

УДК 004.82+004.89

КОНЦЕПЦИЯ УПРАВЛЕНИЯ ИНФОРМАЦИОННЫМИ КОМПОНЕНТАМИ РАЗВИВАЕМОЙ ИНТЕЛЛЕКТУАЛЬНОЙ ИНТЕРНЕТ-СИСТЕМЫ ПО ХИМИИ

И.Л. Артемьева

Институт автоматизации и процессов управления ДВО РАН

Россия, 690041, Владивосток, ул. Радио, 5

E-mail: artemeva@iacp.dvo.ru

Ключевые слова: онтология предметной области, многоуровневая онтология, интеллектуальная система, система, основанная на знаниях

Key words: domain ontology, multilevel ontology, ontology editor, intellectual system, knowledge-based system.

Интеллектуальная система сложной предметной области является системой, основанной на знаниях, которая позволяет накапливать эти знания и предоставляет возможность их использования при решении прикладных задач специалистами области. Практически полезная интеллектуальная система должна быть развиваемой и допускать изменение не только знаний, но и онтологии предметной области и, как следствие, множества классов прикладных задач, для решения которых используется интеллектуальная система. Поэтому важным является сопровождение такой системы, которое для интеллектуальной системы состоит в управлении ее информационными и программными компонентами. Для решения задачи управления некоторым объектом необходимо изучить свойства объекта управления и возникающие подзадачи. В данной работе анализируется структура информационных компонентов практически полезной интеллектуальной Интернет системы в области химии, которая должна охватывать информацию разных разделов данной области. Рассматриваются особенности информационных компонентов такой системы и особенности управления ими. Описывается структура онтологии более высокого уровня общности (метаонтологии), которая может использоваться при управлении онтологиями разных разделов химии и в терминах которой могут быть определены онтологии этих разделов.

MANAGEMENT OF INFORMATION COMPONENTS OF UPGRADABLE INTELLECTUAL INTERNET SYSTEM FOR CHEMISTRY / I.L. Artemieva (Institute of Automation & Control Processes, 5 Radio st., Vladivostok 690041, Russia, E-mail: artemeva@iacp.dvo.ru). An intellectual system for a domain with complicated structures is a knowledge-based system which allows us to accumulate domain knowledge and provides an opportunity for domain specialists to use it to solve applied tasks. Such an intellectual system which will be practically useful for domain specialists should be upgradable and allow users to modify not only knowledge but domain ontology as well and, as a result, sets of applied task classes solved by the intellectual system. Therefore maintenance of an intellectual system is important and consists of its information and program components management. To solve a problem

how one can manage an object, it is necessary to study properties of the object managed and problems to appear while managing. This article analyzes the structure of information components of the intellectual Internet system which should be practically useful for chemists and which should cover information of different chemistry subdomains. The article considers the special features of information components and their management. It describes the structure and terms of ontology with a higher level of generality (metaontology) which can be used to manage chemistry subdomain ontologies and to develop them.

1. Введение

При решении прикладных задач в области химии исследователи используют онтологию и знания разных ее разделов, а задачи, решаемые в этих разделах, являются подзадачами более сложных задач. Поэтому практически полезная компьютерная система должна позволять объединение знаний и онтологии разных разделов химии. Поскольку данная область науки развивается, это требует поддержки процесса развития и программной системы. С одной стороны, такое развитие предполагает добавление новых онтологий и знаний новых разделов области, а с другой стороны, такое развитие предполагает и изменение состава программных компонентов такой системы, предназначенных для решения новых классов прикладных задач. Таким образом, практически полезная интеллектуальная система должна постоянно сопровождаться, а ее разработке должна предшествовать разработка подсистемы сопровождения [1].

Важность задачи сопровождения привела к пониманию необходимости управления развитием программного средства, т.е. его сопровождению с помощью специальных высокоуровневых механизмов управления, сводящих к минимуму изменение его кода, и определению основных подсистем, используемых при управлении: мониторинга, ручного, автоматического и автоматизированного управления [2]. Однако, для разработки таких механизмов и подсистем необходим подробный анализ свойств объектов управления – интеллектуальных систем для сложных предметных областей и их информационных и программных компонентов.

Целью данной работы является анализ структуры информационных компонентов интеллектуальной системы в области химии, которая была бы практически полезной для специалистов данной области и охватывала информацию разных разделов данной области. В статье рассмотрены особенности информационных компонентов такой системы и управления ими. Описана структура метантологии, которая может использоваться при управлении, с использованием терминов которой определяются онтологии новых разделов области.

