ISSN 1064-2307, Journal of Computer and Systems Sciences International, 2011, Vol. 50, No. 3, pp. 448–461. © Pleiades Publishing, Ltd., 2011.
Original Russian Text © E.S. Bolotnikova, T.A. Gavrilova, V.A. Gorovoy, 2011, published in Izvestiya Akademii Nauk. Teoriya i Sistemy Upravleniya, 2011, No. 3, pp. 98–110.

ARTIFICIAL INTELLIGENCE

To a Method of Evaluating Ontologies

E. S. Bolotnikova, T. A. Gavrilova, and V. A. Gorovoy

St. Petersburg State Polytechnical University, Polytechnicheskaya ul. 29, St. Petersburg, 195251 Russia
St. Petersburg State University, Universitetskaya nab. 7–9, St. Petersburg, 199034 Russia

Received July 11, 2010; in final form September 15, 2010

Abstract—The problem of evaluating the quality of ontologies is addressed. A classification of the existing methods of evaluating ontologies is given and a model for evaluating the human perception of ontologies from the cognitive point of view is proposed. In addition, a methodology of application of the proposed model is presented, as well as an example of comparison of two ontologies in the field of artificial intelligence by the given method.

DOI: 10.1134/S1064230711010072

INTRODUCTION

Ontology engineering investigates the problems of design, development, and application of ontologies as universal models of knowledge representation, whose predecessors were hierarchical semantic networks and frames [1, 2]. At the modern stage of development of the Internet and designing distributed intelligent systems, ontologies are becoming to play the key role in artificial intelligence technologies [3]. For example, they are employed as knowledge bases in decision support systems [4].

The problem of estimating the quality of ontologies is one of the topical problems of ontology engineering. This part of the process of development of ontologies is important in practice, which is confirmed by the fact that many different approaches in the field of evaluation of ontologies have been proposed by different groups of researchers. At present more than ten methods are known, and the problem of choosing an appropriate methodology for solving a particular problem is becoming more and more complex. Moderns surveys of the existing methods and approaches in evaluating ontologies were presented in [5, 6]. Qualitative criteria of evaluation from the position of the Gestalt psychology were proposed in [7, 8]. A generalized view of the existing approaches is proposed by the model of classifying methods of evaluating ontologies in Fig. 1 [9].

On the whole, methods of evaluating ontologies are based on problems using one or more following criteria:

(1) Completeness and precision of the dictionary of the subject domain (this problem is solved by the approaches from [10, 11]).

(2) Adequacy of the structure from the point of view of taxonomy, relations, etc. The formal ontology of meta-properties OntoClean [2] is the most famous in this aspect; other methods are presented in [10–12].

(3) The ability of perception (from the cognitive point of view). The Gestalt approach was described in [13]; this aspect is also addressed in [14].

(4) Performance in applications (see [15]).

(5) Choice of the best ontology from several available. As a rule, these works use various metrics, e.g., Ontometric [16].

Ontologies can be evaluated at different stages of development and application:

(1) design and prototyping [2];

(2) testing before release [16];

(3) application [15].

According to the degree of automation, all methods of evaluating ontologies can be split into three groups:

(1) automatic (e.g., EvaLexon [12]),

(2) semi-automatic [14],

(3) manual [16].

The objects of analysis of the existing methods may be one or several of the concepts connected with development of ontologies listed below:

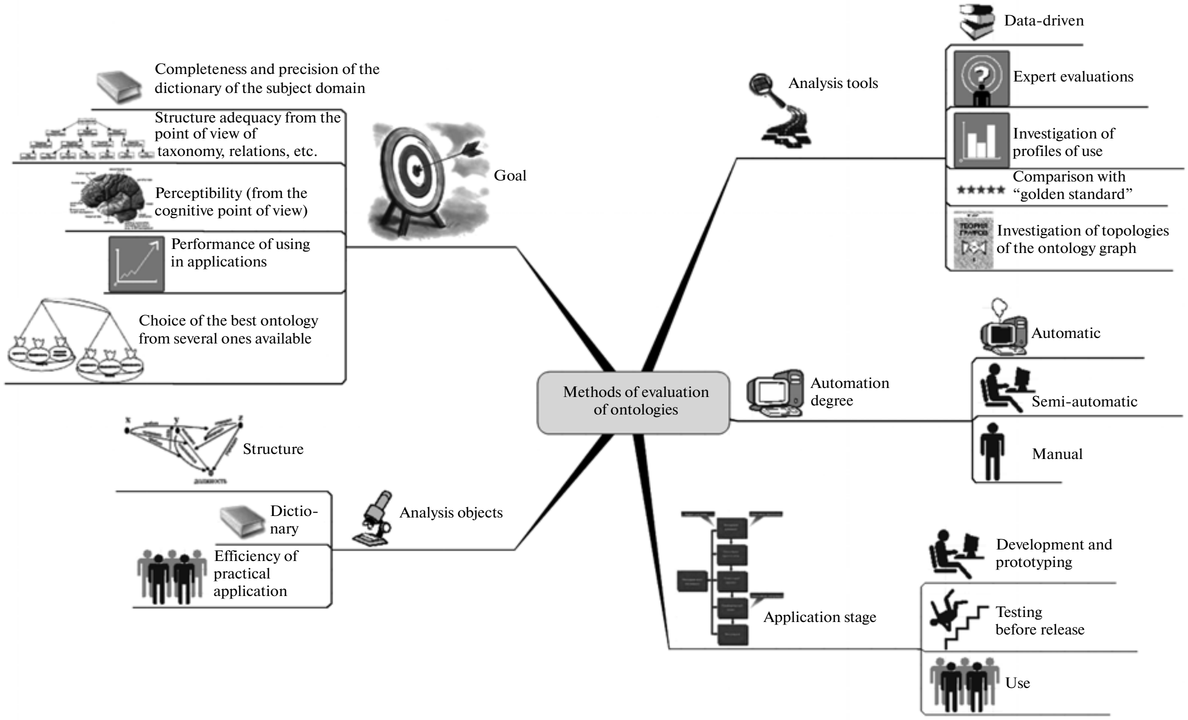

Fig. 1. Model of classification of methods of evaluation of ontologies.

