



OSTIS-2016

(Open Semantic Technologies for Intelligent Systems)

УДК 004.822:514

СРЕДСТВА СТРУКТУРИЗАЦИИ СЕМАНТИЧЕСКИХ МОДЕЛЕЙ БАЗ ЗНАНИЙ

Давыденко И.Т.* , Гракова Н.В.* , Сергиенко Е.С.** , Федотова А.В.**

* *Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь*

ir.davydenko@gmail.com
natalia.grakova@gmail.com

** *Национальный исследовательский ядерный университет «МИФИ»,
г. Москва, Россия*

deav@inbox.ru

*** *Московский государственный технический университет им. Н.Э.Баумана,
г. Москва, Россия*

afedotova.bmstu@gmail.com

В работе рассмотрены средства структуризации семантических моделей баз знаний, основанные на унифицированных семантических сетях. Основное внимание уделяется рассмотрению понятий структура, знание, рассматривается их классификация. Отдельный раздел посвящен описанию таких видов знаний, как предметная область и онтологии.

Ключевые слова: база знаний, структуризация базы знаний, предметная область

Введение

При накоплении больших объемов знаний появляется огромное количество самых различных *понятий* (концептов) и возникает проблема структуризации и систематизации этого многообразия понятий.

Одним из базовых принципов *Технологии OSTIS* (Open Semantic Technology for Intelligent Systems) является принцип унификации структуризации баз знаний [PMS, 2016]. В рамках данного принципа предполагается трактовать семантическую структуру базы знаний компьютерной системы как отражение иерархической системы взаимосвязанных друг с другом *предметных областей*, представляемых в базе знаний. Это предполагает:

- (1) уточнение понятия предметной области;
- (2) разработку языковых средств описания структуры предметных областей с помощью унифицированных семантических сетей;
- (3) разработку языковых средств описания типологии предметных областей и различных видов связей между ними.

Структуризация базы знаний, выделение в ней различных связанных между собой подструктур необходимы по целому ряду причин [Гаврилова и др., 2001]. В частности, это необходимо для дидактических целей (человеку усваивающему некоторые знания, желательно иметь, своего рода оглавление этих знаний, что позволяет планировать их усвоение и рассматривать их с различной степенью детализации), а также для организации распределения работ по проектированию баз знаний (когда разным исполнителям поручается разработка разных фрагментов базы знаний, имеющих достаточно четкие границы) [Голенков, 2012].

Таким образом, в данной работе вводится и уточняется одно из базовых с точки зрения представления знаний в *Технологии OSTIS*, понятие *структуры*, на основе которого далее уточняется понятие *знания*, одним из видов которого является понятие *предметной области*. Для спецификации предметных областей вводится понятие *онтологии* и рассматривается их типология [Tarassov, 2015], [Кудрявцев, 2010], [Федотова, 2015].

В качестве формальной основы для представления знаний в рамках *Технологии OSTIS* используются графодинамические модели

специального вида – семантические модели представления и обработки знаний, в основе которых лежат унифицированные семантические сети с базовой теоретико-множественной интерпретацией. Основным способом кодирования информации для таких сетей является SC-код (Semantic Code) [Голенков, 2012]. Элементы таких сетей, представленных в SC-коде будем называть *sc-элементами*, в свою очередь, узлы таких сетей будем называть *sc-узлами*, связи между ними – *sc-коннекторами*. [IMS, 2016]

Некоторые из рассматриваемых в данной работе понятий будут сопровождаться фрагментами их спецификации на формальном языке SCn (Semantic Code natural). [IMS, 2016]

1. Предметная область структур

В данном разделе вводится понятие *структуры*, рассматривается классификация структур, типология ролей элементов в рамках структур.

1.1. Понятие структуры и типология структур

Одним из наиболее общих (с точки зрения уточнения семантики) понятий при описании свойств какого-либо объекта является понятие *структуры*.

Каждая *структура* представляет собой множество *sc-элементов*, удаление одного из которых может привести к нарушению целостности этого множества.

Пример изображения структуры на языке SCg [IMS, 2016] приведен на рисунке 1.

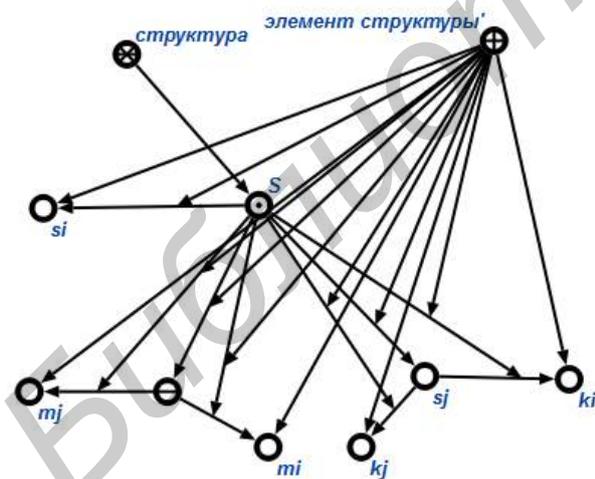


Рисунок 1 - Пример структуры на языке SCg

Выделяются следующие классы структур:

структура

= *sc-структура*

= *структура, представленная в виде текста SC-кода*

<= разбиение*:

{

• *связная структура*

• *несвязная структура*

}

<= разбиение*:

{

• *тривиальная структура*

• *нетривиальная структура*

}

Структуре, представленной в SC-коде, поставим в соответствие орграф, вершинами которого являются *sc-элементы*, а дугами – пары инцидентности, связывающие *sc-коннекторы* с инцидентными им *sc-элементами*, которые являются компонентами указанных *sc-коннекторов*.

Если полученный таким способом орграф является связным орграфом, то исходную структуру будем считать *связной структурой*.

Структуре, представленной в SC-коде, поставим в соответствие орграф, вершинами которого являются *sc-элементы*, а дугами – пары инцидентности, связывающие *sc-коннекторы* с инцидентными им *sc-элементами*, которые являются компонентами указанных *sc-коннекторов*. Если полученный таким способом орграф не является связным орграфом, то исходную структуру будем считать *несвязной структурой*.

Под *тривиальной структурой* понимается *структура*, не содержащая в качестве элементов *связок*.

В свою очередь, под *нетривиальной структурой* понимается *структура*, среди элементов которой есть хотя бы одна *связка*.

1.2. Типология элементов структур

В рамках заданной структуры ее элементы можно классифицировать по следующим признакам:

• насколько полно в рамках заданной структуры представлено множество, обозначаемое заданным sc-элементом вместе с соответствующими дугами принадлежности;

• существуют ли в рамках заданной структуры *sc-элементы*, обозначающие множества, являющиеся надмножествами того множества, которое обозначается заданным sc-элементом;

• уровень («этаж») иерархии перехода от знаков к метазнакам для заданного sc-элемента в рамках заданной *структуры*.

