
- the sample space Ω is mapped in the root;
- each node has 0 or 2 children;
- if nodes A_j and A_k are respectively left and right child of A_i node, then A_j maps the event $\{X_j < x_j^c\}$, whereas A_k maps complementary event $\{X_j \geqslant x_j^c\}$. Thus, we may designate the following equation: $\overline{A_j} = A_k = \{X_j \geqslant x_j^c\}$.

Obviously, the number of the elements in the above-listed structure is $2n + 1$, the number of leaves is $n + 1$, and, furthermore, each element's children are ordered.

The event tree presents the conditions of the decision rules.

DEFINITION 3. The event tree is in *canonical form* if the indices of the random variables – associated with the nodes – aligned from the tree root to the leaves and from left to right, coincide with the first n natural numbers.

From now on we are to consider event trees in canonical form only.

The event trees, presenting the conditions of the decision rules from Case 1 and Case 4 are represented in Fig. 2 and in Fig. 3, respectively.

$$\boldsymbol{\Omega}$$

$$\swarrow \qquad \searrow$$

$$A_1 = \{X_1 < x_1\} \qquad \overline{A}_1 = \{X_1 \geqslant x_1\}$$

$$\swarrow \qquad \searrow$$

$$A_2 = \{X_2 < x_2\} \qquad \overline{A}_2 = \{X_2 \geqslant x_2\}$$

Fig. 2. Event tree, representing the decision rule condition in Case 1.

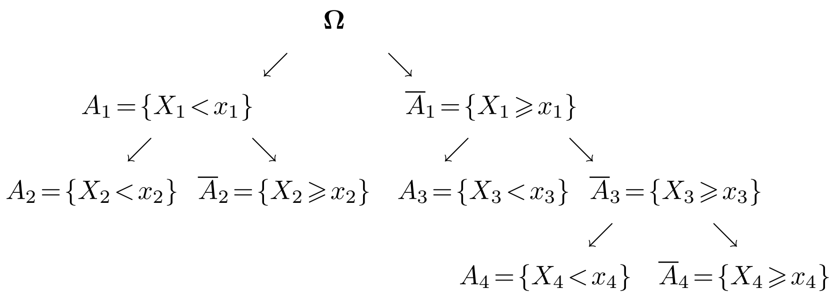


Fig. 3. Event tree, representing the decision rule condition in Case 4.

4.2. *Factors*

DEFINITION 4. The path that goes from the first level to a leaf in the event tree is called a *factor*.

It is obvious that factors' number in the event tree is $n + 1$.

We must bear in mind that we are to interpret the factor as events simultaneously occurring, i.e., as an intersection of the factor's events.

Let $\boldsymbol{F} = \{F_i, i = 1, 2, \ldots, n+1\}$ be the set of the factors in the event tree. It presents the decision rule conditions (see (1)).

For convenience's sake we are to number the factors in event tree from left to right. Further on, we will associate a utility function u_i, $i = 1, 2, \ldots, n + 1$, with each factor, which is to say that each leaf from the event tree will be presumably associated with an utility node. Therefore, a set of utility nodes is associated with a set of factors in the event tree. Consequently, a set of factors, as well as, a set of utility functions is associated with the event tree.

In the light of the ideas, presented in (Shoham, 1997a) and (Shoham, 1997b), we may define a new data structure.

DEFINITION 5. The pair $(\boldsymbol{F}, \boldsymbol{U})$, where $\boldsymbol{F}$ is the set of factors and $\boldsymbol{U}$ is the set of utility nodes – both associated with event tree – is called a *utility network*.

Let us note that to calculate the expected utility we also need to know the probability structure of the random variables set $\boldsymbol{X} = \{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n\}$.

Let us now define the following operations.

The $\psi(F_i) = \psi(A_i)$ operation results in a list of random variables associated with the factor, which factor is further on associated with $\boldsymbol{A}_i$ leaf.

The $\varepsilon(A_i)$ operation results in the index of the element with the greatest index from the random variables list, associated with the factor, which factor is further on associated with $\boldsymbol{A}_i$ leaf.

The $U(A_i)$ operation results in the utility function, associated with $\boldsymbol{A}_i$ leaf.

The $\pi(A_i)$ operation results in $\boldsymbol{A}_i$ node's parent.

Consider the following example.

Let us assume that the factor $F_2 = \{X_1 \geqslant x_1\}, \{X_2 < x_2\}$ is associated with the $\boldsymbol{A}_2$ leaf (see Fig. 2), then $\psi(A_2) = X_1, X_2$, $\varepsilon(A_2) = 2$.

4.3. *Algorithm*

We are now ready to present a propagation algorithm for generation of equations, whose roots are the optimal values x_i^c in the set of factors in the event tree.

Algorithm 2. An algorithm for generating symbolic integral equations.

Input:

- a utility network, associated with the event tree;
- A Bayesian network.

Output: n symbolic integral equations.

Initialization: Each A_i leaf in the event tree is associated with two symbol variables as follows:

$$\boldsymbol{A}_i.\boldsymbol{head} := U(A_i),$$
$$\boldsymbol{A}_i.\boldsymbol{tail} := NIL.$$

Steps:

1. Find $(A_i, \overline{A_i})$ element with the greatest index $\boldsymbol{i}$ from the set of paired leaves, i.e., the leaves having the maximum level number in the event tree.
2. Find $\psi(A_i)$ – the list of the random variables associated with A_i leaf, then find $\boldsymbol{i} = \varepsilon(A_i)$.
3. Find the left hand of the equation:

$$\boldsymbol{left} := \underline{E\{[}\overline{A_i}.\boldsymbol{head}\underline{(T)} - \boldsymbol{A}_i.\boldsymbol{head}\underline{(T)]/}\psi(A_i)\underline{\}+}\overline{A_i}.\boldsymbol{tail}\underline{-}\boldsymbol{A}_i.\boldsymbol{tail}$$