(1) structure [14],
(2) dictionary [10],
(3) efficiency of practical application [15].

According to the tools applied for analysis of the quality and maturity of ontologies, we can split all methods into the following classes:

(1) data-driven methods [10];
(2) methods using expert evaluations [16];
(3) investigation of profiles of use [15];
(4) comparison with a "golden standard" [11, 12];
(5) investigation of the topology of the ontology graph [14, 17].

Design of ontologies of subject domains in order to develop and analyze complex systems (objects) and their control systems is one of the topical problems of the theory and practice of control of complex systems. In evaluating the quality of designed ontologies, the following two aspects are most important: (1) correctness and depth of reflection of the subject domain, and (2) ergonomic aspect of the ontology representation from the point of view of quality and human speed of perception. The problems of ontology completeness and precision are addressed in [10, 11, 18]. In this paper, we consider only the second aspect.

The problems associated with cognitive ergonomics in evaluating ontologies were partially investigated in [13, 14]. By cognitive ergonomics, we mean the field of inter-discipline investigations studying the processes of perception and understanding interfaces, models, and representations from the point of view of ergonomics. The first of the mentioned works was the origin for generating principles, underlying the ability to percept ontologies presented in what follows. In [14], several metrics, which can be applied in evaluating cognitive ergonomic aspect.

From the point of view the classification presented in Fig. 1, the evaluation method proposed in this paper can be described as follows.

Goal: evaluation of the property being perceive (from the cognitive point of view), and the choice of the best ontology among the several ones available.

Analysis object: ontology structure.

Analysis tools: analysis of the ontology graph.

Degree of automation: automatic, and semi-automatic (the final decision is made by an expert based on the calculated automatic model).

Stage of application: development, prototyping (calculation can be performed at each next development iteration), and testing before release.

1. INVESTIGATION OF ONTOLOGY QUALITY BASED ON GRAPH ANALYSIS

The approach to quality evaluation based on the topology of the ontology graph was described in [14, 17]. In [17] several metrics applied for analysis of ontology quality are presented, a part of which is calculated based on the topology of the ontology graph. Let us present here those of them that are referred to metrics of cognitive ergonomics (in the context of these metrics, in what follows, we will consider only is-A arcs as graph edges).

Ontology depth. In [14] Gangemi distinguished three metrics for calculating the depth: absolute depth, calculated as the sum of lengths of all paths of the graph (where a path is any sequence of nodes connected with each other, beginning from the root node and ending by the graph leaf node); mean depth, which is the absolute depth divided by the number of graph paths; maximum depth, which is equal to the maximum path length. The greater is the depth, the more difficult to perceive the graph.

Ontology width. Absolute width represents the sum of the numbers of nodes for each hierarchy level over all levels; mean width, calculated as the absolute width, divided by the number of hierarchy levels; maximum width, which is equal to the number of nodes at the level with a maximum number of nodes. The smaller is the width, the better is ontology from the point of view of cognitive ergonomics.

Ontology tangledness. It is defined as the number of nodes of the ontology graph divided by the number of nodes that have several direct super-classes. Thus in the ontologies in which multiple inheritance is absent (relation is-A), this metric is zero. The smaller is its final value, the better is ontology from the point of view of cognitive ergonomics.

The ratio of the number of classes to the number of properties. The greater is this index, more easily the ontology perceived.

The number of anonymous classes. To improve the ontology quality, it is better to minimize the number of them.

Despite the usability of these metrics, they cover only very small part of factors affecting our perception and the ability to memorize. In the next section, we present the principles that affect human cognitive abilities that provide the ground of the proposed model of metrics.

2. PRINCIPLES THAT PROVIDE THE GROUND OF EVALUATING COGNITIVE ERGONOMIC ASPECT

The basic principle of evaluating the visual perceptibility and understandability of ontologies were presented in [7, 8, 13]. They are based on ideas proposed by Max Wertheimer in the field of Gestalt psychology [19]. He considered all tasks from the point of view of incompleteness or imperfection of structure. For example, the main principle of a good gestalt (good shape) or the law of Pragnanz was formulated as: *"Organization of any structure in nature or in mind should be as good (regular, complete, balanced or symmetric) as the existing conditions make it possible."*

The other cognitive–perceptive principles formulated in the form of laws can also be useful:

proximity: visual stimuli (objects), located close each other are perceived as a whole object;

similarity: things that are similar tend to be grouped together;

V. Kohler *inclusion*: there is a trend to percept only a big figure, rather than a smaller one that belongs to it;

lex parsimoniae (parsimony): the simplest example is the best, known as "Occam's razor": "the simplest explanation is more likely the correct one."

For the purposes of ontology engineering, these laws can be reformulated and made be applicable for a practical knowledge engineer. The main hypothesis can be formulated as: "*Harmony = conceptual balance + clarity.*" Note that the conceptual balance means: concepts of the same hierarchy level are related to the parent concept by the same type of relation (e.g., "class–subclass" or "part–whole"); the depth of branches of the ontology tree has to be approximately the same (±2); the total pattern should be rather symmetric; and if possible cross-references are excluded.

Clarity involves: minimization—the maximum number of concepts of the same level or the branch depth should not exceed the famous Ingve-Miller number (7 ± 2) [20]; transparency for reading—the type of relations has to be obvious, if possible, so that not to overload the scheme of the ontology with excessive information and to omit names of relations.

The model of evaluation of the ontology quality obtained in this paper is based on the principles specified above and in many aspects is their formalization. The goal of application of the model is to evaluate the balance and perceptibility of ontologies by users.

