Structure of knowledge bases of next-generation intelligent computer systems: a hierarchical system of subject domains and their corresponding ontologies

Kseniya Bantsevich
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
Email: ksusha.bantsevich@gmail.com

***Abstract*—The article is dedicated to the ontological approach to the design of knowledge bases of next-generation intelligent computer systems. This approach is based on the representation of the knowledge base as a hierarchical structure of interrelated subject domains and their ontologies built on the basis of top-level ontologies.**

***Keywords*—knowledge base, ontology, top-level ontology, ontological approach to knowledge base design, knowledge, structure, semantic neighborhood, subject domain.**

I. Introduction

The development of information technologies has led to the expansion of the variety of information used and, as a result, to the need to create intelligent systems capable of operating voluminous information resources. The most important types of such resources are knowledge bases.

The knowledge base is a systematized totality of knowledge stored in the memory of an intelligent computer system and sufficient to ensure the purposeful (appropriate, adequate) functioning (behavior) of this system both in its external and internal environment (in its own knowledge base).

An important stage in the development of knowledge bases of intelligent systems is their structuring. Structuring the database, i.e. the allocation of various interconnected substructures in it, is necessary for a number of reasons. In particular, this is necessary to ensure their syntactic compatibility, which implies the unification of the form of knowledge representation.

II. Analysis of existing approaches to solving the problem

To date, there are dozens of models of knowledge representation, each of which is adapted to represent a certain kind of knowledge, while when creating intelligent systems, it often becomes necessary to represent different types of knowledge within a single base. However, currently, none of the existing models, taken separately, can provide this.

In this regard, there is a need to create a universal structured model of knowledge representation, which would allow representing any kind of knowledge in a unified form.

Today, ontologies are the most effective means of structuring various areas of knowledge. The essence of the ontological approach when designing the knowledge base is to consider the structure of the knowledge base as a hierarchical system of allocated subject domains and their corresponding ontologies [1]. However, ontologically, there are many ways in which it is possible to describe the real world as it is. The solution to this problem is the usage of top-level ontologies [2] in the design of knowledge bases of intelligent computer systems.

A competently constructed top-level ontology will allow for broad syntactic compatibility between a large number of ontologies for various subject domains, since the terms of domain-oriented ontologies are subordinate to the terms of higher-level ontology.

At the moment, there are several developed top-level ontologies [3], [4]:

- **Descriptive Ontology for Linguistic and Cognitive Engineering** (DOLCE) [5]
  - The DOLCE ontology is focused on embracing the ontological categories underlying natural language and human common sense.
- **The Standard Upper Ontology** (SUMO) [6]
  - The SUMO ontology was created by combining publicly available ontological contents into a single, comprehensive, and coherent structure.
  - The ontology covers the following areas of knowledge: general types of processes and objects, abstractions (set theory, attributes, relations), numbers and units of measurement, temporal concepts, parts and a whole, agents and intentions.
- **Cyc's upper ontology** (OpenCyc) [7], [8]

- The key concept in the OpenCyc ontology is a collection, which can contain subcollections and instances, which, in turn, can act as any terms of the ontology.
- The OpenCyc knowledge base contains information from various subject domains: Philosophy, Mathematics, Chemistry, Biology, Psychology, Linguistics.

The list represented is not final.

There are more than fifteen top-level ontologies [4], the purpose for creation of which is to use them when creating lower-level ontologies. However, attempts to create a universal top-level ontology capable of ensuring the compatibility of intelligent computer systems have not led to the expected results, as they have a number of key disadvantages:

- Each of the represented ontologies is a monolithic structure in which there is no clear localization into separate small ontologies.
  The main problem in designing fragments of knowledge bases using the ontological approach is to allocate ontologies in such a way that they allow for the relatively independent evolution of each fragment. The data structure of top-level ontologies is a hierarchy consisting of a large number of different concepts. This type of structuring leads to a situation where the need to make changes in one place will necessarily entail the impossibility of editing another part of the ontology. Due to the above, this type of structuring makes ontologies inconvenient for their usage in the development of various intelligent systems.
- The top-level ontologies in question are not part of a complex technology.
  Since the named ontologies are not part of some complex technology, they cannot be considered as part of a library of reusable components, which, in turn, leads to inconveniences in the form of the need to adapt the ontologies used for each specific system.
- There are no knowledge base design technologies based on top-level ontology data.
  The lack of knowledge base design technologies makes it difficult to develop intelligent systems.

The lack of a satisfactory solution to these problems leads to incompatibility of the developed intelligent computer systems. Based on the above, there is a need to build such a system of top-level ontologies that could provide syntactic compatibility between a large number of ontologies of various subject domains in knowledge bases of intelligent computer systems.