4. Generate the equation:

$$\boldsymbol{left}\underline{= 0}$$

5. Find $A_k = \pi(A_i)$, which is the paired leaves parent $(A_i, \overline{A_i})$. If it turns to be the tree's root go to Step 9. Otherwise designate $\overline{A_i^0} = [x_i^c; 1]$ and find the random variables list $\boldsymbol{\psi}(A_k) = \psi(A_i)\backslash \boldsymbol{X}_i$.
6. If $\pi(A_i) = \overline{A_k}$ then accept $\overline{A_k}.\boldsymbol{head} := \boldsymbol{A}_i.\boldsymbol{head}$,

$$\overline{A_k}.\boldsymbol{tail} := \int\limits_{\overline{A_i^0}} \underline{\{}\boldsymbol{left}\underline{\}.}\underline{f_i(}\boldsymbol{x}_i\underline{/}\boldsymbol{\psi}(A_k)\underline{)\mathrm{d}}\boldsymbol{x}_i\underline{+}A_i.\boldsymbol{tail},$$

else accept $\boldsymbol{A}_k$.***head*** $:=$ $\boldsymbol{A}_i$.***head***,

$$A_k.\textbf{\textit{tail}} := \int\limits_{\overline{A}_i^0} \underline{\{}\textbf{\textit{left}}\underline{\}}.\underline{f_i(}\boldsymbol{x}_i\underline{/}\boldsymbol{\psi}(A_k)\underline{)\mathrm{d}}\boldsymbol{x}_i\underline{+}A_i.\textbf{\textit{tail}}$$

7. Remove A_i and $\overline{A_i}$ leaves.
   Hence, we have articulated an equation for x_i^c and a new utility network.
8. Repeat the steps above beginning with Step 1.
9. End.

In the algorithm described above, the $:=$ sign has been used to designate the operation for value acceptance of a symbolic variable; whereas the $=$ sign has been used to note the same operation, but for a numerical variable. The concatenation operation was not explicitly shown. For clarity's sake the symbolic strings in the concatenation operation are underlined; whereas the symbolic variables are in bold.

4.4. *Fundamental Theorem*

Let us consider an arbitrary event tree with $n+1$ leaves. Let us bound the expected utility step-by-step, applying the following operations for each step k, $k = 1, 2, \ldots, n$.

We are to consider a paired leaves and the utility nodes, associated with them. Find an upper bound of the expected utility. Remove the paired leaves, and then associate this upper bound with their parent. However, our aim is not to find an upper bound of the expected utility, but to generate optimal threshold values equations. To find them it is enough to preserve the needed information in two symbolic variables and to associate them with the leave parent. What's more, the information in these two variables is sufficient to restore the upper bound of the expected utility in every step.

First, let us associate two symbolic variables $A_k.head$ and $A_k.tail$, within which the needed information for generating an equation will be preserved, with each leaf $\boldsymbol{A}_k$, $k = 1, \ldots, n+1$.

Let this event tree be initialized in the following manner:

- $A_k.head = u_k(t)$, where $u_k(t)$ is the utility function, associated with $\boldsymbol{A}_k$ leaf,
- $A_k.tail = \phi$, where ϕ is the empty string.

Then, the following theorem holds:

Theorem. *An arbitrary utility network with an initialized event tree is considered. Let* $A_{j-1} = \pi(A_j)$ *and* $\overline{A_{j-1}} = \pi(A_i)$, *and Algorithm* 2 *is applied.*

Then the threshold value x_{j-1}^c *of the random variable, associated with the paired nodes* $(A_{j-1}, \overline{A_{j-1}})$, *satisfies the equation:*

$$E\left\{\overline{[A_{j-1}}.head(T) - A_{j-1}.head(T)]/\boldsymbol{\psi}(A_{j-1})\right\} + \overline{A_{j-1}}.tail - A_{j-1}.tail = 0,$$

where

$$A_{j-1}.head = A_j.head,$$

$$A_{j-1}.tail = \int\limits_{\overline{A_j^0}} \Big\{ E\left\{[\overline{A_j}.head(T) - A_j.head(T)]/\boldsymbol{\psi}(A_j)\right\}$$

$$+\overline{A_j}.tail - A_j.tail \Big\} f_j\left(x_j/\boldsymbol{\psi}(A_{j-1})\right) \mathrm{d}x_j + A_j.tail,$$

where

$$\overline{A_j^0} = \Big\{ x_j \colon E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(A_{j-1})\right\}$$

$$+\overline{A_j}.tail - A_j.tail \geqslant 0 \Big\};$$

$$\overline{A_{j-1}}.head = A_i.head,$$

$$\overline{A_{j-1}}.tail = \int\limits_{\overline{A_i^0}} \Big\{ E\left\{\left[\overline{A_i}.head(T) - A_i.head(T)\right]/\boldsymbol{\psi}(A_i)\right\}$$

$$+\overline{A_i}.tail - A_i.tail \Big\} .f_i\left(x_i/\boldsymbol{\psi}(A_{j-1})\right) \mathrm{d}x_i + A_j.tail,$$

where

$$\overline{A_i^0} = \Big\{ x_i \colon\ E\Big\{ \overline{[A_i}.head(T) - A_i.head(T)]/\boldsymbol{\psi}(A_i) \Big\}$$

$$+\overline{A_i}.tail - A_i.tail \geqslant 0 \Big\},$$

and Algorithm 2 *correctly generates these equations.*

Proof. The basic units a utility network is composed of are called u-units. The three basic units of a utility network are: symmetric, right asymmetric and left asymmetric ones. They are presented in Fig. 4.