3. PRACTICAL APPLICATION OF THE MODEL OF EVALUATION OF ONTOLOGY QUALITY FROM THE POINT OF VIEW OF COGNITIVE ERGONOMICS

Many aspects, partially described in the previous section, affect the quality of an ontology from the cognitive point of view. The proposed evaluation model consists of several metrics, based on measurements of which, we can make conclusions on the ontology quality. The model and complete list of metrics were presented in detail in [21]. Here we place the emphasis on the following: we outline the domain of applicability and possible methodology of dealing with metrics in the course of evaluation of cognitive ergonomic properties of ontologies, as well present an example of using the model in practice.

The proposed metrics can be used as additional metrics, assisting experts in the evaluation of quality of ontologies together with other metrics. In some cases, the model addressed in the paper is not applicable since there is no necessity to evaluate cognitive ergonomic properties of ontologies. For example, if only the computational efficiency is important, and only programs interact with the designed ontology, not people. In addition for the developed model, it is important that samples for comparison are defined (although there are metrics for which we have recommended absolute bounds of values, which are desired to be followed in designing ontologies). Usually such samples are present since the process of designing ontologies, as a rule, has interactive character. Professional designers of ontologies use systems of control of versions. Because of this, we can compare the current version not only with the previous one, but realize how strongly the ontology was changed from the point of view of its users on different intervals of its life cycle. The proposed model of evaluation can help to understand what should be corrected in the description of the subject domain in order to improve it from the point of view of cognitive ergonomics. Thus each next version of the ontology will be better and it can be perceived fast by users.

The model can also be used in evaluating ontologies of the same subject domain produced by different people/teams. The calculated metrics are able to show which of them is better from the point of view of cognitive ergonomics and to choose the best of them if the evaluations of other important criteria differ insignificantly.

In addition, the metrics are efficient in evaluating ontologies since some of them can inform the expert immediately about disadvantages of the ontology from the point of view of human perception. Among these metrics are metrics of cycles and the Ingve–Miller metrics. Moreover the statistical metrics, e.g., the mean-square deviations of the depth or degree of a graph node (the degree is determined by the number of connections of a given node with the other ones) in the case of their big absolute values can help the expert to determine whether the ontology has problematic zones. Then we can consider the nodes and branches of the ontology that mainly contribute the mean-square deviation and make the decision on changing the ontology in these places.

4. METHODOLOGY OF MODEL APPLICATION

First of all, the success in using the mentioned metrics depends on an efficient and clear methodology, which makes it possible to experts not to get entangled among a quite large amount of measurement results available. In what follows, we will give methodological instructions for using the developed model of metrics and present an example of comparison of two ontologies for the same subject domain. It was initially supposed that compared ontologies have the same level of representation of the subject domain, i.e., are semantically equivalent, which can be tested by the corresponding evaluation methods, e.g., [10] or [11].

4.1. Stages of Analysis of the Ontology Size

To provide that the expert is able to have an idea of what values of certain metrics are appropriate, it is necessary first to obtain information about the ontology size. In evaluating two ontologies, it is also necessary to realize to what extent one of them is greater than the other, since even the values of certain metrics are worse for the greater ontology, this does not mean that one of them is designed not as good as the other one. The greater ontology contains more information, and consequently it is more complex to design it well.

At this stage, it is necessary to calculate the following metrics (a detailed description of all metrics can be found in [21]): the number of nodes of the ontology graph; the maximum distance from the root ontology node; the number of leaf nodes of the ontology tree (as arcs of the tree only the main ontology properties distinguished above are considered); the number of nodes of the ontology tree that have leaf nodes in the direct descendants; and the number of arcs of the ontology graph.

In online analysis, the first two metrics are most informative: the number of nodes of the ontology graph and the maximum distance from the root ontology node. If there is no sufficient time for evaluation, the other metrics from the ones listed above can be taken into account if possible.

4.2. Stage of Analyzing the Crucial Ontology Errors

There are errors in ontologies such that it is of no sense to further evaluate ontologies if these errors are not corrected. Among these errors are cycles in ontologies. To investigate cycles, two metrics can be used: the number of different cycles in the graph; and the number of nodes involved in some cycle, divided by the number of graph nodes.

In the majority of cases, if the value of the first metric is nonzero, it is sensible to stop further investigation up to the correction of the ontology. To other crucial errors, we can refer multiple inheritance, since in many cases it is not reasonable. In this case, entanglement measures can be of use:

the number of nodes with several parents; the number of nodes having several parent, divided by the number of graph nodes; and the mean number of parent nodes of a graph node.

It is reasonable to pay special attention of ontology designers to nodes with several parents. The nodes having as descendants both leaf nodes and nodes of the tree that are not leaf nodes deserve special consideration. This analysis is based on estimates of the number of ontology nodes:

having leaf nodes among descendants; having only leaf nodes among descendants; that have leaf nodes and non-leaf nodes as descendants as compared with the whole number of node having leaf nodes among descendants.

If the first and the second numbers do not coincide, it is reasonable to analyze the nodes that provide that these metrics do not coincide.

4.3. Stage of Analysis of the Ingve—Miller Metrics

At this stage, the metrics that show the human ability to perceive an ontology node together with all nodes connected with it by the following properties:

the ratio of the number of nodes with a normal degree to the number of all nodes; and the mean degree of a graph node;

the median of the degree of a graph node; 90 percent confidence intervals for the value of the degree of a graph node; and the mean-square deviation of the degree of a graph node.

The first of the metrics presented above clearly specifies the fraction of concepts with a difficulty perceived number of connections among the ontology concepts. A smaller value of the second metric under other equal values testifies that the ontology can be easier perceived. The last metrics show the balanced nature of the ontology. Even if the degree of certain nodes go beyond the bounds, i.e., is greater than nine, the fact that the most part of concepts (e.g., 90%) satisfy the ergonomics requirements. A smaller mean-square deviation of the degree of a graph node corresponds to a better structure of the ontology, i.e., to a smaller number of concepts going beyond the total pattern of concepts.