2. Модули онтологии химии

Практически полезная интеллектуальная система для химии должна охватывать знания многих ее разделов. Примерами ее разделов являются физическая, органическая и аналитическая химия (рис. 1), раздел катализа. Физическая химия изучает физико-химические процессы [3].

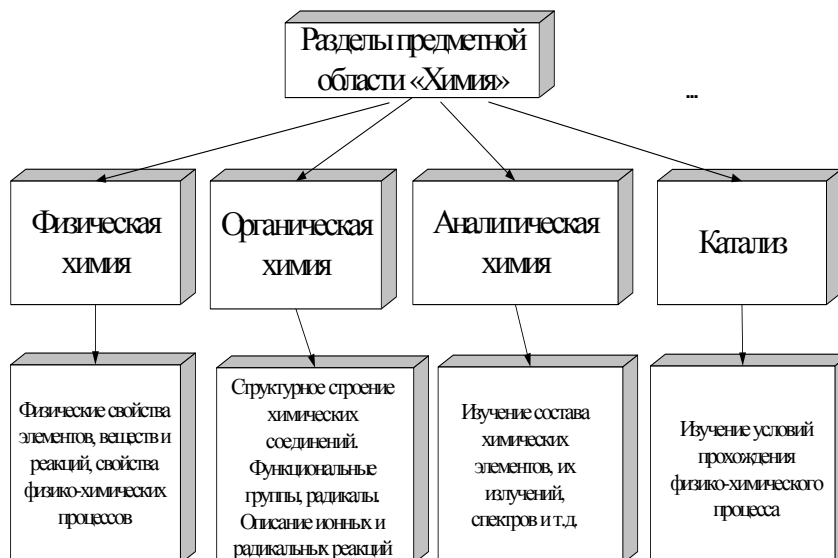


Рис. 1. Примеры разделов предметной области «Химия»

Описание этих процессов дается в терминах свойств участвующих в процессах веществ и реакций. Органическая химия добавляет терминологию, позволяющую говорить о структурных свойствах веществ [4]. Аналитическая химия изучает процессы воздействия на вещества различными излучениями [5]. Раздел катализа добавляет терминологию, с помощью которой могут быть описаны различные условия прохождения физико-химических процессов. Каждому разделу химии соответствует свой модуль онтологии, в котором определяется связанная система понятий. При этом некоторый модуль онтологии может использовать терминологию, определенную в других модулях.

Онтология некоторого раздела может содержать определение большого числа понятий. Это делает ее описание громоздким и трудным для восприятия. Для облегчения понимания такой онтологии она может также быть разбита на модули, в каждом из которых определяется связанная система понятий. Так, модулями онтологии физической химии являются (Рис. 2): "Свойства химических элементов", "Свойства химических веществ", "Свойства химических реакций", "Основы термодинамики", "Термодинамика. Химические свойства", "Термодинамика. Физические свойства", "Термодинамика. Связь физических и химических свойств", "Химическая кинетика". Поскольку вещества строятся из элементов, а реакции – из веществ, модуль "Свойства химических веществ" ссылается на "Свойство химических элементов", а модуль "Свойства

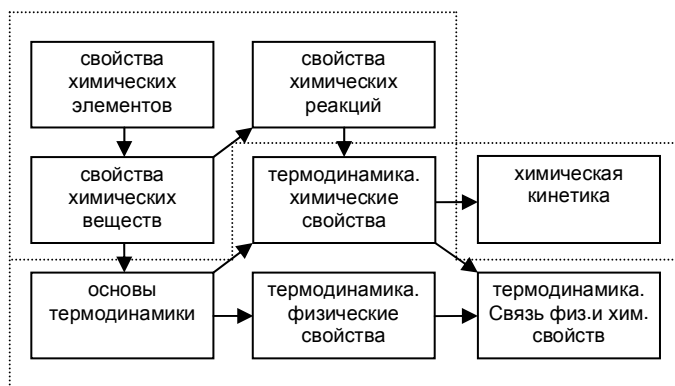


Рис. 2. Модули онтологии физической химии

химических реакций” – на “Свойства химических веществ”. В модуле “Основы термодинамики” определены термины, используемые при описании общих свойств термодинамических систем и их компонентов. Состояния термодинамической системы могут изменяться в ходе физико-химического процесса. Состояния процесса задаются в дискретные моменты наблюдения. В модуле “Термодинамика. Физические свойства” определены термины, используемые при описании фазовых превращений веществ в ходе процесса, без учета химических превращений. В модуле “Термодинамика. Химические свойства” определены термины, используемые при описании химических превращений веществ в ходе процесса без учета фазовых превращений. И, наконец, в модуле “Термодинамика. Связь физических и химических свойств” определены термины, используемые при описании физико-химических процессов. В модуле “Химическая кинетика” определены термины, описывающие динамику прохождения процессов.

