ЮБИЛЕЙНА НАУЧНА СЕСИЯ – 30 години ФМИ
ПУ "Паисий Хилендарски", Пловдив, 3-4.11.2000

AGENT'S BELIEF: A STOCHASTIC MODEL

Veska Pencheva Noncheva

The agent's belief is a threshold function of agent's certainty. However, the agent's certainty is unobservable.
In this paper a stochastic model of agent's belief is presented. This model is based on tests results. An example for determining statistical agent's belief in probability model is discussed.

Key words: agent's belief, utility network, weak monotony rules
AMS (MOS) Subject Classifications: 68T35, 68N05, 62C12.

Introduction

Nowadays Artificial Intelligence aims at creating agents. These agents embody expertise and intelligent behaviour [Russel, Norving,95]. The states of the agents consist of components such as knowledge, belief, intention, obligation.

In this paper a general model for determining the agent's belief is proposed. This model has the following features:

- It is necessary to choose one of several alternative agent's belief states determined in advance.
- The agent's belief problem is formulated through the stochastics terms. One of the random variables, being of particular interest, is unobservable.
- A numeric measure called utility, measuring the profit of every agent's belief state, is given and we aim at maximizing the expected utility.
  One particular case of the problem under discussion is considered in [*Noncheva, 2000*].

The agent's belief can be presented through a threshold function of the degree of the agent's certainty. However, the agent's certainty is unobservable, and because of that testing is needed. The test results are a kind of estimators of the degree of agent's certainty. The information from the tests is used to make decision about the agent's belief. Furthermore, the higher result values received from the first tests will result in lower requirements for the next test results to be obtained.

We are to discuss the decision-making rules having a monotonous form. For instance, the agent is certain that a statement is true, if the value of the variable, representing the certainty, is bigger than the threshold value preliminary determined. The agent, on the other hand, rejects the statement, if the value of the variable is smaller than the threshold value. An algorithm for determining the optimal threshold values is discussed in [*Noncheva, 99*].

General Stochastic Model of the Agent's Belief

The general formulation of the problem for determining the agent's belief state is following:

1. Let $\boldsymbol{X_i}$, *i=1,2,…n* be continious random variables defined on sample spaces $\Omega_i = [0,1], i = 1,...,n$, which random variables can be observed. We interpret $\boldsymbol{X_i}$ as a test result and $\boldsymbol{X_1},\boldsymbol{X_2},...,\boldsymbol{X_n}$ as a sequence of the results from the tests.
2. Let ***T*** be a continious random variable defind on $\Omega_t = [0,1]$, which cannot be observed and is being interpret as *the agent's certainty*.
3. The Bayesian model of the probability structure is known; consequently, the joint probability distribution $\boldsymbol{f}(x_1, \ldots, x_n, t)$ of the random variables $\boldsymbol{X_1},\ldots, \boldsymbol{X_n},\boldsymbol{T}$ is known, as well.
4. The finite set of the possible agent's belief states ***D*** is known, too.
5. The utility function $\boldsymbol{U}(t,d)$: $\Omega_t \times D \to [0,1]$ is also known.

A decision-making rule is the $\delta(x_1,...,x_n)$ rule, which for each possible realization $(x_1,...,x_n)$ of the random vector $(X_1,...,X_n)$ determines which state $\boldsymbol{a_j} \in \boldsymbol{D}$, $j=1,\ldots,k$, will be acquired by the agent's belief. That is, the decision-making rule is a function of random variables $X_1,...,X_n$ defined on $\Omega_1 \times ... \times \Omega_n$ and with range space ***D***. The goal is to find a decision-making rule, which is to maximize the expected utility.

If there are several decisions, resulting in one and the same maximal expected utility, then we can consider each of these decisions as optimal. In this case the randomized decision-making rules are acceptable, but they have no priorities.

It is intuitively obvious that the high value of the result from test *i* will result in low requirements towards the result from test *j*, when *j*>*i*. That is, the preliminary obtained information influences the decision-making rules. The decision-making rules in which the decisions from the test *j* are functions of the obtained result from the test *i*, *i*<*j*, are called *weak rules*.

