ECONTECHMOD. AN INTERNATIONAL QUARTERLY JOURNAL – 2013, Vol. 2, No. 1, 31–37

Design of intelligent decision support systems using ontological approach

V. Lytvyn

Department of Information Systems and Networks, Lviv Polytechnic National University,
e-mail: vasyll@ukr.net

Received July 06 2012: accepted July 25.2012

Abstract. Methods of developing and functioning of intelligent decision support systems based on precedents applying adaptive ontology that are part of intelligent agents are analyzed. Method of distance definition between precedent and current situation based on adaptive ontology is developed. Using mathematical tools of the dynamic programming for modelling of intelligent system functioning is considered. Simplifying the task model is proposed to weigh signs of ontology concepts. Examples of such problems for six processes concerning metal structures protection and maximizing their lifetime are presented.

Key words: intelligent decision support systems, adaptive ontology, intelligent agents, precedents, dynamic programming

INTRODUCTION

The technology of intelligent decision support systems (IDSS) is one of the most developed areas of artificial intelligence. Researches in this area are in the development of automated information systems used in the areas of human activities that require logical reasoning, high skills and experience.

Present level of development of IDSS is two-way development of intelligent agents (IA) [1, 2]:

– IA, based on Case-Based Reasoning, or CBR,
– Agency planning activities (searching the state space).

The choice depends on the IA task. Output methods by precedents are effective when the main source of knowledge about the problem is experience, not theory, solutions are not unique to a particular situation, and can be used in other cases, the purpose of solving the problem is not guaranteed to get the correct solution, and best possible. Output is based on precedents, a method for building IDSS decision on this issue or situation for the consequences of finding analogies that are stored in the precedent. This precedent is called relevant. From a mathematical point of view among elements of the set precedents Pr = $\mathrm{Pr}=\{\mathrm{Pr}_1, \mathrm{Pr}_2,\ldots, \mathrm{Pr}_N\}$, Pr_k is a relevant precedent for which the distance to the current situation S is the smallest, i.e. :

$$\mathrm{Pr}_k = \arg\min_i d(\mathrm{Pr}_i, S).$$

IA planning activities should achieve the target state. First he must build a plan to achieve this state with all possible alternatives. The planning process is based on decomposition. The task of planning ZP contains of 3 components: the set of states "*St*", a set of actions "*A*", a set of target states "*Goal*" (state purpose), i.e.:

$$\mathrm{ZP} = \langle St, A, Goal \rangle.$$

As can be seen for both classes of IDSS is required metrics. In the first case - for assess the relevance of precedents, in the second case - to assess the relevance of states. On way of determining this metric is directly dependent performance of IA. In our view this way should be based on clear and reasonable standard knowledge base. In the field of knowledge engineering, ontology has become such a standard [3]. Therefore, we proposed to build metrics using ontology.

In the ontology model *O* understand the triple form:

$$O = \langle C, R, F \rangle,$$

where: *C* - the notion, *R* - the ratio between the concepts, *F* - interpretation of the concepts and relationships (axioms). Axioms set limits of semantic concepts and relations.

THE MODEL OF ADAPTIVE ONTOLOGY

The effectiveness is considered of adaptation ontology knowledge basis to determine the features of the subject area incorporated in its structure elements and mechanisms to adapt by learning during the operation. One approach to implementing such mechanisms is an automatic weighing concepts Knowledge Base (KB) and semantic connections between them during learning. This takes on the role of factors important concepts and relationships [4-7]. Important factor concept (communication) - a numerical measure which characterizes the importance of certain concepts (communication) in a particular subject area and changes dynamically according to certain rules in service systems. Thus, we expand the notion of ontology by entering into a formal description of factors, important concepts and relations. Therefore, this ontology we define as the top five ones:

$$O = \langle C, R, F, W, L \rangle,$$

where: W - the importance of concepts C, L - the importance of relations R.

Ontology defined in this way will be called adaptive, i.e. one that adapts to domain by modifying the concepts and coefficients of importance of these concepts and relations between them. This ontology is unambiguously represented as a weighed conceptual graph (CG) [8, 9]. So we will build the metrics by using weighed CGs.