Названия разделов химии, модули онтологии разделов, связи между разделами и связи между модулями задают структуру представления информации об онтологии области. По существу, это информация задает знания об устройстве как сложной предметной области, так и ее информационных компонентов, определяя на множестве модулей отношения использования терминологии одного модуля другими. Для сложной предметной области модульными являются как онтология, так и база знаний, причем на множестве модулей базы знаний имеют место те же отношения, которые определены между модулями онтологии.

Система управления информационными компонентами интеллектуальной системы должна позволять задание и добавление разделов области, быть может, сложно структурированных, а также позволять определять на множестве разделов и их подразделов отношения использования.

3. Состав онтологии раздела химии

Рассмотрим теперь устройство каждого модуля онтологии. Каждый раздел химии имеет дело со своим множеством типов химических объектов. Объектами могут быть химические элементы, химические вещества, реакции, оболочки, спектры и т.д., принадлежащие указанным множествам (типам). Объекты каждого типа имеют свой набор свойств. Причем в разных разделах химии изучение объектов одного и того же типа может предполагать изучение разных их свойств. Таким образом, компонентами онтологии любого раздела являются множество

названий типов изучаемых в нем объектов, а также множество названий свойств этих объектов (Рис. 3).

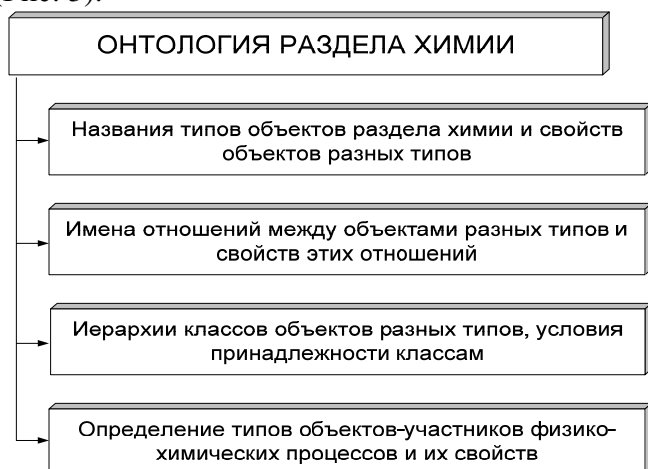


Рис. 3. Состав онтологии раздела

Множество объектов одного типа может делиться на подмножества, называемые классами. Причем таких делений может быть несколько. Например, в основе классификации органических соединений лежат два принципа. Первый принцип классификации это строение углеродного скелета; при этом исходят из возможностей атомов углерода образовывать как цепи, так и циклические структуры. Органические соединения подразделяют на ациклические и циклические, которые, в свою очередь делятся на карбоциклические и гетероциклические и т.д. Второй принцип классификации это наличие функциональных групп. Исходным (материнским) классом являются углеводороды, в которых отсутствуют функциональные группы; все остальные классы можно произвести от углеводородов путем замены атомов водорода (одного или более) на функциональные группы. Примерами классов являются галогенопроизводные, спирты, карбонильные соединения.

Таким образом, онтология раздела может содержать определение различных иерархий классов объектов данного раздела (Рис. 3). Онтологические соглашения могут определять условия принадлежности объектов классам. Если такие условия заданы для всех классов некоторой иерархии, то множества объектов, образующих некоторый класс этой иерархии, могут быть автоматически построены на основе заданных условий.

Между объектами разных типов существуют различные отношения. Примерами отношений между объектами являются отношения, определяющие вхождение химических элементов в состав химического вещества (отношение между веществами и элементами), определяющие оболочки атома химического элемента (отношение между элементами и оболочками), определяющие реагенты и результаты реакций (отношения между реакциями и веществами). Онтология раздела задает названия отношений между объектами разных типов (Рис.3). Отдельную группу терминов онтологии образуют термины, с помощью которых определяются свойства отношений между объектами. Примерами могут служить термины, которые позволяют определить свойства реагентов и результатов некоторой реакции, свойства элементов, входящих в состав вещества, свойства оболочек химического элемента.

