Modeling Uncertain Situations in Decision-Making with Influence Diagrams

Igor Kovalenko[1[0000-0003-2655-6667]], Yevhen Davydenko[1[0000-0002-0547-3689]] and Alyona Shved[1[0000-0003-4372-7472]]

[1]Petro Mohyla Black Sea National University, Mykolaiv, Ukraine
{ihor.kovalenko,davydenko,avshved}@chmnu.edu.ua

Abstract. The modern methods of scenario analysis and forecasting, represented by tree-like graphs, are investigated in this paper. The basic theoretical positions of influence diagrams are considered. Its mathematical apparatus is based on Bayesian networks and supported a combination of graphical representation of analyzed process and its probabilistic behavior. The influence diagrams are an effective tool for modeling uncertain events in a complex decision-making process, which reflect decision-making situations in a more compact form. Influence diagrams also reflect all possible factors of the problem and their relationship in a clear and understandable way; provide all relevant information and implementation of the most appropriate decision-making procedure; allow to perform real-time analysis of possible scenarios. The numerical calculations that illustrate the possibilities of practical application of influence diagrams in complex decision-making problems under uncertainty are provided. The results obtained are aimed at improving the quality and effectiveness of decision-making in scenario analysis process with a large number of interrelated factors.

Keywords: Decision-Making, Influence Diagram, Scenario analysis, Uncertainty.

1 Introduction

The initial situation in decision-making tasks is often characterized by complexity and uncertainty, and various combinations of uncertain events can influence the outcomes of decisions made. These events are known as scenarios [1]. Therefore, there was a need to construct models that should not only reflect real situations, but also be used to formally solve probabilistic inference problems. Which here means calculating the probabilities of scenarios or events based on known probabilities of other related events implementation [1, 2, 4, 10].

The probability trees and decision trees are some of the most widely known methods for modeling uncertain situations, which are successfully used for a small number of uncertain events ($n \leq 4$-5). With the increasing number of uncertain events, probability trees are characterized by a large dimension, which makes it extremely difficult to display the information that they contain (a set of alternative solutions, systems of uncertain events and outcomes of alternative solutions) [4].

In this regard, a new “tool” called the “influence diagram” (ID) appeared to model uncertain events [6-9]. The advantage of IDs in relation to decision trees is especially noticeable for complex decision-making problems, when they are caused by a large number of interrelated factors. In addition, the IDs allow reflecting decision-making situations in a more compact form [1]. At the same time, it should be noted that IDs are currently known only to a fairly narrow circle of specialists and have not received such widespread application as decision trees.

The purpose of the paper is to present the basic theoretical principles of the influence diagrams and propose an example of their practical application in scenarios analysis tasks with a large number of interrelated factors.

2 The Influence Diagram Basics

Let us consider the basics of IDs that reveal the essence and features of them in accordance with the works [1, 4]. The ID is an oriented acyclic graph, the vertices (nodes) of which reflect the set (system) of factors: decisions, events and outcomes. All relevant information is not as closely related to the structure of the ID as is the case for decision trees.

The following types of nodes are distinguished on any ID. The decision node displays a variety of alternative solutions, which are graphically depicted in the form of rectangles or squares on the ID graph. An event node displays a set (complete system) of uncertain events that directly or indirectly determine the outcome of alternative decisions. The event nodes are depicted in the form of circles or ovals in ID graph. The chance node displays a function that evaluates the outcomes of alternative solutions. The chance nodes shown as a diamond in ID graph.

For these types of nodes, the following notation is accepted:

1. *D* used for decision nodes;
2. *C* used for chance nodes;
3. *V* used for value nodes.

An influence diagram, which is presented in the form of a graph and does not contain any other information, is called partially defined. To perform a decision analysis, each node in the graph must be provided with all relevant information. Lists of alternative solutions are associated with each decision node. A list of events and unconditional or conditional probabilities of their implementation are associated with each chance node. A value node is associated with a list of outcomes of all alternative solutions, the probability of their implementation and the value of the assessed function. ID with all information related to the analysis is called completely defined.