1. Let us first consider a partial case in which the leaves of the u-units are leaves in the utility network, too.

1.1. Let us consider the u-unit from Fig. 4a as a utility network. We will bound the expected utility step-by-step.

$$Eu(\overline{A_j}) = E\left\{u_j(T)/\psi(F_{j-1})\right\} + \int\limits_{\overline{A_j}} E\left\{\left[u_{j+1}(T) - u_j(T)\right]/\psi(F_j)\right\}.f_j\left(x_j/\psi(F_{j-1})\right) \mathrm{d}x_j.$$

The expected utility may be presented as follows:

$$Eu(\overline{A_j}) = E\left\{A_j.head(T)/\psi(F_{j-1})\right\} +$$

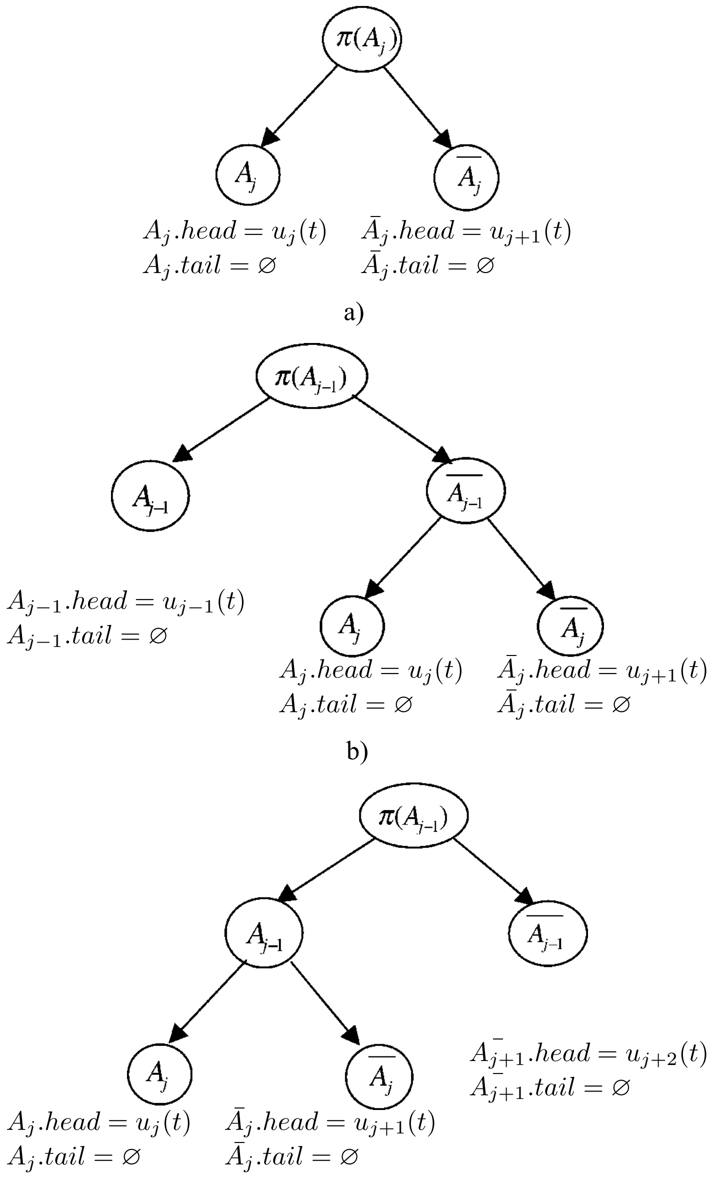


Fig. 4. U-nits: a) symmetric; b) right asymmetric; c) left asymmetric.

$$+\int_{\overline{A_j}} \Big\{ E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(F_j)\right\}$$

$$+\overline{A_j}.tail - A_j.tail \Big\} f_j\left(x_j/\psi(F_{j-1})\right) \mathrm{d}x_j.$$

Then an upper bound of the expected utility is:

$$\begin{aligned} Eu(\overline{A_j^0}) = {} & E\left\{A_j.head(T)/\psi(F_{j-1})\right\} \\ & + \int_{\overline{A_j^0}} \Big\{ E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(F_j)\right\} \\ & + \overline{A_j}.tail - A_j.tail \Big\} f_j\left(x_j/\psi(F_{j-1})\right) \mathrm{d}x_j, \end{aligned}$$

where

$$\begin{aligned} \overline{A_j^0} = \Big\{ x_j\colon\ & E\Big\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_j)\Big\} \\ & + \overline{A_j}.tail - A_j.tail \geqslant 0 \Big\}. \end{aligned} \tag{2}$$

Let us assume that the function in the left hand of the inequation in (2) changes its sign from negative to positive exactly once.

Consequently, x_j^c is the root of the equation:

$$E\Big\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_j)\Big\} + \overline{A_j}.tail - A_j.tail = 0$$

(see Step 3 and Step 4 of Algorithm 2).

1.2. Let us consider the u-unit in Fig. 4b and use the result obtained above. In the variables $\overline{A_{j-1}}.head$ and $\overline{A_{j-1}}.tail$ associated with the node $\overline{A_{j-1}} = \pi(A_j)$ the information is preserved in the following manner:

$$\begin{aligned} \overline{A_{j-1}}.head &= A_j.head = u_j, \\ \overline{A_{j-1}}.tail &= \int_{\overline{A_j^0}} \Big\{ E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(F_j)\right\} \\ & \quad + \overline{A_j}.tail - A_j.tail \Big\} . f_j\left(x_j/\psi(F_{j-1})\right) \mathrm{d}x_j + A_j.tail, \end{aligned}$$

where

$$\overline{A_j^0} = \Big\{ x_j\colon\ E\Big\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_j)\Big\} + \overline{A_j}.tail - A_j.tail \geqslant 0 \Big\}$$

(see Step 6).