Объекты некоторых типов являются участниками физико-химических процессов, в результате которых могут меняться их свойства или образовываться объекты других типов. Онтология раздела определяет, объекты каких типов являются участниками физико-химического

процесса и какие свойства этих объектов характеризуют процесс. Например, в физической химии процесс рассматривается на разных уровнях детализации:

1. в виде последовательности множеств химических веществ;
2. в виде последовательности множеств веществ и реакций, имеющих место на каждом шаге процесса;
3. в виде последовательности множеств веществ и реакций, имеющих место на каждом шаге процесса и фаз системы, причем фазы системы меняются в ходе процесса, также как и их ингредиенты;
4. в виде множества фаз, причем химические реакции протекают в некоторой фазе, а в результате химических реакций меняются ингредиенты фазы;
5. в виде множества фаз с учетом прохождения химической реакции в некоторой фазе.

Органическая химия добавляет следующие уровни рассмотрения:

1. в виде множества фаз с учетом прохождения химической реакции в некоторой фазе и с учетом механизма прохождения каждой реакции на ионном или радикальном уровне;
2. в виде множества фаз с учетом прохождения химической реакции в некоторой фазе и с учетом механизма прохождения каждой реакции на электронном уровне взаимодействия веществ.

Таким образом, определение онтологии раздела включает определение названий типов объектов и их свойств, задание критериев классификации объектов некоторых типов, иерархии классов для каждого вида классификации и условий принадлежности объектам классам, определение названий отношений между объектами разных типов и свойств этих отношений, а также задание уровней рассмотрения физико-химического процесса. Система управления информационными компонентами интеллектуальной системы для химии должна позволять определять все указанные компоненты онтологии при ее создании.

4. Специализированные структурные значения и их онтологии

Значения свойств объектов могут иметь сложную структуру. Например, формула химического соединения представляет собой последовательность компонентов, а каждый компонент – это химический элемент со своими характеристиками в формуле или формула подкомпонента. Правила определения структурного значения задаются специальной онтологией, которая определяет не только структуру, но и накладывает различные ограничения, которым должны удовлетворять компоненты структурного значения.

Различные структурные значения могут быть связаны друг с другом, тогда онтология предметной области содержит дополнительные соотношения, определяющие такие связи. Примерами связей могут служить связи между формулой химического соединения и его структурной формулой, а также связи последней и краткой структурной формулы органических соединений (Рис. 4).

При мониторинге информационных компонентов интеллектуальной системы должны автоматически проверяться условия, задающие связи между соответствующими структурными значениями при их внесении.

Автоматическое управление информационных компонентов интеллектуальной системы состоит в автоматическом построении структурного значения на основе условий связи между заданным и связанным с ним, но отсутствующим структурным значением. Такое построение возможно, если соотношения, задающие связи, могут быть автоматически преобразованы к правилам решения данной задачи.