Let us consider some general approaches to ID construction [1, 4]. First of all, it is an approach for purposefully generating of IDs, the main ideas of which are that by a preliminary analysis of situations the number of decision-making acts and the sequence of their generation for general problem solving are determined. This determines the number of decision nodes on the ID. Then the way of assessment of outcomes of the alternative actions is determined. The value node and associated with it the assessed function is determined in this way. Based on this information, the so-called minimal influence diagram is constructed, which includes only the decision node(s) and the value node. Fig. 1 [1] presents the minimal ID.

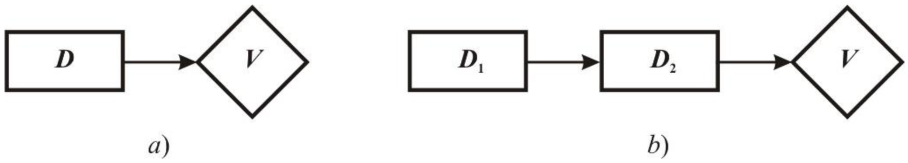

Fig. 1. Minimal influence diagrams: *a*) with one decision node; *b*) with two decision nodes.

Next, by further analysis of the situation, those systems of uncertain events that directly determine the outcome of alternative solutions are determined. The nodes displaying these groups of events are plotted on the ID graph, and the arcs from them are sent to the value node. Fig. 2 [1] presents the minimal ID supplemented by the event nodes C_1 and C_2.

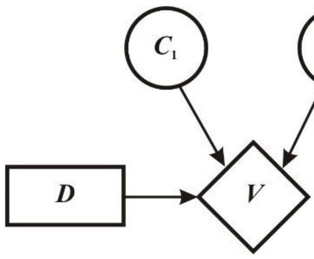

Fig. 2. Minimal influence diagram with two chance nodes C_1 and C_2.

Next, the situation is analyzed to determine whether there are groups of undefined events that affect events in nodes C_1 and C_2.

Suppose the existence of two such systems of uncertain events C_3 and C_4 was established (Fig. 3 [4]). Events of system C_3 can affect the probabilities of events at nodes C_1 and C_2, and events of system C_4 can affect the probabilities of events at node C_1. The nodes corresponding to these groups of undefined events are plotted on the ID and joined by arcs to existing nodes. The expanded ID is shown in Fig. 3 [4].

The analysis continues until all the vague (uncertain) factors that can influence the outcome of the alternative actions are established. The process of analyzing the task and constructing the influence diagram ends when all factors are taken into account. After building the ID, the necessary calculations are performed. The calculation of probabilities of the outcomes of alternative decisions is the most difficult task in ID construction. The outcomes of alternative decisions are determined by combinations of uncertain events in event nodes that are direct predecessors of the value node.

Let us consider the ID construction algorithm, which is based on a sequential calculation of the total probability of events in nodes over the entire set of events in nodes that are direct predecessors of these nodes. Fig. 4 [1] shows the ID consisting of the conditional chance nodes related to the outcomes of alternative decisions.

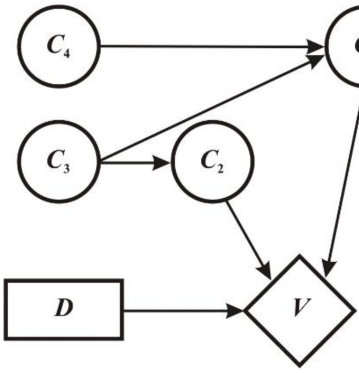

Fig. 3. The ID with additional chance nodes C_3 and C_4.

Since decision nodes are not related to outcomes probabilities, they are not shown in Fig. 4.