4.4. Stage of Analysis of Types of Ontology Relations

The number of different types of relations available in the ontology directly affects its perception. A large number of ontology concepts with different types of outgoing relations can also have negative effect on the cognitive ontology ergonomicity. At this stage, it is necessary to calculate the values of two types of metrics such as the diversity of types of relations of concepts and diversity of the number of relations with the use of the following indices:

the number of different types of relations in the graph; the normalized number of different types of relations, i.e., the number of types of relations, divided by the number of concepts (graph nodes); the ratio of the number of nodes with different types of outgoing relations to the total number of graph nodes; the ratio of the number of nodes with ingoing relations to the number of graph nodes; the mean value of the types of ingoing relations of a graph node; and the mean value of different types of outgoing relations of a graph node.

The third and fourth metrics allow one to reveal graph nodes which have to be analyzed by ontology designers in detail. In ontologies that are appropriate from the point of view of cognitive ergonomics, the last two metrics are close to one.

4.5. Stage of Analysis of the Depth Metrics

If different branches of the ontology tree have different lengths, then this negatively affects the perceptibility of ontologies. Using depth metrics, we manage to understand how uniformly the ontology is developed in different domains. To analyze an ontology at this stage, the following metrics are employed: the minimum depth; the maximum depth; the mean depth; the median of the depth; 90 percent confidence levels of the depth value; mean-square depth deviation; mean-square depth deviation, divided by the mean depth.

If the maximum ontology depth is greater than eight, then it is reasonable for the expert to take into account this fact. When dealing with an ontology, a specialist keeps in mind simultaneously either a concepts with all concepts related with it or a particular ontology branch from the leaf node to the root. Therefore if the branch has more than nine objects (nodes), the ontology is difficult to perceive. Among the presented metrics, two last metrics shows which ontology is better in the balance in depth (naturally, it is necessary to take into account the size of the ontology. The branches that contribute mostly to the biggest mean-square deviation should be analyzed in detail.

4.6. Stage of Analysis of Width Metrics

Width metrics make it possible to evaluate the balance of an ontology based on studying distinctions between the hierarchy levels (in analysis of these metrics, the ontology tree is considered, i.e., only the main distinguished property). If at a lower level there are smaller concepts in the ontology than at the previous one, then this can negatively affect the perception of it by users. It is also important to study how uniformly the ontology is extended from one level to another. The metrics employed for the analysis at this stage are as follows: the mean width; the mean ratio of the widths of adjacent levels; the maximum ratio of the widths of adjacent levels; the result of division of the mean-square deviation of the value of the ratio of the widths of adjacent levels of the graph by the mean value of this ratio.

The least information is provided to the expert by the first metric, since it is difficult to make any conclusions based on it. The second and the last metrics, which make it possible to understand how uniform the ontology is extended in the transition to the new levels of detailing the description of the subject domain, are most informative.

4.7. Stage of Analysis of Branching Metrics

The last, most detailed level deserves special attention in investigating different hierarchy levels in the ontology tree. A part of branching metrics were used at the second methodology stage for analysis of critical ontology errors (the ration of the number of nodes that have both leaf and non-leaf nodes as descendants to the total number of nodes that have leaf nodes among descendants). The remaining metrics can be referred to width metrics, since they study similar ontology properties, but separated in a particular class and are applied at a particular stage of analysis, because of big effect of the last ontology level on its perception. Among them there are indices revealed in the last but one nodes in the graph: the mean number of descendants—leaf nodes; the maximum number of descendants—leaf nodes; the minimum number of descendants—leaf nodes; and the mean-square deviation of descendants—leaf nodes. At this stage of analysis, we can distinguish the graph nodes such that the last level of their hierarchy is not worked out. A smaller value of the last from the listed metrics shows that the last ontology level was not balanced.

4.8. Stage of Decision Making and Producing Recommendations

The aspects of an ontology that affect the quality and the speed of its perception by users analyzed in the described methodology are so diverse that it is very difficult to produce an integral quality index. Nevertheless, let us specify the main factors based on which we can surely speak of what of the considered ontologies is better or worse from the point of view of cognitive ergonomics.

Let us accept the following notation: g is the graph representing the ontology. The concepts (classes and instances) of the ontology are graph nodes, the relations between concepts are represented in the form of graph arcs; G is the set of all nodes of the graph g; and E is the set of all arcs of the graph g.

If we take as an integral criterion a bounded set of metrics, based on which the expert may make conclusions on the quality of ontologies, then it is necessary that it involves the following most significant indices:

(1) The number of different cycles in the graph. In an appropriate ontology, this number has to be zero, since the presence of cycles does not facilitate perception.

(2) The number of nodes involved in a cycle, divided by the number of nodes in the graph. The smaller is the value of this metric, the better is for the ontology, and the optimal value is zero. This index is calculated as

$$m = \frac{N_{v \in C}}{n_G},$$

where n_G is the number of graph nodes, C is the set of graph nodes involved in at least one cycle, and $N_{v \in C}$ is the number of graph nodes involved in a certain cycle.

(3) The ratio of the number of nodes that have both leaf and non-leaf nodes as descendants to the number of nodes that have leaf nodes among descendants

$$m = \frac{N_{v \in S_{LEA\&SIB}}}{N_{v \in S_{LEA}}},$$

where $S_{LEA\&SIB}$ is the set of all nodes that are simultaneously both parents of graph leaf nodes and parents of internal nodes, and $N_{v \in S_{LEA\&SIB}}$ is the number of these nodes; S_{LEA} is the set of all nodes that have among descendants graph leaf nodes, and $N_{v \in S_{LEA}}$ is the number of these nodes.

(4) The ratio of the number of nodes with a normal degree to the total number of nodes. The node degree is the number of arcs for which it is the terminal one. Nodes with normal degree are concepts with the number of relations not exceeding the famous Ingve–Miller number 7 ± 2 [20]. The closer is the value of the metric to one, the better is ontology from the point of view of cognitive ergonomics. To estimate this index, we use the expression

$$m = \frac{N_{v \in GD}}{n_G},$$

where n_G is the number of graph nodes, $GD = \{v \in G | \deg(v) \le 9\}$ is the set of nodes with normal degree, and $N_{v \in GD}$ is the number of graph nodes with a normal degree.