Для формального представления структур были введены понятия, описывающие роли элементов в рамках структуры. Рассмотрим их подробнее.

элемент структуры' – ролевое отношение, указывающее на все элементы каждой структуры.

элемент структуры'

<= разбиение*:

{

• *непредставленное множество'*

- *полностью представленное множество'*
 - *частично представленное множество'*
 - *элемент структуры, не являющийся множеством'*
- }
 <= разбиение*:
 {
 • *максимальное множество'*
 • *немаксимальное множество'*
 }

непредставленное множество' – *ролевое отношение*, связывающее структуру со знаком множества, все элементы которого не являются элементами данной структуры.

непредставленное множество'
 = *множество, не представленное в рамках данной структуры'*
 = *быть знаком множества, элементы которого не являются элементами данной структуры'*
 ∈ *ролевое отношение*

полностью представленное множество' – *ролевое отношение*, связывающее структуру со знаком множества (любого семантического типа – класса, связки или структуры), все элементы которого являются элементами данной структуры.

полностью представленное множество'
 = *множество, полностью представленное в рамках данной структуры'*
 = *множество, все элементы которого являются элементами данной структуры'*
 = *полностью представленный класс'*
 ∈ *ролевое отношение*

частично представленное множество' – *ролевое отношение*, связывающее структуру со знаком множества, не все элементы которого являются элементами данной структуры.

частично представленное множество'
 = *множество, частично представленное в рамках данной структуры'*
 = *множество, некоторые элементы которого являются элементами данной структуры'*
 = *быть знаком множества, некоторые элементы которого являются элементами данной структуры'*
 ∈ *ролевое отношение*

элемент структуры, не являющийся множеством' - *ролевое отношение*, связывающее структуру с знаком некоторой сущности, не являющейся множеством.

максимальное множество' – *ролевое отношение*, связывающее структуру со знаком множества, для которого не существует множества, которое было бы надмножеством указанного множества и знак которого был бы элементом этой же структуры.

немаксимальное множество' – *ролевое отношение*, связывающее структуру со знаком

множества, для которого в рамках данной структуры существует множество, являющееся надмножеством указанного множества.

первичный элемент' – *ролевое отношение*, указывающее на элемент структуры, являющийся либо терминальным элементом, либо знаком множества, такого что не существует другого элемента этой же структуры, который был бы элементом множества, обозначаемого первым из указанных элементов структуры. При этом соответствующая пара принадлежности может существовать, но в состав данной структуры не входит.

первичный элемент'
 = *первичный элемент данной структуры'*
 = *sc-элемент первого уровня в рамках данной структуры'*
 ∈ *ролевое отношение*
 ∈ *семантический уровень структурного элемента*
 ⊂ *элемент структуры'*

вторичный элемент' – *ролевое отношение*, указывающее на элемент структуры, обозначающий множество, все или некоторые элементы которого являются элементами указанной структуры.

вторичный элемент'
 = *вторичный элемент данной структуры'*
 = *элемент данной структуры, имеющий семантический уровень более 2'*
 = *непервичный элемент'*
 ∈ *ролевое отношение*
 ⊂ *элемент структуры'*
 ⊃ *элемент второго уровня'*

элементом второго уровня' в рамках заданной структуры может быть связка первичных элементов, тривиальная структура из первичных элементов или класс первичных элементов.

элемент второго уровня'
 ∈ *ролевое отношение*
 ∈ *семантический уровень структурного элемента*

В зависимости от того, элементы каких уровней представлены в рамках заданной структуры выделяются дополнительные классы структур, в частности, **структура второго уровня** – структура, среди элементов которой есть хотя бы один элемент второго уровня'.

Для того, чтобы описать, элементы каких уровней представлены в рамках структуры, вводится понятие **семантический уровень структурного элемента**, который представляет собой параметр, каждый элемент которого является классом *sc-дуг принадлежности*, связывающих некоторую структуру с теми ее элементами, которые имеют одинаковый семантический уровень в рамках данной структуры. Значением данного параметра является число, обозначающее указанный семантический уровень.

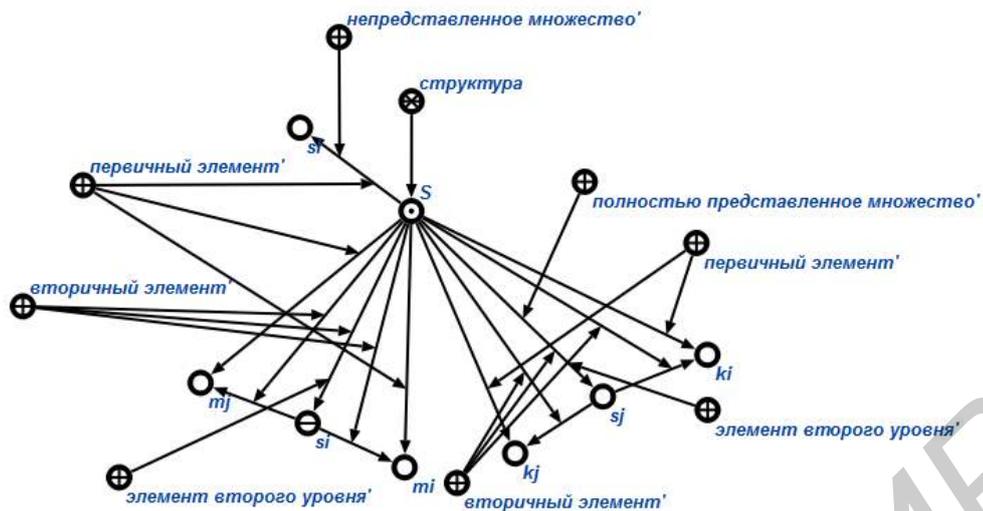


Рисунок 2 – Пример описания элементов структуры на SCg

семантический уровень структурного элемента вычисляется следующим образом:

- элементы структуры, входящие в нее с атрибутом *первичный элемент'* имеют семантический уровень 1;
- уровень элемента, не являющегося *первичным элементом'* структуры, вычисляется путем прибавления 1 к максимальному из уровней элементов этого элемента (множества), входящих в эту же структуру. Например, *sc-дуга*, соединяющая два *первичных элемента'* структуры будет иметь семантический уровень 2, а *sc-элемент*, обозначающий отношение, которому принадлежит указанная *sc-дуга* – семантический уровень 3.