III. Proposed approach

Within this work, it is proposed to take as a basis the approach developed within the *OSTIS Technology* [9] to the development of knowledge bases of next-generation intelligent computer systems. The proposed models are based on the following basic principles of the *OSTIS Technology*:

- the usage of an ontological approach to the design of knowledge bases, which involves structuring the knowledge base grounded on ontologies;
- focus on the possibility of collective design of knowledge bases within the project;
- orientation to the semantic representation of knowledge;
- unification of knowledge base models of intelligent systems.

To solve the above problems, it is necessary:

- to formally clarify and coordinate the interpretation of such concepts as *structure*, *semantic neighborhood*, *subject domain*, *ontology*, since these concepts are the basic classes of entities that form the basis for structuring knowledge bases of intelligent systems;
- to develop top-level ontological models for structuring the knowledge base grounded on the allocated concepts.

The ontological model built on the basis of these concepts will become the Kernel of the knowledge base, ensuring the compatibility of intelligent systems due to the unified representation of knowledge. It should be noted that, depending on the specifics of the systems being developed, their knowledge bases may expand, however, the ontological model underlying the *Kernel*, will ensure further compatibility of the systems being developed.

The approach proposed in this article is based on the ideas of building systems based on semantic networks implemented in the *OSTIS Technology*. This technology is a complex of models, tools, and methods designed for the development of intelligent computer systems, as well as for the constant updating and improvement of the technology itself.

The *OSTIS* Technology is based on the usage of unified semantic networks with a basic set-theoretic interpretation of their elements as a method of knowledge representation. This way of knowledge representation is called an *SC-code*, and the semantic networks, represented in the *SC-code*, are called *sc-graphs* (*sc-texts*, or *texts of the SC-code*). The elements of such semantic networks are called *sc-elements* (*sc-nodes* and *sc-connectors*, which, in turn, can be *sc-arcs* or *sc-edges* depending on their orientation). The *Alphabet of the SC-code* consists of five basic elements, on the basis of which SC-code constructions of any complexity are built, including the introduction of more particular kinds of sc-elements (e.g., new concepts). The memory storing SC-code constructions is called semantic memory, or *sc-memory*.

The key feature of the *SC-code* is the joint usage of the mathematical apparatus of a graph theory and a set theory. This allows, on the one hand, ensuring the strictness and universatility of formalization tools and, on the other hand, ensuring the convenience of storing and processing information represented in this form.

Within the technology, several universal variants of visualization of the *SC-code* constructions are also proposed, such as *SCg-code* (graphic version), *SCn-code* (non-linear hypertextual version), *SCs-code* (linear string version).

The basis of the knowledge base within the *OSTIS Technology* is a hierarchical system of subject domains and ontologies.

Within this article, fragments of structured texts in the SCn-code [10] will often be used, which are simultaneously fragments of source texts of the knowledge base, which are understandable both to a human and to a machine. This allows making the text more structured and formalized while maintaining its readability. The symbol ":=" in such texts indicates alternative (synonymous) names of the described entity, which reveal in more detail some of its features.

Next, we will take a closer look at the fragments of sc-models of these top-level ontologies proposed within the *OSTIS Technology*.

IV. Concept of knowledge and formal models of ostis-systems knowledge bases

Within the model of the ostis-systems knowledge base, syntactically correct (for the corresponding language) and semantically integral information constructions are distinguished. We will call such constructions *knowledge*.

knowledge
⊂ *information construction*
⇒ *coverage**:
  knowledge type
  := [Set of various knowledge types]

The fact that a family of knowledge types is a covering of a Set of various knowledge means that each knowledge belongs to at least one knowledge type that we have identified.

knowledge type
∋ *specification*
  := [description of the specified entity]
  ⊃ *specification of a material entity*
  ⊃ *specification of an inverse entity that is not a set*
    ⊃ *specification of a geometric point*
    ⊃ *specification of the number*
  ⊃ *specification of the set*
    ⊃ *connection specification*
    ⊃ *structure specification*
    ⊃ *class specification*
      ⊃ *specification of a class of entities that are not sets*
      ⊃ *relation specification*
      ⊃ *specification of the class of classes*
      ⊃ *specification of the class of structures*
      ⊃ *specification of concepts*
  ⊃ *semantic neighborhood*
  ⊃ *unique specification*
∋ *formal theory*
∋ *subject domain*
∋ *subject domain and ontology*
  := [subject domain and its ontology]
  := [subject domain and its corresponding unified ontology]
∋ *meta-knowledge*
  := [knowledge specification]
  ⊃ *ontology*
    ⊃ *ontology of the subject domain*
      ⊃ *structural ontology of the subject domain*
      ⊃ *set-theoretic ontology of the subject domain*
      ⊃ *logical ontology of the subject domain*
      ⊃ *terminological ontology of the subject domain*
      ⊃ *unified ontology of the subject domain*
∋ *problem*
∋ *plan*
∋ *protocol*
∋ *method*
∋ *technology*
∋ *knowledge base*

Even a small list of *knowledge types* indicates a huge variety of *knowledge types*.