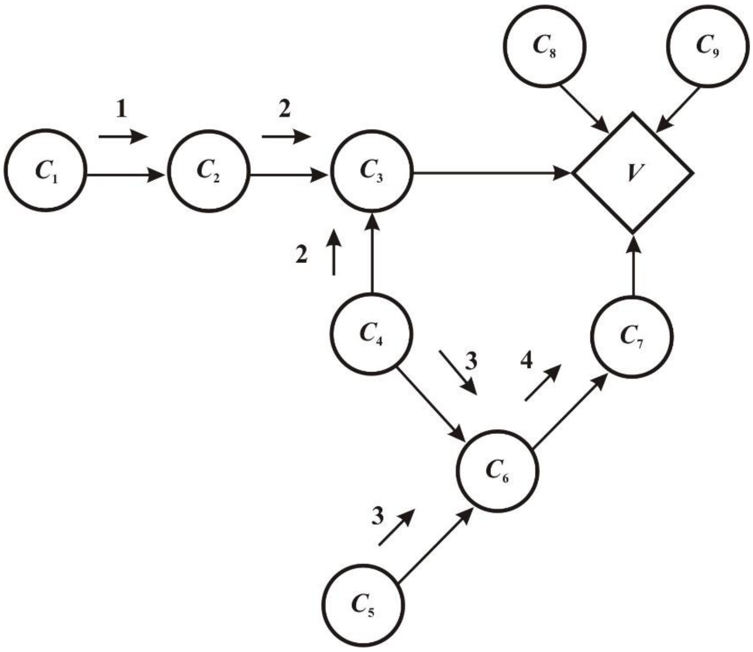

Fig. 4. Fragment of a conditional influence diagram.

The following standard types of connection between chance nodes are distinguished in the ID:

1. linear connection (nodes C_1, C_2);
2. divergent connection (node C_4 and nodes C_3 and C_6);
3. convergent connection (nodes C_2, C_4 and node C_3; nodes C_4, C_5 and node C_6).

Let us consider the procedure for calculation of total probability for various types of connection between nodes [1, 4].

1. Linear connection (Fig. 5 [4]). Let n events C_{pi}, $i=\overline{1,n}$ be connected with the node C_p; m events C_{qj}, $j=\overline{1,m}$ be connected with the node C_q and k events C_{rl}, $l=\overline{1,k}$ be connected with the node C_r. The prior probability $p(C_{pi})$, $i=\overline{1,n}$, of events in the node C_p, the conditional probability $p(C_{qj}|C_{pi})$, $i=\overline{1,n}$; $j=\overline{1,m}$ of events in the node C_q and the conditional probability $p(C_{rl}|C_{qj})$, $j=\overline{1,m}$; $l=\overline{1,k}$ of events in the node C_r are given. The total probability of events C_{qj}, $j=\overline{1,m}$ for the entire set of events $i=\overline{1,n}$ is calculated as follows:

$$p(C_{qj})=\sum_{i=1}^{n}p(C_{pi})\cdot p(C_{qj}\mid C_{pi}),\ \ j=\overline{1,m}. \qquad (1)$$

If necessary, the calculated values of $p(C_{qj})$ are normalized so that the sum of the probability values is equal to 1.

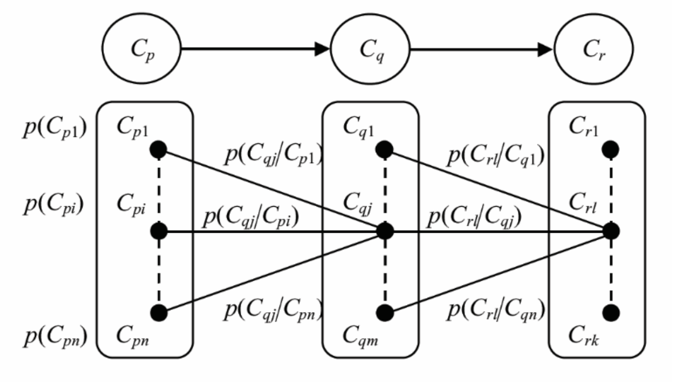

Fig. 5. The scheme of total probability calculation for a linear connection between chance nodes.

Next the total probability values is taken as a priori probability and the total probability of the events in the node C_r is calculated as follows:

$$p(C_{rl})=\sum_{j=1}^{m}p(C_{qj})\cdot p(C_{rl}\mid C_{qj}). \qquad (2)$$

The calculated values $p(C_{rl})$ are taken as the a priori probability of the events in node C_r and, if necessary, are used in further calculations.

2. Divergent connection (Fig. 6 [4]). According to Fig. 6, the nodes C_q and C_r are the proper descendants of a node C_p. The prior probability $p(C_{pi})$, $i=\overline{1,n}$, of events in the node C_p, the conditional probability $p(C_{qj}|C_{pi})$, $i=\overline{1,n}$; $j=\overline{1,m}$ of events in the

node C_q and the conditional probability $p(C_{rl}|C_{pi})$, $i=\overline{1,n}$; $l=\overline{1,k}$ of events in the node C_r are given.