Рисунок 3 – Семантический уровень структурного элемента

В свою очередь, **количество семантических уровней элементов структуры** – параметр, каждый элемент которого представляет собой класс *структур*, у которых совпадает максимальный среди семантических уровней элементов этих структур. Значением данного параметра является число, совпадающее с указанным максимальным семантическим уровнем элементов.

1.3. Соответствия, заданные на структурах

На структурах может быть задан ряд *соответствий**, таких как *гомоморфизм**, *полиморфизм**, *автоморфизм**, *изоморфизм** и т.д.

Отдельное внимание стоит уделить соответствию *аналогичность структур**, которое фиксирует факт наличия некоторой аналогии на подструктурах (подмножествах) указанных структур. Каждой ориентированной паре, принадлежащей *аналогичности структур** может быть поставлено в соответствие множество пар, задающих *сходства** некоторых подструктур и *различия** некоторых подструктур исходных структур.

Пример отношения *аналогичности структур** приведен на рисунке 4.

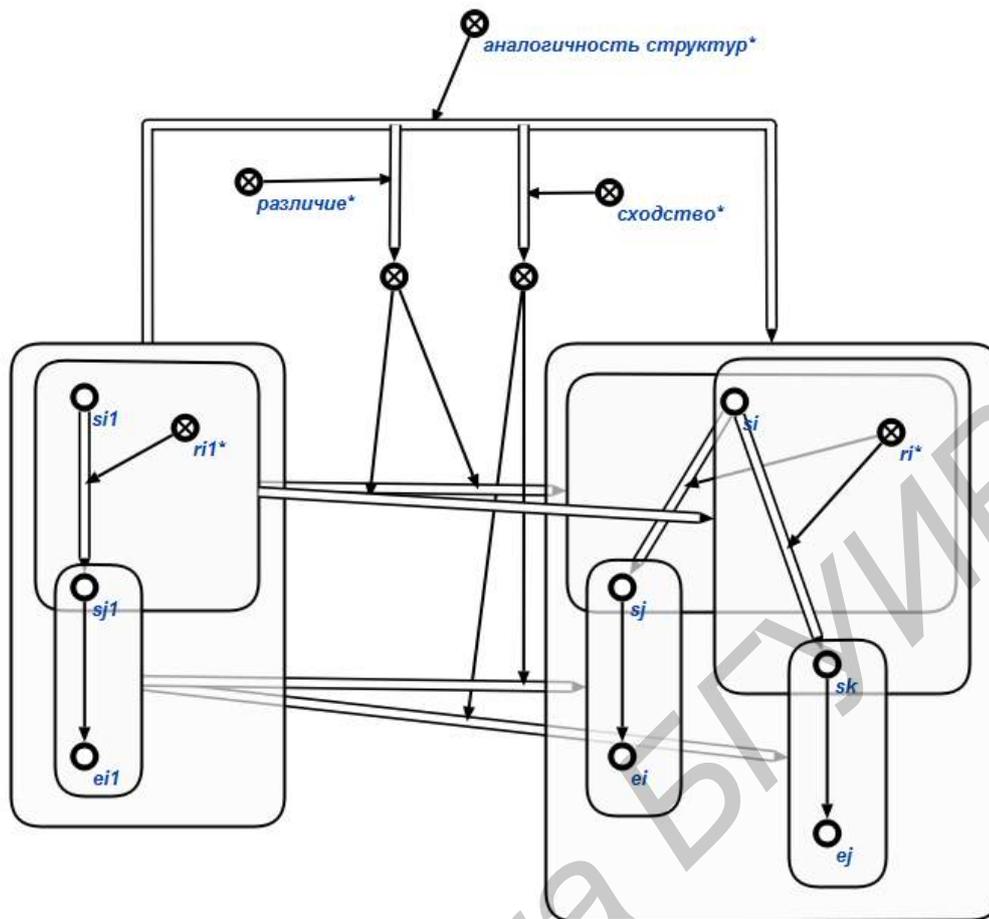


Рисунок 4 - Пример отношения аналогичности структур

2. Предметная область знаний

В рамках базы знаний будем выделять семантически осмысленные *структуры*, обладающие некоторой семантической целостностью. Такие структуры будем называть *знаниями*.

Рассмотрим типологию знаний:

знание

- = *sc-знание*
- = Множество всевозможных знаний
- = *sc-знание* или целостный фрагмент *sc-знания*
- ⊂ структура
- ⊃ семантическая окрестность
- ⊃ сравнение
- ⊃ фактографическое знание
- ⊃ раздел
- ⊃ предметная область
- ⊃ онтология
- ⊃ задача
- ⊃ программа
- ⊃ план
- ⊃ решение
- ⊃ утверждение
- ⊃ определение

Важнейшим отношением, заданным на множестве знаний, является отношение *быть метазнанием**, описывающее переход от знаний к описывающим их метазнаниям [Гаврилова и др., 2001].

метазнание*

- = *быть метазнанием**
- = Бинарное ориентированное отношение, связки которого связывают некоторое исходное знание со знанием, которое является метаотношением исходного знания, его спецификацией, описанием его структуры*

Примером связи между знанием и соответствующим ему *метазнанием** является переход от некоторого исходного знания к описанию его декомпозиции (сегментации) на некоторые части с указанием связей между этими частями.

В *Предметной области знаний* исследуется ряд ролевых отношений, уточняющих семантику *sc-элементов* структур, являющихся *знанием*, в частности, *ключевой sc-элемент* и *главный ключевой sc-элемент*.

ключевой sc-элемент' – ролевое отношение, связывающее знак каждой структуры (текста, знания) со специально выделяемыми (ключевыми) элементами этой структуры.

Каждая структура может иметь либо несколько ключевых элементов, либо один ключевой элемент, либо ни одного.

ключевой sc-элемент'

- = *быть ключевым sc-элементом заданного sc-знания'*
- = *ключевой знак'*
- = *ключевой элемент sc-знания'*
- = *быть ключевым элементом заданной sc-структуры*
- = *быть ключевым элементом заданной sc-знания*
- ⊃ *основной sc-элемент'*
- ∈ *ролевое отношение*

главный ключевой sc-элемент' – *ролевое отношение*, связывающее знак некоторой структуры (текста, знания) со специально выделяемыми главными элементами этой структуры. Главный элемент некоторой структуры является ее *ключевым sc-элементом'*, но имеющим особый статус по каким-либо причинам, зависящим в каждом случае от семантики данного элемента и типа структуры.

главный ключевой sc-элемент'

- = *быть ключевым sc-элементом заданного sc-знания'*
- ⊂ *ключевой sc-элемент'*

Рассмотрим более подробно виды знаний, необходимые для описания структуры базы знаний, такие как *раздел*, *семантическая окрестность*, *предметная область* и *онтология*.