(5) The fraction of nodes with different types of outgoing relations among all graph nodes

$$m = \frac{N_{v \in VD}}{n_G},$$

where $VD = \{v \in G | N_{type(e_v)} > 1\}$ is the set of graph nodes with different types of outgoing relations, and $type(e_v)$ is the set of types of relations outgoing from the node v.

(6) The fraction of nodes with different types of ingoing relations among all graph nodes

$$m = \frac{N_{v \in V\tilde{D}}}{n_G},$$

where $V\tilde{D} = \{v \in G | N_{type(\tilde{e}_v)} > 1\}$ is the set of all graph nodes with different types of ingoing relations, and $type(\tilde{e}_v)$ is the set of types of relations ingoing into the node v.

(7) The mean-square deviation of the degree of a graph node

$$m = \frac{\sum_{v \in G}\left(\deg(v) - \frac{\sum_{v \in G}\deg(v)}{n_G}\right)^2}{n_G - 1} = \frac{\sum_{v \in G}\left(\deg(v) - \frac{2^* n_E}{n_G}\right)^2}{n_G - 1},$$

(8) The ratio of the mean-square deviation of the depth to the mean depth

$$m = \frac{\dfrac{\sum_{j}^{P}\left(N_{j \in P} - \dfrac{\sum_{j}^{P} N_{j \in P}}{n_{P \subseteq g}}\right)}{n_{P \subseteq g} - 1}}{\dfrac{\sum_{j}^{P} N_{j \in P}}{n_{P \subseteq g}}},$$

(9) The mean-square deviation of descendants-leaf nodes at the last but one graph nodes

$$m = \frac{\sum_{j \in SIB_{LEA}}\left(N_{j \in SIB}^{j \subseteq LEA} - \dfrac{\sum_{j \in SIB_{LEA}} N_{j \in SIB}^{j \subseteq LEA}}{n_{SIB_{LEA}}}\right)}{n_{SIB_{LEA}} - 1},$$

where $n_{SIB_{LEA}}$ is the number of all leaf nodes having the same parent, and $N_{j \in SIB}^{j \subseteq LEA}$ is the number of leaf nodes in the group of descendants having the same common parent node.

(10) The maximum ratio of the widths of adjacent levels

$$m = \frac{N_{l_i \in L}}{N_{l_{i-1} \in L}} \quad \forall k\left(\frac{N_{l_i \in L}}{N_{l_{i-1} \in L}} > \frac{N_{l_k \in L}}{N_{l_{k-1} \in L}}\right),$$

where $N_{li \in L}$ is the number of nodes at the level i of the graph g.

For each of the listed indices, the best values are clear, but for some of them, namely, characteristics 7–10, we should keep in mind the ontology size, since in large-size ontologies, somewhat greater values of these metrics are possible. In the cases when an ontology has all ten indices better than another ontology, we can state that the first ontology is better from the point of view of cognitive ergonomics if the sizes of ontologies are comparable. If not all is so definite, then the expert should take into account other metrics and make a decision keeping in mind which metrics in this case have greater importance for him, and which are of less importance. Unfortunately, there are no universal recommendations on this subject.

In the course of analysis of an ontology according to the presented methodology at each stage, potential bottlenecks in ontology perception are revealed. The obtained list of concepts and problems can be efficiently used for further work in improving the ontology quality.

5. EXAMPLE OF COMPARISON EVALUATION OF TWO ONTOLOGIES IN THE FIELD OF ARTIFICIAL INTELLIGENCE (AI)

We compared two ontology in the field of AI designed based on the ASM classification in AI and based on Steven Russell and Peter Norvig textbook "Artificial Intelligence: A Modern Approach". First two mind-maps (see Figs. 2 and 3) were developed. From them ontologies in the format owl were made, which made it possible to automatize the calculation of metrics for analysis using the tool COAT (Cognitive Ontology AssesmenT), whose operation principle and structure were described in [21]. In the considered ontologies, only one type of relation is used, is-part-of; therefore the ontology graph and ontology tree coincide in this case, which somewhat facilitates the problem of analysis.

5.1. Stage of Analysis of the Ontology Size

Table 1 shows that the Norvig ontology is two times greater than the ASM ontology in the number of objects. This should be taken into account in the comparison of different metrics that are not normalized to the ontology size.