It is natural to discuss the decision-making rules, having a *monotonous form*, i.e. the rules with threshold points $x_i^c, i = 1,2,..., n$, forming partions of the sample spaces $\Omega_i = [0,1]$ in the following manner $\Omega_i = A_i + \overline{A_i}, i = 1,..., n$, where $A_i = \{x_i : x_i < x_i^c\}$ and $A_i = \{x_i : x_i \geq x_i^c\}$. Therefore, the problem for determining the agent's belief state means that we are to find *n* threshold points $x_i^c, i = 1,2,..., n$, for each $\boldsymbol{X}_i$, $i = 1,2,..., n$, which points are optimal in accordance with the Bayesian approach.

Hence the purpose is to find a *weak monotonous Bayesian rule* for determining the agent's belief state.

Example

Consider statistical agent's belief in the probability model i.e. in the type of population distribution.

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

The agent's belief in the probability model must be based on the mechanism knowledge of the phenomenon under investigation. But if the phenomenon is unknown, the agent can make its own choice about the probability distribution after it has tested statistical hypotheses. It can also ask the user for his opinion and make use of his expertise.

In order to form its belief in the type of population distribution, the agent can start with a statistical test for symmetry. When the hypothesis for symmetry cannot be rejected, the agent has to continue with tests for normality. Usually a Goodness-of-fit test is first used and after that the user opinion is asked. The results from these three tests are respectively 1-p_1, 1-p_2, where p_1 and p_2 are the p-values of the statistics of the two statistical tests, and the degree of the user certainty of normality, represented as numbers in the interval [0,1].

Designate with $\boldsymbol{X_1}$ the observed value 1-p_1, where p_1 is p-value of the statistics of the test for symmetry. Designate with $\boldsymbol{X_2}$ the observed value 1-p_2, where p_2 is p-value of the statistics of the test for normality. Designate with $\boldsymbol{X_3}$ the degree of the user's certainty of normality. Designate with $\boldsymbol{T}$ the agent's certainty of normality, which cannot be observed.

Assume that $\boldsymbol{X_1}$, $\boldsymbol{X_2}$, $\boldsymbol{X_3}$ and $\boldsymbol{T}$ are continuous random variables with a joint probability density function $\boldsymbol{f}(x_1,x_2,x_3,t)$.

The decision rule $\delta(x_1,x_2,x_3)$ determines the state a_j, $j=0,1,2,3$, of the agent's belief for each possible realization (x_1,x_2,x_3) of the random vector (X_1,X_2,X_3).

The weak decision rule δ in this case has the form:

$$\{(x_1,x_2,x_3):\delta(x_1,x_2,x_3)=a_0\}=A_1\times[0,1]\times[0,1]$$

$$\{(x_1,x_2,x_3):\delta(x_1,x_2,x_3)=a_1\}=\overline{A_1}\times A_2(x_1)\times[0,1]$$

$$\{(x_1,x_2,x_3):\delta(x_1,x_2,x_3)=a_2\}=\overline{A_1}\times\overline{A_2}(x_1)\times A_3(x_1,x_2)$$

$$\{(x_1,x_2,x_3):\delta(x_1,x_2,x_3)=a_3\}=\overline{A_1}\times\overline{A_2}(x_1)\times\overline{A_3}(x_1,x_2),$$

where A_1 and $\overline{A_1}$ are the sets of values of the test leading respectively to the rejection and to the acceptance of the statement for symmetry, $A_2(x_1)$ and $\overline{A_2}(x_1)$ are sets of values of the pre-test leading respectively to the rejection and to the acceptance of the statement for normality, $A_3(x_1,x_2)$ and $\overline{A_3}(x_1,x_2)$ are sets of values of the post-test leading respectively to the rejection and to the acceptance of the statement for normality.

The weak monotonous rule for making a decision about the agent's belief is defined as follows:

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

where x_1^c, x_2^c, x_3^c are the threshold values for X_1, X_2, X_3, and a_j, $j=0,1,2,3$ are the following agent's belief states:

- $\boldsymbol{a_0}$ - the agent rejects the assumption for the symmetry of the distribution, describing the population under investigation. In the process of the statistical analysis the agent will use the median as the best estimate of the “center” of the distribution since the mean is strongly influenced by outliers in the data.

- $\boldsymbol{a}_1$ - the agent rejects the assumption for normality of the distribution. It will make only use of the assumption for symmetry in the statistical analysis.
- $\boldsymbol{a}_2$ - the agent supposes (suspects) that the distribution of the population being investigated is normal. In the process of the statistical analysis it will make use only of tests which are not sensitive to moderate deviations from the assumption for normality. An example of such a robust test is the ***t***-test.
- $\boldsymbol{a}_3$ – the agent convinced that the distribution describing the population is normal. It will also use statistical tests, which are sensitive to deviations from the assumption for normality. Such tests are, for example, Pearson’s, Fisher’s and Bartlett’s tests for equality of variances.