In that case the upper bound of the expected utility may be presented as follows:

$$Eu(\overline{A_j^0}) = E\left\{\overline{A_{j-1}}.head(T)/\psi(F_{j-1})\right\} + \overline{A_{j-1}}.tail\,. \tag{3}$$

We remove $(A_j, \overline{A_j})$ leaves (see Step 7 of Algorithm 2) and in that way, it is within the $\pi(A_j) \equiv \pi(\overline{A_j})$ node where the all information necessary for generating the next equations is preserved.

We continue to bound the expected utility.

$$\begin{aligned} Eu(\overline{A_{j-1}}, \overline{A_j^0}) &= E\{u_{j-1}(T)/\psi(F_{j-2})\} \\ &\quad + \int\limits_{\overline{A_{j-1}}} \Big\{ E\{[u_j(T) - u_{j-1}(T)]/\psi(F_{j-1})\} \\ &\quad + \int\limits_{\overline{A_j^0}} E\{[u_{j+1}(T) - u_j(T)]/\psi(F_j)\} \\ &\quad .f_j(x_j/\psi(F_{j-1}))\,\mathrm{d}x_j \Big\} .f_{j-1}(x_{j-1}/\psi(F_{j-2}))\,\mathrm{d}x_{j-1} \\ &= E\{A_{j-1}.head(T)/\psi(F_{j-2})\} \\ &\quad + \int\limits_{\overline{A_{j-1}}} \Big\{ E\{[A_{j-1}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} \\ &\quad + \overline{A_{j-1}}.tail - A_{j-1}.tail \Big\} .f_{j-1}(x_{j-1}/\psi(F_{j-2}))\,\mathrm{d}x_{j-1}. \end{aligned}$$

An upper bound of the expected utility is:

$$\begin{aligned} Eu(\overline{A^0_{j-1}}, \overline{A_j^0}) &= E\{A_{j-1}.head(T)/\psi(F_{j-2})\} \\ &\quad + \int\limits_{\overline{A^0_{j-1}}} \Big\{ E\{[A_{j-1}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} \\ &\quad + \overline{A_{j-1}}.tail - A_{j-1}.tail \Big\} .f_{j-1}(x_{j-1}/\psi(F_{j-2}))dx_{j-1}, \end{aligned}$$

where

$$\begin{aligned} \overline{A^o_{j-1}} = \Big\{ x_{j-1}\colon\ & E\Big\{ [\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1}) \Big\} \\ & + \overline{A_{j-1}}.tail - A_{j-1}.tail \geqslant 0 \Big\}. \end{aligned} \tag{4}$$

Let us assume that the function in the left hand of the inequation in (4) changes its sign from negative to positive exactly once.

Then x^c_{j-1} is the root of the equation:

$$E\left\{ \left[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)\right]/\psi(F_{j-1}) \right\} + \overline{A_{j-1}}.tail - A_{j-1}.tail = 0.$$

In the variables $\pi(\overline{A_{j-1}}).head$ and $\pi(\overline{A_{j-1}}).tail$ associated with the node $\overline{A_{j-1}} = \pi(A_j)$ the information is preserved in the following manner:

$$\pi(\overline{A_{j-1}}).head = A_{j-1}.head,$$

$$\pi(\overline{A_{j-1}}).tail = \int\limits_{\overline{A^0_{j-1}}} \Big\{E\left\{\left[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)\right]/\psi(F_{j-1})\right\}$$
$$+\overline{A_{j-1}}.tail - A_{j-1}.tail\Big\}.f_{j-1}(x_{j-1}/\psi(F_{j-2}))dx_{j-1} + A_{j-1}.tail,$$

where

$$\overline{A^0_{j-1}} = \Big\{x_{j-1}\colon\ E\Big\{\overline{[A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\Big\}$$
$$+\ \overline{A_{j-1}}.tail - A_{j-1}.tail \geqslant 0\Big\}$$

(see Step 6).

We remove the paired leaves $(A_{j-1}, \overline{A_{j-1}})$ (see Step 7 of Algorithm 2) and in such way the $\pi(A_{j-1}) \equiv \pi(\overline{A_{j-1}})$ node carries the all information required for generating the next equations.

1.3. Let us now consider the u-unit in Fig. 4c.

At the first step the expected utility may be presented as follows:

$$\begin{aligned} Eu(\overline{A_j}) &= E\left\{u_j(T)/\psi(F_{j-1})\right\} \\ &\quad + \int\limits_{\overline{A_j}} E\left\{\left[u_{j+1}(T) - u_j(T)\right]/\psi(F_j)\right\} f_j(x_j/\psi(F_{j-1}))\mathrm{d}x_j \\ &= E\left\{A_j.head(T)/\psi(F_{j-1})\right\} \\ &\quad + \int\limits_{\overline{A_j}} \Big\{E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(F_{j-1})\right\} \\ &\quad + \overline{A_j}.tail - A_j.tail\Big\} f_j(x_j/\psi(F_{j-1}))\mathrm{d}x_j. \end{aligned}$$

In the variables, associated with $A_{j-1} = \pi(A_j)$ node we preserve the needed information:

$$A_{j-1}.head = A_j.head$$

and

$$A_{j-1}.tail = \int\limits_{\overline{A^0_j}} \Big\{E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(F_j)\right\}$$
$$+\overline{A_j}.tail - A_j.tail\Big\}.f_j(x_j/\psi(F_{j-1}))\mathrm{d}x_j + A_j.tail,$$

where

$$A^0_j = \Big\{x_j\colon\ E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(F_j)\right\}$$
$$+\ \overline{A_j}.tail - A_j.tail \geqslant 0\Big\}$$

(see Step 6).

As a result we obtain the upper bound for the expected utility:

$$Eu(\overline{A_j^0}) = E\left\{A_{j-1}.head(T)/\psi(F_{j-1})\right\} + A_{j-1}.tail, \tag{5}$$

and x_j^c is the root of the equation:

$$E\left\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_{j-1})\right\} + \overline{A_j}.tail - A_j.tail = 0.$$

We remove $(A_j, \overline{A_j})$ leaves (see Step 7 of Algorithm 2) and associate the obtained upper bound $Eu(\overline{A_j^0})$ with their parent A_{j-1}.

