Journal of the Applied Mathematics, Statistics and Informatics (JAMSI), **1** (2005), No. 1

Decision Trees and Influence Diagrams in Decision Analysis[1]

MILAN TEREK

Abstract

The article deals with the possibilities of using of decision trees and influence diagrams in the solution of some complicated decision problems in which the uncertainty is taken into account. We present one from the algorithms of the solution of influence diagrams. We illustrate the procedures of solution on one example from the area of capital market. The comparison of the decision trees and influence diagrams is also given.

Mathematics Subject Classification 2000: 62C10 Bayesian problems; characterization of Bayes procedures. 62C12 Empirical decision procedures; empirical Bayes procedures.
Additional Key Words and Phrases: Statistical Decision Theory, Decision Analysis, Decision Trees, Influence Diagrams

1. INTRODUCTION

Decision analysis provides structure and guidance for thinking systematically about hard decisions (complexity, inherent uncertainty, multiple objectives, different perspectives lead to different conclusions). Along with a conceptual framework for thinking about hard problems, decision analysis provides analytical tools that can make the required hard thinking easier [4]. It allows people to make effective decisions more consistently and is intended to help people deal with difficult decisions.

Decision analysis provides guidance and structure for systematic thinking in difficult decision situations. Applying the decision analysis techniques can lead to better decisions (the one that gives the best outcome).

We will present two effective tools of decision analysis: decision trees which are well known also in our literature and less known influence diagrams. We will try to compare them.

2. DECISION ANALYSIS AND STATISTICAL DECISION THEORY

According to [11] decision analysis is a prescriptive approach designed for normally intelligent people who want to think hard and systematically about some important real problems.

[1] This article was elaborated with support of the grant agency VEGA in the framework of the project number 1/1247/04: Advanced Statistical Techniques and Decision Making in the Quality Improving Process.

Decision analysis is concerned with the decision making under uncertainty. Personal judgments about uncertainty are important inputs for decision analysis. Decision-analysis approach allows the inclusion of subjective judgments. Really, decision analysis requires personal judgments; they are important components for making good decisions.

Decision analysis is based on the knowledge of Statistical Decision Theory.

Statistical decision theory is concerned with the making of decisions in the presence of statistical knowledge about some of uncertainties involved in the decision problem. We will assume that these uncertainties can be considered to be unknown numerical quantities and will represent them by θ [1].

In the statistical decision theory experimental information (for example sample information) is combined with other types of information.

The first is a knowledge of the possible consequences of the decisions. Often this can be quantified by determining the loss (or the gain (utility)) that would be incurred for each possible decision and for the various possible values of θ. Note that a gain is just a negative loss, so there is no real difference between the two approaches.

The second source of nonexperimental information is called prior information. This is information about θ arising from sources other than the statistical investigation.

Decision analysis can be also defined as the methods of decision making based on the statistical decision theory. It includes mainly the methods of Bayesian decision making which are based on the Bayes decision theory (utilizes nonexperimental sources of information).

A. Actions will be denoted by *a*, while the set of all possible actions under consideration will be denoted *A*.

B. The unknown quantity θ which affects the decision process is obviously called state of nature. The symbol Θ we will use to denote the set of all possible states of nature. When experiments are performed to obtain information about θ, the experiments are designed so that the observations are distributed according to some probability distribution which has θ as an unknown parameter. In such situations θ we will call parameter and Θ the parameter space.

C. The key element of decision theory is the loss function. We will assume a loss function $L(\theta, a)$ is defined for all $(\theta, a) \in \Theta \times A$. If a particular action a_j is taken and θ_i turns out to be the true state of nature, then a loss $L(\theta_i, a_j)$ will be incurred. Only loss functions satisfying $L(\theta, a) \geq -K > -\infty$ are obviously considered. Generally we can consider the payoffs that can be stated in terms of profits, costs, or any other measure of output may be appropriate for the particular situation being analyzed. If the output is of type “the smaller the better” we will use the loss function $L(\theta, a)$, if of the type “the larger the better” we will use the gain function $V(\theta, a)$ defined for all $(\theta, a) \in \Theta \times A$. It is clear that $V(\theta, a) = -L(\theta, a)$.

D. We will assume that we know the prior information concerning θ in terms of a probability distribution on Θ. The symbol $\pi(\theta)$ will be used to represent a prior probability distribution function of θ.