2.1. Понятие раздела базы знаний и их типология

База знаний каждой *ostis-системы* декомпозируется на *разделы*.

Понятие *раздела* представляет собой знак множества всевозможных разделов, входящих в состав различных баз знаний. Каждый раздел является условно дидактически выделяемым фрагментом базы знаний, обладающим логической целостностью и завершенностью. Вся база знаний конкретной *ostis-системы* также может рассматриваться как один неатомарный *раздел*.

На понятии раздела заданы отношения *декомпозиция раздела** и *базовый порядок разделов**.

декомпозиция раздела* – это *квазибинарное отношение* между разделом и множеством его подразделов. Данное отношение задает дидактическую структуру раздела. В отличие от отношения *базовая декомпозиция**, которое связывает некоторую сущность с другими сущностями, являющимися её частями, отношение **декомпозиция раздела*** связывает раздел с его подразделами, т.е. сужается область определения отношения *базовая декомпозиция**.

базовый порядок разделов* – это бинарное отношение между разделами, определяющее

порядок их следования в рамках декомпозиции раздела.

Данное отношение задает дидактический порядок следования подразделов в рамках декомпозиции более общего раздела.

Выделяется следующая типология разделов базы знаний:

раздел

- = *раздел базы знаний*
- = *sc-модель раздела базы знаний*
- <= *разбиение**:

- {
- *атомарный раздел*
- *неатомарный раздел*
- }

- ⊃ *раздел-обоснование*
- ⊃ *раздел-описание принципов*
- ⊃ *раздел-документация*
- ⊃ *раздел-описание предметной области*
- ⊃ *раздел-описание семантической окрестности*
- ⊃ *раздел-теория предметной области*

Под **неатомарным разделом** понимается знак множества всевозможных неатомарных разделов, входящих в состав различных документаций, то есть разделов, которые декомпозируются на более частные разделы.

Соответственно, под **атомарным разделом** понимается знак множества всевозможных атомарных разделов, входящих в состав различных документаций, то есть разделов, не декомпозируемых на более частные разделы.

Любое понятие, входящее в состав атомарного раздела и не являющееся ключевым в рамках этого раздела должно стать ключевым в рамках какого-либо другого атомарного раздела рассматриваемой базы знаний.

Для каждого раздела необходимо явно указать принадлежность к множеству атомарных или неатомарных разделов.

С точки зрения описываемой информации можно выделить следующие классы разделов:

- **раздел-обоснование** представляет собой формальный текст, обосновывающий разработку чего-либо, то есть постановку проблемы, указание недостатков существующих решений такого рода, достоинства предлагаемых решений, предполагаемый эффект от применения разработанных моделей, методов и средств и т.д.

- **раздел-документация** содержит документацию к чему-либо, как правило – к *ostis-системе* или ее компонентам. Документация описывает назначение системы или подсистемы, принципы работы с той или иной системой или подсистемой, ее структуру и состав. В случае многократно используемого компонента документация также содержит руководство по доработке и использованию такого компонента.

- **раздел-описание предметной области** содержит знак самой этой *предметной области*, знаки всех ее элементов, а также всю спецификацию этой *предметной области*, включая все ее онтологии, в том числе – логическую онтологию, содержащую описание формальной теории, соответствующей данной *предметной области*.

- **раздел-описание семантической окрестности** содержит знак *структуры*, которая является семантической окрестностью некоторого *sc-элемента*, знаки всех ее элементов, а так же спецификацию этой *структуры*, например, указание принадлежности более частному классу семантических окрестностей.

- **раздел-описание принципов** содержит описание основополагающих принципов устройства либо функционирования какой-либо технологии, подсистемы, компонента и т.д.

2.2. Отношения, задающие дополнительную спецификацию разделов базы знаний

Любой *раздел* базы знаний может дополнительно быть специфицирован с использованием следующих отношений:

- **комментарий*** – бинарное отношение, связывающее некоторую сущность (*sc-текст* или знак файла) со знаком *sc-текста*, являющегося комментарием к этой сущности.

- **аннотация*** – бинарное отношение, связывающее некоторый *раздел* и *sc-текст*, являющийся аннотацией к данному *разделу*. Как правило, аннотация содержит краткое (недетализированное) описание того, чему посвящен данный *раздел*.

- **введение*** – бинарное отношение, связывающее некоторый *раздел* и *sc-текст*, являющийся введением к данному *разделу*. Как правило, введение содержит информацию о том, как сущности, рассматриваемые в данном *разделе*, связаны с сущностями, описанными в других *разделах*.

- **заключение*** – это бинарное отношение, связывающее некоторый *раздел* и *sc-текст*, являющийся заключением к данному *разделу*. В заключении, как правило, подводятся итоги и указываются основные положения, описанные в *разделе*, а также рассматриваются направления развития самого *раздела* и связанных с ним.

- **предисловие*** – это бинарное отношение, связывающее некоторый *раздел* и *sc-текст*, являющийся предисловием к данному *разделу*.

- **послесловие*** – это бинарное отношение, связывающее некоторый *раздел* и *sc-текст*, являющийся послесловием к данному *разделу*.

- **эпиграф*** – это бинарное отношение, связывающее некоторый *раздел* и *sc-текст*, являющийся эпиграфом к данному *разделу*.

2.3. Предметная область семантических окрестностей

Для спецификации отдельных *сущностей* в рамках базы знаний вводится понятие **семантической окрестности**.

семантическая окрестность – это знание, являющееся спецификацией (описанием) некоторой *сущности*, знак которой является **ключевым элементом** указанного знания. Заметим, что каждая **семантическая окрестность** в отличие от знаний других видов имеет только один ключевой элемент (ключевой знак, знак описываемой сущности). Также отметим, что многообразие видов семантических окрестностей свидетельствует о многообразии семантических видов описаний различных сущностей.

семантическая окрестность

= *sc-окрестность*

= *семантическая окрестность, представленная в виде sc-текста*

= *sc-текст, являющийся семантической окрестностью некоторого sc-элемента*

= *спецификация заданной сущности, знак которой указывается как ключевой элемент этой спецификации*

= *описание заданной сущности, знак которой указывается как ключевой элемент этой спецификации*

⊆ *знание*

⊃ *семантическая окрестность по инцидентным коннекторам*

⊃ *полная семантическая окрестность*

⊃ *базовая семантическая окрестность*

⊃ *специализированная семантическая окрестность*

Перечислим основные виды *семантических окрестностей*.

семантическая окрестность по инцидентным коннекторам – это вид *семантической окрестности*, в которую входят знаки всех коннекторов, инцидентных заданному элементу, а также знаки всех элементов, инцидентных указанным коннекторам.

