

УДК 81'37:[002:681.3.016:54]

М. Ф. Мизинцева

СПЕЦИАЛЬНОЕ СТРУКТУРИРОВАНИЕ ХИМИЧЕСКОЙ ИНФОРМАЦИИ

Рассматриваются вопросы, касающиеся особенностей структуры конкретной предметной области одного из важнейших разделов современного естествознания — химии. Обсуждаются примеры построения семантических сетей по химии высокомолекулярных соединений, силикатных материалов, координационных соединений. Формулируются принципы построения семантических сетей по химии и химической технологии.

Химия — хорошо структурированная область естествознания, изучающая материю и химическую форму движения на уровне атомов и молекул. Современная химия описывает многие явления законами, выраженными в математической форме, сложность которой затрудняет точное решение. В то же время количественные физико-химические характеристики весьма точно определяются из экспериментальных данных. Эта особенность приводит к необъятности химической информации.

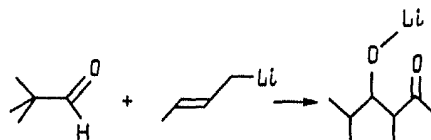
1. ХАРАКТЕРИСТИКА ПРЕДМЕТНОЙ ОБЛАСТИ ХИМИИ

Основные химические термины и законы в совокупности с химической символикой и номенклатурой составляют основу химического языка, который отражает логические и математические связи исследуемого химического явления и представляет его в виде модели. Названия и химические формулы соединений отражают их качественный и количественный состав, являются источником информации и систематизированы по различным принципам. В настоящее время общепризнанная система номенклатуры для различных классов соединений утверждена международной химической организацией IUPAC [1, 2, 3]. Большое значение для структурирования химической науки имеют также фундаментальные химические теории и законы. Одной из основополагающих теорий химии является теория строения вещества, главный тезис которой — наличие обратной связи между строением и свойствами вещества (А. М. Бутлеров [4]). На основании этого фундаментального положения возможно глубокое структурирование практически любого раздела или области химии.

Глубокие внутренние связи всех химических элементов и их соединений отражает Периодический закон Д. И. Менделеева. Графическое изображение периодического закона — система химических элементов — также приводит к естественному структурированию химии по различным признакам. Использование химических формул и символов дает возможность кодировать вещества (химические соединения).

Большие возможности структурирования открывает наличие генетических связей между классами различных соединений. Так, в неорганической химии — это связь: «простое вещество — оксид — гидроксид — соли», в органической химии — «кислородсодержащие производные углеродов: спирты — альдегиды — кислоты — эфиры».

Иерархическое деление на классы справедливо в большинстве случаев для всех разделов химии. Примером иерархической организации химической науки может служить один из современных разделов органической химии — химия высокомолекулярных соединений (ВМС) или полимеров. Органическая химия использует в качестве основного признака ВМС структуру и последовательность присоединения фрагментов макромолекулы — так называемых функциональных групп, что и определяет класс полимера и его свойства. Следует также отметить как важную особенность органической химии использование графического изображения структуры, т. е. языка графики, для представления знаний. Например, сущность реакции альдольной конденсации может быть изображена графически:



Прикладное значение химической науки, тесно связанное с проблемами синтеза и изучения свойств новых химических соединений, может быть отражено при структурировании химии как предметной области [5].

Одним из наиболее эффективных способов представления химической информации являются семантические сети [6].

2. ПОСТРОЕНИЕ СЕМАНТИЧЕСКИХ СЕТЕЙ ПО ХИМИИ

Используя естественное структурирование химической науки, можно представить химию как предметную область с помощью семантических сетей. В настоящей работе построены семантические сети для различных разделов и направлений современной химии, таких, как «Химия высокомолекулярных соединений», «Химия координационных соединений», «Теория электролитической диссоциации», «Химия силикатных материалов» и др.

Каждый из перечисленных разделов представляет собой конкретную предметную область, обособленную по определенным признакам. Для построения семантических сетей был использован метод Куиллиана [7], модель которого имитирует естественное понимание языка.

