

Decision-making System Based on The Ontology of The Choice Problem

Oleksandr Stryzhak^{*}, V'yacheslav Horborukov, Vitalii Prychodniuk, Oleg Franchuk and Roman Chepkov

National Center "Junior Academy of Sciences of Ukraine", 38-44, Dehtyariivska St., Kyiv, 04119, Ukraine

*Email: sae953@gmail.com

Abstract. The complexity of the decision-making process for experts in various topics is in the problem of initial processing of unstructured information. This problem of Big data requires the latest solutions that provide a structural reflection and integrative use of information descriptions under consideration by experts. It takes the process of information and analytical evaluation and decision-making beyond one subject area and raises questions about the ontological consideration of the problem itself. The ontological principles of such consideration are based on multiple hyperproperties of information for the realization of the categories of integration, systematicity and continuity, as well as implementation of structures and their functionality reflection and transformation. For this, there is the ontology of the problem of rational choice based on means of a ranking of alternatives on a set of indicators. It provides the creation of a system designed to solve the ranking problem based on the ontological model of the subject area. Hyperproperties of ontological representation of large amounts of information ensure the implementation of meta-tasks of its analysis, its structuring, synthesis, and support the process of rational choice. Such an approach creates the conditions for optimal processing of a Big Data.

1. Introduction

The decision-making process is based on the analysis of information and requires structured information. Moreover, this structure must meet the conditions of systematology, i.e. to include processing rules. This presentation of information is the most effective when based on the ontological approach. Ontology [1-6] is an attempt for a comprehensive and detailed formalization of some knowledge area using a conceptual model. In general, the structure of an arbitrary conceptual model is strictly hierarchical. All its concepts (objects) can be divided and grouped into certain classes based on their properties. The hierarchy is then determined by the relationships between classes. These relationships are also defined between concepts belonging to certain classes. If such connections have functional interpretation, the conceptual model has an ontology format.

Ontology is a structural model of the subject area, that is formed by expert using a word-class scientific classification. It is this classification that determines the conditions for the functional interpretation of the relationships between concepts. This interpretation defines certain formalisms for classes of concepts representation, that can be used to form decision-making support systems (hereinafter – DMSS) based on the ontology.

The links between concepts have particular importance in the ontology. Due to them, the ontology is not just a structure of concepts, but also reflects the complex relations between them and



comprehensively represents the subject area. There are three main types of connections between concepts:

- Taxonomic relations – express "is a species of" relations or the "general/specific" relations, i.e. relations that put the two entities in "general - partial" relations;
- Compositional connections – express "is a part of" relations;
- Topological connections – reflect how different components of the system are connected to each other through certain connections, or show the "paths" of physical interactions between components, as well as provide information about the spatial location of these components.

Consider the following problems of construction and use of ontologies [1, 6]:

1) The conceptual unit structure – the choice problem and level of detail of units (concepts), the status of service units, the boundary between concepts and lexical variants, the recognition of units in texts.

2) Formal model – formalized, using a certain language of knowledge, a description of the conceptual system, which specifies:

- scientifically substantiated classification;
- set of binary relations between concepts;
- axiomatics and rules of logical formal inference.

3) The set of interpretive functions, the definition of which requires the following conditions:

- axiomatics definition;
- scientifically substantiated terminologies;
- predicative representation of taxonomies;
- definition of an interpretive function for each binary relationship (at least one);
- logical rules formation for determining concepts and classes compatibility/incompatibility;
- quantitative data functional interpretation;
- determination of classes taxonomies representation levels;
- rules for determining the connectivity of the each concept contexts set.

Taking into account that the ontology is an object-oriented functional resource, which is based on expert knowledge and qualitatively represents the subject area and can be used by any intelligent information technology, in particular, as a source of information. Then objects (alternatives) and their properties can be separated from it for the direct solution of the multicriteria ranking problem. Thus, the use of ontology as a basis for presenting alternatives is a very important issue in the creation of decision-making support systems.

In the first section the technology of constructing the information environment of the multiple criteria ranking problem is described. Technology carries out the transformation of the ontological model of the subject area on the basis of the interpretive selection functions constructed using hyper-relation over elements of the taxonomic structure of the ontology and properties of its objects.

In the second section an informational model of the system for solving the problem of ranking alternatives is described. The information model of the system is represented by a set of modules, which provided by "Polyhedron" information technology.