We propose to determine the distance between the annotations as a distance between the major concepts (weight center) of these annotations. The center of the CG weight is a concept, an average distance from which to other concepts is the least. Obviously, thus determined distance will depend on the way we determine the distance between two adjacent CG vertices. For this purpose it is proposed to determine the distance between adjacent vertices, as:

$$d_{ij} = \frac{Q}{L_{ij}\left(W_i + W_j\right)}, \tag{1}$$

where: W_i and W_j are the coefficients of importance of vertices C_i and C_j respectively; L_{ij} is coefficient of importance of the relation between vertices; Q is constant which depends on the particular ontology.

Let us assume that $L_{ij} = \infty$, then $d_{ij} = 0$.

Then lets us find the weight centers of the conceptual graph. The vertices for which the average distance is the smallest:

$$\overline{d}_{i^*} = \min_i \overline{d}_i. \tag{2}$$

The average distance $\overline{d}_i$ for the vertex C is calculated according to the formula:

$$\overline{d}_i = \frac{\sum_{j=1, j\neq i}^{n} d_{ij}^*}{n-1}, \tag{3}$$

where: n is a number of graph vertices; d_{ij}^* is the shortest path between vertices C_i and C_j is calculated using known algorithms, such as Bellman–Ford, Dijkstra, and Floyd–Warshall.

Then, according to its conceptual graphs, let us find the distance between the abstracts. Note that the proposed distance thus satisfies all three metric axioms.

Example. We will show the developed distance efficiency on the example of abstracts of scientific papers.

Let us consider three abstracts of papers.

1. The work is carried out in the direction of information technologies development focused on the natural language information processing. The author proposes a model for the problem: putting a language material in order with the help of the uniform standard, which is considered rather significant for the given class of technologies. It is one of the central problems concerning the development of the effective technologies for language information processing.

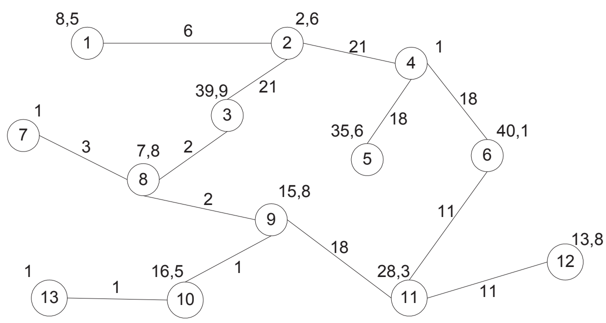
The corresponding conceptual graph of the first abstract is presented in Fig. 1.

2. The article presents the basic concepts on researching and solving the problem of automatic knowledge retrieval from text documents. Industrial system that solves the stated- above as well as one of its main application tasks are described.

The corresponding conceptual graph of the second abstract is presented in Fig. 2.

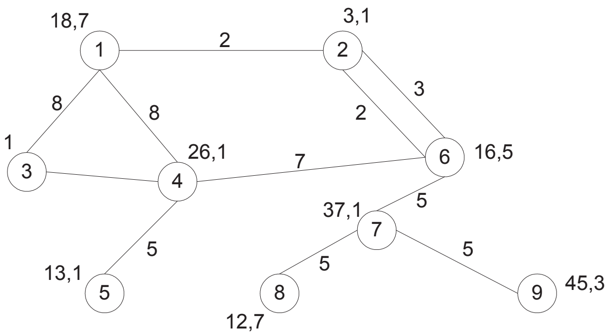
3. The paper deals with the problem of automated text consistency analysis. It is proposed to implement text consistency via text logic analysis with the involvement of the knowledge of application domain, natural language and normative base.

The corresponding conceptual graph of this abstract is presented in Fig. 3.



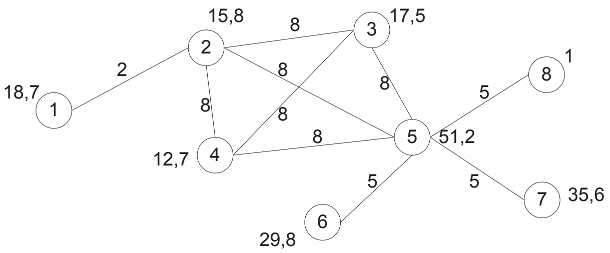
1 - work
2 - direction
3 - information technologies
4 - focus
5 - natural_language
6 - information
7 - author
8 - model
9 - problem
10 - material
11 - technologies
12 - language
13 - standard

Fig. 1. Conceptual graph of the first abstract



1 - article
2 - concepts
3 - application
4 - system
5 - indistrial
6 - task
7 - knowledge
8 - automatic
9 - text_document