At the second step the expected utility may be presented as follows:

$$\begin{aligned} Eu(\overline{A_{j-1}}, \overline{A_j^0}) = {} & E\left\{A_{j-1}.head(T)/\psi(F_{j-2})\right\} + A_{j-1}.tail \\ & + \int\limits_{\overline{A_{j-1}}} \Big\{E\left\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\right\} \\ & + \overline{A_{j-1}}.tail - A_{j-1}.tail\Big\} f_{j-1}\left(x_{j-1}/\psi(F_{j-2})\right) \mathrm{d}x_{j-1}. \end{aligned}$$

Then x_{j-1}^c is the root of the equation:

$$E\left\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\right\} + \overline{A_{j-1}}.tail - A_{j-1}.tail = 0.$$

We appropriate

$$\begin{aligned} & \pi(A_{j-1}).head = A_{j-1}.head, \\ & \pi(A_{j-1}).tail = \int\limits_{\overline{A_{j-1}^0}} \Big\{E\left\{\left[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)\right]/\psi(F_{j-1})\right\} \\ & \quad + \overline{A_{j-1}}.tail - A_{j-1}.tail\Big\} f_{j-1}\left(x_{j-1}/\psi(F_{j-2})\right) \mathrm{d}x_{j-1} + A_{j-1}.tail, \end{aligned}$$

where

$$\begin{aligned} \overline{A_{j-1}^0} = \Big\{ & x_{j-1}\colon E\left\{\left[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)\right]/\psi(F_{j-1})\right\} \\ & + \overline{A_{j-1}}.tail - A_{j-1}.tail \geqslant 0\Big\}. \end{aligned}$$

An upper bound of the expected utility is

$$Eu(\overline{A_{j-1}^0}, \overline{A_j^0}) = E\left\{\pi(A_{j-1}).head(T)/\psi(F_{j-2})\right\} + \pi(A_{j-1}).tail\,.$$

We remove $(A_{j-1}, \overline{A_{j-1}})$ leaves (see Step 7 of Algorithm 2). Then $\pi(A_{j-1}) = \pi(\overline{A_{j-1}})$ node carries the all information required for generating the next equations.

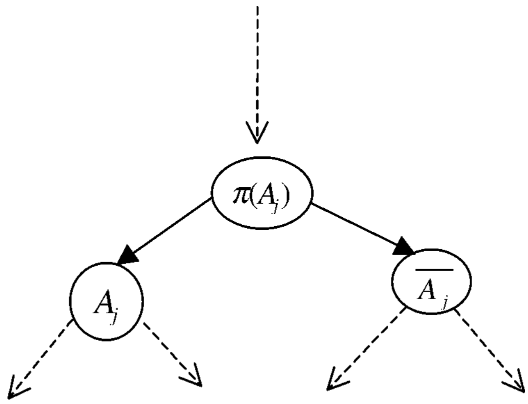


Fig. 5. Symmetric u-unit in an utility network.

2. In the general case herein the unit leaves are nodes in the utility network (see Fig. 5). Let us assume that the utility function

$$u^*(t) = E\left\{\overline{A_j}.head(T)/\psi(F_j)\right\} + \overline{A_j}.tail$$

is associated with $\overline{A_j}$ leaf and that the other utility function

$$u^{**}(t) = E\left\{A_j.head(T)/\psi(F_j)\right\} + A_j.tail$$

is associated with A_j leaf.

Then the expected utility is

$$\begin{aligned} Eu(\overline{A_j}) \;&=\; E\left\{A_j.head(T)/\psi(F_{j-1})\right\} + A_j.tail \\ &\quad+ \int\limits_{\overline{A_j}} \Big\{E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(F_j)\right\} \\ &\quad+\overline{A_j}.tail - A_j.tail\Big\} f_j\left(x_j/\psi(F_{j-1})\right) \mathrm{d}x_j, \end{aligned}$$

and an upper bound of the expected utility is:

$$\begin{aligned} Eu(\overline{A_j^0}) \;&=\; E\left\{A_j.head(T)/\psi(F_{j-1})\right\} + A_j.tail \\ &\quad+ \int\limits_{\overline{A_j^0}} \Big\{E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(F_j)\right\} \\ &\quad+\overline{A_j}.tail - A_j.tail\Big\} f_j\left(x_j/\psi(F_{j-1})\right) \mathrm{d}x_j, \end{aligned}$$

where

$$A_j^0 = \Big\{x_j\colon\, E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(F_j)\right\}$$

$$+ \overline{A_j}.tail - A_j.tail \Big\} \geqslant 0.$$

Therefore, x_j^c is the root of the equation:

$$E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(F_j)\right\}\overline{A_j}.tail - A_j.tail = 0$$

(see Step 3 and Step 4 of Algorithm 2).

2.1. Let $\pi(A_j) = A_{j-1}$.

In the variables associated with $A_{j-1} = \pi(A_j)$ node we preserve the needed information:

$$\begin{aligned} A_{j-1}.head &= A_j.head \quad \text{and} \\ A_{j-1}.tail &= \int\limits_{\overline{A_j^0}} \Big\{E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(F_j)\right\} \\ &\quad + \overline{A_j}.tail - A_j.tail\Big\}.f_j(x_j/\psi(F_{j-1}))\mathrm{d}x_j + A_j.tail \end{aligned}$$

(see Step 6).

Then, we receive the following upper bound of the expected utility:

$$Eu(\overline{A_j^0}) = E\{A_{j-1}.head(T)/\psi(F_{j-1})\} + A_{j-1}.tail$$

(see (5)).