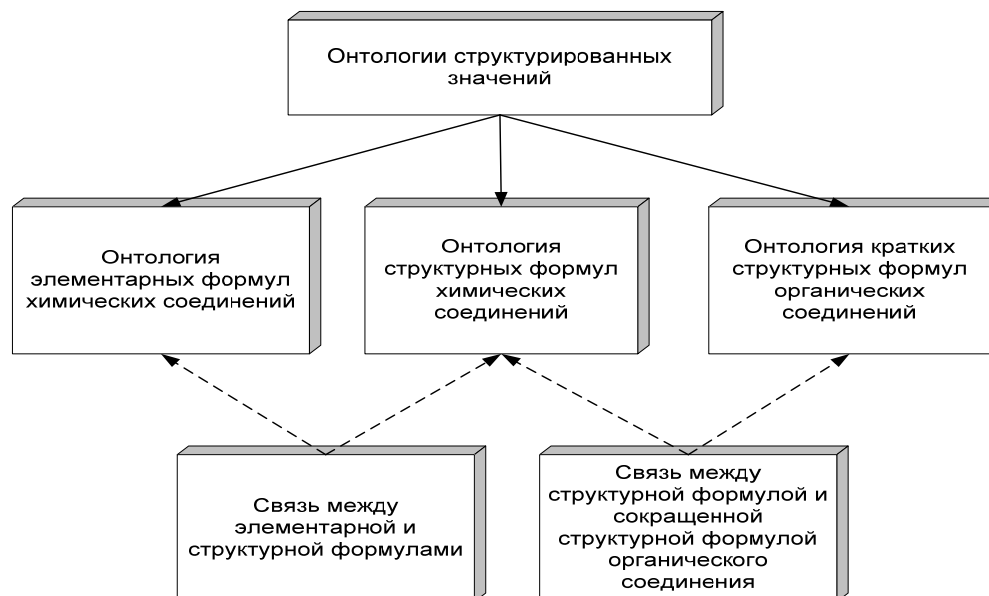


Рис. 4. Примеры онтологий структурных значений и связей между ними

Для некоторых специализированных структурных значений в предметной области существует принятый графический способ представления. Например, каждая связь между компонентами структурной формулы обозначается своим типом линии. Графическое представление информации о структуре более удобно для восприятия специалисту. Это приводит к необходимости использования специализированных графических редакторов, которые должны «знать» онтологию представления структурных значений и автоматически выполнять преобразование введенного графического представления к тому представлению, которое поддерживается оболочкой информационных компонентов интеллектуальной системы.

Наличие специализированных редакторов влечет за собой усложнение управления процессом редактирования информации. Вызов специализированного редактора должен происходить автоматически под управлением онтологии раздела.

Таким образом, система управления информационными компонентами интеллектуальной системы для химии должна позволять определение структурных значений, которые могут использоваться в любом разделе химии, должна обеспечивать проверку согласования связанных между собой значений и автоматическое построение отсутствующих по заданным, а также должна производить автоматический вызов специализированных редакторов структурных значений под управлением онтологии.

5. Знания и архивы данных

Модуль знаний химии содержит знания раздела этой области. Он может состоять из двух частей. Первая часть модуля знаний – это значения терминов онтологии раздела химии. В ней задаются объекты, изучаемые разделом, их компоненты, отношения между объектами, свойства объектов, их компонентов и т.д.

Вторая часть модуля знаний представляется множеством формул некоторого логического языка. В ней задаются дополнительные соотношения между объектами и их свойствами, которые

не могут быть заданы как значения терминов онтологии. При записи этих формул используются термины онтологии раздела.

Онтология, используемая для создания интеллектуальных систем, предназначенных не только для хранения, поиска и редактирования онтологий и знаний, а также для решения других задач профессиональной деятельности, содержит определение системы понятий, при помощи которой задаются исходные данные таких задач, а также представляются результаты решений. Такая система понятий для химии позволяет описать различные свойства конкретных физико-химических процессов (Рис. 5).

Онтология любого раздела может быть разбита на две части: первая содержит термины, значения которых используются при задании знаний раздела, а вторая – термины, которые используются при описании свойств

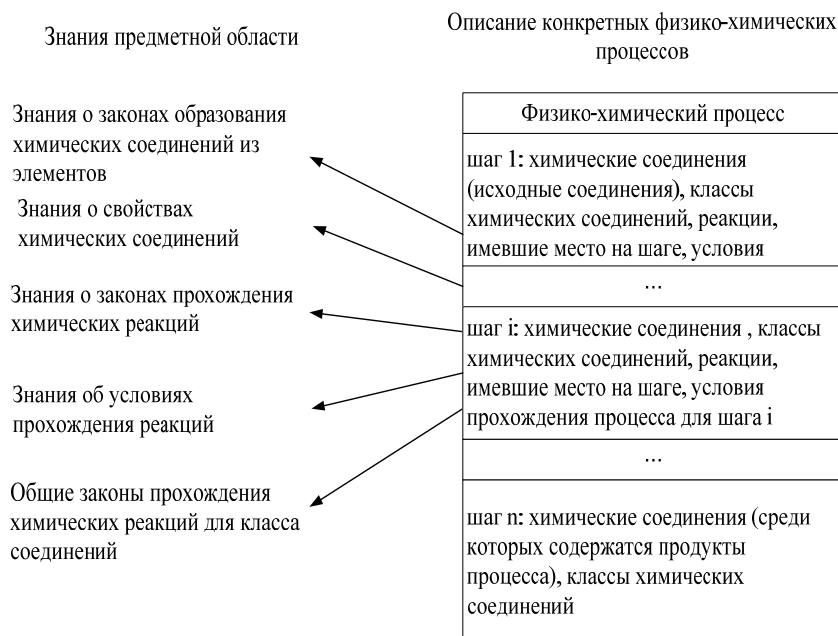


Рис. 5. Две части онтологии раздела

конкретных физико-химических процессов – описаний результатов проводимых исследователями экспериментов. Множество описаний таких экспериментов образуют архивы данных, которые могут входить в состав информационных компонентов интеллектуальных систем. Архивы данных позволяют проверить правильность знаний, хранящихся в интеллектуальной системе, при мониторинге информационных компонентов.