семантическая окрестность по инцидентным коннекторам

⊃ *семантическая окрестность по выходящим дугам*

⊃ *семантическая окрестность по входящим дугам*

семантическая окрестность по выходящим дугам – это вид *семантической окрестности*, в которую входят знаки всех дуг, выходящих из заданного *sc-элемента*, а также знаки их вторых компонентов, также указывается факт принадлежности этих дуг каким-либо отношениям.

семантическая окрестность по выходящим дугам

⊃ *семантическая окрестность по выходящим дугам принадлежности*

семантическая окрестность по выходящим дугам принадлежности – это вид *семантической окрестности*, в которую входят знаки всех дуг принадлежности, выходящих из заданного *sc-*

элемента, а также знаки их вторых компонентов. При необходимости может указываться факт принадлежности этих дуг каким-либо ролевым отношениям.

семантическая окрестность по входящим дугам – это вид *семантической окрестности*, в которую входят знаки всех дуг, входящих в заданный *sc-элемент*, а также знаки их первых компонентов, также указывается факт принадлежности этих дуг каким-либо отношениям.

семантическая окрестность по входящим дугам
▷ *семантическая окрестность по входящим дугам принадлежности*

семантическая окрестность по входящим дугам принадлежности – это вид *семантической окрестности*, в которую входят знаки всех дуг принадлежности, входящих в заданный *sc-элемент*, а также знаки их первых компонентов. При необходимости может указываться факт принадлежности этих дуг каким-либо ролевым отношениям.

полная семантическая окрестность – это вид *семантической окрестности*, включающий описание всех связей описываемой сущности.

Структура *полной семантической окрестности* определяется прежде всего семантической типологией описываемой сущности.

Так, например, для *понятия* в полную семантическую окрестность необходимо включить следующую информацию (при наличии):

- варианты идентификации на различных внешних языках;
- принадлежность некоторой предметной области с указанием роли, выполняемой в рамках этой предметной области;
- теоретико-множественные связи заданного понятия с другими *sc-элементами*;
- определение или пояснение;
- высказывания, описывающие свойства указанного понятия;
- задачи и их классы, в которых данное понятие является ключевым
- описание типичного примера использования указанного понятия;
- экземпляры описываемого понятия.

Для *понятия*, являющегося отношением дополнительно указываются:

- домены;
- область определения;
- схема отношения;
- классы отношений, которым принадлежит описываемое отношение.

базовая семантическая окрестность – это вид *семантической окрестности*, содержащий минимальную (краткую) информацию об описываемой сущности

Структура *базовой семантической окрестности* определяется прежде всего семантической типологией описываемой сущности.

Так, например, для *понятия* в базовую семантическую окрестность необходимо включить следующую информацию (при наличии):

- варианты идентификации на различных внешних языках;
- принадлежность некоторой предметной области с указанием роли, выполняемой в рамках этой предметной области;
- определение или пояснение.

Для *понятия*, являющегося отношением дополнительно указываются:

- домены;
- область определения;
- описание типичного примера использования указанного отношения.

базовая семантическая окрестность

= минимально достаточная семантическая окрестность

= минимальная спецификация описываемой сущности

= сокращенная спецификация описываемой сущности

= основная семантическая окрестность

специализированная семантическая окрестность – это вид *семантической окрестности*, набор связей для которой уточняется отдельно для каждого класса такой окрестности.

специализированная семантическая окрестность

▷ *пояснение*

▷ *примечание*

▷ *правило идентификации экземпляров*

▷ *терминологическая семантическая окрестность*

▷ *теоретико-множественная семантическая окрестность*

▷ *логическая семантическая окрестность*

▷ *описание типичного экземпляра*

▷ *описание декомпозиции*

пояснение – знак *sc-текста*, поясняющего описываемую сущность.

примечание – знак *sc-текста*, являющегося примечанием к описываемой сущности. В примечании обычно описываются особые свойства и исключения из правил для описываемой сущности.

правило идентификации экземпляров – это *sc-текст* являющийся описанием правил построения идентификаторов элементов заданного класса.

терминологическая семантическая окрестность – семантическая окрестность, описывающая идентификацию указанной сущности.

теоретико-множественная семантическая окрестность – описание связи описываемого

понятия с другими понятиями с помощью теоретико-множественных отношений.

описание декомпозиции – семантическая окрестность, описывающая декомпозицию некоторой сущности на частные сущности.

логическая семантическая окрестность – семантическая окрестность, описывающая семейство высказываний, описывающих свойства данного понятия.

описание типичного экземпляра – это *sc-текст* являющийся описанием типичного примера использования рассматриваемого класса.

сравнительный анализ – описание сравнительного анализа некоторой сущности с другими сущностями.

3. Предметная область предметных областей

Предметная область – это важнейший вид знаний, входящих в состав смыслового пространства [Голенков, 2015]. Каждая предметная область фокусирует внимание на описание связей соответствующего класса объектов исследования. Каждый знак, входящий в состав базы знаний, должен принадлежать (входить в состав) хотя бы одной предметной области, выполняя в ней ту или иную роль. Каждой предметной области можно поставить в соответствие:

- множество семантических окрестностей, описывающих объекты исследования этой предметной области;
- семейство различного вида онтологий, описывающих свойства понятий этой предметной области.

Предметные области являются основой структуризации смыслового пространства, средством локализации, фокусирования внимания на свойства наиболее важных классов описываемых сущностей, которые становятся классами объектов исследования в *предметных областях*.

Понятие базы знаний тесно связано с понятием *предметной области*. Соотношение между базой знаний и описываемой ею предметной областью задает семантику базы знаний компьютерной системы.

Рассмотрение структуры базы знаний во взаимосвязи с предметной областью позволяет рассматривать исследуемые объекты на разных уровнях детализации. Детализацию рассмотрения исследуемых объектов можно осуществлять как в рамках исходной предметной области, так и в системе самостоятельных, но связанных между собой предметных областей.

При переходе от предметной области к ее модели, представленной в виде семантической сети, выполняются следующие условия:

- каждому элементу предметной области взаимно однозначно соответствует обозначающий его элемент семантической сети;

- каждому сигнатурному элементу предметной области взаимно однозначно соответствует либо обозначающий его ключевой узел семантической сети, либо обозначающий элемент алфавита семантической сети.

В состав **Предметной области предметных областей** входят структурные спецификации всех *предметных областей*, входящих в состав базы знаний *ostis-системы*, в том числе, самой **Предметной области предметных областей**. Таким образом, **Предметная область предметных областей** является, во-первых, *рефлексивным множеством*, во-вторых, рефлексивной предметной областью, то есть *предметной областью*, одним из объектов исследования которой является она сама.