2.1. Построение семантической сети по химии высокомолекулярных соединений

Химия высокомолекулярных соединений выделена нами в предметную область на основании классификации полимеров. Возможны различные способы классификации ВМС. На рис. 1 представлена семантическая

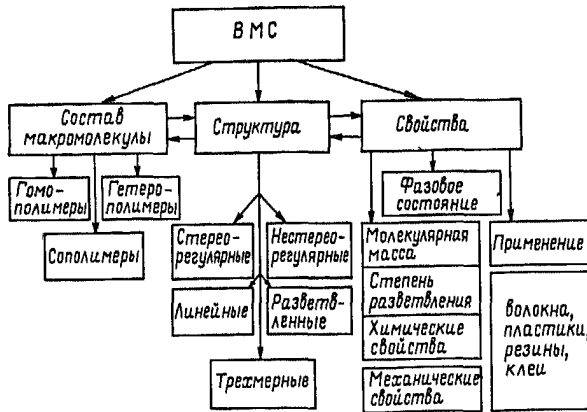


Рис. 1. Семантическая сеть «Химия высокомолекулярных соединений»

сеть ВМС, в которой отражены различные подходы к классификации полимеров. Основой семантической сети является цикл «состав — структура — свойства — применение». По химическому составу макромолекулы ВМС делятся на гомо- и гетерополярные, а также сополимеры, по характеру присоединения макромолекулы различают цепные (линейные и разветвленные) и сетчатые полимеры, по фазовому состоянию — аморфные и кри-

сталлические, по происхождению — природные и синтетические. Состав и структура определяют свойства и области применения полимеров. Так, линейные полимеры, способные к высокоэластичному состоянию, используются в качестве волокон, пленок, клеев и т. п. Сетчатые или сшитые полимеры в большей степени проявляют механические свойства (прочность и упругость), что позволяет использовать их как каучукоподобные и резиновые пластики. Важной характеристикой ВМС является молекулярная масса (M), величина которой колеблется от нескольких тысяч до млн атомных единиц массы (а. е. м.), причем, чем больше M , тем прочнее полимеры.

Особый класс составляют биополимеры, которые являются основой всех живых организмов и участвуют практически во всех процессах жизнедеятельности [8].

Как видно из рис. 1, в качестве вершин данной семантической сети использованы понятия «состав макромолекулы», «структура», «свойства». В то же время эти вершины являются именами объектов, которые образуют узлы сети. В соответствии с иерархией на более низких уровнях расположены имена свойств и значений. Между уровнями с помощью стрелок установлены функциональные отношения. Например, рассмотрим цикл, образованный объектом «структура», свойствами и значениями, которые зависят от характера структуры и являются одновременно параметрами, определяющими конкретные качества ВМС, представленные в узле, образованном объектом «свойства». Объединив все циклы с помощью так называемых «сценариев», получаем семантическую сеть.

Таким образом, при составлении семантической сети «Химия полимеров» использованы формальные признаки этой области, связанные с терминологией, а также проявления внутренних связей между составом и свойствами полимеров.