Они также могут использоваться при анализе и обобщении результатов экспериментов с целью получения новых знаний. В этом случае архивы данных либо являются входными данными системы, автоматически производящей такие обобщения, либо используются для проверки правильности обобщений, выполненных вручную и внесенных в информационные компоненты системы.

Среди онтологических соглашений предметной области существуют соглашения, имеющие форму равенств и задающие связи между значениями нескольких терминов. Такие равенства могут использоваться при автоматическом управлении информационными компонентами, позволяя вычислять значения одних терминов при задании значений всех других терминов, входящих в равенство.

Система управления информационными компонентами интеллектуальной системы по

химии должна позволять определение терминов онтологии любого раздела, используемых для описания свойств конкретных физико-химических процессов, должна позволять создание различных архивов таких описаний и их использование для проверки непротиворечивости имеющихся в системе знаний, а также должна позволять автоматическое вычисление значений одних терминов на основании значений других терминов.

6. Метаонтологии для химии

Метаонтология химии определяет систему понятий, в терминах которой описываются метаонтологии для разных разделов химии. Часть терминов метаонтологии химии играют роль параметризованных конструкторов: при задании значений параметров такого конструктора автоматически строится множество терминов метаонтологии раздела. Метаонтология раздела содержит систему понятий, используемую уже при создании всех модулей онтологии раздела. Приведем примеры некоторых терминов метаонтологии химии.

Типы объектов – этот термин позволяет определить множество типов химических объектов, изучаемых некоторым разделом химии. Онтологическое соглашение говорит о том, что множества объектов различных типов не пересекаются. Каждый тип объектов есть множество объектов, каждый из которых может быть представлен некоторым обозначением, числом либо структурным значением. Способ представления объектов каждого типа определяется при задании метаонтологии раздела. При ее определении создаются онтологии новых структурных значений, если это требуется для определяемого раздела химии. Конкретные объекты для каждого типа задаются знаниями раздела.

Термин **собственные свойства объектов** определяет имя конструктора, аргументом которого является тип объекта t , а результатом – множество конструкторов, у каждого из которых аргумент есть множество значений или множество кортежей m , а результат – множество функций, аргументом каждой из которых является объект типа t , а результатом – элемент множества m . В модели метаонтологии конструктор задается λ -термом. Задание онтологии второго уровня состоит в конкретизации конструкторов, которые будут использоваться при определении терминов онтологии раздела – имен и областей значений для функций, используемых при описании различных собственных свойств объектов. Значения свойств для каждого объекта фиксируются в знаниях предметной области.

Любой объект некоторого типа может иметь компоненты, которые являются объектами других типов. Термин **типы компонентов объектов** обозначает функцию, которая сопоставляет типу объектов t не пустое множество названий типов объектов - его компонент. Метаонтология раздела определяет значение данного термина. Термины, при помощи которых определяются компоненты объектов каждого типа, определяются онтологией раздела. Состав каждого объекта задается в знаниях раздела.

Метаонтологии химии также содержит понятия, позволяющие определить терминологию, используемую при описании свойств конкретных физико-химических процессов. Приведем примеры терминов.

Термин **число шагов процесса** обозначает количество шагов конкретного физико-химического процесса. Термин **типы объектов процесса** обозначает множество типов объектов, которые рассматриваются как участники физико-химического процесса. Термин **свойства компонентов процесса** обозначает конструктор, аргументом которой является тип объектов t , а результатом – множество конструкторов, аргументом каждого из которых является множество значений или кортежей значений m , а результатом – функция, аргументами которой

являются номер шага процесса и участник этого шага (объект типа t), а результатом – элемент множества m .

Метаонтология предназначена для управления процессом создания и редактирования онтологии. Конструкторы метаонтологии могут использоваться для автоматического формирования терминов метаонтологии или онтологии раздела при задании параметров конструкторов. Однако при развитии информационных ресурсов изменяться может не только метаонтология и онтология раздела, но и метаонтология химии. В этом случае система управления информационными компонентами интеллектуальной системы для химии должна обеспечивать согласованное редактирование всех связанных информационных компонентов интеллектуальной системы.

