

УДК 541.6

КОМПЬЮТЕРНАЯ ГЕНЕРАЦИЯ ТОПОЛОГИЧЕСКИХ ОБРАЗОВ ОРГАНИЧЕСКИХ СОЕДИНЕНИЙ ИСХОДЯ ИЗ ИХ ГРАФИЧЕСКОГО ПРЕДСТАВЛЕНИЯ

А.А. Репин, В.В. Туровцев, Ю.Д. Орлов
Кафедра общей физики

Разработан интерфейс для ввода и отображения структурных формул химических соединений. Созданы и реализованы алгоритмы генерации их топологических образов.

Использование компьютерных этих технологий привело к созданию электронных баз данных, в том числе и по термодинамическим свойствам органических соединений. Однако сегодня базы термодинамических данных для органических соединений не полностью удовлетворяют современным потребностям. Решение этой проблемы мы видим в создании баз термодинамических данных нового поколения [1; 2]. Нами в течение ряда лет ведется работа по созданию баз термодинамических данных нового поколения с элементами искусственного интеллекта. Полученные к настоящему времени результаты опубликованы в [1-4]. Эта работа включает ряд этапов, одним из которых является разработка удобного для пользователя способа ввода информации и приведения ее к виду, оптимальному для выполнения расчетов.

В настоящее время на рынке программных продуктов имеются такие пакеты программ, как HyperChem, ChemOffice, ISIS_Draw и др. Получение исходных кодов данных разработок сопряжено с большими финансовыми затратами, если вообще возможно. Поэтому необходима разработка собственных средств ввода и представления информации. Решению этой задачи посвящена данная статья.

Наиболее привычным и общепринятым способом изображения химических соединений является структурная формула, которая и была выбрана в качестве исходной для графического изображения органических соединений на экране. Одним из наиболее удобных способов отображения химических структур является их представление в виде топологических образов. Такой образ – матрицу смежности молекулярного графа мы использовали в качестве основного [3]. Рассматриваемый этап включает построение структурных формул на экране компьютера с последующим созданием модифицированной матрицы смежности молекулы или радикала.

Построение графического изображения. Структурная формула представляет собой изображение атомов в виде символов, обозначающих химические элементы, химических связей между ними, которые задаются несколькими линиями в зависимости от типа химической связи (одинарные, двойные и тройные), и бензольных колец, представленных правильными шестиугольниками с вписанными в них окружностями.

Информация о рисуемых химических элементах вносится в записи (record) атомов по мере их отображения. Запись атома имеет следующие поля:

- символьное представление атома;
- номер, идентифицирующий химический элемент;
- координаты изображения атома;
- валентность;
- количество неспаренных электронов;
- количество химически связанных с ним атомов водорода;
- число, показывающее принадлежность этого атома к бензольным кольцам.

Все записи атомов включаются в список (list) указателей на соответствующую запись. Для изображения на экране химического элемента, необходимо активизировать его запись. В результате вызывается процедура внесения активизированного химического элемента (записи атома) в список атомов с координатами курсора в момент щелчка. Параметром процедуры является запись активизированного атома. В процедуре выполняется проверка на незанятость окрестной области курсора изображением других атомов и бензольных колец.

Информация о рисуемых химических связях вносится в запись связей, которая имеет следующую структуру:

- тип химической связи,
- указатель на 1-й атом,
- указатель на 2-й атом.

Записи связей вносятся в список связей, являющийся списком указателей на соответствующую запись. Для изображения на экране необходимой химической связи, ее следует активизировать. Далее щелкнуть левой кнопкой мыши на атомы, между которыми требуется нарисовать связь. В результате этих действий выполняется вызов функции, которая возвращает указатель на атом, если курсор находился над этим атомом и нет ограничений (учитываются свойства химического элемента, наличие атомов водорода) на возможность существования активизированной связи с выбранным атомом. Если при первом щелчке мышью возвращается непустой указатель, он записывается в промежуточную запись. При втором щелчке проверяется возможность попадания в первый атом, если этого не произошло и возвращаемый указатель непустой, он вносится в промежуточную запись связи. После этого в записях атомов выполняется уменьшение связанных с ними атомов водорода на порядок связи. Затем промежуточная запись связи вносится в список связей, и рисуется данная связь.

Информация о бензольных кольцах представляется записями бензольных колец, которая имеет следующую структуру:

- массив указателей на записи атомов размерностью 6;
- координаты левого верхнего и правого нижнего углов квадрата, в который вписывается окружность.

При изображении бензольных колец возникает проблема не только проверки на свободное место, но и возможности построения систем

конденсированных колец. Для ее решения нами разработаны оригинальные алгоритмы, которые были реализованы в процедурах построения бензольного кольца и систем бензольных колец. Для первой процедуры параметрами являются координаты в момент щелчка левой кнопкой мыши, а для второй - координаты первого атома (левый верхний, обход по часовой стрелке) в бензольном кольце.