We remove the paired leaves $(A_j, \overline{A_j})$ (see Step 7 of Algorithm 2). And it is with their parent A_{j-1} we associate the upper bound $Eu(\overline{A_j^0})$, obtained above.

2.2. Let $\pi(A_j) = \overline{A_{j-1}}$.

In the variables associated with $\overline{A_{j-1}} = \pi(A_j)$ nodes, the necessary information is preserved in the following manner:

$$\begin{aligned} \overline{A_{j-1}}.head &= A_j.head, \\ \overline{A_{j-1}}.tail &= \int\limits_{\overline{A_j^0}} \Big\{E\left\{\left[\overline{A_j}.head(T) - A_j.head(T)\right]/\psi(F_j)\right\} \\ &\quad + \overline{A_j}.tail - A_j.tail\Big\}.f_j(x_j/\psi(F_{j-1}))\mathrm{d}x_j + A_j.tail, \end{aligned}$$

where

$$\overline{A_j^0} = \Big\{x_j\colon\, E\Big\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_j)\Big\} + \overline{A_j}.tail - A_j.tail \geqslant 0\}$$

(see Step 6).

In that case an upper bound of the expected utility is:

$$Eu(\overline{A_j^0}) = E\left\{\overline{A_{j-1}}.head(T)/\psi(F_{j-1})\right\} + \overline{A_{j-1}}.tail \quad \text{(see (3))}.$$

We remove the paired leaves $(A_j, \overline{A_j})$ (see Step 7 of Algorithm 2). And it is again with their parent $\overline{A_{j-1}}$ we associate the upper bound, obtained above.

Therefore, the theorem is proved.

4.5. *Example*

The problem for decision making about the agent belief state, as presented in Case 4, is considered herein. It may be represented by both the utility network in Fig. 6 and Bayesian network in Fig. 7. The integral equations will be generated by means of Algorithm 2.

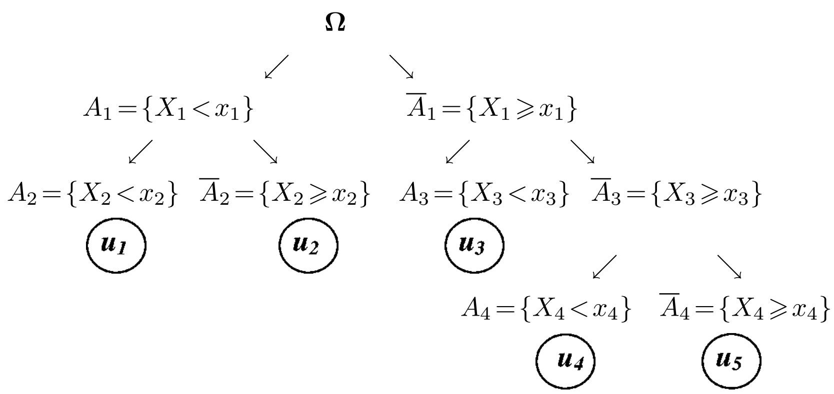
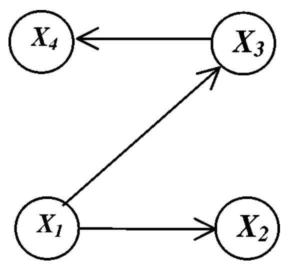


Fig. 6. Utility network presenting the utility structure in Case 4.



$$f(x_1, x_2, x_3, x_4) = f_1(x_1).f_2(x_2/x_1).f_3(x_3/x_1).f_4(x_4/x_3).$$

Fig. 7. Bayesian network presenting the probability structure in Case 4.

Initialization:
$\boldsymbol{A}_2$.***head*** $:= u_1$, $\boldsymbol{A}_2$.***tail*** $:= NIL$,
$\overline{A_2}$.***head*** $:= u_2$, $\overline{A_2}$.***tail*** $:= NIL$,
$\boldsymbol{A}_3$.***head*** $:= u_3$, $\boldsymbol{A}_3$.***tail*** $:= NIL$,
$\boldsymbol{A}_4.head := u_4$, $\boldsymbol{A}_4$.***tail*** $:= NIL$,
$\overline{A_4}$.***head*** $:= u_5$, $\overline{A_4}$.***tail*** $:= NIL$.
Steps:

1. The paired leaves $(A_i, \overline{A_i})$ have the greatest index from the highest level of the event tree.
2. $\boldsymbol{\psi}(A_4) = \boldsymbol{X}_4, \boldsymbol{X}_3, \boldsymbol{X}_1$ and $\boldsymbol{i} = \varepsilon(A_4) = 4$.
3. ***left*** $:= E\{[\boldsymbol{u}_5(T) - u_4(T)]/\boldsymbol{X}_4, \boldsymbol{X}_3, \boldsymbol{X}_1\}$.
4. The equation is
$$E\left\{[\boldsymbol{u}_5(T) - \boldsymbol{u}_4(T)]\,/\boldsymbol{X}_4, \boldsymbol{X}_3, \boldsymbol{X}_1\right\} = 0.$$
5. $\overline{A_4^0} = [x_4^c, 1]$, $\overline{A_3} = \pi(A_4)$, and $\Psi(\overline{A_3}) = X_3, X_1$.
6. $\overline{A_3}$.***head*** $:= \boldsymbol{u}_4$,
$$\overline{A_3}.\textbf{\textit{tail}} := \int\limits_{\overline{A}_i^0} \Big\{E\left\{[\boldsymbol{u}_5(T) - \boldsymbol{u}_4(T)]\,/\boldsymbol{X}_4, \boldsymbol{X}_3, \boldsymbol{X}_1\right\}\Big\}.f_4\left(\boldsymbol{x}_4/\boldsymbol{X}_3, \boldsymbol{X}_1\right)\mathrm{d}\boldsymbol{x}_4.$$
7. We remove A_4 and $\overline{A_4}$ leaves.
Hence, the first equation $E\{[\boldsymbol{u}_5(T) - \boldsymbol{u}_4(T)]/\boldsymbol{X}_4, \boldsymbol{X}_3, \boldsymbol{X}_1\} = 0$ is obtained.
8. The steps above are repeated beginning with Step 1.

