



УДК 519.68:15:54

© 2004 г. **И.Л. Артемьева**, канд. техн. наук,  
**Н.В. Рештаненко**

(Институт автоматизации и процессов управления ДВО РАН, Владивосток)

## **МОДУЛЬНАЯ МОДЕЛЬ ОНТОЛОГИИ ОРГАНИЧЕСКОЙ ХИМИИ<sup>1</sup>**

В работе описаны принципы построения модульной модели онтологии органической химии. Описана структура онтологии и ее модели, которые охватывают вузовский курс органической химии. В работе также приведен обзор существующих онтологий, описывающих термины химии, и указаны их основные недостатки.

### **Введение**

В настоящее время в связи с развитием компьютерных технологий появилась потребность в программных системах, автоматизирующих деятельность в таких сложных предметных областях как химия.

Органическая химия является частью общей химии. Она тесно связана с неорганической, физической и биологической химией и вместе с тем имеет свою специфику. Поэтому решение задач в области органической химии требует большого объема информации и знания всех закономерностей. Программная система, позволяющая решать задачи без физического вмешательства, дает возможность сэкономить время и средства, необходимые для решения этих же задач в лабораторных условиях.

Для создания сложных программных систем, решающих задачи в области органической химии, необходимо описать онтологию предметной области и построить ее модель.

Согласно работе [1] онтология предметной области содержит:  
систему понятий действительности;

---

<sup>1</sup> Работа выполнена в рамках программы № 16 фундаментальных исследований Президиума РАН на 2004 г. Проект "Теоретические основы интеллектуальных систем, основанных на онтологиях, для интеллектуальной поддержки научных исследований" (№ 10002-251/П-16/021-387/110504-266).

систему понятий знаний;  
связь между системой понятий знаний и системой понятий действительности.

Данная работа посвящена разработке онтологии «Органическая химия» в объеме вузовского курса обучения.

Вопрос создания онтологии в различных частях предметной области "Химия" является одним из наиболее рассматриваемых. Классический подход заключается в том, что нужно перейти от химии к модели химии. Онтологии используются как механизм выражения и распределения знаний сообщества для определения общего словарного запаса и поддержки интеллектуальных запросов в разнообразных базах данных. Онтологии и метаданные описывают организацию и содержание ресурсов. Такие определения и описания необходимы для операций и слияния ресурсов. Для конструирования онтологий используются различные средства.

В настоящее время построены онтологии некоторых разделов молекулярной биологии, которые предлагают терминологию для определения множества химических элементов, описания процессов внутри клетки. Онтология TAMBIS (TaO) [2] описывает биоинформатику, покрывает основные понятия молекулярной биологии и биоинформатики: макромолекулы, их предназначение, структуру, функции, клеточное расположение и процессы, в которых они взаимодействуют. TaO онтология построена с использованием языка OIL.

Существует также экспериментальная онтология для бионеорганических центров, известная под именем COME (<http://www.ebi.ac.uk/~kirill/come/COME.xml>). COME состоит из сущностей трех видов: Молекула (MOL), Бионеорганический Мотив (BIM) и Бионеорганические Протеины (PRX).

Также построены онтологии, представляющие понятия и отношения в более узко направленных областях – таких как химические кристаллы, керамические материалы, биоэнергетические центры. Примером таких онтологий может служить онтология Chemical-Crystals ([www-ksl-svc.stanford.edu:5915/doc/hpkb/query-results/real-ontolingua/chemical-crystals](http://www-ksl-svc.stanford.edu:5915/doc/hpkb/query-results/real-ontolingua/chemical-crystals)). Онтология Chemical-Crystals описывает различные типы кристаллической структуры веществ. Эта онтология построена с использованием методологии, известной как *методология* [3], которая является точным руководством по созданию новых онтологий.

Один из представителей онтологии химии – Plinius [4], онтология керамических материалов. Для того чтобы вместить много уровней обобщения, онтология Plinius толкует материалы как абстракцию от образца. Каждый образец является уникальным. При анализе входного текста Plinius интерпретирует каждую ссылку на материал как ссылку на образец и связывает с ним уникальную метку. Эта метка сама интерпретируется как характеристика образца.

Другой пример онтологии – онтология чистых веществ [5]. Определение чистых веществ дано через химический состав, т.е. через структурные правила, которые определяют чистые вещества в терминах химических веществ и натуральных чисел.