Since the events in nodes C_q and C_r are conditionally independent of each other, until a certain event occurs in node C_p (here we are talking about a priori event probabilities), the total event probability in nodes C_q and C_r is calculated as follows:

$$p(C_{rl}) = \sum_{j=1}^{m} p(C_{pi})\cdot p(C_{rl} \mid C_{pi}), \quad l=\overline{1,k}; \tag{3}$$

$$p(C_{qj}) = \sum_{j=1}^{m} p(C_{pi})\cdot p(C_{qj} \mid C_{pi}), \quad j=\overline{1,m}. \tag{4}$$

In the case when the node C_p has more than two proper descendants, the total probability calculation is carried out in exactly the same way.

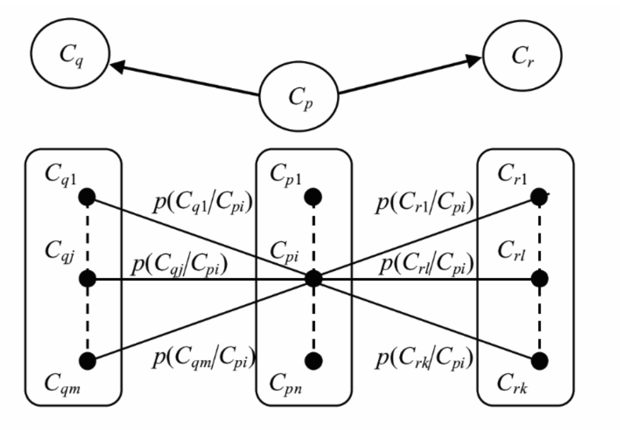

Fig. 6. The scheme of total probability calculation for a divergent connection between chance nodes.

3. Convergent connection (Fig. 7 [4]). According to Fig. 7 the nodes C_q and C_r are the proper ancestors of a node C_p. The prior probability $p(C_{qj})$, $j=\overline{1,m}$ of events in the node C_q, the prior probability $p(C_{rl})$, $l=\overline{1,k}$ of events in the node C_r, and the conditional probability $p(C_{pi}|C_{qj}, C_{rl})$, $i=\overline{1,n}$; $j=\overline{1,m}$; $l=\overline{1,k}$ of events in the node C_p are given for all possible combinations of events in nodes C_q and C_r (Fig. 7).

The total event probability in nodes C_p is calculated as follows:

$$p(C_{pi}) = \sum_{j=1}^{m}\sum_{l=1}^{k} p(C_{qj})\cdot p(C_{rl})\cdot p(C_{pi} \mid C_{qj}, C_{rl}), \quad i=\overline{1,n}. \tag{5}$$

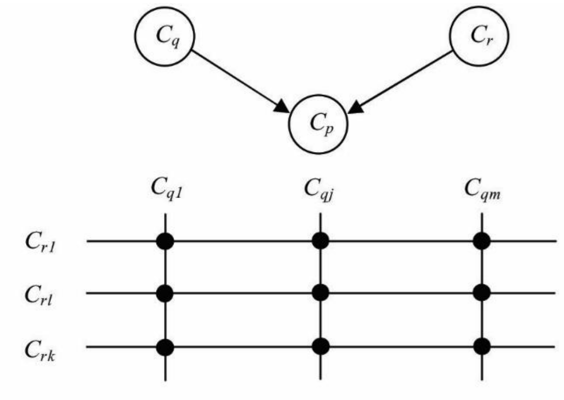

Fig. 7. The scheme of total probability calculation for a convergent connection between chance nodes.

Such calculations can be performed for the entire ID, starting with the initial nodes and ending with nodes that are proper ancestors of the chance node. The sequence of procedures for the total probability calculation for the nodes of the diagram is determined by the structure of the relations (connections) between the nodes of the events.