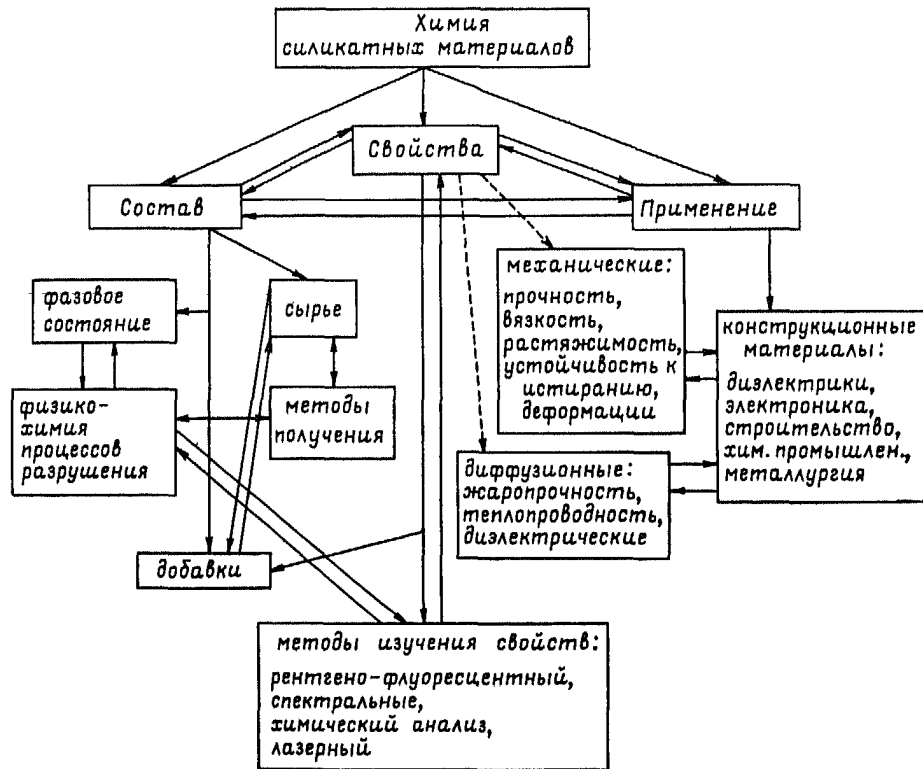


Рис. 2. Семантическая сеть «Химия силикатных материалов»

2.2. Построение семантической сети «Химия силикатных материалов»

Аналогичный подход применен при построении семантической сети «Химия силикатных материалов» (рис. 2). Однако основным признаком данного раздела ПО был не только тот факт, что все эти соединения образованы или получены на основе одного химического элемента — кремния (Si), но и то, что все эти вещества являются конструкционными материалами с уникальными комплексами физико-химических свойств, т. е. по принципу их применения. Развитие электронной и аэрокосмической промышленности стало возможным благодаря появлению многих новых керамических материалов. Солнечные батареи, атомные и химические реакторы — примеры устройств и оборудования из керамики и силикатов [9]. Силикатные материалы чрезвычайно широко распространены в природе (кварц, каолин, полевой шпат и т. п.), составляют до 75% земной коры и находят широкое применение в промышленности. Учитывая эти особенности и признаки ПО, в качестве вершин семантической сети были выбраны объекты, которым присвоены имена «состав», «применение», «свойства».

Семантическая сеть на рис. 2 раскрывает важные циклические связи между вершинами семантической сети «состав — применение», а также между вершинами, соответствующими методам изучения свойств и методам получения. Таким образом, семантическая сеть «Химия силикатных материалов» отражает особенности данной предметной области.

2.3. Построение семантической сети «Теория электролитической диссоциации»

Особый интерес представляло построение семантической сети для ПО, содержанием которой были бы теоретические разделы современной и классической химии. В том числе те теории, которые могут быть описаны с использованием не только химического языка, но также и математического аппарата. Для примера был выбран раздел физической химии «Растворы», в частности, классическая теория электролитической диссоциации

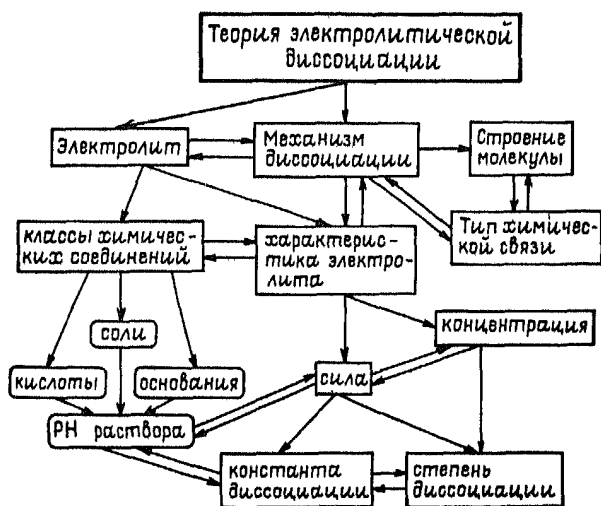


Рис. 3. Семантическая сеть «Теория электролитической диссоциации»

[10]. Как известно, ее основные положения состоят в следующем:

все вещества делятся на электролиты; электролиты в водном растворе под действием молекул воды распадаются с образованием катионов и анионов;

диссоциация — обратимый процесс. Развитие этих основных положений определило выявление существенных закономерностей и правил, применяя которые можно получить важные физико-химические характеристики растворов электролитов — например, силу электролитов, pH раствора, константу и степень диссоциации, а в связи с учетом обратимости процесса диссоциации возможно приложение к его описанию законов химической кинетики.