Knowledge is divided into *declarative* and *procedural*. *Declarative knowledge* is understood as *knowledge* that has only *denotational semantics*, which is represented as a semantic *specification* of the concepts system used in this *knowledge*. *Procedural knowledge* is meant as *knowledge* that has not only *denotational semantics* but also *operational semantics*, which is represented as a family of *agents specifications* that interpret *procedural knowledge* aimed at solving some initiated problem.

Within the *OSTIS Technolog*y, relations defined on a set of knowledge are also distinguished.

relation defined on a set of knowledge
∋ *child knowledge**
  := [knowledge that inherits from the "maternal" knowledge all the properties of the research objects described there]
  ⊃ *child section**

⊃ *private subject domain and ontology**

∋ *specification**

:= [be knowledge, which is a specification (description) of a given entity]

∋ *ontology**

:= [be a semantic specification of a given knowledge*]

∋ *semantic equivalency**

∋ *therefore**

∋ *logical equivalency**

V. Concept of a structure

Existing approaches to the development of *knowledge bases* are grounded on considering specific elements of the knowledge base (classes, instances, relations, etc.) as specification objects. However, with the accumulation of large amounts of information in the *knowledge base*, it becomes necessary to allocate entire fragments of the *knowledge base* and be able to specify them, considering them as separate entities. The concept of a *structure* is the basis for the representation of *knowledge*, *meta-knowledge*, and their structuring.

The concept of a *structure* is one of the most general concepts (from the point of view of clarifying semantics) when describing the properties of an object.

By *structure* we mean a set of *sc-elements*, the removal of one of which may lead to a violation of the integrity of this set.

Let us consider the typology of the *structures* described in the *knowledge base*:

structure

⇒ *subdividing**:

{• *connected structure*

• *disconnected structure*

}

The structure represented in the SC-code will be matched with an orgraph whose vertices are sc-elements and arcs are connectives of incident relations connecting sc-connectors with incident sc-elements, which are components of these sc-connectors. If the orgraph obtained in this way is a connected orgraph, then the initial structure will be considered a *connected structure*. If the orgraph obtained in this way is not a connected orgraph, then the initial structure will be considered a *disconnected structure*.

structure

⇒ *subdividing**:

{• *trivial structure*

• *non-trivial structure*

}

A *trivial structure* is a *structure* that does not contain connectives as elements. In turn, a *non-trivial structure* means a *structure*, among the elements of which there is at least one connective.

On the basis of stationarity/nonstationarity, *dynamic structures* (processes) are distinguished – *structures* whose composition changes over time, as well as *static structures* – *structures* whose composition does not change over time.

structure

⇒ *subdividing**:

{• *process*

:= [dynamic structure]

• *static structure*

:= [stationary structure]

}

structure

⇒ *subdividing**:

{• *temporal structure*

• *permanent structure*

}

For the formal representation of *structures*, concepts describing elements within the structure were introduced:

structure element′

⇒ *subdividing**:

{• *unrepresented set′*

:= [set that is not represented within this structure′]

• *fully represented set′*

:= [set fully represented within this structure′]

• *partially represented set′*

:= [set partially represented within this structure′]

• *structure element that is not a set′*

}

A number of correspondences can be determined between structures, such as *homomorphism*, *polymorphism*, *automorphism*, *isomorphism*, as well as *similarity of structures*, which allows fixing the fact that there is some analogy, similarities, and differences of some substructures of the *structures* under consideration.

VI. Concept of a semantic neighborhood

For the specification of particular entities within the knowledge base, the concept of a *semantic neighborhood* is introduced. A *semantic neighborhood* is a specification (description) of a given entity, the sign of which is indicated as a key element of this specification.

The set of attributes by which entities can be specified is different. In addition, it may be necessary to specify the same entity in different aspects and explicitly record these aspects in the knowledge base.

A *semantic neighborhood* is a specification of a given entity, the sign of which is indicated as a key element of this specification. Unlike other knowledge types, semantic neighborhood has only one key element.

There are complete and basic semantic neighborhoods.

complete semantic neighborhood
:= [full specification of some described entity]

The structure of the *full semantic neighborhood* is determined primarily by the semantic typology of the entity being described. So, for example, for a concept, it is necessary to include the following information in the *full semantic neighborhood* (if available):

- identification options in various external languages (sc-identifiers);
- belonging to a certain subject domain with an indication of the role performed within this subject domain;
- set-theoretic connections of a given concept with other sc-elements;
- definition or explanation;
- propositions describing the properties of the specified concept;
- problems and their classes, in which this concept is a key element;
- description of a typical example for using this concept;
- instances of the described concept.