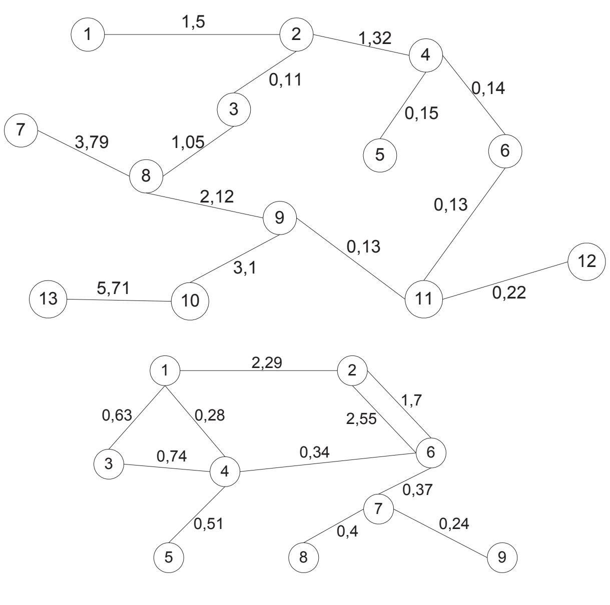
Fig. 2. The graph of the second abstract



1 - paper
2 - problem
3 - text
4 - automated
5 - knowledge
6 - domain
7 - natural_language
8 - normative_base

Fig. 3. The graph of the third abstract

The value of weights of concepts and relations is taken from the test ontology known as "computer sciences direction" for the subject area "artificial intelligence". Using formula (1) in which Q was accepted to be 100, we obtain the weighted graphs, shown in Fig. 4.



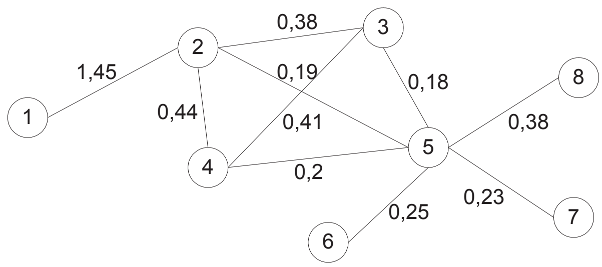


Fig. 4. The weighted graphs for three abstracts

Using Dijkstra's algorithm and making the appropriate calculations by formulas (2) and (3), we obtain that the weight centers of relevant graphs are:

$$\overline{d}^1 = \{11\} = \{\text{'technologies'}\},\ \overline{d}^2 = \{4\} = \{\text{'system'}\},$$
$$\overline{d}^3 = \{5\} = \{\text{'knowledge'}\}.$$

The 1st and the 2nd texts are easily linked using vertex 'natural_language' (the 5th in the 1st text and the 7th in the 3rd text). Then $d^1_{5,11} = 0,42$; $d^3_{5,7} = 0,23$. And the distance between the 1st and the 3rd texts equals 0.65.

$$\overline{d}^{13} = 0,42 + 0,23 = 0,65.$$

The 2nd and the 3rd texts are easily linked using vertex 'knowledge' (the 7th in the 2nd text and the 5th in the 3rd text). Then $d^2_{4,7} = 0,71$, which is the distance between the 2nd and 3rd texts, because 'knowledge' is the center of the 3rd weighted graph. So $\overline{d}^{23} = 0,71$.

Since in the first two texts there are no common vertices we use the 3rd text to calculate the distance between the 1st and the 2nd texts. Then:

$$\overline{d}^{12} = \overline{d}^{13} + \overline{d}^{23} = 0,65 + 0,71 = 1,36.$$

Thus, the 1st and 3rd texts are the closest in content and the 1st and 2nd texts are the farthest in content.

We proposed a method for calculating the weights of classes:

1. Total weight W^i_j of ontology class is equal to its own weight Wo^i_j, weight of sub-classes Ws^i_j and adjacent classes Wn^i_j (classes that are related with this class with no IS-A link):

$$W^i_j = Wo^i_j + Ws^i_j + Wn^i_j, \tag{4}$$

where: $Ws^i_j = \sum_k Wc^{i+1}_k \cdot L_{j,k}$ is weight k of sub-classes of the j–class of i–level, moreover, for a root class the i–level equals 0, $Wc^{i+1}_k = Wo^{i+1}_k + Ws^{i+1}_k$ is weight of class C^{i+1}_k; $L_{j,k}$ is weight of the relation between class C^i_j and C^{i+1}_k.