Example. An investor with 1 000 000,- EUR to invest knows that he can:

1. Give all amount to bank account (BA) at 8 % and earn 80 000,- EUR in one year, but after tax his profit would be 68 000,- EUR.

2. Invest the entire amount at 10 % interest in municipal bonds and make a profit 100 000,- EUR in one year, since interest earned from municipal bonds is tax exempt. The actual value of the bonds varies with the prime interest rate set by the Central Bank. For example, an increase of two percentage points may decrease the value of the bonds by 80 000,- EUR and his net profit would be 100 000,-EUR – 80 000,- EUR = 20 000,- EUR if he sells the bonds in the money market at the end of the year. A tax adjustment of 14 000,- EUR on this loss leaves a profit 20 000,- EUR + 14 000,- EUR = 34 000,- EUR. On the other hand, a decrease of two percentage points in the prime rate may increase the value of the bonds by 100 000,- EUR. The investor will have to pay 20 000,- EUR in taxes on this 100 000,- EUR increase if he sells the bonds in the money market and leaving a profit of 100 000,- + (100 000,- EUR – 20 000,- EUR) = 180 000,-EUR.

3. Split his investment evenly between BA and municipal bonds.

Suppose the investor has be able to assign the following subjective probabilities to each state of nature: $P(\theta_1)$ = 0,1, $P(\theta_2)$ = 0,6, $P(\theta_3)$ = 0,3.

Of course, the prime interest rate is not restricted to these three changes, and the investor could alter the splitting of his investment into any ratio. For simplicity we will consider only three changes in the prime interest rate. The after-taxes position of the investor is given in Figure 1 for each combination of action and state of nature (change in the prime interest rate) θ_i (*i*, *j* = 1, 2, 3) in the payoff table.

$\pi(\theta)$	a_j / θ_i	100 % in BA	50 % in BA and 50 % in municipal bonds	100 % in municipal bonds
0,1	$-2\,\%$	68	124	180
0,6	$0\,\%$	68	84	100
0,3	$+2\,\%$	68	51	34

Figure 1 *Payoff table*

Generally in the payoff table the entries represent the payoff for each combination of action and state of nature.

The payoff tables represent one possibility for structuring some decision problems.

3. INFLUENCE DIAGRAMS

An influence diagram (firstly published in [8]) provides a simple graphical representation of a decision problem. The elements of a decision problem – decisions to make, uncertain events, and the value of outcomes - show up in the influence diagram as different shapes. These shapes are linked with arrows or arcs in specific ways to show the relationship among the elements.

According [4], squares represent decisions, circles represent chance events, and rectangles with rounded corners represent values. These shapes generally are referred to as nodes: decision nodes, chance nodes, and value nodes [4]. In [5] the author introduced another shapes and corresponding node names, useful in the cases of multi objective decision making. Nodes are put together in a graph, connected by arrows or arcs. A node at the beginning of an arc we will call a predecessor and one at the end of an arc a successor.

The rules for using arcs to represent relationships among the nodes are shown in Figure 2 (according to [4]).

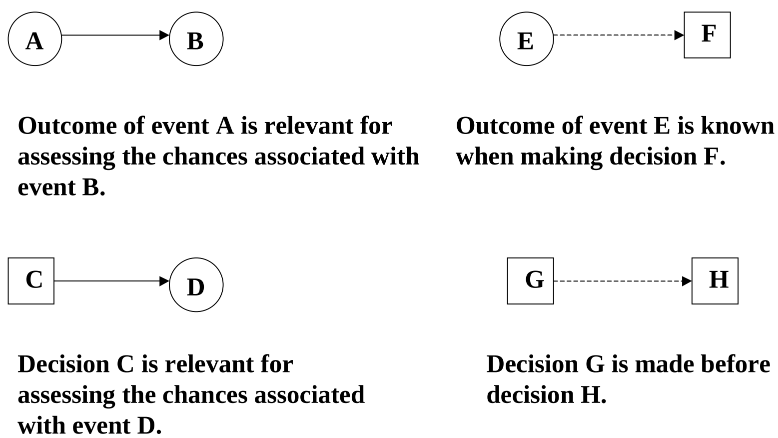


Figure 2 *Representing Influence with Arrows*

Generally, two kinds of arcs exist; these are represented by solid and dashed arrows, respectively [4] (in [5] only solid arrows are used).

A solid arrow or arc pointing into a chance node designates relevance, which indicates that the predecessor is relevant for assessing the chances associated with the uncertain event.