For a concept that is a relation, the following is additionally specified:

- domains;
- scope of definition;
- relation diagram;
- classes of relations to which the described relation belongs.

basic semantic neighborhood
:= [minimally sufficient semantic neighborhood]
:= [minimum specification of the described entity]

The structure of the *basic semantic neighborhood* is determined primarily by the semantic typology of the entity being described. For example, for a concept, the following information should be included in the *basic semantic neighborhood* (if available):

- identification options in various external languages (sc-identifiers);
- belonging to a certain subject domain with an indication of the role performed within this subject domain;
- definition or explanation.

For a concept that is a relation, the following is additionally specified:

- domains;
- scope of definition;
- description for a typical example of a connective of the specified relation (specification of a typical instance).

Also, a *specialized semantic neighborhood* is distinguished – a type of the *semantic neighborhood*, the set of relations for which is specified separately for each type of such a neighborhood.

specialized semantic neighborhood
$\supset$ *explanation*
$\supset$ *note*
$\supset$ *description of a typical instance*

The concept of a *semantic neighborhood*, supplemented by the clarification of such concepts as semantic distance between signs (semantic proximity of signs), the radius of the semantic neighborhood, is a promising basis for the study of the properties of semantic space.

VII. Concept of a subject domain

The most important stage in the development of knowledge bases is the process of identifying the *subject domains* described and their representation in the knowledge base.

The concept of the ***subject domain*** is the most important methodological technique that allows distinguishing only a certain class of entities under study and only a certain family of relations set on the specified class from the whole variety of the World, that is, localization is carried out, focusing attention only on this, abstracting from the rest of the studied World.

Each *subject domain* can be matched to:

- a family of corresponding ontologies of different types;
- a set of semantic neighborhoods describing the research objects in this subject domain.

A ***Subject domain*** is a ***structure***, which includes:

- the main research (description) objects – primary and secondary ones;
- various classes of research objects;
- various connectives whose components are the research objects (both primary and secondary ones), and possibly other such connectives, that is, the connectives (as well as the research objects) may have different structural levels;
- different classes of the above connectives (i.e., relations);
- different classes of objects that are neither research objects nor the above-mentioned connectives, but are components of these connectives.

At the same time, all classes declared by the concepts under study must be fully represented within this subject domain together with their elements, elements of elements, etc. up to terminal elements.

Each knowledge type can be matched with a subject domain, which is the result of integrating all knowledge of this type. This knowledge becomes the research object within the specified subject domain.

subject domain
:= [connections system of a certain set of research objects, the *key elements* of which are:

- classes (more precisely, class signs) of research objects (objects described by this subject domain);
- specific research objects with special properties;
- classes of connections that are part of the system under consideration – relations defined on the set of elements of the system under consideration;
- parameters specified on a set of elements of the system under consideration;
- classes of structures that are fragments of the system under consideration.

]
:= [structure representing a set of connections (more precisely, the signs of connections) and the corresponding set of components of these connections, which include:

- elements (instances) of some specified classes of research objects (primary entities under study);
- the connections themselves that are part of the specified structure;
- introduced classes of research objects;
- introduced relations (connection classes);
- introduced parameters (classes of equivalent entity classes);
- parameter values (and, in particular, values for the measured parameters);
- introduced structures that are fragments (substructures) of the structure under consideration;
- introduced classes of substructures of the structure under consideration.

]

The *subject domains* allocated within the *knowledge base* of the intelligent system and their corresponding *ontologies* are a kind of semantic strata, clusters that allow "decomposing" all *knowledge* stored in memory on "semantic shelves" in the presence of clear criteria that allow unambiguously determining on which "shelf" should this or that *knowledge* be placed.

According to the level of research attention, concepts within the subject domain can perform the following roles:

role of the subject domain element
:= [role relation that links subject domains with their key signs]
:= [role of the key element (the sign of the key entities) of the subject domain]
:= [role of the key sign of the subject domain]
∋ *class of research objects′*
  := [be a class of primary (for a given subject domain) research objects′]
∋ *maximum class of research objects′*
  := [class of research objects for which in the specified (!) subject domain there is no other class of research objects that would be its superset′]
∋ *key research object′*
  := [special research object′]
  := [be a sign of a special research object within a given subject domain′]
  := [research object with special properties′]
∋ *concept used in the subject domain′*
  := [concept used in a given subject domain not as one of the research objects but as a key concept′]
∋ *primary research element of the subject domain′*
  := [sign of the primary research object within a given subject domain′]
∋ *secondary research element of the subject domain′*
  := [sign of the secondary research object within the subject domain′]
∋ *non-investigated element of the subject domain′*
  := [auxiliary element of a subject domain being investigated in another (adjacent) subject domain′]