7. Общие и частные онтологии и базы знаний

Информационные компоненты интеллектуальной Интернет системы содержат общие информационные ресурсы, доступные всем пользователям системы, а также индивидуальные ресурсы, доступные одному пользователю или ограниченной группе пользователей. Доступ к информационным ресурсам происходит в соответствии с полномочиями пользователей, определяемыми при их регистрации в системе [6]. Индивидуальные ресурсы могут присутствовать в составе интеллектуальной системы по двум причинам.

Во-первых, индивидуальные ресурсы являются следствием необходимости защиты информации, хранимой в общем ресурсе. Все пользователи системы изменяют информацию из своего ресурса, а администратор системы производит объединение информации в общем ресурсе.

Во-вторых, индивидуальные ресурсы позволяют хранить частные онтологии и знания, используемые при научных исследованиях группой исследователей. В этом случае можно считать, что такие онтологии или знания относятся к своему разделу области, доступному только этой группе пользователей. Если группа пользователей решила перенести часть хранящейся информации в общий информационный ресурс, она должна сообщить об этом администратору.

При переносе информации производится мониторинг, который позволяет установить отсутствие нарушений ограничений целостности, а также проверить отсутствие противоречий между добавляемой и хранимой информацией.

Система управления информационными компонентами интеллектуальной системы по химии должна поддерживать создание и изменение общей информационной базы, баз разделов химии и баз исследователей (Рис. 6). Пунктирные стрелки на рис. 6 используются для указания того факта, что информационный компонент, в который входит стрелка, создается под управлением информационного компонента, из которого выходит стрелка. Общая информационная база содержит онтологию и знания, определяющие свойства объектов, используемых во всех разделах химии, а также онтологии структурных значений и связи между ними. База некоторого раздела содержит онтологию и знания этого раздела и имеет описанные выше компоненты. База исследователя содержит некоторую частную онтологию и знания и имеет ту же структуру, что и база раздела. При переносе информации в общую базу или базу раздела из базы исследователя система управления должна автоматически производить требуемый мониторинг.