2. Ontological Representation of the Choice Problem Based on Ranking Alternatives

By definition, an ontology is formed on the basis of term fields. Each term field is an object-oriented presentation of the subject area concepts. The identified relationships between term fields provide a definition of interpretive functions set, which are a functional manifestation of the objects (concepts) properties of the ontology, and which control the process of information resource delivery on all stages of decision-making. Therefore, it's quite useful to present an information model in the DMSS environment as a certain ontology [6-8].

The methodology of ontological systems uses the object representation of concepts (objects) in the form of term fields. This provides the definition of semantic connections between all contextual descriptions of concepts (objects) that define the ontology [5, 7, 9, 10].

A concept (object) is a certain entity that has a certain structural representation and is characterized by many properties that determine its functional behaviour.

A concept can have a set of states that is determined by the set of its properties functional values.

The functionality of each concept (object) of the ontology is determined by a set of interpretive functions. They include a certain class of functions that provide interaction between the concepts of ontology. This class is defined by the binary relationships that exist between the term fields of the ontology concept classes. This class of functions will be called methods. The methods ensure full use of the concept structure. Moreover, they determine the conditions for the implementation of interfaces with each concept.

A separate set should include individual properties of the concept, which are unique to it and define this concept separately from others.

The operability of the DMSS environment is determined by two types of interaction between classes of concepts that have a hierarchical structure:

- binary relationships that can be defined on concepts and that actually determine the methods of interaction;
- aggregators (aggregation), that determine the hierarchy of the type part-whole, whole-part, and provide a scientifically sound systemology of the ontological system.

An ontological system that is able to reflect the operability of the particular subject area (SA) environment is implemented using the properties of a non-empty concepts set that satisfies following requirements:

- objects are organized in a hierarchical structure of a finite concepts set that describe a given structure of its concepts interaction and can be represented in the form of the taxonomy;
- the taxonomy can be represented as an oriented graph without cycles, where the vertices define the concept and the arcs define the relationship between the concepts;
- concepts are agents for the descriptions of certain network documents that reflect the passive knowledge systems of a particular SA;
- all concepts and connections between them can be defined by a certain system of axioms;
- there is a formal theory that can define all concepts, their properties and functionality.

2.1. Taxonomy of the Choice Problem

In general, the ontology of the subject area is formally represented by an ordered trio [1, 6-8, 11] of the form (1).

$$O = \langle X, R, F, \tilde{A}, D, R_s \rangle \quad (1)$$

where X is a set of ontological system concepts and forms the basis of the SA operability; R is the set of connections between concepts, including their contexts; F is a set of interpretive functions and methods; \tilde{A} is a set of axioms defined by the SA systemology, that provides the formation of true predicative expressions, the arguments of which are concepts, their properties, relationships, functions and methods of ontology; D is a set of additional descriptions of SA concepts, that take into account their operability; R_s is a set of constraints that determines the scope of particular subject area conceptual structures.

One of the system components of the ontological system is a taxonomy [12], which reflects a certain hierarchy of concepts interaction. In this case, the hierarchy itself is set using binary relations that are interpreted by certain methods defined for this ontology.

Taxonomy \tilde{T} is a nonempty subset of the set of concepts X of the ontology (1) between which the nonempty set of ordering relations $\tilde{\rho}$ is established. To form a model of the ranking problem based on ontology, it's necessary to determine a transformation of the form (2).

$$\tilde{H}_{MR} : \langle X, R, F, \tilde{A}, D, R_s \rangle \rightarrow M_R \quad (2)$$

where M_R is the ranking problem model.

The ranking problem is to order alternatives with the assignment of a rating to each alternative – an integral assessment for all criteria (indicators). The rating determines a linear ordering that also allows to solve a choice problem (determining the best alternatives).

The formalization of the ranking problem can be given by the following mathematical expression (3).

$$M_R = \langle A, K, F_{crit}, G_{ra} \rangle \quad (3)$$

where $A = \{x_1, x_2, \dots, x_n\}$ is an ordered set of alternatives; K is the set of criteria; $F_{crit} : A \times K \rightarrow Q$ is a function that determines the values Q of alternatives by a certain criterion; G_{ra} is a ranking rule that allows setting of a linear order for alternatives.