Calculation of certain components of the class total weight is shown in the diagram (Fig. 5).

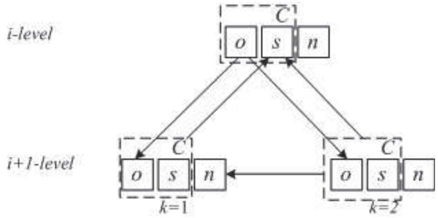


Fig. 5. Calculation scheme of certain components of the class total weight

2. At the moment of introduction of the new sub-class to (i+1)-level it gets its own weight Wo_j^{i+1}, which is equal to half of its own weight of the higher i-level class:

$$Wo_j^{i+1} = \frac{1}{2} Wo_j^i. \quad (5)$$

The weight of class Wc_j^i and all parent classes up to the root class increases by the value of the weight of newly created sub-class:

$$Wc_j^m = Wc_j^m + Wo_j^{i+1}, \forall m \le i. \quad (6)$$

3. When relation is established between the concepts k_1 and k_2 an edge appears between corresponding vertices of the graph ontology, and to the weight of adjacent classes Wn_1 the weight Wn_2 is added and vice versa – to Wn_2 the weight of new adjacent class Wn_1 is added, so:

$$Wn_j = \sum_k Wc_k \cdot L_{j,k}. \quad (7)$$

Re-establishing of relations leads to the appearance of multiple edges in the graph.

4. Multiplicity of edges represents the frequency of appearance of a V pair of semantically related concepts $L_{i+1} = V \cdot L_i$. After recalculation multiple edges do not increase the vertex valence.
5. Weight of ontology instance equals to the total weight of appropriate class.

So, the model of the ontology allows us to calculate the weight coefficients of their components in the process of their insertion, removal and use during system exploitation, thus realizing the mechanism of adaptation to the user domain [10].

FUNCTIONING OF INTELLIGENT AGENTS IN STATE SPACE ON BASED ADAPTIVE ONTOLOGY

Let the IA must decide solution for protection structures of metal from corrosion. Structures of metal is in some state, which called the initial state $St(0)$. This state is characterized by a set of attributes $X = \{x_1,\ldots,x_n\}$ (as example for pipeline it may be - lifetime, the diameter of the pipe, material for realizing processing, etc. The values which take these signs $z_i = z(x_i)$, $x_i \in X$ define the value function of utility $f = f(Z)$, $Z = (z_1,\ldots,z_n)$ this structures of metal. Obviously, that not all signs have the same effect on the value f. Therefore it is appropriate to consider a subset of attributes $X_W \subset X$ where W is the weight of importance of attributes (as example for pipeline it may be – lifetime). Obviously, that the values of attributes are interdependent. The goal of IA is processing structures of metal, that is put it in some state (call it the target state Goal), for getting maximum value f.

From now we will consider that the state of goal is single. If several states of goal are present, it means that goal can be written as a disjunction of these states. Achieving this state is the solution of some subtasks therefore assumption about singling of target state is normal.

Action A consists of four parts: the name of action, the list of parameters, preconditions and results. Plan is defined as a tuple of four elements – <set of actions, set of limitation of order , set of causal relations, the set of open preconditions> [2, 11, 12]. We propose to use graphs model of diagram of UML (Unified Modeling Language) [13] for врахування decomposition and / or dependence between states and transitions, for showing alternatives achieve the target states. Example of this oriented graph with target state *Goal* is showing in Fig. 6.

Moreover, the essence of rational functioning IA is transition to the final state *Goal* with minimum expenditure of resources G (resources can be specified in the form of cash equivalent, man-hours, time, etc.), namely the correct allocation of resources for action. For determination values of attributes depending on the performed actions we will use the ontology O of materials science [3]. The resulting task will be formally written in the form:

$$P:\ St(0) \xrightarrow{G,O} Goal. \quad (8)$$

The dynamic programming (DP) [2] is most often used in the role of mathematical tools, that is allowing to plan multistep controlled processes , and processes which are developing in time. We will show that our task of rational planning of IDS can be reduced to DP.