1. Now $(A_3, \overline{A_3})$ have the biggest index from the highest level of the event tree.
2. $\boldsymbol{\psi}(A_3) = \boldsymbol{X}_3, \boldsymbol{X}_1$ and $\boldsymbol{i} = \varepsilon(A_4) = 3$.
3. ***left*** $:= E\{[\boldsymbol{u}_4(T) - \boldsymbol{u}_3(T)]/\boldsymbol{X}_3, \boldsymbol{X}_1\} + \int_{\overline{A}_4^0}\{E\{[\boldsymbol{u}_5(T) - \boldsymbol{u}_4(T)]/\boldsymbol{X}_4, \boldsymbol{X}_3, \boldsymbol{X}_1\}\}.f_4(\boldsymbol{x}_4/\boldsymbol{X}_3, \boldsymbol{X}_1)\mathrm{d}\boldsymbol{x}_4$.
4. We obtain the following equation:
$$E\left\{[\boldsymbol{u}_4(T) - \boldsymbol{u}_3(T)]\,/\boldsymbol{X}_3, \boldsymbol{X}_1\right\} + \int\limits_{\overline{A}_4^0}\Big\{E\left\{[\boldsymbol{u}_5(T) - \boldsymbol{u}_4(T)]/\boldsymbol{X}_4, \boldsymbol{X}_3, \boldsymbol{X}_1\right\}\Big\}.f_4(\boldsymbol{x}_4/\boldsymbol{X}_3, \boldsymbol{X}_1)\mathrm{d}\boldsymbol{x}_4 = 0.$$
5. $\overline{A_3^0} = [x_3^c, 1]$, $\overline{A_1} = \pi(A_3)$, and $\Psi(\overline{A_1}) = X_1$.
6. $\overline{A_1}$.***head*** $:= \boldsymbol{u}_3$ and
$$\overline{A_1}.\textbf{\textit{tail}} := \int\limits_{\overline{A}_3^0}\Bigg\{E\left\{[\boldsymbol{u}_4(T) - \boldsymbol{u}_3(T)]\,/\boldsymbol{X}_3, \boldsymbol{X}_1\right\} + \int\limits_{\overline{A}_4^0}\Big\{E\left\{[\boldsymbol{u}_5(T) - \boldsymbol{u}_4(T)]\,/\boldsymbol{X}_4, \boldsymbol{X}_3, \boldsymbol{X}_1\right\}\Big\}.f_4 (\boldsymbol{x}_4/\boldsymbol{X}_3, \boldsymbol{X}_1)\mathrm{d}\boldsymbol{x}_4\Bigg\}f_3(\boldsymbol{x}_3/\boldsymbol{X}_1)\mathrm{d}\boldsymbol{x}_3.$$

7. We remove A_3 and $\overline{A_3}$ leaves.
   Now we obtained the equation as below:
   $$E\left\{[\boldsymbol{u}_4(T)-\boldsymbol{u}_3(T)]\,/\boldsymbol{X}_3,\boldsymbol{X}_1\right\} + \int\limits_{\overline{A}_4^0}\Big\{E\left\{[\boldsymbol{u}_5(T)-\boldsymbol{u}_4(T)]\,/\boldsymbol{X}_4,\boldsymbol{X}_3,\boldsymbol{X}_1\right\}\Big\}.f_4(\boldsymbol{x}_4/\boldsymbol{X}_3,\boldsymbol{X}_1)\mathrm{d}\boldsymbol{x}_4 = 0.$$
8. We repeat the steps above beginning with Step 1.

1. $(A_2,\overline{A_2})$ have the biggest index from the highest level of the event tree.
2. $\boldsymbol{\psi}(A_2)=\boldsymbol{X}_2,\boldsymbol{X}_1$ and $\boldsymbol{i}=\varepsilon(A_4)=2$.
3. ***left*** $:= E\{[\boldsymbol{u}_2(T)-\boldsymbol{u}_1(T)]/\boldsymbol{X}_2,\boldsymbol{X}_1\}$.
4. The equation $E\{[\boldsymbol{u}_2(T)-\boldsymbol{u}_1(T)]/\boldsymbol{X}_2,\boldsymbol{X}_1\}=0$ is generated.
5. $\pi(A_2)=\overline{A_1}$, $\Psi(\overline{A_1})=X_1$.
6. $\overline{A_2^0}=[x_2^c,1]$, A_1.***head*** $:= \boldsymbol{u}_1$,
   $$A_1.\textbf{\textit{tail}} := \int\limits_{\overline{A}_2^0}\Big\{E\left\{[\boldsymbol{u}_2(T)-\boldsymbol{u}_1(T)]\,/\boldsymbol{X}_2,\boldsymbol{X}_1\right\}\Big\}.f_2(\boldsymbol{x}_2/\boldsymbol{X}_1)\mathrm{d}\boldsymbol{x}_2.$$
7. We remove A_2 and $\overline{A_2}$ leaves.
   Hence, the equation $E\{[\boldsymbol{u}_2(T)-\boldsymbol{u}_1(T)]/\boldsymbol{X}_2,\boldsymbol{X}_1\}=0$ is generated.
8. We repeat the steps above starting with Step 1.