Предметная область – это результат интеграции (объединения) частичных семантических окрестностей, описывающих все исследуемые сущности заданного класса и имеющих одинаковый (общий) предмет исследования (то есть один и тот же набор отношений, которым должны принадлежать связи, входящие в состав интегрируемых семантических окрестностей).

Предметная область представляет собой *структуру*, в состав которой входят:

- (1) основные исследуемые (описываемые) объекты – первичные и вторичные;
- (2) различные классы исследуемых объектов;
- (3) различные связи, компонентами которых являются исследуемые объекты (как первичные, так и вторичные), а также, возможно, другие такие связи – то есть связи (как и объекты исследования) могут иметь различный структурный уровень;
- (4) различные классы указанных выше связей (то есть отношения);
- (5) различные классы объектов, не являющихся ни объектами исследования, ни указанными выше связями, но являющихся компонентами этих связей.

При этом все классы, объявленные исследуемыми понятиями, должны быть полностью представлены в рамках данной предметной области вместе со своими элементами, элементами элементов и т.д. вплоть до терминальных элементов.

Понятие **предметной области** является важнейшим методологическим приемом, позволяющим выделить из всего многообразия исследуемого Мира только определенный класс исследуемых сущностей и только определенное семейство отношений, заданных на указанном классе. То есть осуществляется локализация, фокусирование внимания только на этом, абстрагируясь от всего остального исследуемого Мира. [Голенков, 2012]

Во всем многообразии **предметных областей** особое место занимают:

(1) **Предметная область предметных областей**, объектами исследования которой являются всевозможные **предметные области**, а предметом исследования – всевозможные **ролевые отношения**, связывающие предметные области с их элементами, отношения, связывающие предметные области между собой, отношения, связывающие предметные области с их онтологиями

(2) **Предметная область сущностей**, являющаяся предметной областью самого высокого уровня и задающая базовую семантическую типологию *sc-элементов* (знаков, входящих в тексты *SC-кода*)

(3) Семейство **предметных областей**, каждая из которых задает семантику и синтаксис некоторого *sc-языка*, обеспечивающего представление онтологий соответствующего вида (например, *теоретико-множественных онтологий*, *логических онтологий*, *терминологических онтологий*, *онтологий задач и способов их решения* и т.д.)

(4) Семейство **предметных областей** верхнего уровня, в которых классами объектов исследования являются весьма «крупные» классы сущностей. К таким классам, в частности, относятся:

- класс всевозможных *материальных сущностей*,
- класс всевозможных *множеств*,
- класс всевозможных *связей*,
- класс всевозможных *отношений*,
- класс всевозможных *структур*,
- класс всевозможных *временных (нестационарных) сущностей*,
- класс всевозможных *действий* (акций),
- класс всевозможных *параметров* (характеристик),
- класс *знаний* всевозможного вида
- и т.п.

Каждой **предметной области** можно поставить в соответствие:

(1) семейство соответствующих ей *онтологий* разного вида;

(2) некий *язык* (в нашем случае – язык, построенный на основе *SC-кода*), тексты которого представляют различные фрагменты соответствующей предметной области.

Указанные языки будем называть *sc-языками*. Их синтаксис и семантика полностью задается *SC-кодом* и *онтологией* соответствующей **предметной области**. Очевидно, что в первую очередь нас должны интересовать те *sc-языки*, которые соответствуют **предметным областям**, имеющим общий (условно говоря, предметно независимый) характер. К таким предметным областям, в частности, относятся:

- **Предметная область множеств**, описывающая множества и различные связи между ними
- **Предметная область графовых структур**
- **Предметная область чисел** и числовых структур
- и т.д.

Каждому типу знаний можно поставить в соответствие предметную область, которая является результатом интеграции всех знаний данного типа. Эти знания и становятся объектами исследования в рамках указанной предметной области

Понятие **предметной области** может рассматриваться как обобщение понятия алгебраической системы. При этом семантическая структура базы знаний может рассматриваться как иерархическая система различных **предметных областей**.

3.1. Типология предметных областей

Можно говорить о типологии **предметных областей** по разным структурным признакам:

- наличие метасвязей;
- наличие исследуемых структур, входящих в состав предметной области;
- наличие исследуемых (смежных, дополнительных) объектов, которых исследуются в других предметных областях.

Одним из основных критериев классификации **предметных областей** является зависимость от времени связей между сущностями, входящими в состав рассматриваемой **предметной области**.

предметная область

= *sc-модель предметной области*

= *sc-текст предметной области*

= *sc-граф предметной области*

= *представление предметной области в SC-коде*

⊆ *знание*

⊆ *бесконечное множество*

⊆ *разбиение**:

- {
- *стационарная предметная область*
- *нестационарная предметная область*
- }

стационарная предметная область - это **предметная область**, в которой связи между сущностями, входящими в ее состав, не зависят от времени (не меняются во времени). При этом некоторые из указанных сущностей могут иметь конечное время "жизни" (конечное время существования).

Таким образом, элементами **стационарной предметной области** не могут быть **временные сущности**.

нестационарная предметная область - это **предметная область**, в которой некоторые связи между сущностями, входящими в ее состав,

меняются со временем (то есть носят ситуационный, нестационарный характер, другими словами, являются *временными сущностями*).

3.2. Отношения, задающие роли элементов предметной области

Любая *предметная область* с формальной точки зрения является *структурой*. Уточним роли элементов, входящих в состав *предметной области*, которые являются частным случаем ролей элементов структур, рассмотренных выше.

понятие предметной области' – это *ролевое отношение*, указывающее в рамках *предметной области* на знак множества, являющегося классом некоторых объектов.

понятие предметной области'

\leq разбиение*:

- {
- *исследуемое понятие'*
- *понятие, исследуемое в частной предметной области'*
- *понятие, исследуемое в более общей предметной области'*
- *понятие, исследуемое в неродственной предметной области'*
- }

\in *неосновное понятие*

понятие – класс, являющийся исследуемым понятием хотя бы для одной *предметной области*.

понятие

\leq *строгое включение**:

класс

\leq *второй домен**:

понятие предметной области'

исследуемое понятие' – это *ролевое отношение*, указывающее в рамках *предметной области* на знак множества, являющегося классом объектов исследования данной *предметной области*, то есть такого множества, все элементы которого являются элементами данной *предметной области*.

исследуемое понятие' может быть:

(1) классом *первичных элементов'* этой *предметной области*;

(2) отношением (классов связей), связывающих элементы этой *предметной области* – уровень иерархии этих элементов может быть различным;

(3) классом *структур*, все элементы которых являются элементами заданной *предметной области*, т.е. классом подструктур заданной *предметной области*.