Также разработана иерархическая модель онтологии физической химии [6 – 10]. Модульная онтология физической химии определяет множество разделов предметной области и связи между ними, описывает систему понятий каждого раздела и задает связи между понятиями разделов. Онтология физической химии состоит из восьми связанных друг с другом разделов: «Элементы», «Вещества», «Реакции», «Основы термодинамики», «Термодинамика. Химические свойства», «Термодинамика. Физические свойства», «Термодинамика. Связь физических и химических свойств», «Химическая кинетика». В основе онтологии данной предметной области лежит метаонтология, которая определяет метапонятия, используемые при определении систем понятий каждого раздела.

Но модель физической химии определяет только часть понятий и законов органической химии. Свойства элементов, веществ и реакций, рассматриваемых в онтологии физической химии, в основном позволяют определить физические характеристики элементов, веществ и реакций. Помимо этого, упомянутая онтология не содержит терминов, описывающих классификации химических веществ и реакций, что является неотъемлемой частью органической химии. Поэтому можно считать, что ни одна из упомянутых работ не покрывает основные химические представления и законы органической химии.

Описание онтологии и ее модели для предметной области «Органическая химия» приведено в работах [11 – 14].

Модель представлена средствами многосортного языка прикладной логики, разработанного в ИАПУ ДВО РАН [15 – 17]. Модель онтологии доступна на странице <http://www.iacp.dvo.ru/es/>.

Целью данной работы является описание структуры модели онтологии органической химии.

### **Структура онтологии органической химии и ее модели**

Онтология органической химии является смешанной непримитивной онтологией, и ее модель представляется в классе смешанных необогащенных систем логических соотношений с параметрами [15 – 16]. Структура модели онтологии соответствует структуре онтологии. Как упоминалось выше, онтологию органической химии, как и любую онтологию, можно представить в виде двух взаимосвязанных систем понятий [1] – системы понятий знаний и системы понятий действительности (рис. 1).

Термины системы понятий знаний органической химии используются при описании значений свойств химических элементов, химических ве-

ществ и реакций. Множество значений терминов системы понятий знаний входит в базу знаний органической химии, которая, кроме этих значений, содержит также множество химических законов, представленных в виде логических утверждений.



Рис. 1. Представление онтологии органической химии в виде двух систем понятий.

Термины системы понятий действительности включают все термины, необходимые для описания химического процесса.

Химический процесс может рассматриваться на разных уровнях детализации (см. рис. 1):

- 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; фазы системы меняются в ходе процесса, также как и их ингредиенты;
- 4) в виде множества фаз. Химические реакции протекают в некоторой фазе, а в результате химических реакций меняются ингредиенты фазы;
- 5) в виде множества фаз с учетом времени прохождения химической реакции в некоторой фазе;

6) в виде множества фаз с учетом времени прохождения химической реакции в некоторой фазе и с учетом механизма прохождения каждой реакции на ионном или радикальном уровне;

7) в виде множества фаз с учетом времени прохождения химической реакции в некоторой фазе и с учетом механизма прохождения каждой реакции на электронном уровне взаимодействия веществ.

Онтология содержит понятия, необходимые для описания каждого из уровней детализации представления химического процесса.

Так как органическая химия тесно связана с неорганической и физической химией, онтология органической химии использует некоторую терминологию онтологии физической химии [6 – 10].

Так как предметная область «Органическая химия», даже ограниченная пределами вузовского курса обучения, является сложной и объемной, то при определении онтологии было принято решение разбить ее на модули. Рассмотрим подробнее представление системы понятий знаний и системы понятий действительности в виде взаимосвязанных модулей онтологии органической химии.

Система понятий знаний онтологии органической химии состоит из 14 модулей: «Электронное строение», «Структурная формула», «Тип гибридизации», «Виды связей», «Функциональные группы», «Углеродный скелет», «Структурная формула вещества», «Номенклатура», «Изомерия», «Реакционные центры», «Пространственное расположение», «Интермедиаты», «Радикальные реакции», «Ионные реакции». Каждый модуль описывает систему понятий одного раздела предметной области. В каждом модуле все термины описаны неформально на русском языке.