Depending on the degree of formation of the ranking problem model, the following cases can be distinguished:

- 1) $M_R^{1A} (A \neq \emptyset, K = \emptyset, F_{crit} = \cdot, G_{ra} = \cdot)$ – only alternatives are defined.
- 2) $M_R^{1K} (A = \emptyset, K \neq \emptyset, F_{crit} = \cdot, G_{ra} = \cdot)$ – only criteria are defined.
- 3) $M_R^2 (A \neq \emptyset, K \neq \emptyset, F_{crit} = \cdot, G_{ra} = \cdot)$ – sets of alternatives and criteria are defined.
- 4) $M_R^3 (A \neq \emptyset, K \neq \emptyset, F_{crit} \neq \cdot, G_{ra} = \cdot)$ – values for criteria are set for alternatives. In this model, it's assumed that the ranking problem is formed.
- 5) $M_R^4 (A \neq \emptyset, K \neq \emptyset, F_{crit} \neq \cdot, G_{ra} \neq \cdot)$ – the ranking rule is set for the formed ranking model.

In fact, the mechanism of modelling the ranking problem and its solution based on ontology is a multi-stage process, each one requires separate procedures. The general scheme of such a transformation can be represented by the following process:

$$O \rightarrow M_R^{1A} \text{ \AA } M_R^2 \rightarrow M_R^3 \rightarrow M_R^4 \rightarrow O^* \quad (4)$$

where O^* is an extension of the ontology O , which defines new relations of advantages and properties for the concepts of ontology (alternatives) $A \subseteq X$. It should be also noted that the new properties can be established as a result of post-analysis based on the solution of the inverse ranking problem.

Thus, the first stage of forming an ontological model of the ranking problem is to distinguish elements that can be considered as alternatives among the concepts X . Alternatives can be considered homogeneous objects with common properties. Such choice can be based on the choice function (5).

$$F_{sel}(X) = Y \quad (5)$$

where Y is a set of selected objects, which is a subset of X .

The classical rational choice functions include [8, 13-15] those with the following properties:

- 1) Condition of imitation (H):

$$X' \subseteq X \Rightarrow F_{sel}(X') \supseteq F_{sel}(X) \cap X' \quad (6)$$

- 2) Condition of constancy (K), strict imitation:

$$\forall X' \subseteq X, X' \cap F_{sel}(X) \neq \emptyset \Rightarrow F_{sel}(X') = F_{sel}(X) \cap X' \quad (7)$$

- 3) Condition of consent (C):

$$X' \cup X'' = X \Rightarrow F_{sel}(X) \supseteq F_{sel}(X') \cap F_{sel}(X'') \quad (8)$$

- 4) Condition of independence or rejection (R):

$$F_{sel}(X) \subseteq X' \subseteq X \Rightarrow F_{sel}(X') = F_{sel}(X) \quad (9)$$

To construct a class of choice mechanisms that generate a characteristic region of rejection (B) in the function space, a *hyperrelation* $Y \gamma Z$ is required [14, 16], which connects pairs of sets or their elements. Then the choice rule on such a structure is determined by the expression (10).

$$\forall Z \subseteq X, F_{sel}(X) \gamma Z \quad (10)$$

If the hyperrelation γ has the property of hypertransitivity, then the selection function also belongs to the characteristic region of inheritance. Property of hypertransitivity means that for any sets Y and X , and for any element z a condition (11) is met.

$$\forall x \in X, Y \gamma x \wedge X \gamma z \Rightarrow ((Y \cup X) \setminus \{x\}) \gamma z \quad (11)$$

The main feature of alternatives is their homogeneity. Therefore, one of the ways to define choice function is using the hyperdominant mechanism for selecting optimal sets. In this case the hyperrelation γ can be defined by a function Φ , which returns a scalar value for a subset of objects (12).

$$Y \gamma Z \Leftrightarrow \Phi(Y) \geq \Phi(Z) \quad (12)$$

Then the selection function will have a form (13)–(14).

$$\Phi(Y) = \max_{Z \subseteq Y} \Phi(Z) \quad (13)$$

where Y is the result of choice function (5);

$$\Phi(Z) = GH_1 \left(|Z|, \left| \bigcap_{z \in Z} P_z \right| - \left| \bigcup_{z \in Z} P_z \right| + \hat{q}, \left| \bigcap_{z \in Z} P_z \right| - \tilde{q} \right), \quad Z \subseteq X \quad (14)$$

where function $GH_1(x, y, z) = x(\text{sign}(y) + \text{sign}(z))$, \hat{q} , \tilde{q} are integer threshold values, P_z is a set of element's properties z ($z \in Z$), X is a set of objects.

