KNOWLEDGE REPRESENTATION AND REASONING. MIVAR TECHNOLOGIES

O. O. Varlamov[1], I. A. Danilkin[2]
MIVAR LLC
Moscow, RUSSIA
e-mail: [1]o.varlamov@mivar.ru, [2]i.danilkin@mivar.ru

Abstract

The technology of knowledge representation based on mivar network and its matrix representation are considered. The mechanism that allows to design computational algorithms of solutions of assigned tasks on the basis of data from mivar networks is described.

1 Introduction

Development of artificial intelligence (AI) systems is an important and quite a topical task of today. Such system types as expert systems, text meaning understanding, image recognition, robotic systems aim to change the life of a modern man. However, it should be noted that currently available intelligent systems are designed to support regular user groups to solve highly specialized tasks. Thus, design of theory, methods and technologies of AI remains an urgent task for development of intelligent systems. Moreover, it becomes increasingly important [3].

As mentioned above, available intelligent systems aim to solve highly specialized tasks, since building and using knowledge bases that provide the foundation for such systems requires costly human and material resources. Thus, the developers of intelligent systems face a wide range of difficulties. Knowledge representation and search are the two fundamental problems that still occupy developers of intelligent systems.

Knowledge representation models should provide a simple mechanism of data description and development of the knowledge bases which are required to implement intelligent behavior of such systems. On the one hand, representation method should make the knowledge understandable for machine, on the other hand it should ensure easy description of knowledge structure for its developer. Therefore, in the process of development different knowledge representation models, two aims are pursued: expressiveness and efficiency. Moreover, such systems should ensure the most natural way of knowledge representation. Nowadays, there is a large number of approaches to data representation: predicate description, semantic networks, production rules, as well as neural-networking, evolutionary, agent-oriented and stochastic approaches to representation, and many others [2]. All these approaches aim to reach a fair compromise between efficiency and expressiveness of representation.

In this paper mivar-based technology of knowledge representation is considered. It is aimed to simplify knowledge acquisition and accumulation since there is no necessity in experts involvement and logical inference methods changing [1, 4]. More than that, mivar-based technology of data processing is proposed, which aims to increase the speed and quality of acquiring results. Since the problems of knowledge representation and search are interconnected, an intelligent system should not only be aware of the subject, but also be able to solve tasks set in the subject domain.

Mivar technologies were used in development of such AI systems as: text meaning understanding, image recognition, robotic systems and expert systems.

2 Mivar knowledge representation technologies

The concept of mivar network is one of the basic concepts of the proposed mivar-based approach to data representation. It is mivar network that ensures formalization and representation of human knowledge. Let us consider a subject domain M.

Mivar network - the method of representing objects of the subject domain and rules of their processing in the form of a bipartite directed graph consisting of objects (P) and rules (R). These objects and rules together form the model of the subject domain.

Mivar network has the following significant properties:

1. The network consists of the elements of two types (two partitions of the graph): the nodes – objects (P) and the arcs – rules (R).

2. For each variable all the information is stored in explicit form about all the rules R, for which it is an input variable X or an output variable Y with the indication of that;

3. For each rule R information about all its input and output variables P is stored in explicit form, including the information about the number of input (X) and output (Y) variables;

4. The storage of all the necessary information of such a network is organized on the basis of database technologies adapted to operate with mivar networks;

5. In each element of the mivar network, being node or arc, all the adjacent arcs and nodes are determined coherently and completely. Being in any place of mivar network, it is always obvious from where can we move to it and to where we can move from it, which eliminates brute forcing while searching for logical inference on the mivar network.

Bipartite graph of mivar network can be represented in the form of two-dimensional matrix $(P) \times (R)$, where n is the number of parameters (objects) of the subject domain, m is the number of rules connecting objects (see Table 1).

Mivar network is constructed by binding aggregates of different types according to the following: "object-rule" and "rule-object" . Interconnection such as "object-object" and "rule-rule" are forbidden. In general form the interconnection is as follows: "object(s)-rule-object(s)" . The element from which the interconnection goes is indicated first. The element to which the interconnection moves is indicated second. It allows us to determine direction of the graph and exclude possibilities of misinterpreting or converting objects through backtracking. Designed in such a way, mivar network

Table 1: Matrix representation of mivar network

Rules \ Parameters	1	2	3	4	5	...	$n-2$	$n-1$	n
1	**X**	**X**	**X**					**Y**	**Y**
2			**X**	**Y**	**Y**			**X**	**X**
...						...			
m		**X**		**X**	**X**		**Y**		

is scalable, as at any time it is possible to add elements of aggregates of any types available without the necessity to change processing methods. Moreover, describing mivar network does not require the involvement of an expert. In most cases it is sufficient to move objectively existing objects and connections (rules) in the mivar form.