So, for the ID fragment showed in Fig. 4, the first procedure is a calculation of the total probability of events in node C_2 for the entire set of events in node C_1. There is a linear type of connection between nodes. Next, the total probability of events in node C_3 can be calculated over the entire set of events in nodes C_2 and C_4, since there is a converging type of connection. The next procedure is a calculation of the total probability of events in node C_6 for the entire set of combinations of events in nodes C_4 and C_5. The calculated total probability of events in node C_6 are used as a priori probability for calculating the total probability of events in node C_7. The sequence of procedures is shown by arrows and numbers in Fig. 4. At the same time, the calculation procedures order is generally arbitrary. So, the procedure 3 could be performed first, then procedure 4, and only after this the procedures 1 and 2. The following calculation procedures order is also acceptable: 1–3–4–2 or 1–3–2–4. Thus, the execution order for these calculation procedures is determined only by the structure of the influence diagram and the convenience and simplicity of calculations

3 Application of Influence Diagrams for Decision Support

Let us consider an example of a scenario analysis of the ship construction costs coordination in the pre-contract stage using an influence diagram.

Suppose it is planning to purchase a combined tanker with the following characteristics: the deadweight (DWT) is 40 000 tons; the cost of one DWT ton is 1250 cu; the vessel contract price is 50.0 million cu. The initial information for the analysis is the cost structure, presented in Table 1, based on which the wholesale price was 50.4 million cu. In order for the contract price to be 50.0 million cu, it is necessary to establish, with VAT (20%), the required value of the wholesale price equal to 50.0 / 1.2 = 41.7

million cu. To reduce the wholesale price by 50.4 – 41.7 = 8.7 million cu, it is necessary to reduce the cost of a number of articles costing.

Table 1. Ship Construction Cost Structure

Designation	Name	Cost (million cu)
S_1	Materials	10.0
S_2	Ship equipment	14.4
S_3	Counterparty service	1.5
S_4	Rigging	1.0
S_5	Labor input	1.5
S_6	Factory-wide overhead rate	3.0
S_7	Factory workshop-wide overhead rate	1.9
S_8	Salary	0.7
S_9	Additional salary	0.3
S_{10}	Transport and procurement service	1.1
S_{11}	Equipment maintenance and operating costs	6.6
S_{12}	Cost price	42.0
S_{13}	Profit (20%)	8.4
S_{14}	Wholesale price	50.4

It can be done by reducing the cost of two groups of indicators: articles costing that depend on market conditions (materials, ship equipment, counterparty service) and articles costing that can be managed by a shipbuilding company (factory-wide and factory workshop-wide overhead rates, equipment maintenance and operating costs). With all that said, let us construct an influence diagram for the task of ship construction cost reduction (Fig. 8 [3]).

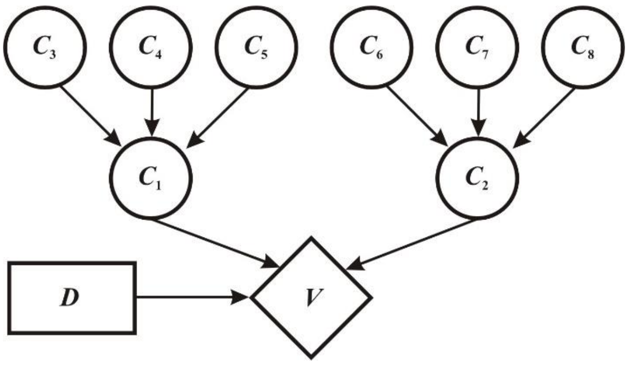

Fig. 8. Influence diagram for the task of ship construction cost reduction.

ID nodes carry the following information:

$$\text{Decision node } D \quad \begin{cases} \text{to enter into a contract for the purchase of a ship;} \\ \text{not enter into a contract for the purchase of a ship.} \end{cases}$$

Chance nodes are:

C_1 characterizes the probability of cost reduction of indicators that depends on market conditions.

C_2 characterizes the probability of reducing costs which are managed by a shipbuilding company.

C_3, C_4, C_5 characterize the probability of cost reduction of materials, ship equipment and counterparty service, respectively.

C_6, C_7, C_8 characterize the probability of reducing costs associated with factory-wide and factory workshop-wide overhead rates, equipment maintenance and operating costs.

Node V is represented by the following evaluation expression:

$$V = V_1+V_2 = p(C_1) \cdot (S_1+S_2+S_3)+p(C_2) \cdot (S_6+S_7+S_{11}),$$

where p is the probability of the implementation of relevant events C.