The function GH_1 (14) can be defined in another way, which would take into account not only the presence of common properties, but also their quantity. This selection function focuses on finding the maximum number of homogeneous objects. If there is more than one group of such objects in the ontology, then due to the consistent application of this mechanism, they can also be identified.

2.2. Ontological Representation of Alternatives

After defining choice function and selecting alternatives, the next step in forming a model of the ranking problem is to establish criteria for alternatives (15).

$$K = \bigcup_{z \in A} P_z \quad (15)$$

where A is a set of alternatives, P_z is the set of properties of the alternative z ; K – a set of criteria.

Thus the values of attributes should be represented as the ternary relation in the ontological model (16).

$$(x, y, F_{crit}(x, y)) \in R_v \quad (16)$$

where x is an alternative from the set A ; y is a criterion from the set K ; F_{crit} is a function, which defines the criterion value for given alternative; R_v are the trinary relations given for the set X .

After defining a set of criteria and alternatives, and their criteria values, it can be assumed that the model of the ranking problem is actually set. Further refinement and resolution depend on the direct involvement of the statement.

Based on the formal definition of ontology (1) and the category of choice, which is represented by expression (10), we present the ontology of the choice problem [7, 8,] as the expression (17).

$$O_{TPCh} = \langle \tilde{T}, (\tilde{\rho} \cup R_v), F_{sel}(\tilde{T}) \rangle \quad (17)$$

where \tilde{T} is a taxonomy; $\tilde{\rho}$ is a partial order relation over the set \tilde{T} ; R_v are the properties of objects of the ontology given by relation "object"- "attribute"- "value"; F_{sel} is a choice function.

The object components that determine the ontology of the choice problem, allow us to interpret them as different information resources, the concepts of which are associated with certain binary relations of partial order and may have unary properties that characterize them in a certain qualitative form.

Now the mathematical model of the ranking problem (3) can be directly specified. Each alternative $x_i \in A$ is defined by the values of some set of indicator functions $\{f_i(x)\}$. Each such function specifies the value of a criterion [8, 13-15, 17], which belongs either to a predetermined set or is calculated according to certain mathematical rules [14, 15, 18]. In the first case, the options are possible: the set of values is given by a point or linguistic scale [19] or as a numerical interval. An example of the second case is the synthesis of local priorities in the method of analysis of hierarchies [8, 14, 15]. So, the problem of ranking alternatives by a set of indicators is to establish a certain order (18), based on the calculation of the some general indicator $G_{ra}(x)$ values (19).

$$xGy \Leftrightarrow G_{ra}(x) \geq G_{ra}(y) \quad (18)$$

$$G_{ra}(x) = G_{ra}(f_1(x), \dots, f_m(x), \omega_1, \dots, \omega_m) \quad (19)$$

where G is a reflexive, antisymmetric and transitive binary relation defined for the set A ; $G_{ra}(x)$ is a general indicator; $f_i(x)$ are criterion indicator function; ω_i are weights, indicating an importance of certain criterion [13-15, 17].

Defining global indicator allows defining linear order between alternatives, which allows to determine the best alternative.

3. Ontological Decision-Making Support System Based on Solving the Rational Choice Problem

The solution of rational choice problems based on means of a ranking of alternatives on a set of indicators is the system for the solution of a ranking problem based on the ontological model of the subject area.

From the point of view of software engineering, a software system is considered as a set of descriptions presented as mathematical models, formalisms and modelling techniques.

The general information model of the ontological system for solving the ranking problems S_A looks like (20). It's represented by some finite set of integrated S_{A_i} software modules.

$$S_A = \sum_{i=1}^n S_{A_i} \quad (20)$$

The set of modules used by this system is a subset of modules, provided by "Polyhedron" information technology [7]. This set has a form (21).

$$S_A = \{S_{MO}, S_{MT}, S_{OA}, S_{AL}, S_{CL}\} \quad (21)$$

where S_{MO} is a module responsible for representing the ontology and performing basic operations on it [4,6,7]; S_{MT} is a module responsible for representing the model of the problem of ranking alternatives, main parts of which are alternatives and criteria; S_{OA} is a module responsible for the analysis of the ontology based on its information model; S_{AL} is a module responsible for actual solving of a ranking problem; S_{CL} is a module responsible for preparing data for display in the user's web interface and for processing his requests, if necessary – with usage of other auxiliary service modules.