На рис. 3 представлена разработанная нами семантическая сеть «Теория электролитической диссоциации» (ТЭД). Данная семантическая сеть не только отражает терминологию теории, но и показывает сущность основных положений ТЭД. Например, из сети ясно, что электролитами могут являться определенные классы химических соединений: соли, кислоты, основания. Строение молекулы электролита связано с типом химической связи конкретных соединений, что определяет механизм диссоциации. Кроме того, как видно из рис. 3, семантическая сеть показывает и возможность количественной характеристики растворов электролитов.

2.4. Построение семантической сети «Химия координационных соединений»

ПО, обозначенная нами как «Комплексные или координационные соединения», — направление неорганической химии, основной проблемой которого является синтез и изучение новых комплексных соединений. Этот раздел химии как ПО имеет ряд особенностей. С одной стороны, ее структура носит общий характер, и в то же время существует специфическая терминология, которая во многом связана с современными представлениями о координационной связи. В принципе можно выделить некоторые основные проблемы, которые будут соответствовать другим вершинам семантической сети, помимо стандартных для предметных областей химии («состав — свойства»). Это — классификация, методы синтеза и исследования, основы технологических процессов (экстракция, осаждение и т. д.) [11].

Семантическая сеть «Комплексные или координационные соединения» выявляет не только сущность этой предметной области, но и рассматривает ее методологическую часть (рис. 4). Например, рассмотрим фрагмент данной семантической сети, описывающей цикл «свойства — физико-химические характеристики — константы устойчивости — методы расчета — методы исследования». Выделим в данном фрагменте один из важнейших циклов, отражающий основную проблему химии координационных соединений «методы расчета — константы устойчивости — методы исследования». Функциональная зависимость этих элементов семантической сети заключается в том, что методы исследования, рассматривающие различные параметры синтезированных соединений, определяют конкретные методы расчета, использующие данные параметры (условия синтеза, pH раствора, концентрации взаимодействующих реагентов, ионную силу раствора, природу элемента комплексообразователя и т. д.) [12]. В то же время, и выбор измеряемого параметра исследуемого комплекса определяет метод расчета. В связи с этим возникает необходимость расширить фрагмент семантической сети фреймовыми структурами, что позволяет использовать экспериментальные данные.