Как видно на рис. 2, модули «Электронное строение», «Виды связей» и «Тип гибридизации» базируются на разделах иерархической онтологии физической химии «Элементы», «Вещества» [6 – 7] и содержат термины для описания разных видов связей и типов гибридизации соответственно. Модули «Интермедиаты», «Радикальные реакции» и «Ионные реакции» базируются на разделах иерархической онтологии физической химии «Элементы», «Вещества», «Реакции» [6 – 7] и содержат термины для описания ионов, радикалов, а также особенностей ионных и радикальных реакций. Модули «Функциональные группы» и «Углеродный скелет» базируются на модуле «Структурная формула» и содержат термины для описания функциональных групп и углеродных скелетов органических веществ. Так как структурную формулу можно представить в виде углеродного скелета с функциональными группами, то модуль «Структурная формула вещества» базируется на модулях «Функциональные группы» и «Углеродный скелет». Модули «Номенклатура», «Изомерия» и «Реакционные центры» базируются на модуле «Структурная формула вещества».

Так как пространственное расположение молекулы вещества зависит от его структурной формулы и типов гибридизации всех атомов углерода

органического соединения, то модуль «Пространственное расположение» базируется на модулях «Структурная формула вещества» и «Тип гибридизации».

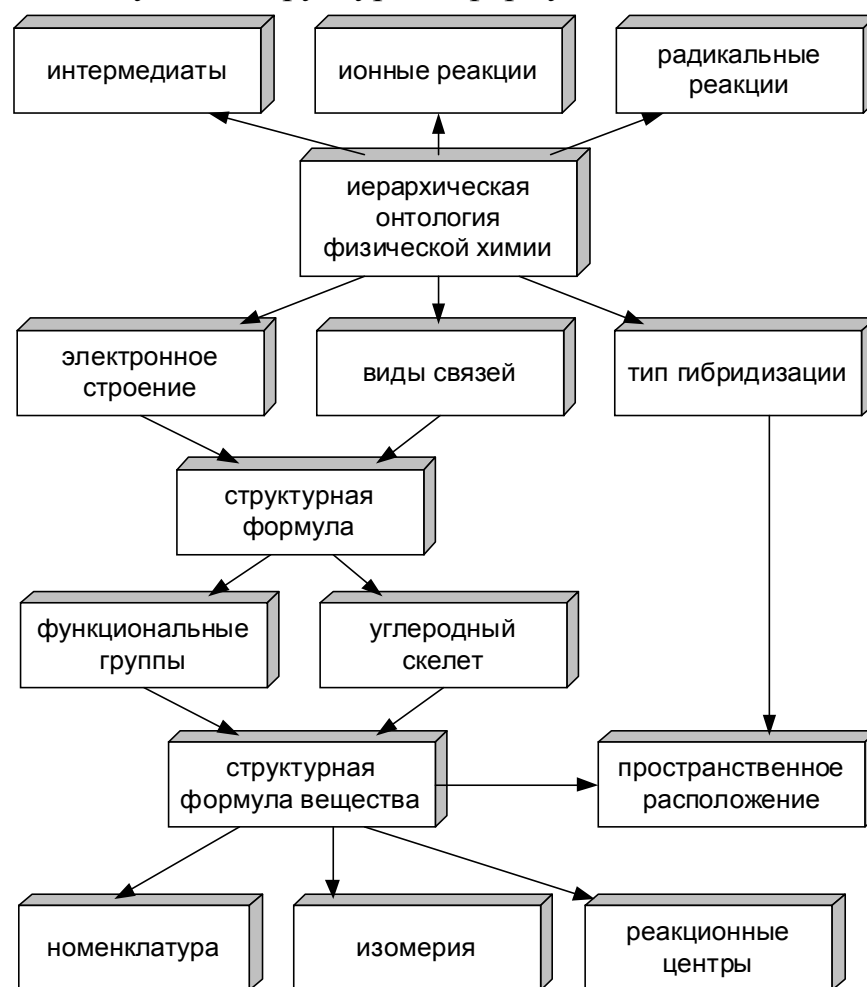


Рис. 2. Структура модульной онтологии органической химии.

дизации».

При определении терминов системы понятий действительности онтологии органической химии используются термины системы понятий действительности онтологии физической химии [6 – 10].

Как указано выше, система понятий действительности включает термины, необходимые для описания химического процесса. Иерархическая онтология физической химии содержит термины, позволяющие определять химический процесс на уровнях детализации с первого по пятый [8 – 9]. Онтология органической химии использует все эти термины и описывает новые термины, необходимые для описания химического процесса на шестом и седьмом уровнях детализации.

В онтологии также рассмотрены различные виды классификаций объектов предметной области. Так как классификация – это некоторое обобщение над системой понятий знаний, то задачи классификации рассматриваются отдельно от основной онтологии, хотя неотрывно связаны с

ней. Задача классификации состоит в нахождении класса, к которому относится некоторый объект предметной области по некоторым известным свойствам этого объекта. Знания из базы знаний органической химии определяют для каждого класса объектов: какими свойствами обладают объекты, принадлежащие этому классу.