For example, an arrow from Event A to Event B means that the chances (probabilities) associated with B are different for different outcomes of A.

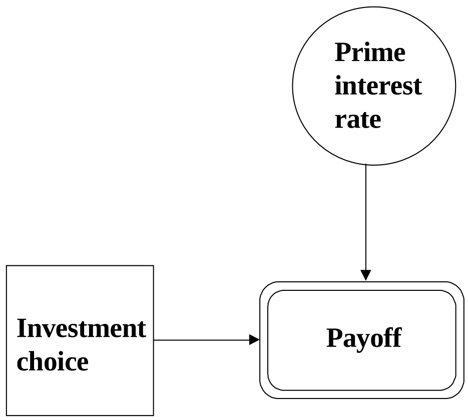
If an arrow points from a decision node to a chance node, then the specific chosen action (decision alternative) is relevant for assessing the chances associated with the succeeding uncertain event.

Dashed arrows point to decision nodes and indicate that the decision is made knowing the outcome of the predecessor node. A dashed arrow from a chance node to a decision means that the outcome of the chance event is known when the decision is made. This is the specific information available to the decision maker.

An arrow from one decision node to another means that the first decision is made before the second. The sequential ordering of decisions is shown in influence diagram by the path through the decision nodes.

Influence diagrams that are correctly constructed have no cycles; regardless of the starting point, there is no path following the arrows that leads back to the starting point.

We will illustrate the construction of influence diagram. The simplest decision problem is one in which there is a single decision to make, one uncertain event, and one outcome that is affected by both the decision and the uncertain event. This is the case of our example. In the Figure 3 there is the influence diagram of investor decision problem.



States of nature (probabilities):

1. θ_1 – – 2 % (0,1)
2. θ_2 – 0 % (0,6)
3. θ_3 – + 2 % (0,3)

Actions:

1. a_1 – 100 % in BA
2. a_2 – 50 % in BA and 50 % in municipal bonds
3. a_3 – 100 % in municipal bonds

Action:	**State of nature:**	**Payoff:**
a_1	θ_1	68
	θ_2	68
	θ_3	68
a_2	θ_1	124
	θ_2	84
	θ_3	51
a_3	θ_1	180
	θ_2	100
	θ_3	34

Figure 3 Influence Diagram of Investor Decision Problem

4. DECISION TREES

Influence diagrams are good for displaying a decision` s structure, but they hide many of the details. To reveal more details, a decision tree, another decision-modeling approach, is used.

As in influence diagram, squares represent decisions to be made, and circles represent chance events.

The branches emanating from a square correspond to the actions available to the decision maker.

The branches from a circle represent the possible outcomes – states of nature of a chance event (with the probabilities of the prior probability distribution in the brackets). The payoff is specified at the end of the branches. In the Figure 4 is the decision tree of the investor decision problem.

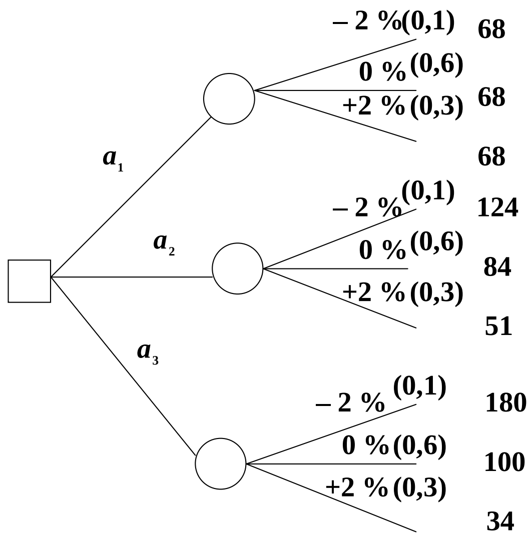


Figure 4. Decision Tree of the Investor Decision Problem

As you see, the tree flows from left to right, and the decision is represented by the square on the left side. The three branches represent the actions.

The branches from the chance nodes represent the outcomes – states of nature.

The interpretation of decision trees requires explanation.

First, the options represented by branches from a decision node must be such that the decision maker can choose only one option.

Second, each chance node must have branches that correspond to a set of mutually exclusive and collectively exhaustive outcomes. Mutually exclusive means that only one of the outcomes can happen and collectively exhaustive means that no other possibilities exist; one of the specified outcomes must occur.

Third, a decision tree represents all possible paths that the decision maker can follow through time, including all possible decisions and outcomes of chance events.