Assume that utility structure has the following form:

$$\boldsymbol{u}(t) = \begin{cases} \boldsymbol{u}_0(t), & \text{if } \boldsymbol{a}_0 \\ \boldsymbol{u}_1(t), & \text{if } \boldsymbol{a}_1 \\ \boldsymbol{u}_2(t), & \text{if } \boldsymbol{a}_2 \\ \boldsymbol{u}_3(t), & \text{if } \boldsymbol{a}_3, \end{cases}$$

where $\boldsymbol{u}_i(t)$, $i=1,2,3,4$ are continuous, monotonic and bounded functions.

It is well known that if the data is lognormal, that means that it will be normally distributed after a log transformation. Therefore, this model could be improved by adding possibility for data transformation, thus trying to obtain normally distributed responses.

Data Structures

Let us assume that $\boldsymbol{X}=\{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n\}$ is a finite set of continuous random variables. The ***Event Tree*** is a binary treelike structure having 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 $\{X_j \geq x_j^c\}$. Thus, we may designate the following equation: $\overline{A_j} = A_k = \{X_j \geq x_j^c\}$.

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 tree, presenting the conditions of the decision rule from *Example* is represented in *Figure 1*.

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

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 ***F***={***F****ᵢ*, *i=1,2,…,n+1*} be the set of the factors in the event tree. It presents the decision rule conditions. For convenience's sake we are to number the factors in event tree from left to right.

It is with each factor $\boldsymbol{F}_i$ that one of the agent's belief states is associated. In such case we say that the set of agent's belief states is associated with the set of factors from the event tree. The decision-making rule for the agent's belief state can be presented by the above mentioned sets.

Further on, we will associate a utility function u_i with each factor $\boldsymbol{F}_i$, $i=1,2,\ldots,n+1$, which is to say that each leaf from the event tree is associated with a utility node. Therefore, a set of utility nodes is associated with the 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.

The pair (***F***, ***U***), where ***F*** is the set of factors and ***U*** is the set of utility nodes - both associated with event tree - is called a *utility network*.

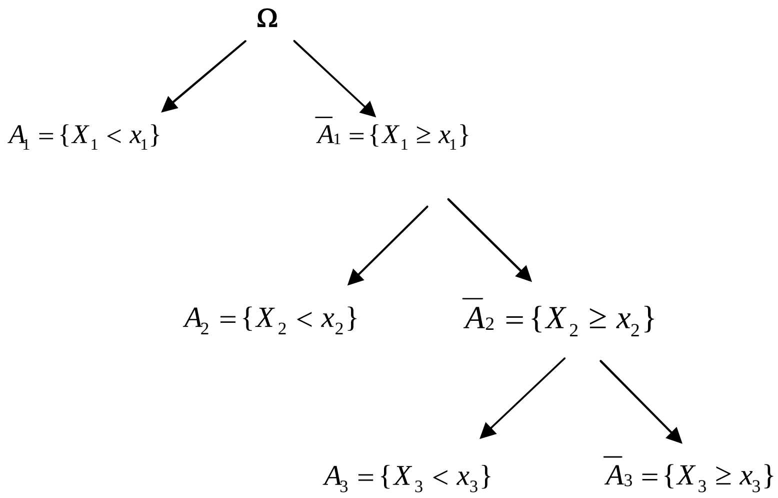

Figure 1. Event tree from *Example*.

REFERENCE

[*Noncheva,99*] Noncheva V. Agent Belief: Presentation, Propagation, and Optimization, *INFORMATICA, Vol.10, No.4,* (1999), 403-426.

[*Noncheva,2000*] Noncheva V. A Statistical Agent's Belief in Gaussian Model, *Mathematica Balkanica, Vol. 14, Fasc. 3-4,* (2000) (to appear).

[Russel, Norving,95] Russel S., P. Norving Artificial Intelligence. A modern Approach, Prentice Hall, 1995.

Faculty of Mathematics and Informatics, University of Plovdiv,
24 Tsar Assen St., 4000 Plovdiv, Bulgaria
E-mail: nonchev@plovdiv.techno-link.com