исследуемое понятие'

= *класс объектов исследования данной предметной области'*

= *понятие, исследуемое в данной предметной области'*

\in *ролевое отношение*

\leq *разбиение**:

- {
- *максимальный класс объектов исследования'*
- *немаксимальный класс объектов исследования'*
- }

\leq *разбиение**:

- {
- *исследуемый класс первичных элементов'*
- *исследуемое отношение'*
- *класс исследуемых структур'*
- }

\leq *включение**:

полностью представленное множество'

максимальный класс объектов исследования'

это *ролевое отношение*, указывающее в рамках *предметной области* на множество, являющееся максимальным классом объектов исследования данной *предметной области*, то есть на такое *исследуемое понятие'*, для которого в рамках данной *предметной области* не существует другого *исследуемого понятия'*, которое бы являлось надмножеством для данного.

немаксимальный класс объектов исследования'

это *ролевое отношение*, указывающее в рамках *предметной области* на такое *исследуемое понятие'*, для которого в рамках данной *предметной области* существует другое *исследуемое понятие'*, являющееся надмножеством первого.

исследуемый класс первичных элементов'

такое *исследуемое понятие'* для данной *предметной области*, что все его элементы являются ее первичными элементами'.

исследуемое отношение'

это *ролевое отношение*, указывающее в рамках *предметной области* на множество связей, являющееся исследуемым отношением данной *предметной области*, то есть таким отношением, все связки которого являются элементами этой *предметной области*.

При этом элементы таких связей также входят в данную *предметную область*, но в общем случае могут не являться элементами *исследуемых понятий'* данной *предметной области*.

класс исследуемых структур' – это *ролевое отношение*, указывающее в рамках *предметной области* на множество *структур*, знак каждой из которых принадлежит данной *предметной области*.

В общем случае в данную *предметную область*, могут входить не все элементы таких *структур*, а только некоторые из них (хотя бы один для каждой *структуры*).

понятие, исследуемое в частной предметной области'

это *понятие предметной области'* данной *предметной области*, которое является *исследуемым понятием'* для какой-либо из *частных предметных областей** относительно данной.

понятие, исследуемое в предметной подобласти'
= класс исследуемых объектов, который детально исследуется в предметной области, которая является частной по отношению к заданной предметной области'

∈ ролевое отношение

≤ включение*:

частично представленное множество'

понятие, исследуемое в более общей предметной области' – это понятие предметной области', которое является исследуемым понятием' для какой-либо из предметных областей, для которых данная предметная область является частной предметной областью*.

понятие, исследуемое в более общей предметной области'

≤ включение*:

частично представленное множество'

понятие, исследуемое в неродственной предметной области' – это понятие предметной области', которое является исследуемым понятием' для какой-либо из предметных областей, являющихся неродственными предметными областями* для данной.

понятие, исследуемое в неродственной предметной области'

≤ включение*:

частично представленное множество'

3.3. Отношения, заданные на предметных областях

На множестве предметных областей могут быть заданы теоретико-множественные отношения: *включение**, *объединение**, *пересечение**, *декомпозиция**, *гомоморфизм**, *изоморфизм**, а также специальные отношения, область определения которых является множеством предметных областей. Рассмотрим некоторые из них.

частная предметная область* – бинарное ориентированное отношение, с помощью которого задается иерархия предметных областей путем перехода от менее детального к более детальному рассмотрению соответствующих исследуемых понятий.

частная предметная область*

= дочерняя предметная область*

= быть частной предметной областью*

= предметная область, детализирующая описание одного из классов объектов исследования другой (более общей) предметной области*

∈ бинарное отношение

∈ ориентированное отношение

∈ неролевое отношение

⊃ частная предметная область по классу первичных элементов*

⊃ частная предметная область по исследуемым отношениям*

частная предметная область по классу первичных элементов*

= сужение предметной области по классу первичных элементов*

частная предметная область по исследуемым отношениям*

= частная предметная область по предмету исследования*

= сужение предметной области по предмету исследования*

Связки отношения **неродственные предметные области*** связывают две предметные области, имеющие общие элементы, однако не связанные отношением частная предметная область*.

неродственные предметные области*

∈ бинарное отношение

∈ неориентированное отношение

Рассмотренные отношения являются частным случаем отношения **метазнание***.

Каждая предметная область задает соответствующий специализированный язык (**sc-язык**), представляющий собой множество всевозможных фрагментов (подструктур), включаемых в состав этой предметной области.

sc-язык – это подязык (подмножество) **SC-кода**, ориентированный на представление **sc-текстов**, являющихся фрагментами некоторой предметной области. Таким образом, каждому **sc-языку** взаимно однозначно соответствует некоторая предметная область (точнее, **sc-модель** некоторой предметной области).

Отношение **sc-язык и соответствующая предметная область*** – это бинарное ориентированное отношение, каждая связка которого связывает знак некоторого **sc-языка** (под первым атрибутом) и знак соответствующей этому **sc-языку** предметной области.

4. Предметная область онтологий

Для спецификации (описания свойств) соответствующей предметной области, ориентированной на описание свойств и взаимосвязей понятий, входящих в состав указанной предметной области используется такой вид знаний, как **онтология** [Кудрявцев, 2010], [Федотова, 2015].

онтология

= система понятий соответствующей предметной области

= концептуальный каркас (скелет) описания некоторой предметной области

= концептуальная (семантическая) основа различных языков, обеспечивающих описание объектов исследования, принадлежащих заданной предметной области

= семантический интерфейс для интеграции знаний по заданной предметной области и для согласованного понимания различными

- субъектами этих знаний
- = онтология соответствующей предметной области
- = описание концептов и отношений заданной предметной области
- <= включение*:
- знание
- <= разбиение*:
- {
- интегрированная онтология
- структурная спецификация
- теоретико-множественная онтология
- логическая иерархия понятий
- логическая онтология
- логическая иерархия высказываний
- терминологическая онтология
- онтология задач и решений задач
- онтология классов задач и способов решения задач
- }

Связь между предметной областью и ее онтологией задается отношением *онтология** (*быть онтологией**), которое является частным видом отношения *метазнание** (*быть метазнанием**).

онтология* – это бинарное отношение, связывающее некоторую предметную область с ее онтологией (спецификацией).