Онтология описывает классификацию химических элементов, функциональных групп, реакционных центров, органических веществ, видов номенклатуры органических веществ, типов изомерии органических веществ и химических реакций, проходящих с участием органических веществ.

Существуют два вида классификации органических веществ (рис. 3):  
 классификация по строению углеродного скелета;  
 классификация по наличию функциональных групп.



Рис. 3. Классификация органических веществ.

Существуют два вида классификации органических реакций (рис. 4):  
 классификация по типу разрыва связи;  
 классификация по конечному результату.

Описание классов позволяет определять свойства не только каждого объекта предметной области, но и свойства множеств этих объектов. Так, например, соединения одного класса проявляют схожие физико-химические свойства, могут быть участниками реакций определенного класса и т.д.

Рассмотрим один из модулей модели онтологии, содержащий термины для описания структурной формулы органических соединений. Структурная формула является одним из ключевых понятий органической хи-

мии, позволяющих условно изображать строение молекул органических соединений с учетом геометрического расположения атомов.



Рис. 4. Классификация органических реакций.

В отличие от эмпирических формул структурные формулы показывают порядок связи атомов в молекуле и валентность каждого атома, входящего в состав молекулы, что однозначно определяет молекулы разных органических соединений.

### Модуль "Структурная формула"

Структурная формула(ST, математические кванторы, интервалы) =  $\langle (\{ \text{электронная конфигурация, элементный состав, константы онтологии} \}), SS \rangle$ , где SS – предложения, описанные ниже.

Описание значений имен.

1.1. *Тип связи*  $\equiv \{ \text{простая, двойная, тройная} \}$

*Вспомогательный термин «тип связи» обозначает множество возможных типов связи по кратности.*

1.2. *Номера элементов*  $\equiv I[1, \infty]$

*Вспомогательный термин «номера элементов» обозначает множество возможных номеров элементов.*

1.3. *Множество связей*  $\equiv (\cup (n: [1, \text{максимальное количество связей}])) (\times \text{химические элементы, номера элементов, тип связи}) \uparrow n$

*Вспомогательный термин «множество связей» обозначает множество связей, которые могут быть у одного химического элемента. Каждая связь представляет собой тройку: химический элемент, с которым образована связь; номер этого химического элемента; тип связи между химическими элементами.*



1.4. Возможные компоненты структурной формулы  $\equiv (\times \text{ химические элементы, номера элементов, } \{\text{множество связей}\})$

Вспомогательный термин «компонента структурной формулы» обозначает множество компонент структурной формулы. Каждая компонента представляет собой тройку: химический элемент; номер этого химического элемента; множество связей, которые образует этот химический элемент.

1.5. Возможная структурная формула  $\equiv (\cup (n: [1, \infty]) - \text{возможные компоненты структурной формулы} \uparrow n)$

Вспомогательный термин «возможная структурная формула» обозначает множество возможных компонент структурной формулы.

1.6. Компонента структурной формулы  $\equiv (\lambda(v: \{(\times \text{ возможная структурная формула, } I) \pi(2, v1) \in I[1, \text{length}(\pi(1, v1))]\}) \pi(1, \pi(\pi(2, v), \pi(1, v))))$

Вспомогательный термин «компонента структурной формулы» обозначает функцию, которая ставит в соответствие структурной формуле и номеру компонента в этой структурной формуле химический элемент, имеющий данный номер в структурной формуле.

1.7. Связи элемента в компоненте структурной формулы  $\equiv (\lambda(v: \{(\times \text{ возможная структурная формула, } I) \pi(2, v1) \in I[1, \text{length}(\pi(1, v1))]\}) \pi(3, \pi(\pi(2, v), \pi(1, v))))$

Вспомогательный термин «связи элемента в компоненте структурной формулы» обозначает функцию, которая ставит в соответствие структурной формуле и номеру компонента в этой структурной формуле множество связей химического элемента структурной формулы, имеющего данный номер в структурной формуле.

Описание сортов имен.

2.1.  $\chi(\text{структурная формула}) = (\text{органическое соединение} \rightarrow \text{возможная структурная формула})$

Термин «структурная формула» обозначает функцию, которая ставит в соответствие возможной формуле вещества множество его возможных структурных формул.

Онтологические соглашения.