Finally, it is useful to think of the nodes as occurring in a time sequence. Beginning on the left side of the tree, a decision typically happens first, and then it is followed by other decisions or chance events in chronological order.

5. THE CONDITIONAL BAYES DECISION PRINCIPLE

As mentioned before, we will be involved with decision making in the presence of uncertainty. A natural method of proceeding in the face of this uncertainty is to consider the expected loss of making a decision, and then choose an optimal decision with respect to this expected loss.

We will consider Bayesian expected loss (the non-Bayesian school of decision theory (frequentist or classical school) adopts a quite different expected loss – see for example in [1]).

If $\pi^*(\theta)$ is the believed probability distribution of θ at the time of decision making, the Bayesian expected loss of an action *a* is

$$\rho(\pi^*, a) = E^{\pi^*}[L(\theta, a)]$$

We use π^* rather then π, because π will usually refer to the initial prior distribution for θ, while π^* will typically be the final (posterior) distribution of θ after seeing the data. Likewise we can define Bayesian expected gain

$$v(\pi^*, a) = E^{\pi^*}[V(\theta, a)]$$

In [1] the Conditional Bayes Principle is defined as follows. Choose an action $a \in A$ which minimizes $\rho(\pi^*, a)$ (maximizes $v(\pi^*, a)$). Such an action will be called a Bayes action and will be denoted a^{π^*}.

The application of this principle is appropriate when the repeated decisions are made in the relatively similar conditions.

In decision theory there are also several possible frequentist decision principles. The three most important are Bayes risk principle, the minimax principle, and the invariance principle [1]. Their applications in the practice are not frequent.

6. SOLVING DECISION TREES AND INFLUENCE DIAGRAMS

Firstly we will describe the algorithm for solving decision trees. This algorithm is called folding back the tree (backward induction analysis) and it is well known also in our literature.

We start at the endpoints of the branches on the far right-hand side and move to the left, (1) calculating Bayesian expected gain (Bayesian expected loss) when we encounter a chance node, or (2) choosing the branch with the highest Bayesian expected gain (lowest Bayesian expected loss) when we encounter a decision node. In the Figure 5 you can see the solved decision tree of our investor decision problem.

We started from the right by calculation of Bayesian expected gains (using prior distribution $\pi(\theta)$) for all chance nodes (the nodes were evaluated)

$$v(\pi, a_1) = 68 \times 0{,}1 + 68 \times 0{,}6 + 68 \times 0{,}3 = 68$$

$$v(\pi, a_2) = 124 \times 0{,}1 + 84 \times 0{,}6 + 51 \times 0{,}3 = 78{,}1$$

$$v(\pi, a_3) = 180 \times 0{,}1 + 100 \times 0{,}6 + 34 \times 0{,}3 = 88{,}2$$

In step 2 we choose the branch with the highest Bayesian expected gain. It is the branch corresponding to action a_3. All others branches are pruned (two parallel vertical lines on the branch).

In our example $a^{\pi^*} = a_3$ is Bayes action.

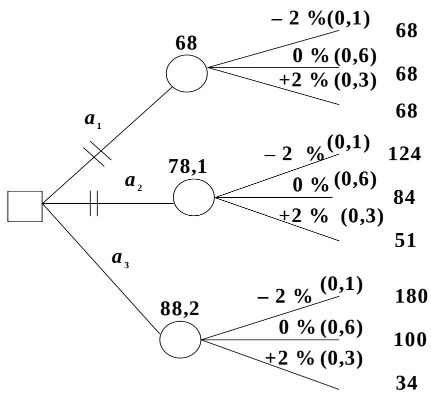


Figure 5 Solving Decision Tree of the Investor Decision Problem

We will describe an algorithm for solving influence diagrams (according to [4]).

1. Check to make sure the influence diagram has only one value node and that there are not cycles. If any nodes other then the value node have arrows into but not out of them, they can be eliminated.

2. Look for any chance nodes that

(a) directly precede the value node and

(b) do not directly precede any other node.

Any such chance node found should be reduced by calculating Bayesian expected gain. The value node inherits the predecessors of the reduced nodes. (That is any arrows that went into the node just reduced should be redrawn to go into the value node).

This step is just like calculating Bayesian expected gains for chance nodes at far right-hand side of a decision tree. We will see it in our example.

3. Look for a decision node that

(a) directly precedes the value node and

(b) has as predecessors all other direct predecessors of the value node.