The following types of subject domains are distinguished:

subject domain
⇒ *subdividing**:
  {• *static subject domain*
    := [stationary subject domain]
    := [*subject domain*, in which the relations between the entities that are part of it do not depend on time (do not change in time); *temporal entities* cannot be the elements of the ***static subject domain***]
  • *quasi-static subject domain*
    := [*subject domain*, the solution of problems in which does not require taking into account the temporal properties of research objects]
  • *dynamic subject domain*

:= [non-stationary subject domain]
:= [*subject domain*, which describes a change in the state (including the internal structure) of research objects and/or a change in the configuration of connections between research objects]
:= [*subject domain*, in which some relations between entities that are part of it change over time (that is, they are situational, non-stationary in nature, in other words, they are *temporal entities*)]
}
⇒ *subdividing**:
{• *primary subject domain*
:= [*subject domain*, the research objects of which are external entities (denoted by primary *sc-elements*)]
• *secondary subject domain*
:= [meta-subject domain]
:= [*subject domain*, the research objects of which are *sc-sets* (relations, parameters, structures, classes of structures, knowledge, languages, etc.)]
}

In all the variety of subject domains, a special place is occupied by:

- the ***Subject domain of subject domains***, the research objects of which are all kinds of subject domains and the research subject are all kinds of role relations linking subject domains with their elements, relations linking subject domains with each other, relation linking subject domains with their ontologies;
- the ***Subject domain of entities***, which is the subject domain of the highest level and defines the basic semantic typology of sc-elements (signs included in the texts of the SC-code);
- a family of *subject domains*, each of which defines the semantics and syntax of some *sc-language* that provides a representation of *ontologies* of the appropriate type (for example, set-theoretic ontologies of terminological ontologies);
- a family of *top-level subject domains*, in which the classes of research objects are very "large" entity classes. Such classes, in particular, include:
  – class of various material entities,
  – class of various sets,
  – class of various connections,
  – class of various relations,
  – class of various structures,
  – class of various temporal (non-stationary) entities,
  – class of various actions (influences),
  – class of various parameters (characteristics),
  – class of various knowledge,
  – etc.

It is important to note that a *subject domain* can also be considered as a *semantic neighborhood* if we consider its center to be the sign of an entity that is the maximum class of research objects.

VIII. Concept of an ontology

For the formal specification of the corresponding subject domain, focused on the description of the properties and relations of the concepts that make up the specified subject domain, such a knowledge type as *ontology* is used.

Ontologies are the most important *knowledge type*, providing semantic systematization of *knowledge* stored in memory of *intelligent computer systems* (including *ostis-systems*) and, accordingly, semantic structuring of *knowledge bases*.

ontology
:= [sc-ontology]
:= [semantic specification of any knowledge having a sufficiently complex structure, of any integral fragment of the knowledge base — a subject domain, a method for solving complex problems of a certain class, a description for the history of a certain activity type, a description for the scope of a certain set of actions (problem-solving areas), a representation language for problem-solving methods, etc.]
:= [semantic *specification* of some enough informative resource (knowledge)]
⊂ *specification*
⊂ *meta-knowledge*
∈ *knowledge type*
:= [most important *meta-knowledge* type included in the knowledge base]
:= [specification (clarification) of the *concepts* that system used in the corresponding (specified) *knowledge*]

The ontology includes:

- the typology of the specified knowledge;
- connections of the specified knowledge with other knowledge;
- the specification of key concepts used in the specified knowledge, as well as key instances of some such concepts.

It is important to note that if a *specification* can specify (describe) any *entity*, then an *ontology* specifies only various *knowledge*. At the same time, the most important objects of such a specification are *subject domains*.

The main *purpose* of building *ontology* is semantic clarification (explanation, and ideally definition) of such a family of *signs* used in given *knowledge*, which are sufficient to understand the meaning of all specified *knowledge*. As it turns out, the number of characters whose meaning determines the meaning of all specified *knowledge* is not large.

ontology
⇒ *subdividing**:
{• *informal ontology*
• *formal ontology*
:= [ontology represented in a formal language]
:= [formal description of the denotational semantics (semantic interpretation) of the specified knowledge]
}

Obviously, in the absence of sufficiently complete formal ontologies, it is impossible to ensure semantic compatibility (integrability) of various knowledge stored in the knowledge base, as well as acquired from the outside.