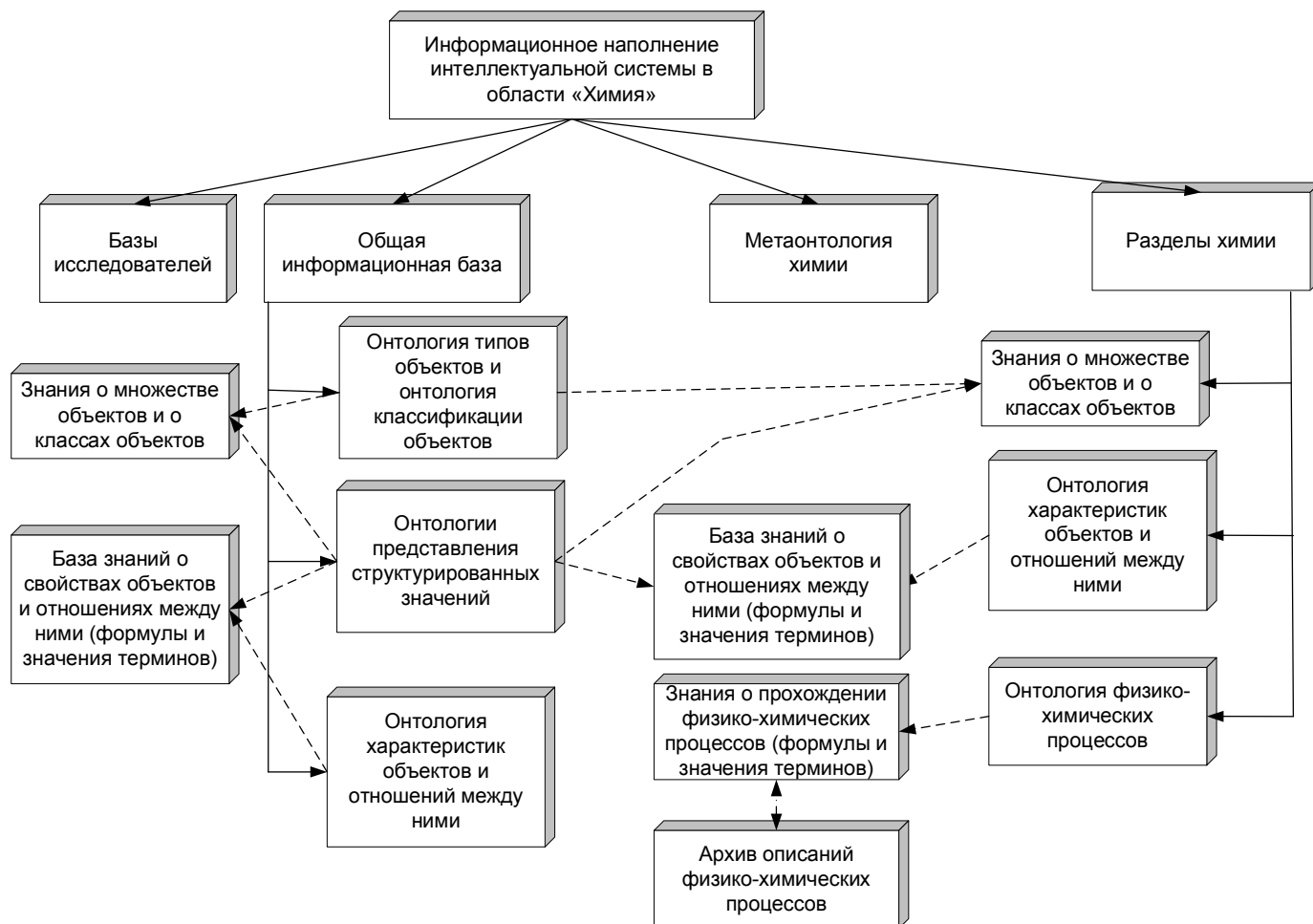


Рис. 6. Структура информационного наполнения интеллектуальной системы по химии

Заключение

В данной работе проанализирована структура информационных компонентов практически полезной развиваемой интеллектуальной Интернет системы в области химии, которая должна охватывать информацию разных ее разделов. Рассмотрены особенности информационных компонентов такой системы и управления ими. Описана структура онтологии раздела и состав метаонтологий, используемых при создании онтологий. Особенностью онтологии раздела является наличие онтологий структурированных значений, с которыми связаны специализированные средства редактирования информации. Определяется состав системы управления развитием информационных ресурсов интеллектуальной системы по химии.

Список литературы

- [1] Artemieva I.L. Specialized shells of intelligent systems for domains with complicated structures // International Book Series "Information Science and Computing". - N 11. Intelligent Engineering: Papers are selected from Proc. of the Intern. Conf. of the Join International Events of Informatics

- “ITA 2009” (KDS 2009), Varna, Bulgaria, 2009. P.123-130. ISSN 1313-0455. – Suppl. to the International Journal "Information Technologies & Knowledge". 2009. Vol. 3. Sofia, Bulgaria (Inst. of Information Theories and Applications FOI ITHEA).
- [2] Грибова В.В., Клещев А.С., Шалфеева Е.А. Системы управления интеллектуальными Интернет-приложениями. Владивосток: ИАПУ ДВО РАН, 2010. 31 с.
- [3] Артемьева И.Л., Цветников В.А. Фрагмент онтологии физической химии и его модель // Исследовано в России [Электронный ресурс]: многопредметн. научн. журн. / Моск. физ.-техн. ин-т. - Долгопрудный: МФТИ. 2002. № 5. С.454-474. - <http://zhurnal.ape.relarn.ru/articles/2002/042.pdf>
- [4] Артемьева И.Л., Высоцкий В.И., Рештаненко Н.В. Модель онтологии предметной области (на примере органической химии) // Научно-техническая информация, сер.2.- 2005. - № 8. - С.19-27.
- [5] Артемьева И.Л., Мирошниченко Н.Л. Модель онтологии рентгенофлуоресцентного анализа // Информатика и системы управления. 2005. № 2. С.78-88. – ISSN 1814-2400.
- [6] Kleshchev A.S., Orlov V.A. Computer knowledge banks. The universal way to solve the problem of information editing // Information technologies. 2006. №5. PP. 25-31.
- [7] Kleshchev A.S., Artemjeva I.L. A Mathematical Apparatus for Domain Ontology Simulation. An Extendable Language of Applied Logic. Int. J. Inf. Theories and Appl., 12(2), 149-157 (2005).