If any such node was not found, go directly to step 5.

If such a decision node was found, it shall be reduced by choosing the Bayes action. When decision nodes are reduced, the value node does not inherit any new predecessors.

This step is like folding a decision tree back through a decision node at the far right-hand side of the tree. We will see it in our example.

4. Return to step 2 and continue until the influence diagram is completely solved (all nodes reduced).

5. One from arrows between chance nodes have to be reversed. This is the procedure that requires using of Bayes` theorem.

Finding an arrow to reverse is a delicate process. First, the finding of the correct chance node is necessary. The criteria are that

(a) it directly precedes the value node and

(b) it does not directly precede any decision node.

Call the selected node A. Look at the arrows out of the node A. Find an arrow from A to chance node B (call it $A \longrightarrow B$) such that there is no other way to get from A to B by following arrows. The arrow $A \longrightarrow B$ can be reversed using Bayes` theorem. Both nodes inherit each other `s direct predecessors and keep their own direct predecessors. After reversing an arrow, return to step 2 and continue until the influence diagram is solved.

Now we will solve the influence diagram of our investor decision problem.

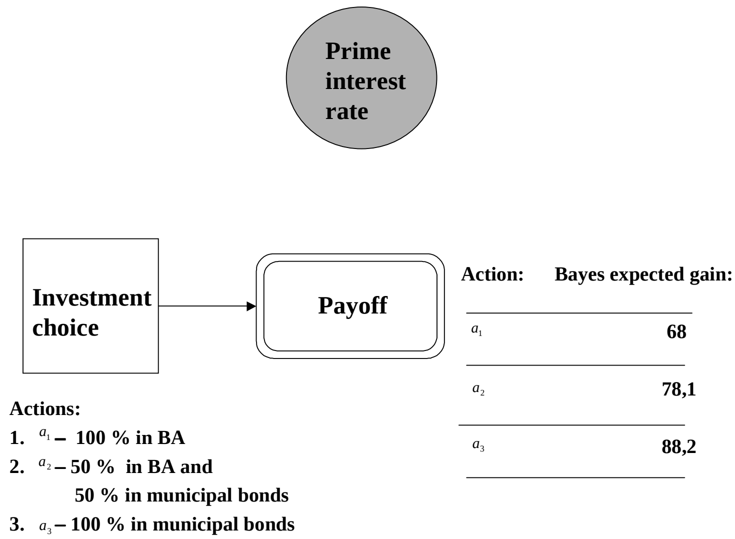


Figure 6 First Step in solving the Influence Diagram of Investor Decision Problem

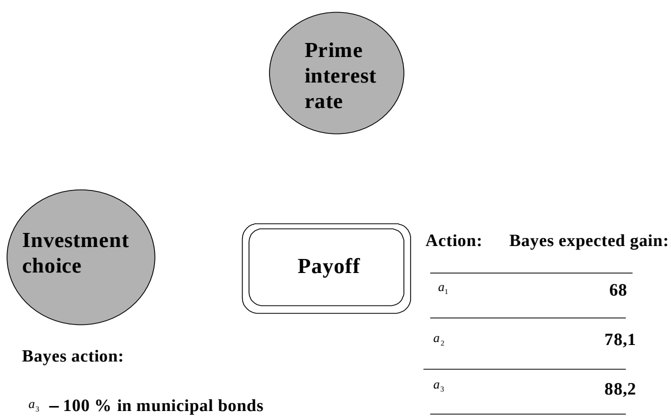


Figure 7 Second Step in solving the Influence Diagram of Investor Decision Problem

7. COMPARISON OF DECISION TREES AND INFLUENCE DIAGRAMS

The example has shown that, decision trees display considerably more information than the influence diagrams. It should also be obvious, however, that decision trees get “messy” much faster than do influence diagrams as decision problems become more complicated [5].

The level of complexity of the representation is very important. When it is necessary to present the results of a decision analysis to upper-level managers, their understanding of the graphical presentation is crucial. Influence diagram is superior in this regard; it is easy for people to understand.

Should we use decision tree or influence diagram? They complement each other very well. Influence diagram is particularly useful for the structuring phase of problem solving and form representing large problems. Decision tree is very useful for displaying the details of a problem.

It is clear that any properly built influence diagram can be converted into a decision tree, and vice versa.

One strategy of using them is to start by using an influence diagram to help understand the major elements of the problem and then convert to a decision-tree approach to fill in details [4], [5].