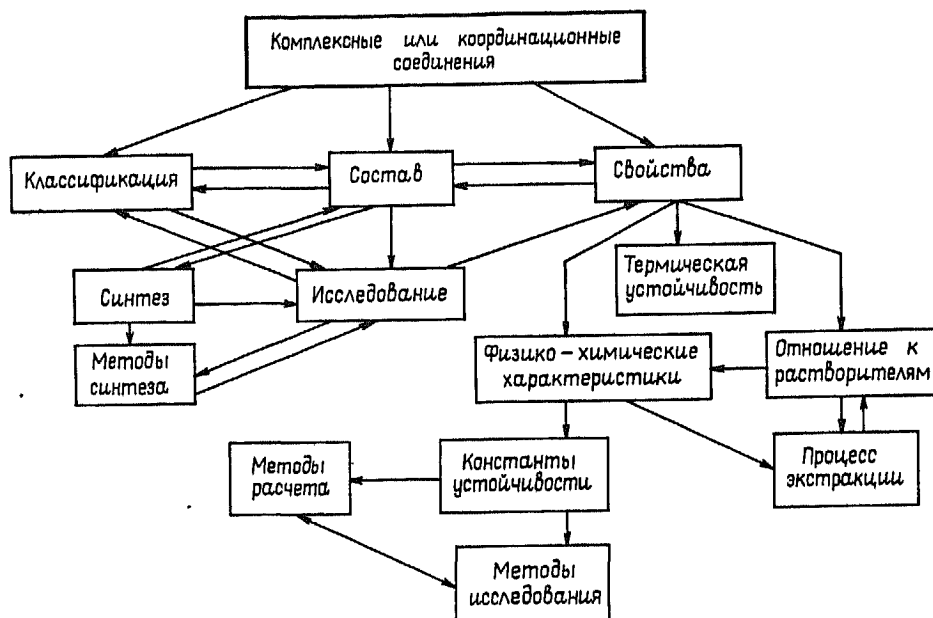


Рис. 4. Семантическая сеть «Комплексные или координационные соединения»

3. ПРИНЦИПЫ ПОСТРОЕНИЯ СЕМАНТИЧЕСКИХ СЕТЕЙ ПО ХИМИИ

Таким образом, при составлении семантических сетей по конкретным предметным областям химии были использованы формальные признаки ПО (термины, имена понятий), теоретические положения и закономерности, а также установлены внутренние связи семантической сети.

Так, при построении семантической сети «Химия ВМС» использованы различные подходы к классификации полимеров; в основе семантической сети «Химия силикатных материалов» — особенности направления применения керамики; семантическая сеть «Химия координационных соединений» аккумулировала разнообразие проблем этой области неорганической химии; семантическая сеть теоретического раздела «ТЭД» базировалась на основных положениях и закономерностях теории, в том числе и выраженных математически. На основании изложенного выше и учитывая анализ ПО химии, способы выделения основных понятий и отношений, возможность дальнейшей детализации семантической сети, а также специфику химии как экспериментальной науки, были сформулированы некоторые принципы построения семантических сетей.

Построение семантических сетей возможно на основе:
 — классификации и номенклатуры веществ и соответствующей терминологии, что отражает естественные генетические связи классов химических соединений;
 — теоретических положений, закономерностей, особенностей направления научной области, т. е. тех научных фактов, которые аккумулируются экспертами;
 — теории строения химических соединений, отражающей функциональную связь основных химических понятий «состав — структура — свойства». Естественным развитием семантических сетей, построенных на основе этого принципа, является рассмотрение конкретных областей применения химических веществ.

Сформулированные принципы могут быть использованы для построения семантических сетей различных современных областей химии.

СПИСОК ЛИТЕРАТУРЫ

1. Гордон А., Форд Р. Спутник химика.— М.: Мир, 1976.
2. Химия и периодическая таблица / Под ред. К. Сайото.— М., 1982.— С. 366.
3. J. Chem. Education.— 1971.— № 48.— Р. 433.
4. Попов Э. В. Экспертные системы. Решение неформализованных задач в диалоге с ЭВМ.— М.: Наука, 1982.— 288 с.
5. Коршак В. В. Химия высокомолекулярных соединений.— М.: Высшая школа, 1951.
6. Мизинцева М. Ф., Ананьева Т. Н., Бондарь В. В., Лангбайн Д. Концепция лингвистического обеспечения проблемноориентированной базы знаний // НТИ. Сер. 2.— 1991 — № 6.— С. 9—14.
7. Представление и использование знаний / Под ред. Х. Уэно, М. Исидзука.— М.: Мир, 1989.— 220 с.
8. Химический энциклопедический словарь.— М.: Сов. энциклопедия, 1983.
9. Juillard J. Modern Ceramics // Impact of Science on Society.— 1988.— № 149.— Р. 26—39.
10. Курс физической химии / Под ред. А. В. Киреева — М.: Высшая школа, 1972.
11. Современная химия координационных соединений / Под ред. Дж. Льюиса, Р. Уилкинса.— М.: Ин. лит-ра, 1963.
12. Альберт А., Сергент Е. Константы конденсации кислот и оснований.— М.: Ин. лит-ра, 1972.

Материал поступил в редакцию 05.01.95.