An *ontology* is most often interpreted as a specification of conceptualization (specification of a *concepts* system) of a given *subject domain*. Here we mean a description of the set-theoretic relations (first of all, the classification) of the *concepts* used, as well as a description of various regularities for entities belonging to these *concepts*. However, important types of the *subject domain* specification are also:

- a description of the relations of the specified *subject domain* with other *subject domains*;
- a description of the terminology of the specified *subject domain*.

ontology of the subject domain
:= [description of the *denotational semantics* of the language being defined (set) by the corresponding (specified) *subject domain*]
:= [information suprastructure (meta-information) over the corresponding (specified) *subject domain*, describing various aspects of this *subject domain* as a sufficiently large, self-sufficient, and semantically integral fragment of the *knowledge base*]
:= [meta-information (meta-knowledge) about some *subject domain*]

The *ontology of the subject domain* can be interpreted, on the one hand, as a *semantic neighborhood* of the corresponding *subject domain* and, on the other hand, as a *combination* of a certain type of *semantic neighborhoods* of all *concepts* used within the specified *subject domain*, as well as possibly key instances of the specified *concepts*, if there are any.

Each specific ontology of a given type is a semantic neighborhood of the corresponding (specified) subject domain. Each *ontology type* uniquely corresponds to a *subject domain*, fragments of which are specific *ontologies* of this type. Consequently, each *ontology type* has its own specialized sc-language that provides a representation of *ontologies* of this type.

ontology of the subject domain
⇒ *subdividing**:
{• *particular ontology of the subject domain*
:= [*ontology* representing the specification of the relevant subject domain in one aspect or another]
• *unified ontology of the subject domain*
:= [ontology of the *subject domain*, which is the result of combining all known *particular ontologies* of this subject domain]
}

Each *particular ontology* is a fragment of a *subject domain*, which includes all (!) particular ontologies belonging to the corresponding *ontology type*. At the same time, the specified *subject domain*, in turn, also has a corresponding *ontology*, which is no longer a meta-knowledge (like any ontology) but a meta-meta-knowledge (a specification of meta-knowledge).

particular ontology of the subject domain
⇒ *subdividing**:
{• *structural specification of the subject domain*
:= [*meta-knowledge* type describing the properties of *subject domains* corresponding to this meta-knowledge type]
:= [scheme of the subject domain]
• *set-theoretic ontology of the subject domain*
:= [sc-specification of a given subject domain within the *subject domain of sets*]
• *logical ontology of the subject domain*
:= [sc-text of the formal theory of a given subject domain]
• *terminological ontology of the subject domain*
}

structural specification of the subject domain
:= [structural ontology of the subject domain]
:= [role structure of the key elements of the subject domain]

:= [scheme of the concepts roles of the subject domain and its relation with related subject domains]
:= [scheme of the subject domain]
:= [specification of the subject domain from the point of view of graph theory and theory of *algebraic systems*]
:= [description of the internal (role) structure of the *subject domain*, as well as its external relations with other *subject domains*]
:= [description for the roles of the key elements of the subject domain (first of all, concepts), as well as the "place" of the specified subject domain in the set of similar ones]
:= [*semantic neighborhood* of the *subject domain* sign within this *subject domain* itself, which includes all *key signs* that are part of the *subject domain* (key concepts and key objects of subject domain research) with an indication of their roles (properties) within this *subject domain*, and the *semantic neighborhood* of the sign of the specified *subject domain* within the *Subject domain of subject domains*, including the relations of the specified *subject domain* with other semantically close to it *subject domains* (private and maternal, similar in one sense or another (for example, isomorphic), having the same *classes of research objects* or the same sets of *relations under study*)]

set-theoretic ontology of the subject domain
:= [*semantic neighborhood* of the specified *subject domain* within the *subject domain of sets*, describing the set-theoretic relations between *concepts* of the specified *domain*, including the relations of *relations* with their *definition areas*, and *domains*, the relations of the *parameters* used, and the classes of their *areas of definition*]
:= [ontology that describes:
- □ a classification of research objects of the specified subject domain;
- □ the correlation of the areas of definition and domains of the relations used with the selected classes of research objects, as well as with the selected classes of auxiliary (adjacent) objects that are not research objects in the specified subject domain;
- □ a specification of the relations used and, in particular, an indication of whether all connectives of these relations are part of the specified subject domain.

]

The set-theoretic ontology of the subject domain includes:
- set-theoretic connections (including taxonomy) between all the concepts used, which are part of the specified subject domain;
- a set-theoretic specification of all *relations* that are part of the specified subject domain (orientation, arity, area of definition, domains, etc.);
- a set-theoretic specification of all parameters used in the subject domain (parameter definition areas, scales, units of measurement, reference points);
- a set-theoretic specification of all classes of structures used.

logical ontology of the subject domain
:= [formal theory of a given (specified) domain, describing various properties of concepts instances used in the specified subject domain with the help of variables, quantifiers, logical connectives, formulas]