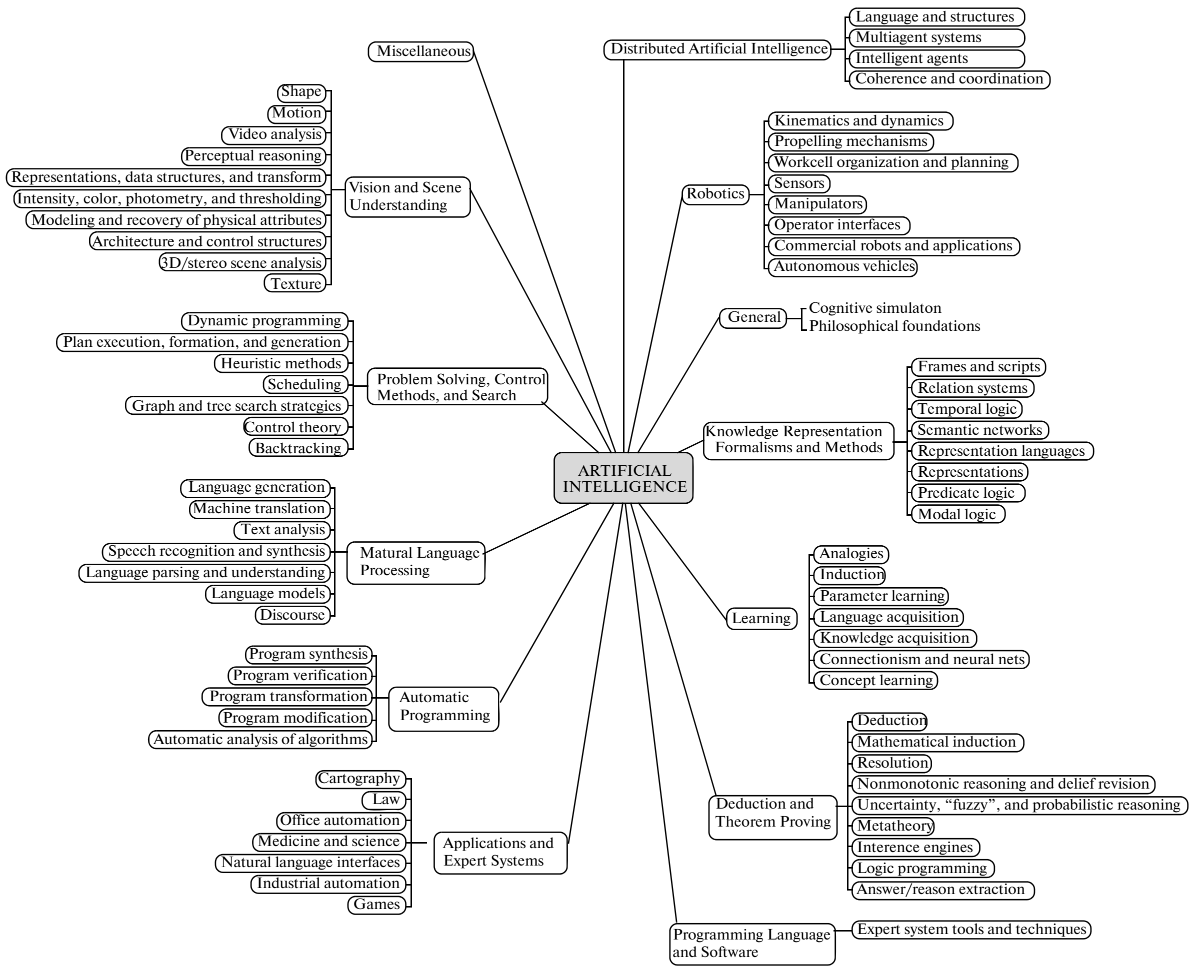

Fig. 2. Mind–map of the ASM classification in artificial intelligence.

5.2. Stage of Analysis of Crucial Ontology Errors

In connection with the fact that in both ontologies there are no crucial errors connected with multiple inheritance and presence of cycles, a part of metrics in Table 2 are omitted. At this stage, we can reveal one problematic node in both ontologies: in the ASM ontology, this is the connective of the central concept "Artificial Intelligence" and "Miscellaneous", and in the Norvig ontology this is the concept "General", which among descendants has both the non-leaf node "Intelligent Agents" and two leaf nodes, "The Foundations of Artificial Intelligence" and "The History of Artificial Intelligence".

Table 1. Values of metrics of ontology size

Metric	Ontology	
	ASM	Norvig
Number of nodes of the ontology graph	89	163
Maximum distance to the ontology root	2	3
Number of leaf nodes of the ontology tree	76	131
Number of nodes of the ontology tree that have leaf nodes in the direct descendants	13	25
Number of arcs of the ontology tree	88	162