1. Finally, we work with the paired leaves $(A_1,\overline{A_1})$.
2. $\boldsymbol{\psi}(A_1)=\boldsymbol{X}_1$ and $\boldsymbol{i}=\varepsilon(A_4)=1$.
3. ***left*** $:=$
   $$E\left\{[\boldsymbol{u}_3(T)-\boldsymbol{u}_1(T)]\,/\boldsymbol{X}_1\right\} + \int\limits_{\overline{A}_3^0}\Big\{E\left\{[\boldsymbol{u}_4(T)-\boldsymbol{u}_3(T)]\,/\boldsymbol{X}_3,\boldsymbol{X}_1\right\} + \int\limits_{\overline{A}_4^0}\Big\{E\big\{\big[\boldsymbol{u}_5(T)-\boldsymbol{u}_4(T)\big]/\boldsymbol{X}_4,\boldsymbol{X}_3,\boldsymbol{X}_1\big\}\Big\}.f_4(\boldsymbol{x}_4/\boldsymbol{X}_3,\boldsymbol{X}_1)\mathrm{d}\boldsymbol{x}_4\Big\} \times f_3(\boldsymbol{X}_3/\boldsymbol{X}_1)\mathrm{d}\boldsymbol{x}_3 - \int\limits_{\overline{A}_2^0}\Big\{E\left\{[\boldsymbol{u}_2(T)-\boldsymbol{u}_1(T)]\,/\boldsymbol{X}_2,\boldsymbol{X}_1\right\}\Big\}.f_2(\boldsymbol{x}_2/\boldsymbol{X}_1)\mathrm{d}\boldsymbol{x}_2.$$
4. The equation is:
   $$E\left\{[\boldsymbol{u}_3(T)-\boldsymbol{u}_1(T)]\,/\boldsymbol{X}_1\right\} + \int\limits_{\overline{A}_3^0}\Big\{E\left\{[\boldsymbol{u}_4(T)-\boldsymbol{u}_3(T)]\,/\boldsymbol{X}_3,\boldsymbol{X}_1\right\} + \int\limits_{\overline{A}_4^0}\Big\{E\big\{[\boldsymbol{u}_5(T) -\boldsymbol{u}_4(T)]/\boldsymbol{X}_4,\boldsymbol{X}_3,\boldsymbol{X}_1\big\}\Big\}.f_4(\boldsymbol{x}_4/\boldsymbol{X}_3,\boldsymbol{X}_1)\mathrm{d}\boldsymbol{x}_4\Big\} f_3(\boldsymbol{X}_3/\boldsymbol{X}_1)\mathrm{d}\boldsymbol{x}_3 - \int\limits_{\overline{A}_2^0}\Big\{E\left\{[\boldsymbol{u}_2(T)-\boldsymbol{u}_1(T)]\,/\boldsymbol{X}_2,\boldsymbol{X}_1\right\}\Big\}.f_2(\boldsymbol{x}_2/\boldsymbol{X}_1)\mathrm{d}\boldsymbol{x}_2 = 0.$$
5. $\pi(A_2)=\Omega$. Go to Step 9.

9. The end.

Conclusions

The agent belief in the probability model may be presented via the following three items $(\mathfrak{B}, \mathfrak{D}, \mathfrak{U})$, where:

- $\mathfrak{B}$ is a Bayesian network, presenting the probability structure of the agent belief problem;
- $\mathfrak{D}$ is a decision network of the agent belief state.
- $\mathfrak{U}$ is a utility network, presenting the utility structure of the agent belief problem.

A decision for the agent belief state can be made via propagation in $\mathfrak{D}$.
The agent belief state can be optimized via propagation in $\mathfrak{U}$, using at the time $\mathfrak{B}$.

References

Noncheva, V. (1998). An approach to optimizing the agent's belief about the probability distribution of a population. *Scientific works of Plovdiv University "Paissii Hilendarski"*, **33**(3) (to appear).
Noncheva, V., and S. Stojanov (1996). What should a future multiagent system for statistical investigations be? In *Statistical data analysis '95*, *Proc. of SDA'95 and SDA'96*, pp. 69–75.
Prat, A., J. Lores, K. Sanjeevan, I. Sole, and J. Catot (1994). Application of a multiagent distributed architecture for time series forecasting, COMPSTAT-94, 301–316.
Shachter, R.D. (1986). Evaluating influence diagrams. *Operations Research*, **34**, 871–882.
Shachter, R.D. (1988). Probabilistic inference and influence diagrams, *Operations Research*, **36**, 591–604.
Shoham, Y. (1997). A symmetric view of probabilities and utilities. In *Proc. of IJCAI-97*.
Shoham, Y. (1997). Two senses of conditional utility. In *Proc. of UAI-97*.

V. Noncheva is a professor assistant in probability, applied statistics and artificial intelligence at University of Plovdiv. She received her M. Sc. degree in probability and statistics in University of Sofia in 1981. In 1984 she specialized microcomputer software in Computer Education Center in Budapest. She is finishing her Ph. D. work on an approach to building of intelligent statistical software. She is interested in the two-side problem: what can modern statistical software benefit from artificial intelligence, and what can stochastics offer artificial intelligence systems.

Agento įsitikinimas: prezentacija, sklidimas ir optimizavimas

Veska Noncheva

Straipsnio tikslas yra parodyti stochastinį būdą agentų įsitikinimui statistiškai modeliuoti ir optimizuoti naudojantis tikimybiniu modeliu. Apibrėžti du tinklai: agento įsitikinimo būklės sprendimo tinklas $\mathfrak{D}$ ir naudos tinklas $\mathfrak{U}$, pateikiantis agento įsitikinimo uždavinio naudos struktūrą. Agento įsitikinimas yra atvaizduotas naudojantis trijų dydžių kombinacija $(\mathfrak{B}, \mathfrak{D}, \mathfrak{U})$. Čia $\mathfrak{B}$ yra Bajeso tinklas, nusakantis agento įsitikinimo uždavinio tikimybinę struktūrą. Taip pat pasiūlyti du sklidimo tinklu algoritmai.