- онтология***
- = *sc-онтология**
- = *быть онтологией предметной области**
- = *sc-онтология, специфицирующая заданную предметную область**
- => *первый домен**:
- предметная область*
- => *второй домен**:
- онтология*

В зависимости от рассматриваемых свойств понятий предметной области, которые описываются в онтологии, выделяют следующие типы онтологий:

- **структурная спецификация** – это онтология, в которой описываются роли понятий, входящих в состав предметной области, а также связи специфицируемых предметных областей с другими предметными областями.
- **теоретико-множественная онтология** – это онтология, описывающая теоретико-множественные связи между понятиями заданной предметной области (включение, разбиение, объединение, пересечение, разность множеств, область определения, домен, функция)
- **логическая онтология** – это онтология, описание системы высказываний заданной предметной области.
- **логическая иерархия понятий** – это онтология, являющаяся надстройкой над логической онтологией, включающая описание системы определений понятий заданной предметной области с указанием набора понятий,

через которые определяется каждое определяемое понятие рассматриваемой предметной области.

- **используемые константы*** – это отношение, связывающее понятие со множеством понятий, на основании которых оно определяется в рамках рассматриваемой предметной области.

- **логическая иерархия высказываний** – это онтология, являющаяся надстройкой над логической онтологией и включающая описание системы утверждений рассматриваемой предметной области с указанием набора утверждений, через которые доказывается каждое утверждение.

- **используемые утверждения*** – это отношение, связывающее утверждение со множеством утверждений, на основании которых оно доказывается в рамках рассматриваемой предметной области

- **терминологическая онтология** – это онтология, описывающая систему основных и неосновных терминов (имен, внешних обозначений), соответствующих концептам и отношениям заданной предметной области, а также описание правил построения терминов для сущностей, являющихся элементами (экземплярами) указанных концептов и отношений.

- **онтология задач и решений задач** – это онтология, описывающая задачи и их классы, решаемые в рассматриваемой предметной области

- **онтология классов задач и способов решения задач** – это онтология, описывающая способы решения задач и их классов в рамках предметной области. Является надстройкой над онтологией задач и классов задач.

- **интегрированная онтология** – это онтология, объединяющая все онтологии различного вида некоторой предметной области.

Заключение

Рассмотренные в работе средства позволяют осуществлять семантическую структуризацию баз знаний любой системы, построенной по *Технологии OSTIS* унифицированным образом.

Семантическая структура базы знаний представляет собой иерархическую систему описываемых ею предметных областей, надстраиваемых над заданной основной предметной областью.

Построение семантической структуры базы знаний требует не только явного представления спецификации каждой описываемой предметной области в виде формального текста, но и явного описания всевозможных связей между этими предметными областями.

В работе представлены результаты, полученные в научных группах МГТУ им. Н.Э.Баумана (кафедра «Компьютерные системы автоматизации производства») и БГУИР (кафедра интеллектуальных информационных технологий) совместно.

Работа выполнена при поддержке гранта БРФФИ-РФФИ-М «Методы и средства онтологического моделирования для семантических технологий проектирования интеллектуальных систем» (Ф15PM-074, 15-57-04047).

Библиографический список

[Гаврилова и др., 2001] Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. Учебник / Гаврилова Т.А. [и др.]; – СПб.: Изд-во «Питер», 2001.

[Голенков, 2012] Голенков, В.В., Гулякина Н.А. Графодинамические модели параллельной обработки знаний: принципы построения, реализации и проектирования. – В кн. Междунар. научн.-техн. конф. «Открытые семантические технологии проектирования интеллектуальных систем» (OSTIS-2012). Материалы конф. [Минск, 16-18 февр. 2012 г.]. – Минск: БГУИР, 2012.

[Голенков, 2015] Голенков, В.В. Семантическая технология компонентного проектирования систем, управляемых знаниями. Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2015): материалы V Междунар. научн.-техн. конф. / В. В. Голенков, Н.А Гулякина // Мн.: БГУИР, 2015

[Кудрявцев, 2010] Кудрявцев Д.В. Системы управления знаниями и применение онтологий: Учеб. пособие / Д.В. Кудрявцев. – СПб.: Изд-во Политехн. ун-та, 2010.

[Тузовский и др., 2005] Тузовский А. Ф., Чириков С. В., Ямпольский В. З. (2005) Системы управления знаниями. — Томск: Изд-во науч.-техн. литературы.

[Федотова, 2015] Федотова А.В., Давыденко И.Т. Онтологическое моделирование технического обслуживания // Материалы V Международной научно-технической конференции OSTIS-2015 (Минск, 19–21 февраля 2015 г.) – Минск: БГУИР, 2015. – С. 429-438.

[Tarassov, 2015] Tarassov V.B., Fedotova A.V., Stark R., Karabekov B.S. Granular Meta-Ontology and Extended Allen's logic: Some Theoretical Background and Application to Intelligent Product Lifecycle Management Systems // Proceedings of the 4th International Conference on Intelligent Systems and Applications (INTELLI'2015, St. Julians, Malta, October 11-16, 2015). – Copenhagen: IARIA XPS Press, 2015. – P.86-93. ISBN: 978-1-61208-437-4.

[IMS, 2016] Метасистема IMS.OSTIS [Электронный ресурс]. Минск, 2016. – Режим доступа: <http://ims.ostis.net/>. – Дата доступа: 15.01.2016.

THE TOOLS OF STRUCTURING THE SEMANTIC MODELS OF KNOWLEG BASES

Davydenko I.T. *, Grakova N.V. *,
Sergienko E.S. **, Fedotova A.V.***

*Belarusian State University of Informatics and
Radioelectronics, Minsk, Republic of Belarus

ir.davydenko@gmail.com

natalia.grakova@gmail.com

** National Research Nuclear University MEPHI
(Moscow Engineering Physics Institute), Moscow,
Russia Federation

deav@inbox.ru

*** Bauman Moscow State Technical University,
Moscow, Russia

afedotova.bmstu@gmail.com

The paper presents the tools of structuring the semantic models of knowleg bases, based on unified semantic networks. The main attention is paid for such concepts as structure, knowledge, considered their classification. Separately section describes the types of knowledge such as ontologies and subject domain.

Key words: knowledge base, knowledge base structuration, subject domain.

Introduction

Structuring of knowledge bases, highlighted in it various interrelated substructures needed for a variety of reasons. It is necessary for didactic aims, and for organizing the distribution of work on the knowledge bases design.

Main part

In this work introduced and refined one of the basic terms of knowledge representation OSTIS technology, the concept of structure, which is the basis for the concept knowledge, one type of which is the concept of subject domain. For the subject domain specification introduces the concept of ontology and considered their topology.

Conclusion

The tools, considered in this work, allows realised the semantic structuring of knowledge base of each system, based on OSTIS technology unified way. The semantic structure of knowledge base is hierarchical system of subject domains, describe by it.

Construction of the semantic structure of the knowledge base requires not only the explicit representation of the specifications described each subject domain in the form of a formal text, but the explicit description of all possible connections between these domains.