3.1. ( $v$ : органическое соединение)

( $i$ :  $I[1, \text{length}(\text{структурная формула}(v))]$ )

( $j$ :  $I[1, \text{length}(\text{структурная формула}(v))]$ )

$i \neq j \Rightarrow (\pi(2, \pi(i, \text{структурная формула}(v)))) \neq (\pi(2, \pi(j, \text{структурная формула}(v))))$

Каждый атом химического элемента структурной формулы имеет

свой уникальный номер.

3.2. ( $v$ : органическое соединение)

( $i$ :  $I[1, \text{length}(\text{структурная формула}(v))]$ )

$(\pi(2, \pi(i, \text{структурная формула}(v)))) \leq \text{length}(\text{структурная формула}(v))$

Номера элементов в структурной формуле не превышают количества компонент структурной формулы, входящих в нее.

3.3. ( $v$ : возможная формула вещества)

( $v1$ :  $\{(v2$ : химические элементы) принадлежит соединению( $v, v2$ )\})

$(\forall (i$ :  $I[1, \text{length}(\text{структурная формула}(v))]$ )  $v1 = \pi(1, \pi(i, \text{структурная формула}(v)))$ )

Структурная формула содержит описание каждого химического элемента, входящего в состав данного вещества.

3.4. ( $v$ : возможная формула вещества)

( $v1$ :  $\{(v2$ : химические элементы) принадлежит соединению( $v, v2$ )\})

$(\mu (\{(i$ :  $I[1, \text{length}(\text{структурная формула}(v))]$ )  $\pi(1, \pi(i, \text{структурная формула}(v))) = v1\}) = \text{индекс}(v, v1)$ )

Структурная формула вещества содержит описание каждого атома каждого химического элемента, входящего в состав данного химического соединения.

3.5. ( $v$ : органическое соединение)

( $i$ :  $I[1, \text{length}(\text{структурная формула}(v))]$ )

$\mu(\text{связи элемента в компоненте структурной формулы}(\text{структурная формула}(v), i)) \leq \text{число электронов для присоединения у элемента}(\pi(1, \pi(i, \text{структурная формула}(v))))$

Число связей каждого атома некоторого химического элемента соединения не может превышать максимально возможное число электронов для присоединения у этого химического элемента.

3.6. ( $v$ : органическое соединение)

( $i$ :  $I[1, \text{length}(\text{структурная формула}(v))]$ )

$\text{связи элемента в компоненте структурной формулы}(\text{структурная формула}(v), i) \neq \emptyset$

В соединении не может быть несвязных атомов.

## Заключение

Таким образом, в работе описано устройство модульной модели онтологии органической химии в пределах вузовского курса обучения. В работе приведены существующие онтологии, описывающие химические понятия, указаны их основные недостатки и описана структура модели.

Созданная модель онтологии предоставляет возможность постановки

любых задач предметной области «Органическая химия» в пределах вузовского курса обучения и будет использована при создании компьютерного банка знаний по химии.