Influence diagrams and decision trees give two approaches to modeling a decision problem. Because the two approaches have different advantages, one may be more appropriate then the other, depending on the modeling conditions and requirements of the concrete decision situation.

For example, if it is important to communicate the overall structure of the model to other people, an influence diagram could be more appropriate.

Using both approaches together may prove useful; the goal, after all, is to make sure that the model accurately represents the decision situation [5].

In general there is no particular advantage in using a decision tree or an influence diagram when there is only a single decision to make, as in our example. In this situation a payoff table is fully sufficient. These tools are very useful when there is a sequence of decisions to make.

Generally, it is possible to realize also the sensitivity analysis that is obviously very useful and take in consideration more then one criteria (see for example in [5]).

The including of the risk attitude of the decision maker is also possible with aid of using the utility theory (see for example [4], [5], [13]).

REFERENCES

[1]. Berger, J. O.: *Statistical Decision Theory and Bayesian Analysis*. New York: Springer – Verlag 1993.
[2]. Bonitzer, J.: *Reflexions sur les Modeles Statistiques de Decision*. Revue de Statistique Appliquée, vol. XXXII, no 1, 1984.
[3]. Burns, W. J. – Clemen, R. T.: *Covariance Structure Models and Influence Diagrams*. Management Science, Vol. 39, No. 7, 1993.
[4]. Clemen, R. T.: Making Hard Decisions: An Introduction to Decision Analysis. Boston: PWS – Kent 1991.
[5]. Clemen, R. T. – Reilly, T.: *Making Hard Decisions with Decision Tools.* Pacific Grove: Duxbury Thomson Learning 2001.
[6]. Covaliu, Z. – Oliver, R. M.: Representation and Solution of Decision Problems Using Sequential Decision Diagrams. Management Science, Vol. 41, No. 12, 1995.
[7]. Fishburn, P. C.: *Foundations of Decision Analysis: Along the Way*. Management Science, Vol. 35, No. 4, 1989.
[8]. Howard. R. A.– Matheson, J. E.: *Influence Diagrams.* In: R. Howard and J. Matheson (eds.), The Principles and Applications of Decision Analysis, Vol. II, Polo Alto, CA: Strategic Decision Group 1984.
[9]. Howard, R. A.: *Decision Analysis: Practice and Promise*. Management Science, Vol. 34, No. 6, 1988.
[10]. Iman, R. L. – Conover, W. J.: Modern Business Statistics. New York: Wiley and Sons 1989.
[11]. Keeney, R. L. – Raiffa, H.: *Decisions with Multiple Objectives*. New York: J. Wiley and Sons, 1978.
[12]. Kirkwood, C. W.: An algebraic Approach to Formulating and Solving Large Models for Sequential Decision Under Uncertainty. Management Science, Vol. 39, No. 7, 1993.
[13]. Raiffa, H.: Decision Analysis, Introductory Lectures on Choices under Uncertainty. Massachusetts: Addison – Wesley 1970.
[14]. Sarin, R. – Wakker, P.: *Folding Back in Decision Tree Analysis*. Management Science, Vol. 40, No. 5, 1994.
[15]. Schachter, R. D.: *Automating Probabilistic Inference*. Technical Report, Stanford University, Stanford 1984.
[16]. Schachter, R. D.: *Evaluating Influence Diagrams*. Operations Research, Vol. 34, No. 6, 1986.
[17]. Schachter, R. D.: *Probabilistic Inference and Influence Diagrams*. Operations Research, Vol. 36, No. 4, 1988.
[18]. Schachter, R. D. – Kenley, C. R.: *Gaussian Influence Diagrams*. Management Science, Vol. 35, No. 5, 1989.
[19]. Smith, J. E. – Holtzman, S. – Matheson, J. E.: *Structuring Conditional Relationships in Influence Diagrams*. Operations Research, Vol. 41, No. 2, 1993.
[20]. Smith, J. Q.: *Influence diagrams for Bayesian decision analysis*. European Journal of Operational Research 40, 1989.
[21]. Terek, M.: Úvod do analýzy rozhodovania a bayesovskej indukcie. Bratislava: Ekonóm 2000.

Milan Terek,
Department of Statistics, Unoversity of Economics in Bratislava,
Dolnozemská cesta 1, 852 35 Bratislava, Slovak Republic
tel.: 02-67295713, e-mail: terek@dec.euba.sk

Received November 2004