As an example let us consider the subject domain "Geometry. Triangles" . Here as variables are any sides, angles, segments of the triangle. Rules are different interconnections between these variables such as definitions, theorems, axioms, etc. The part of mivar network of the considered subject domain is represented in the form of the matrix M (see Table 2).

Table 2: Part of the mivar network. Geometry

Rules \ Parameters	side AB	side BC	side AC	Perimeter	...	AB is .. greater than BC	AC is .. greater than BC
The perimeter of the triangle (using the lengths of three sides)	**X**	**X**	**X**				**Y**
The side BC (using the perimeter and the ratios of the sides)		**Y**		**X**		**X**	**X**
The side AC (using the ratio of the sides)		**X**	**Y**				**X**
The side AB (using the ratio of the sides)	**Y**	**X**				**X**	

3 Reasoning

On the basis of subject domain representation described above it is possible to design an algorithm allowing us to implement information search inside the mivar network, set open and hidden interconnections between data inside mivar network, construct computational algorithms of the set task in corresponding subject domain on the basis of data from mivar network. Such a methodology for searching logical inference path allows us to avoid brute forcing of all possible rules on each step. This algorithm includes forming aggregates of known parameters and setting one or more required parameters, then the processing for each known parameter (which was not processed before) is carried out to find required parameters. This processing involves the following stages:

– determines rules (which were not launched before) in which, firstly, known parameter serves as the input variable and secondly, all other variables are known;

– launch these found rules and add output variables of launched rules to the aggregate of known parameters; moreover, if all the required parameters are found, the processing is stopped;

– design the sequence of launched rules in the order of launching, thus, the designed sequence is logical inference path.

Let us illustrate the basic variant of the scheme described above using a simple example from the subject domain "Geometry. Triangles" . To do this are given steps of the solution of the following task from this subject domain: *It is needed to find the lengths of the sides AB and AC of the triangle ABC, if the perimeter of triangle is 28 cm, the side BC is 4 cm less than AB and 9 cm less than AC.* It should be noted that in the task described above input parameters (Z) are the perimeter of the triangle and the ratios between the sides of the triangle. It is required to find (W) – the sides AB and AC of the triangle ABC. This information can be presented in the matrix M – to do this let us add an additional service row and a service column to the matrix. The row is designed to track the changes in known data. The column is designed to track the rules used. The result of the first step of working with mivar matrix is represented in Table 3.

Table 3: The mivar matrix after first step.

Param. / Rules	side AB	side BC	side AC	Perimeter	...	AB is .. greater than BC	AC is .. greater than BC	service column
Perimeter of triangle (using lengths of three sides)	**X**	**X**	**X**				**Y**	
Side BC (using perimeter and ratios of sides)		**Y**		**X**		**X**	**X**	
Side AC (using ratio of sides)		**X**	**Y**				**X**	
Side AB (using ratio of sides)	**Y**	**X**				**X**		
Service row	**W**		**W**	**Z**		**Z**	**Z**	

According to the mechanism described above, on the second step the rule "The side BC using the perimeter and the ratios of the sides" can be launched, which is indicated in the corresponding cell of the service column. Output of the rule after launching indicated as known variable Z. The results of this step are represented in Table 4.

Table 4: The mivar matrix after second step.

Rules \ Param.	side AB	side BC	side AC	Perimeter	...	AB is .. greater than BC	AC is .. greater than BC	service column
Perimeter of triangle (using lengths of three sides)	**X**	**X**	**X**				**Y**	
Side BC (using perimeter and ratios of sides)		**Y**		**X**		**X**	**X**	✓
Side AC (using ratio of sides)		**X**	**Y**				**X**	
Side AB (using ratio of sides)	**Y**	**X**				**X**		
Service row	**W**	**Z**	**W**	**Z**		**Z**	**Z**	

In another words, by doing corresponding actions, we obtain the following algorithm to solve the set task: "The side BC using the perimeter and the ratios of the sides" $\longrightarrow$ "The side AC using the ratio of the sides" $\longrightarrow$ "The side AB using the ratio of the sides" .

The described mechanisms of information representation and information processing allow us to reduce the number of experts involved to develop AI systems and simplify subject domain descriptions, which allows to describe them more fully.

References

[1] Chibirova M.O. (2015). Analysis of existing approaches: ontology, cognitive maps and mivar nets. *Radio Industry*. Vol. **3**, pp. 55-66. (In Russian)

[2] Lakemeyer G., Nebel B. (1994). *Foundation of knowledge representation and reasoning*. Springer-Verlag, Berlin.

[3] Luger G.F. (2009). *Artificial intelligence: structures and strategies for complex problem solving*. Pearson Education, Boston.

[4] Varlamov O.O. (2003). Multidimensional informational space of data and rules representing overview. *Information technologies*. Vol. **5**, pp. 42-47. (In Russian)