The effectiveness U of the whole process P can be presented as the sum of efficiencies $U_j\left(j=\overline{1,n}\right)$. of separate steps, that is: $U = \sum_{j=1}^{n} U_j$, entitled additive criterion. The adopting some decisions is relating with each stage (Step) of task, so-called stepper control $q_j\left(j=\overline{1,n}\right)$, that is passing as efficiency of this stage and the whole process. Solving the task of dynamic programming is lying in finding a management $Q = (q_1, q_2,\ldots,q_n)$ of the entire process, which is maximizing the overall effectiveness $\max U = \sum_{j=1}^{n} U_j$. The optimal solution of this task is management Q^* that is consisting of a set of optimal turn-based managements $Q^* = \left(q_1^*, q_2^*, \ldots, q_n^*\right)$ and allows for the achievement of maximal effectiveness: $U^* = \max_{q \in Q} \left\{U(q)\right\}$.

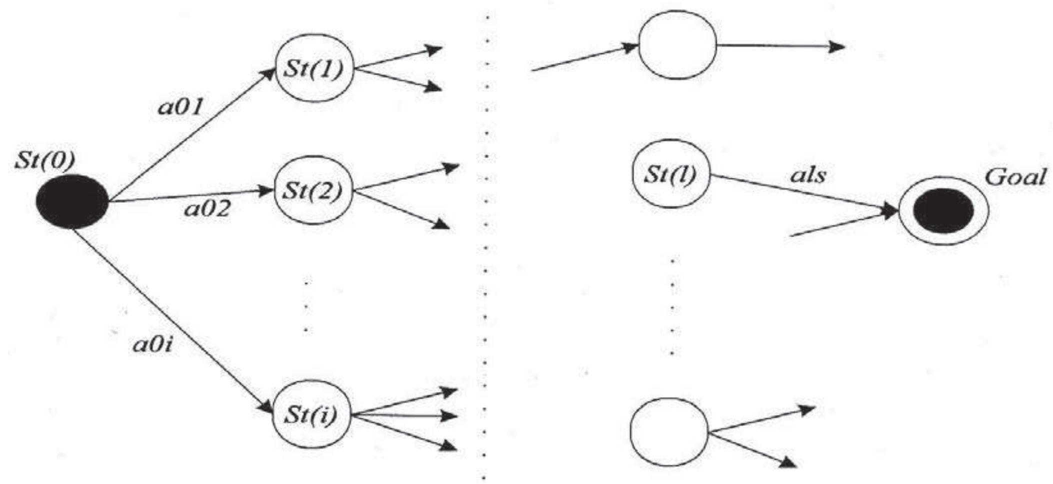


Fig. 6. Example of diagram states of UML modeling behavior of intelligent agent

IDS should be able to assess the conditions and actions for selecting the necessary actions. It is easier to make with states in which structures of metal already was presented. It is more difficult to estimate future states. The everistic functions or meta knowledge that are storing in the ontology of materials science are using for estimating. Therefore, first we will consider the estimation of pasted states, then action, and finally a combination of them, that is leading to a new state [14-15].

Let $v(St(i))$ is valuation of state $St(i)$. State of goal *Goal* is defining by necessity of some set of attributes X_W to reach some values of $z(x, Goal)$ $\forall x \in X_W$. Any state $St(i)$ is given by its set of attributes Y_i, which are taking the values $z(y, St(i))$ $\forall y \in Y_i$.

For evaluation state $St(i)$ set of attributes and their values of state $St(i)$ need to display ψ into set of attributes and values of state *Goal*. Clearly, this displaying is using a basis of knowledge of software, namely bonus pluging of ontology Semantic Web Rule Language:

$$\psi : Y_i \xrightarrow{O} X. \quad (9)$$

Then the estimation of state $v(St(i))$ is calculated:

$$v(St(i)) = d(St(i), Goal) = \\ = \sum_{x \in X_W} \varphi\big(z(\psi(y), St(i)), z(x, Goal)\big), \quad (10)$$

where: X_W is a set of attributes with the largest weights W in ontology, φ is metric (as example Euclidean, Manhattan, Lemmings, Zhuravleva, etc., the choice of which is depending on the type of attributes of the problem (quantitative, qualitative, mixed) [16, 17].