The logical ontology of the subject domain includes:
- a formal definitions of all concepts that are definable within the specified subject domain;
- informal explanations and some formal specifications (at least examples) for all concepts that are indefinable within the specified subject domain;
- a hierarchical concepts system, in which for each concept studied in the specified subject domain, either the fact of the indefinability of this concept is indicated, or all the concepts on the basis of which the definition of this concept is given are indicated. As a result, the set of concepts under study is divided into a number of levels:
  - undefined concepts;
  - concepts of the 1st level, defined only on the basis of undefined concepts;
  - concepts of the 2nd level, defined on the basis of concepts that change the 1st level ones and below;
  - etc.
- a formal record of all axioms, i.e. propositions that do not require proof;
- a formal record of propositions whose truth requires justification (proof);
- formal texts of proving the truth of propositions, which are a specification for the sequence of steps of the corresponding reasoning (steps of logical inference, the application of various logical inference rules);
- a hierarchical system of propositions, in which for each proposition, true in relation to the specified subject domain, either the axiomaticity of this proposition is indicated, or all propositions are listed, on the basis of which this proposition is proved. As a result, the set of propositions that are true in relation to the specified subject domain is divided into a number of levels:
  - axioms;

- propositions of the 1st level, proved only on the basis of axioms;
- propositions of the 2nd level, proved on the basis of propositions that are at the 1st level and below.

- a formal record of hypothetical propositions;
- a formal description of the logical-semantic typology of propositions – propositions about existence, non-existence, uniqueness, propositions of a defining type (which can be used as definitions of the corresponding concepts);
- a formal description of various types of logical-semantic relations between propositions (for example, between an utterance and its generalization);
- a formal description of the analogy:
  - between definitions;
  - between propositions of any kind;
  - between proofs of different propositions.

terminological ontology of the subject domain
:= [ontology describing rules for constructing terms (sc-identifiers) corresponding to sc-elements belonging to the specified subject domain, as well as describing various kinds of terminological relations between the terms used, characterizing the origin of these terms]
:= [system of terms of a given subject domain]
:= [thesaurus of the relevant subject domain]
:= [dictionary of the relevant (specified) subject domain]
:= [fragment of the global *Subject domain of sc-identifiers* (external identifiers of sc-elements), providing a terminological specification of some subject domain]

Now let us take a closer look at the concept of a *unified ontology of the subject domain*.

unified ontology of the subject domain
:= [combination of all particular ontologies corresponding to the same subject domain]
⇐ *generalized combination**:
{• *structural specification of the subject domain*
• *set-theoretic ontology of the subject domain*
• *logical ontology of the subject domain*
• *terminological ontology of the subject domain*
}

subject domain and ontology
:= [integration of some *subject domain* with the corresponding *unified ontology*]
:= [subject domain & ontology]
⇐ *generalized combination**:
{• *subject domain*
• *unified ontology of the subject domain*
}
:= [sc-text that is a combination of some subject domain represented in the SC-code and the combined ontology of this subject domain, also represented in the SC-code]
:= [integration of the subject domain and all ontologies specifying this subject domain]
:= [set of various *facts* about the structure of some activity domain for some *subjects*, as well as various types of *knowledge* specifying this field of activity]
:= [facts and knowledge about a certain field of activity]
:= [sc-model of the subject domain and various onthologies specifying this subject domain (and, first of all, its key concepts) from different angles]
:= [coherent fragment of the ostis-system knowledge base from a logical and semantic point of view, focusing on a specific class of research objects and on a specific aspect of their consideration]

Subject domains and ontologies are the main *type of knowledge base sections*, having a high degree of their independence from each other and clear rules for their coordination, which ensures their semantic (understandable) compatibility within the entire *knowledge base*.

IX. Subject domains of the represented concepts

Each of the represented concepts corresponds to *subject domains and ontologies*, in which this concept is the maximum class of research objects:

- **Subject domain of knowledge and ostis-systems knowledge bases**

Subject domain of knowledge and ostis-systems knowledge bases
∈ *subject domain*
∋ *maximum class of research objects'*:
knowledge
∋ *class of research objects'*:
• *knowledge type*
• *relation defined on a set of knowledge*

- **Subject domain of structures**

Subject domain of structures
∈ *subject domain*
∋ *maximum class of research objects'*:
structure
∋ *class of research objects'*:
• *connected structure*

- *disconnected structure*
- *trivial structure*
- *nontrivial structure*

∋ *relation under study'*:
- *structure element'*
- *unrepresented set'*
- *fully represented set'*
- *structure element that is not a set'*
- *polymorphism**
- *homomorphism**
- *isomorphism**
- *similarity of structures**

- **Subject domain of semantic neighborhoods**

Subject domain of semantic neighborhoods
∈ *subject domain*
∋ *maximum class of research objects'*:
semantic neighborhood
∋ *class of research objects'*:
- *full semantic neighborhood*
- *basic semantic neighborhood*
- *specialized semantic neighborhood*
- *terminological semantic neighborhood*
- *explanation*
- *note*
- *set-theoretic semantic neighborhood*
- *logical semantic neighborhood*

- **Subject domain of subject domains**

The *Subject domain of subject domains* includes the structural specifications of all *subject domains* that are part of the *ostis-system* knowledge base, including the *Subject domain of subject domains* itself. Thus, the *Subject domain of subject domains* is, firstly, a *reflexive set* and, secondly, a reflexive subject domain, that is, a *subject domain*, one of the research objects of which is itself.