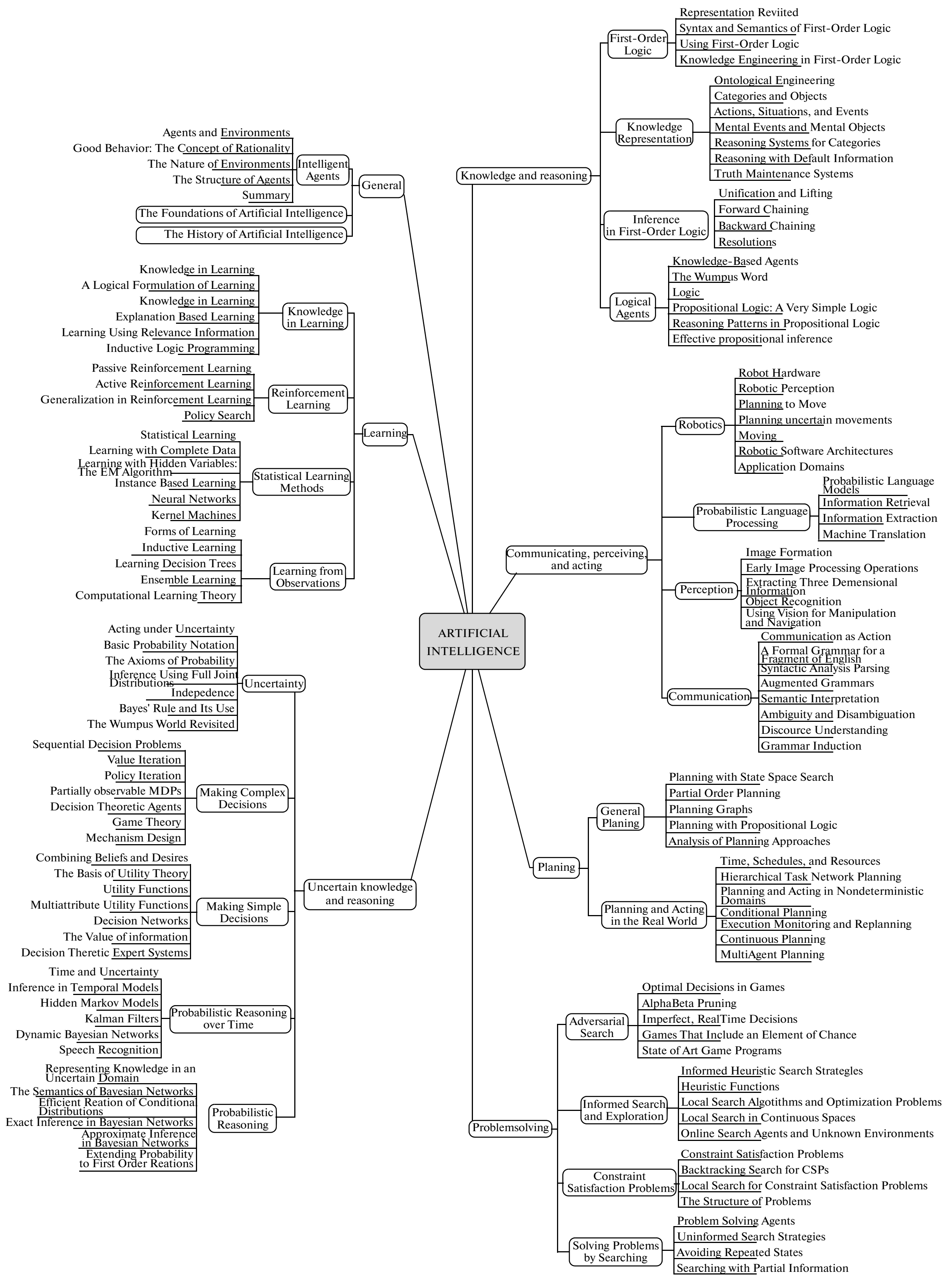

Fig. 3. Mind-map in artificial intelligence based on the textbook by Steven Russell and Peter Norvig "Artificial Intelligence: A Modern Approach."

Table 2. Values of metrics revealing crucial errors in the ontology

Metric	Ontology	
	ASM	Norvig
Number of nodes of the ontology graph	0	0
Number of nodes with several parents	0	0
Number of ontology nodes that have leaf nodes among descendants	13	25
Number of ontologies that have only descendants-leaf nodes	12	24
Ratio of number of nodes that have leaf and non-leaf nodes as descendants to the total number of nodes that have leaf nodes among descendants	0.077	0.04

Table 3. Values of the Ingve-Miller metrics

Metric	Ontology	
	ASM	Norvig
Ratio of the number of nodes with normal degree to the number of all nodes	0.615	1.0
Mean degree of a graph node	7.69	6.03
Median of the degree of a graph node	8.0	6.0
90 percent confidence bound for the degree of a graph node	12.2	8.0
Mean-square quadratic deviation of the degree of a graph node	0.94	0.125

Table 4. Values of depth metrics

Metric	Ontology	
	ASM	Norvig
Minimum depth	1	2
Maximum depth	2	3
Mean depth	1.987	2.985
Depth median	2.0	3.0
90 percent confidence bound for the depth	2.0	3.0
Mean-square deviation of the depth	0.013	0.015
Mean-square deviation of the depth divided by the mean depth	0.007	0.005

5.3. Stage of Analysis of Ingve—Miller Metrics

From the presented values in Table 3, we immediately see that the Norvig ontology outperforms the ASM ontology in Ingve—Miller metrics, event despite the difference in sizes. In the Norvig ontology, there are no nodes with degree nine at all. In the ASM classification, there are following problematic points: the central concept "Artificial Intelligence" has 13 descendants simultaneously; so many concepts are difficult to be perceived simultaneously; and concept "Vision and Scene Understanding" has ten descendants—leaf nodes. At this stage, we can distinguish nodes whose degrees differ from the mean value maximally and to draw increased attention to them.

5.4. Stage of Analysis of Types of Ontology Relations

In the addressed ontologies, there is only one type of relation, so this stage can be eliminated.

5.5. Stage of Analysis of Depth Metrics

The results of Table 4 testify that the ontologies are comparable according to the depth metrics, and it is difficult to choose the best one. At a first glance, the mean-square deviation of the depth of the ASM

Table 5. Values of width metrics

Metric	Ontology	
	ASM	Norvig
Mean width	29.667	40.75
Mean ratios of widths of adjacent levels	9.385	5.223
Maximal ratio of widths of adjacent levels	13.0	7.0
Mean-square deviation of the ratio of widths of adjacent levels of the graph to the mean ratio of widths of adjacent levels	13.07	1.83

Table 6. Values of branching metrics

Metric	Ontology	
	ASM	Norvig
Mean number of descendants—leaf nodes of the last but one in the graph	5.84	5.24
Maximal number of descendants—leaf nodes of last but one in the graph	10	7
Minimal number of descendants—leaf nodes of the last but one in the graph	1	2
Mean-square deviation of descendants—leaf nodes of the last but one in the graph	8.974	2.107

Table 7. Values of main metrics for two ontologies

Metric	Ontology	
	ASM	Norvig
Number of different cycles in the graph	0	0
Number of nodes having several parents divided by the number of nodes in the graph	0	0
Fraction of the number of nodes that have both leaf and non-leaf nodes as descendants among all numbers that have leaf-nodes among descendants	0.077	0.04
Ratio of the number of nodes with normal degree to the number of all nodes	0.615	1.0
Mean-square deviation of the degree of a graph node	0.94	0.125
Mean-square deviation of the depth divided by the mean depth	0.007	0.005
Mean-square deviation of descendants—leaf nodes of the last but one nodes in the graph	8.974	2.107
Maximum ratio of widths of adjacent levels	13.0	7.0

ontology is a bit smaller, but if we normalize the metrics by the ontology size (the last metric), then the Norvig ontology is slightly better.

5.6. Stage of Analysis of Width Metrics

The obtained values in Table 5 show that the ASM ontology adds information from level to level faster, and that, probably, it lacks intermediate hierarchy levels, which may facilitate the perception.

5.7. Stage of Analysis of Branching Metrics

It is obvious from the results of Table 6 that the Norvig ontology is much more balanced in the domain of leaf nodes, despite the fact that the total number of leaf nodes in it is considerably greater.