In our investigation for selecting the actions of IA we will rely on rational adopting decision by system, namely on the correct allocation of available resources G on actions for the achievement of maximum efficiency U. Therefore, we will assume that every action a_{ij} is uniquely determined by the expenditure of resources g_{ij}^k (cost of transition from state to state), where $k = 1,2,\ldots,n_i$, n_i is the quantity of alternatives α_k for doing transition a_{ij}. Therefore, further action will be marked by three indexes a_{ij}^k: transition from state $St(i)$ in state $St(j)$, with using an alternative α_k. For example, three alternatives: mechanical, chemical and thermal withdrawal can be used for removing the protective coating from the surface of the pipeline.

Each alternative is characterized by the cost of resources and term of the exploiting. Information about alternatives and the costs of resources, prize from transition of state (time of exploitation , etc.) is also stored in the ontology. Obviously, that new alternatives can be appearing, therefore it is necessary to monitor new information in scientific journals. The task of automatic replenishment of the ontology which is based on the analysis of scientific texts is describing in [5]. Effectiveness of action a_{ij}^k is the meaning of function which is dependent on an assessment of the transition to a new state $v(St(j))$, the cost of resources g_{ij}^k and the prize $f(Z_j)$:$U(a_{ij}^k) = \delta(St(j), g_{ij}^k, f(Z_j))$ where Z_j is the importance of attributes in the state $St(j)$. Then the action of IDS is determining the management Q. Thus, the task (9) is reduced to the task of dynamic programming.

We will consider the example of adopting decisions by IA for modernizing the pipeline (see Fig. 7). Initial state: unprocessed. The final state (state of goal): processed:

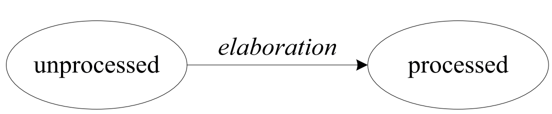


Fig. 7. The general task of modernization the pipeline

The task is divided into three subtasks (preparation, coating, protection), the first of which is divided into 4 subtasks (disclosure of tube surface, removal of protective coating, decreasing, priming) as showing in Fig. 8. The alternative solutions are used for solving of each subtasks. So for subtasks of removal of protective coating, one of three alternatives: mechanical, chemical or thermal can be used. All this information is stored in an appropriate ontology (ontology of PO modernizing the oil and gas

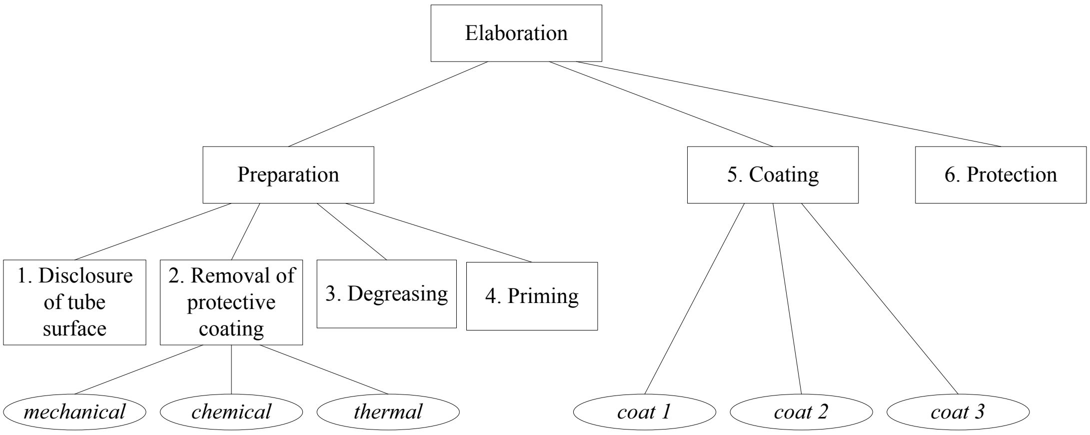


Fig. 8. Processing task decomposition “elaboration”

pipelines have been in the process of elaboration in the laboratory of system analysis of scientific and technical information of Physical-Mechanical Institute of Ukrainian NAN G.V. Karpenko) [18-20].

Thus in general it is necessary to successively solve the six subtasks $P_1, P_2, \ldots, P_6$. For each task it is necessary to choose the method of solution (alternatives) [21]. If G is the available resource, r_e is the desired lifetime of the pipeline, then rationality of adopting decisions will consist in:

$$\begin{cases} U = \sum_{i=0}^{N-1} U\left(a_{ij}^{k}\right) \rightarrow \max, \\ r \geq r_e, \\ \sum_{i=0}^{N-1} g_{ij}^{k} \leq G. \end{cases} \qquad (11)$$

Let the resource consist of 6 units: $G = 6$. Example of possible costs g and prizes U depending on the number of the process and alternatives are presented in Table 1.