## ЛИТЕРАТУРА

1. *Kleshchev A.S., Artemjeva I.L.* A structure of domain ontologies and their mathematical models/ // In the Proceeding of The Pacific Asian Conference on Intelligent systems 2001 (PAIS 2001), Korea Intelligent Information Systems Society, 2001. P.410-420.
2. *Robert S., Carole G., Horrocks I., Bechhofer S.* Building a Bioinformatics Ontology Using OIL. // Department of Computer Science University of Manchester. Oxford Road Manchester UK M13 9PL, August 1, 2001.
3. *Mariano Fernández-López, Asunción Gómez-Pérez.* Methontology. // Laboratorio de Inteligencia Artificial, Facultad de Informática, Universidad Politécnica de Madrid, Campus de Montegancedo sn. Boadilla del Monte, 28660. Madrid, Spain.  
<http://delicias.dia.fi.upm.es/ontoweb/sig-tools/meetings/sig-amsterdam/methontology.doc>.
4. *Paul E. van der Vet Piet-Hein Speel, Nicolaas J.I. Mars.* The Plinius ontology of ceramic materials. // Memoranda Informatica. 1995. P.95-35.
5. *Paul E. van der Vet Piet-Hein Speel, Nicolaas J.I. Mars.* Bottom-up construction of ontologies: the case of an ontology of pure substances. // Memoranda Informatica. 1995. P.95-31.
6. *Артемьева И.Л., Цветников В.А., Реутов В.А.* Иерархическая модель онтологии физической химии. Ч. 1. Модель метаонтологии "Сущности". Препр. Владивосток: ИАПУ ДВО РАН, 2001. (<http://www.iacp.dvo.ru/es/>).
7. *Артемьева И.Л., Цветников В.А., Реутов В.А.* Иерархическая модель онтологии физической химии. Ч. 2. Модели систем понятий "Вещества" и "Реакции". Препр. Владивосток: ИАПУ ДВО РАН, 2001. (<http://www.iacp.dvo.ru/es/>).
8. *Артемьева И.Л., Цветников В.А., Реутов В.А.* Иерархическая модель онтологии физической химии. Ч. 3. Модели онтологий "Основы термодинамики" и "Термодинамика. Физические свойства". Препр. Владивосток: ИАПУ ДВО РАН, 2001. (<http://www.iacp.dvo.ru/es/>).
9. *Артемьева И.Л., Цветников В.А., Реутов В.А.* Иерархическая модель онтологии физической химии. Ч. 4. Модели онтологий "Термодинамика. Химические свойства", "Термодинамика. Связь физических и химических свойств" и "Химическая кинетика". Препр. Владивосток: ИАПУ ДВО РАН, 2001. (<http://www.iacp.dvo.ru/es/>).
10. *Артемьева И.Л., Цветников В.А.* Фрагмент онтологии физической химии и его модель // Электронный журнал "Исследовано в России": 2002. 3. <http://zhurnal.ape.relarn.ru/articles/2002/042.pdf>.
11. *Артемьева И.Л., Высоцкий В.И., Рештаненко Н.В.* Модель онтологии предметной области "Органическая химия". Электронное строение химических элементов. Препр. Владивосток: ИАПУ ДВО РАН, 2003. (<http://www.iacp.dvo.ru/es/>).
12. *Артемьева И.Л., Высоцкий В.И., Рештаненко Н.В.* Модель онтологии предметной области "Органическая химия". Органические соединения: функциональные группы, элементный состав, тип гибридизации, агрегатное состояние, растворимость. Препр. Владивосток: ИАПУ ДВО РАН, 2003. (<http://www.iacp.dvo.ru/es/>).
13. *Артемьева И.Л., Высоцкий В.И., Рештаненко Н.В.* Модель онтологии предметной области "Органическая химия". Органические соединения: структурная формула, классификация по расположению атомов углерода, типы связи между химическими элементами, пространственное расположение, изомерия. Препр. Владивос-



- ток: ИАПУИАПУ ДВО РАН, 2003. (<http://www.iacp.dvo.ru/es/>).
14. *Артемова И.Л., Высоцкий В.И., Реитаненко Н.В.* Модель онтологии предметной области "Органическая химия". Органические реакции: классификация реакционных центров, радикалы, ионы, основные свойства реакций. Препр. Владивосток: ИАПУ ДВО РАН, 2003. (<http://www.iacp.dvo.ru/es/>).
  15. *Клещев А.С., Артемова И.Л.* Математические модели онтологии предметной области. Ч. 2. Компоненты модели // Научно-техническая информация. Сер. 2. 2001. № 3. С.19-28.
  16. *Клещев А.С., Артемова И.Л.* Необогатенные системы логических соотношений. Ч. 1. // Научно-техническая информация. Сер. 2. 2000. №7. С.18-28.
  17. *Клещев А.С., Артемова И.Л.* Необогатенные системы логических соотношений. Ч. 2. // Научно-техническая информация. Сер.2. 2000. №8. С.8-18.
  18. *Майерс Г.* Надежность программного обеспечения. // М.: Мир, 1980.

*Статья представлена к публикации членом редколлегии А.С. Клещевым.*

УДК 519.68:15:54

© 2004 г. **Н.О. Лицис**

(Институт автоматизации и процессов управления ДВО РАН, Владивосток)

## **ПРОТОТИП ИНТЕРНЕТ-СИСТЕМЫ КЛАССИФИКАЦИИ ОРГАНИЧЕСКИХ РЕАКЦИЙ<sup>1</sup>**

Описывается прототип Интернет-системы, решающей задачу классификации органических реакций на основе построенной модели онтологии соответствующего раздела химии. Приведена модель онтологии, а также краткое описание архитектуры программной системы.

### **Введение**

При построения модели онтологии рассматриваемого раздела предметной области использовались источники [1 – 9].

Органическая химия в настоящее время переживает период бурного развития. Использование новейших физико-химических методов исследования приводит к углублению теоретических представлений в органиче-

---

<sup>1</sup> Работа выполнена в рамках программы № 16 фундаментальных исследований Президиума РАН на 2004 г. Проект № 10002-251/П16/021-387/110504-266.