5.8. Stage of Decision Making and Making Recommendations

Consider in Table 7 the most significant metrics for both ontologies, based on which we can conclude which of them is better from the point of view of cognitive ergonomics.

According to this not big set of results, we can unambiguously conclude that the Norvig ontology in the majority of parameters outperforms the ASM ontology from the point of view of cognitive ergonomics, despite the difference in sizes. In the course of analysis, particular bottlenecks were also revealed, which most significantly negatively affected the perception of the ontologies by users.

CONCLUSIONS

Evaluation of cognitive ergonomicity of ontologies is important in the cases when the ontology is aimed at learning or passing on knowledge. Note that the models of evaluating ontologies existing at present do not give an opportunity completely analyze them from the point of view of quality and human speed of perception. The proposed model fills this gap. The majority of metrics employed within its scope can be calculated automatically, which substantially reduces the load on the expert making the decision on the final evaluation of the ontology quality.

The proposed model also makes it possible to designers of ontologies to understand whether their ontology becomes better from the point of view of cognitive ergonomics, comparing the subsequent iterations with the results of the preceding ones or to choose a better ontology from a set of ontologies describing the same subject domain. The presented results may be efficiently applied in systems of knowledge management in enterprises and in education and to assist specialists and teachers in designing more qualitative ontologies.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research, project no. 08-07-00062_a.

REFERENCES

1. T. R. Gruber, "A Translation Approach to Portable Ontologies," Knowledge Aquisition **5** (2), 199–220 (1993).
2. N. Guarino and C. Welty, "Evaluating Ontological Decisions with OntoClean," Commun. ACM **45** (2), 61–65 (2002).
3. A. S. Kleshchev and E. A. Shalfeeva, "Defining Structural Properties of Ontologies," Izv. Ross. Akad. Nauk, Teor. Sist. Upr., No. 2, 69–78 (2008) [Comp. Syst. Sci. **47** (2), 226–234 (2008)].
4. V. V. Andreev, V. A. Vittikh, S. V. Batishchev, et al., "Methods and Tools for Designing Open Multiagent Systems for Supporting Decision-Making Processes," Izv. Ross. Akad. Nauk, Teor. Sist. Upr., No. 1, 126–137 (2003) [Comp. Syst. Sci. **42** (1), 122–131 (2003)].
5. J. Brank, M. Grobelnik, and D. Mladenic, "A Survey of Ontology Evaluation Techniques," in *Proceedings of Conference on Data Mining and Data Warehouses, SiKDD'2005, Ljubljana, Slovenia, 2005*, http://kt.ijs.si/dunja/sikdd2005/Papers/BrankEvaluationSiKDD2005.
6. J. Hartmann, Y. Sure, A. Giboin, et al., Methods for Ontology Evaluation, Knowledge Web Deliverable D1.2.3, 2005, http://www.starlab.vub.ac.be/research/projects/knowledgeweb/KWeb-Del-1.2.3-Revised-v1.3.1.
7. T. A. Gavrilova, "To an Approach to Ontological Engineering," Novosti Iskusstvennogo Intellekta, No. 3, 25–31 (2005).
8. T. A. Gavrilova, "Gestalt-Principles of Designing Ontologies," in *Proceedings of 2nd Conference on Cognitive Science, St. Petersburg, Russia, 2006*, Vol. 1, pp. 240–242 [in Russian].
9. V. A. Gorovoy, "A Model of Classification of Methods of Evaluating Ontologies," in *Proceedings of 2nd International Youth Conference on Artificial Intelligence: Philosophy, Methodology, and Innovations*, St. Petersburg, Russia, 2007, pp. 307–310 [in Russian].
10. C. Brewster, H. Alani, S. Dasmahapatra, et al., "Data Driven Ontology Evaluation," in *Proceedings of International Conference on Language Resources and Evaluation, Lisbon, Portugal, 2004.*
11. A. Maedche and S. Staab, "Measuring Similarity between Ontologies," in *Proceedings of CIKM, LNAI, Siguenza, Spain, 2002*, Vol. 2473, pp. 251–263.
12. P. Spyns, *EvaLexon: Assessing Triples Mined from Texts. Technical Report 09, STAR Lab* (Belgium, Brussels, 2005).
13. T. Gavrilova and V. Gorovoy, "Technology for Ontological Engineering Lifecycle Support," Information Theories & Applications **14**, 19–25 (2007).

14. A. Gangemi, C. Catenacci, M. Ciaramita, et al., "Ontology Evaluation and Validation, An Integrated Formal Model for the Quality Diagnostic Task," http://www.loa-cnr.it/Files/ OntoEval4OntoDev_Final.
15. R. Porzel and R. Malaka, "A Task-Based Approach for Ontology Evaluation," in *Proceedings of ECAI, Valencia, Spain, 2004.*
16. A. Lozano-Tello and A. Gomez-Perez, "Ontometric: A Method to Choose the Appropriate Ontology," J. Datab. Mgmt **15** (2), 1–18 (2004).
17. A. S. Kleshchev and E. A. Shalfeeva, *Catalog of Ontology Properties, Principles of Organization of the Catalog, Preprint, 2007* (IAPU DVO RAN, Vladivostok, 2007) [in Russian].
18. B. V. Dobrov, N. V. Lukashevich, O. A. Nevzorova, et al., "Methods of Automated Design of Application Ontology," Izv. Ross. Akad. Nauk, Teor. Sist. Upr., No. 2, 58–68 (2004) [Comp. Syst. Sci. **43** (2), 213–222 (2004)].
19. M. Werthheimer, *Productive Thinking* (Progress, Moscow, 1987) [in Russian].
20. G. Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," The Psychological Review **63**, 81–97 (1956).
21. T. A. Gavrilova, V. A. Gorovoy, and E. S. Bolotnikova, "Evaluation of Cognitive Ergonomicity of Ontologies Based on Analysis of a Graph," Iskusstvennyi Intellekt i Prinyatie Reshenii, No. 3, 33–41 (2009).