Using the method of functional equations (2), which is designed for solving tasks of dynamic programming, we will obtain the optimal path, which is shown in Fig. 9.

Table 1. Table of costs and prizes

№ altern.	Process 1		Process 2		Process 3		Process 4		Process 5		Process 6	
	Cost	Prize	Cost	Prize	Cost	Prize	Cost	Prize	Cost	Prize	Cost	Prize
a_1	0	5	0	8	0	3	0	4	0	2	0	3
a_2	1	7	1	9	1	4	1	7	1	3	1	7
a_3	2	8	2	12	-	-	-	-	2	6	-	-

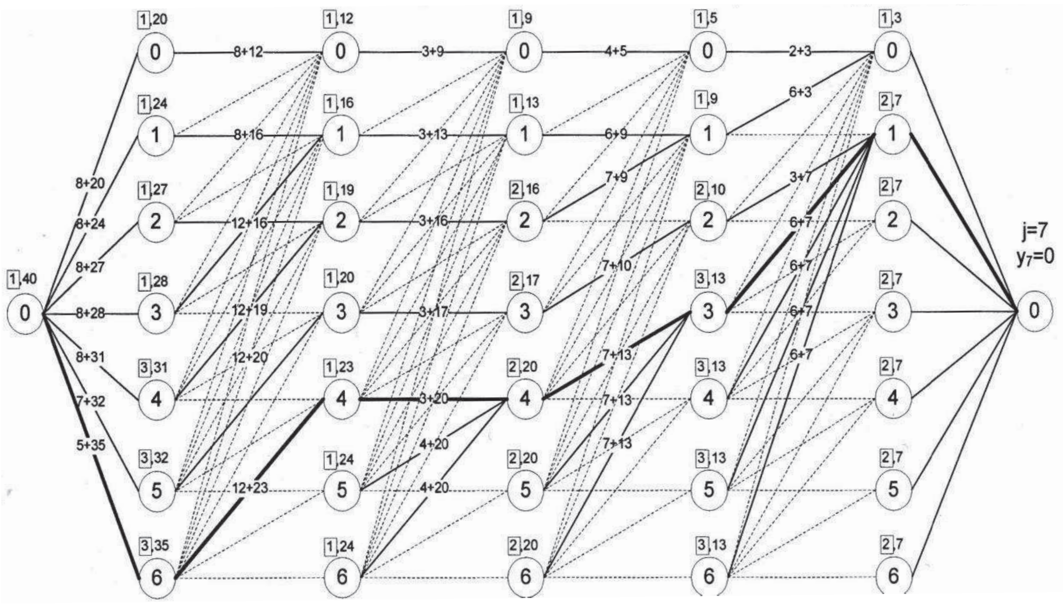


Fig. 9. The process of solving the task of dynamic programming

The optimal plan of distribution of resources among the processes of processing the pipeline is showing in Table 2. Revenue from the operation of the pipeline will be 40 units.

Table 2. Table of adopting by IDS decisions

Process	№ alternatives	Resources	Prize
Process 1	1	0	5
Process 2	3	2	12
Process 3	1	0	3
Process 4	2	1	7
Process 5	3	2	6
Process 6	2	1	7

CONCLUSIONS

The mathematical model of intelligent decision support systems, depends on the class of problems. All these models use metric for finding the relevant precedents or relevance of states. To construct such metrics ontology is used. For this purpose, generally to three elements of procession, which sets the ontology (set of concepts, relationships and their interpretation) we add two scalar values (the importance of concepts and relations) which are used to calculate the necessary distances.

Consider the use of mathematical tools of dynamic programming for solving the task of planning protection of metal from corrosion which is based on intellectual diagnostic systems. The core of knowledge basis is material ontology. For determining the effectiveness of intellectual diagnostic system's work we proposed to use rules that are set in ontology. Significance of characteristics from ontology of concepts for simplifying the model of tasks was set. The example of functioning of intellectual diagnostic system, which is based on the modernization of pipeline in order to maximize its lifetime is shown there.

Consider the four classes of problems. The list of tasks is included in these classes. The examples of the functioning of some intelligent systems is based on the developed models.