Subject domain of subject domains
:= [Subject domain, the research objects of which are subject domains]
∈ *reflexive set*
∋ *maximum class of research objects'*:
subject domain
∋ *class of research objects'*:
- *static subject domain*
- *dynamic subject domain*
- *concept*
- *nontrivial structure*

∋ *relation under study'*:
- *concept under study'*
- *maximum class of research objects'*
- *non-maximum class of research objects'*
- *class of structures under study'*
- *private subject domain**
- *concept studied in the private subject domain'*

- **Subject domain of ontologies**

Subject domain of ontologies
∈ *subject domain*
∋ *maximum class of research objects'*:
ontology
∋ *class of research objects'*:
- *structural specification of the subject domain*
- *particular ontology of the subject domain*
- *unified ontology of the subject domain*
- *set-theoretic ontology of the subject domain*
- *logical ontology of the subject domain*
- *ontology of the subject domain*

X. Conclusion

In the article, an ontological approach to the design of knowledge bases of next-generation intelligent computer systems is considered. This approach is based on the representation of the knowledge base of intelligent computer systems based on the OSTIS Technology as a hierarchical structure of interrelated subject domains and their ontologies built on the basis of top-level ontologies.

The formal interpretation of such concepts as knowledge, structure, semantic neighborhood, subject domain, ontology has been clarified, which together made it possible to determine on their basis the ontological model of knowledge bases of next-generation intelligent computer systems.

This model can form the Kernel of the knowledge base, which will ensure the compatibility of intelligent systems due to the unified representation of knowledge.

The results obtained make it possible to increase the efficiency of the development of next-generation intelligent systems due to the component approach to the development of knowledge bases and automation tools for their development.

Acknowledgment

The author would like to thank the research group of the Departments of Intelligent Information Technologies of the Belarusian State University of Informatics and Radioelectronics for its help in the work and valuable comments.

References

[1] I. Davydenko, "Ontology-based knowledge base design," *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]*, pp. 42–50, 2018.
[2] S. J.F, "Top-level ontological categories," *International Journal of Human-Computer Studies*, vol. 43, no. 5-6, pp. 669–685, 1995.
[3] C. Partridge, *A survey of Top-Level Ontologies*. The Construction Innovation Hub, 2020.

[4] V. Mascardi, *A Comparison of Upper Ontologies*. DBLP, 2007.
[5] (2022, Nov) DOLCE: Descriptive Ontology for Linguistic and Cognitive Engineering. [Online]. Available: http://www.loa.istc.cnr.it/dolce/overview.html
[6] (2022, Nov) Suggested Upper Merged Ontology (SUMO). [Online]. Available: https://github.com/ontologyportal/sumo
[7] (2022, Nov) Cycorp – cycorp making solutions better. [Online]. Available: https://cyc.com/
[8] (2022, Nov) Github: Opencyc. [Online]. Available: https://github.com/asanchez75/opencyc
[9] V. Golenkov, N. Gulyakina, and D. Shunkevich, *Otkrytaja tehnologija ontologicheskogo proektirovanija, proizvodstva i jekspluatacii semanticheski sovmestimyh gibridnyh intellektual'nyh komp'juternyh sistem [Open technology of ontological design, production and operation of semantically compatible hybrid intelligent computer systems]*. Bestprint [Bestprint], 2021.
[10] (2022, Nov) IMS.ostis Metasystem. [Online]. Available: https://ims.ostis.net

Структура баз знаний интеллектуальных компьютерных систем нового поколения: иерархическая система предметных областей и соответствующих им онтологий

Банцевич К.А.

В работе рассмотрен онтологический подход к проектированию баз знаний интеллектуальных компьютерных систем нового поколения. Данный подход основан на представлении базы знаний интеллектуальных компьютерных систем на основе Технологии OSTIS как иерархической структуры взаимосвязанных предметных областей и их онтологий, построенных на базе онтологий верхнего уровня.

Уточнена формальная трактовка таких понятий, как знание, структура, семантическая окрестность, предметная область, онтология, что в совокупности позволило определить на их основе онтологическую модель баз знаний интеллектуальных компьютерных систем нового поколения.

Данная модель может составить Ядро базы знаний, которое позволит обеспечить совместимость интеллектуальных системы за счет унифицированного представления знаний.

Полученные результаты позволяют повысить эффективность разработки интеллектуальных систем нового поколения за счет компонентного подхода к разработке баз знаний и средств автоматизации их разработки.

Received 17.11.2022