Let us make the necessary calculations on the influence diagram and set next expert estimates of unconditional and conditional probabilities of all events:

$p(C_3) = 0.2;$ $p(C_6) = 0.3;$

$p(C_4) = 0.1;$ $p(C_7) = 0.25;$

$p(C_5) = 0.3;$ $p(C_8) = 0.4;$

$p(C_1|C_3,C_4,C_5) = 0.3;$ $p(C_2|C_6,C_7,C_8) = 0.4.$

Since events represent a converging type of connection, based on (2), the total probabilities of events C_1 and C_2 been calculated.

So, we have:

$$p(C_1) = p(C_3) \cdot p(C_4) \cdot p(C_5) \cdot p(C_1| C_3, C_4, C_5) = 0.002;$$

$$p(C_2) = p(C_6) \cdot p(C_7) \cdot p(C_8) \cdot p(C_2| C_6, C_7, C_8) = 0.012.$$

Next, taking into account the numerical values presented in Table 1 ($S_1 = 10.0$; $S_2 = 14.4$; $S_3 = 1.5$; $S_6 = 3.0$; $S_7 = 1.9$; $S_{11} = 6.6$), the value node V value has been calculated: $V = (10.0 + 14.4 + 1.5) \cdot 0.002 + (3.0 + 1.9 + 6.6) \cdot 0.012 = 0.1898$ million cu.

Thus, the given probability of cost reduction for the planned costing items make it possible to expect a price reduction of only 0.188 million cu. Therefore, the decision to purchase of a ship is not made.

4 Conclusions

The ID satisfies a number of practical requirements: in a clear and understandable form it reflects all the many factors of the problem and their relationship; provides all relevant information; allows the implementation of the most appropriate decision-making procedure. However, the most effective may be the complex application of ID and decision tree. Any decision-making situation at the initial stage can be modeled by an influence diagram. For a better understanding of the situation, certain ID fragments can be transformed into a decision tree, which gives a more detailed presentation of the problem under consideration.

Acknowledgment. This research was partially supported by the state research project "Development of information and communication decision support technologies for strategic decision-making with multiple criteria and uncertainty for military-civilian use" (research project no. 0117U007144, financed by the Government of Ukraine).

References

1. Borisov, A. N., Uzga-Rebrovs, O. I., Savenko, K. I.: Probabilistic inference in intelligent systems [in Russian]. Riga (2002)
2. Kovalenko, I., Davydenko, Ye., Shved, A.: Development of the procedure for integrated application of scenario prediction methods. Eastern-European Journal of Enterprise Technologies, 2/4(98), 31–37 (2019). doi: 10.15587/1729-4061.2019.163871
3. Kovalenko, I. I., Chernov, S. K., Shved, A. V. at el.: System analysis methods in marine cluster tasks [in Russian]. Novoe slovo, Kharkov (2017)
4. Uzga–Rebrovs, O.: Modern concepts and applications of probability theory [in Russian]. RA Izdevnieciba, Rezekne (2004)
5. Cooper, G. F.: A method for using belief networks as influence diagrams. In R. D. Schachter, T. S. Levitt, L. N. Kanal, and J. F. Lemmer (eds), Proceedings of the 4th Workshop on Uncertainty in Artificial Intelligence, pp. 55–63. Elsevier Science Publishers, New York (1988)
6. Howard, R. D., Matheson, J. E.: Influence Diagrams. Decision Analysis, 2(3), 127–143 (2005). doi: 10.1287/deca.1050.0020
7. Pearl, J.: Probabilistic reasoning in intelligent systems. 1st Edition. Networks of plausible inference. Morgan Kaufmann Publishers, San Mateo, California (2014)
8. Schachter, R. D.: Evaluation influence diagrams. Operation research, 34(6), 871–882 (1986)
9. Sweatman, W., Wake. G., Cooper, H.: Using influence diagrams as a tool for decision making. ANZIAM Journal, 52, M147–M170 (2010). doi: 10.21914/anziamj.v52i0.3572
10. Walley, P.: Statistical reasoning with imprecise probabilities. Chapman and Hall, London (1990)