REFERENCES

1. **Aamodt A. and Plaza E. 1994.** Case-based reasoning: Foundational issues, methodological variations, and system approaches AI Communications, Vol. 7, Issue 1, 39-59.
2. **Russel S., Norvig P. 2003.** Artificial intelligence. A modern approach, Prentice-Hall, Upper Saddle River, N.J.
3. **Gruber T. R. 1993.** A translation approach to portable ontologies Knowledge Acquisition. No. 5 (2), 199-220.
4. **Dosyn D., Darevych R., Lytvyn V. and Dalyk U. 2006.** New knowledge evaluation using massage model of NLT document Proceedings of the International Conference on Computer Science and Information Technologies, 118-119.
5. **LytvynV., Shakhovska N., PasichnykV. and Dosyn D. 2012.** Searching the Relevant Precedents in Dataspaces Based on Adaptive Ontology. Computational Problems of Electrical Engineering. – Volume 2, Number 1. – Lviv, 75-81.
6. **Dosyn D. and Lytvyn V. 2012.** Planning of Intelligent Diagnostics Systems Based Domain Ontology The VIIIth International Conference Perspective Technologies and Methods in MEMS Design. - Polyana, Ukraine, 103.
7. **Lytvyn V., Dosyn D., Medykovskyj M. and Shakhovska N. 2011.** Intelligent agent on the basis of adaptive ontologies construction Signal Modelling Control. – Lodz.
8. **Montes-y-Gómez M., Gelbukh A. and López-López A. 2000.** Comparison of Conceptual Graphs Lecture Notes in Artificial Intelligence Vol. 1793. – Springer-Verlag: http://ccc.inaoep.mx/~mmontesg/publicaciones/ 2000/ ComparisonCG.
9. **Sowa J. 1976.** Conceptual graphs for a database interface IBM Journal of Research and Development. – Vol. 20. – № 4, 336–357.
10. **Knappe R., Bulskov H. and Andreasen T. 2004.** Perspectives on Ontology-based Querying International Journal of Intelligent Systems: http://akira.ruc.dk/~knappe/publications/ijis2004.pdf
11. **Frederick S. and Gerald J., 2005.** Introduction to Operations Research McGraw-Hill: Boston MA; 8th. (International) Edition.
12. **Callı A., Gottlob G., Pieris A. 2010.** Advanced processing for ontological queries Proc. of the 36th Intl Conf. on Very Large Databases, 554–565.
13. **Booch G., Jacobson I. and Rumbaugh J. 2000.** OMG Unified Modelling Language Specification.
14. **Gladun A. and Rogushina J. 2006.** Ontological Approach to Domain Knowledge Representation for Informational Retrieval in Multiagent Systems International Jornal "Information Theories & Applications". - V.13. - № 4, 354-362.
15. **Guarino N. 1995.** Formal Ontology, Conceptual Analysis and Knowledge Representation International Journal of Human-Computer Studies. – № 43(5-6), 625–640.
16. **Boris M., Motik B., Sattler U. and Studer R. 2005.** Query Answering for OWL-DL with Rules Journal of Web Semantics. – № 3(1), 41–60.
17. **Saleh M.E. 2011.** Semantic-Based Query in Relational Database Using Ontology Canadian Journal on Data, Information and Knowledge Engineering. – Vol. 2. – № 1, 1–16.
18. **Dosyn D., Lytvyn V. and Yatsenko A. 2012.** DP-optimization of steel corrosion protection techniques in the intelligent diagnostic system FHMM. - №9, 329-333.
19. **Shakhovska N., Lytvyn V. and Medykovskyj M. 2012.** Dataspace Class Algebraic System for Modelling Integrated Processes Journal of applied computer science vol. 20 no. 1, 69-80.
20. **Dosyn D., Darevych R. and Lytvyn V. 2004.** Modelling of the intelligent text recognition agents based on dynamic ontology Proceedings of the 4th International Conference „Internet – Education – Science – 2004". – Baku-Vinnytsia-Veliko Turnovo. – V. 2, 577–579.
21. **Kuzmin O. Ye., Melnyk O. H., Shpak N. O., Mukan O. V. 2012.** The concept of creation and use of the polycriterial diagnostics systems of enterprise activity, Econtechmod. An international quarterly journal on economics in technology, new technologies and modelling processes. Vol. I, No 4, 23-28.