The main user of the system is decision maker, who is provided with the means to form and solve the ranking alternatives problems.

4. Conclusions

An ontological methodology based on an object-oriented approach represents the subject area to which specific application tasks belong, in the form of a set of objects with certain attributes and relationships. Therefore, the effective use of the decision support system to solve, in particular, the problems of selection and ranking of alternatives, is not possible without the development and use of formal mechanisms for building an information environment of the system from the ontological model

of the subject area. This possibility is achieved under the conditions of non-empty transformation of the taxonomy into a set of alternatives and non-empty reflection of the set of properties of objects, which make up the ontology, into a set of criteria. Hyperproperties of ontological representation of large amounts of information ensure the implementation of meta-tasks of its analysis, its structuring, synthesis, and support the process of rational choice. Such an approach creates the conditions for optimal processing of a large number of diverse information arrays. Usage of provided models and described information system allows to reduce the complexity of the decision-making process for experts, including the initial processing of unstructured information and following process of selecting the best alternatives.

References

- [1] Guarino N 1997 Understanding, Building and Using Ontologies *Human-Computer Studies* **46** pp 293–310
- [2] Sismondo S 2015 Ontological turns, turnoffs and roundabouts *Social Studies of Science* **45** 441–448tba
- [3] Buckingham Shum S, Motta E and Domingue J 2000 ScholOnto: an ontology-based digital library server for research documents and discourse *Int. J. Digit. Libr.* **3** 237–248
- [4] Gladun V 1994 *Processes of Formation of New Knowledge* (Sofia: SD Pedagogue) p 192
- [5] Dovgyi S, Stryzhak O 2021 Transdisciplinary Fundamentals of Information-Analytical Activity. In: Ilchenko M, Uryvsky L, Globa L (eds) *Advances in Information and Communication Technology and Systems. MCT 2019. Lecture Notes in Networks and Systems*, vol 152. Springer, Cham. https://doi.org/10.1007/978-3-030-58359-0_7
- [6] Palagin A, Rippa S and Sachenko A 2008 Conceptualization and problems of ontologies *Artificial Intelligence* **3**, 374
- [7] Stryzhak O, Prykhodniuk V and Podlipaiev V 2019 Model of transdisciplinary representation of GEOspatial information *Advances in Information and Communication Technologies* ed Ilchenko M, Uryvsky L et al. (Springer) **560** 34–75
- [8] Gorborukov V, Stryzhak O and Franchuk O 2013 Use of ontologies in decision-support systems. *Mathematical Modeling in Economics: Collection of Scientific Works* **3** 33–39
- [9] Booch G, Jakobson I and Rumbaugh J 2007 *Object-Oriented Analysis and Design with Applications* (Boston: Addison-Wesley) p 717
- [10] Weisfeld M 2013 *The Object-Oriented Thought Process* (London: Pearson education) p 307
- [11] Gruber T R 1993 A translation approach to portable ontology specifications *Knowledge Acquisition* **5** 199–220
- [12] Shatalkin A I 2012 *Taxonomy. Grounds, Principles and Rules* (Moscow: Association of scientific publications of KMK) p 600
- [13] Aizermann M A and Aleskerov F T 1990 Selection of variants: theory basics (Moscow: Science) p 240
- [14] Malyshevsky A 1998 *Qualitative Models in the Theory of Complex Systems* (Moscow: Fizmatlit) p 528
- [15] Saaty T L 2001 *Decision Making with Dependence and Feedback. The Analytic Network Process* (Pittsburgh, Pennsylvania: RWS Publications) p 370
- [16] Barendregt H 2012 *The Lambda Calculus. Its Syntax and Semantics* (London: King's College) p 611
- [17] Steuer R 1986 *Multiple Criteria Optimization: Theory, Computation, and Application* (Hoboken, New Jersey: Wiley) p 546
- [18] Keeney R L and Raiffa H 1993 *Decisions with Multiple Objectives: Preferences and Value Trade-Offs* (Cambridge: Cambridge University Press) p 593
- [19] Gray B 2015 *Linguistic Variation in Research Articles. When Discipline Tells Only Part of the Story* (Ames: Iowa State University) p 222