Процедура построения бензольного кольца:

1. Вызываем функцию определения выбранной связи.
2. Если функция возвратила непустой указатель, то:
 - 2.1. Если функция возвратила указатель не на связь в бензольном кольце, то выход из процедуры.
 - 2.2. Если атомы выбранной связи не связаны с атомами водорода и входят только в одно бензольное кольцо, то выход из процедуры.
 - 2.3. Если $Atom1Y > Atom2Y$ (координаты 1-го и 2-го атомов связи), то:
 - 2.3.1. $Y1 := Atom1Y$ ($Y1, Y2$ – промежуточные переменные)
 - 2.3.2. $Y2 := Atom2Y$
 - 2.3.3. $X1 := Atom1X$ ($X1, X2$ – промежуточные переменные)
 - 2.3.4. $X2 := Atom2X$
 - 2.4. иначе
 - 2.4.1. $Y1 := Atom2Y$
 - 2.4.2. $Y2 := Atom1Y$
 - 2.4.3. $X1 := Atom2X$
 - 2.4.4. $X2 := Atom1$
 - 2.5. Если $X1 = X2$, то:
 - 2.5.1. Вызов процедуры построения смежного бензольного кольца с параметрами $X2$ и $Y2$
 - 2.5.2. Вызов процедуры построения смежного бензольного кольца с параметрами $X2-70$ и $Y2$
 - 2.6. иначе
 - 2.6.1. Если $X1 > X2$, то
 - 2.6.1.1. Вызов процедуры построения смежного бензольного кольца с параметрами $X1-70$ и $Y1$
 - 2.6.1.2. Вызов процедуры построения смежного бензольного кольца с параметрами $X2$ и $Y2-40$
 - 2.6.2. иначе
 - 2.6.2.1. Вызов процедуры построения смежного бензольного кольца с параметрами $X1$ и $Y1$
 - 2.6.2.2. Вызов процедуры построения смежного бензольного кольца с параметрами $X1-70$ и $Y1-40$
3. Иначе
 - 3.1. Проверка нахождения атомов в окрестности курсора по всем элементам списка атомов
 - 3.2. Если место занято, то выход из процедуры
 - 3.3. Добавление 6 записей атомов в список атомов и в массив записи бензольного кольца
 - 3.4. Добавление 6 записей связей в список связей
 - 3.5. Добавление записи бензольного кольца в список бензольных колец
 - 3.6. Изображение бензольного кольца на экране
4. Конец процедуры.

Процедура построения бензольного кольца выполняется, когда активизировано рисование бензольного кольца после щелчка левой мыши в определенном нами месте. Далее функция определения выбранной связи возвращает указатель на связь, если мы попали на изображение связи, иначе - пустой указатель. Если указатель непустой, то определяем направление наклона связи. В зависимости от последнего передаем координаты 1-го атома бензольного кольца в процедуру построения смежного кольца. Причем эта процедура вызывается два раза, так как смежное кольцо можно нарисовать как слева от выбранной связи, так и справа. Процедура построения смежного бензольного кольца проверяет наличие в данном месте другого бензольного кольца и других атомов за исключением бензольных атомов, с которыми может быть выполнено сопряжение. В случае их обнаружения выполнение процедуры прерывается и действие передается далее по ходу вызывающей процедуры, иначе производится проверка на возможность сопряжения с соседними бензольными кольцами. В случае их отсутствия добавляются недостающие атомы и связи, которые вносятся в соответствующие списки. Если функция определения выбранной связи возвращает пустой указатель, то выполняется проверка на свободное место. В случае положительного результата список атомов пополняется 6 записями атомов, расположенных в вершинах правильного шестиугольника, список связей - 6 записями связей между этими атомами и список бензольных колец - новой записью бензольного кольца.

Создание модифицированной матрицы смежности. После построения структурной формулы для выполнения расчетов выбирается химическое соединение щелчком правой кнопки. В результате этих действий строится матрица смежности. Размерность матрицы определяется количеством записей в списке атомов. На главной диагонали записывается код, несущий информацию об идентификационном номере химического элемента, количестве химически связанных с ним водородов, валентности и количестве неспаренных электронов. Недиагональные элементы матрицы первоначально заполняются нулями, а затем по списку связей (по номерам атомов, между которыми имеется химическая связь) определяется недиагональный элемент, в который записывается порядок связи. По координатам курсора в момент щелчка выбираются ближайший атом и его номер в списке. По этому номеру находятся номера атомов, входящих в выбранную структурную формулу по алгоритму, реализованному в процедуре построения матрицы выбранного химического соединения. Параметром этой процедуры является номер выбранного атома.

Процедура построения матрицы выбранного химического соединения

1. В первый элемент массива выделяемых номеров атомов заносим параметр процедуры.
2. Запоминаем индекс массива.
3. Выбираем все элементы строки, кроме равных нулю и диагонального,

- номер которой указан в элементе массива с помеченным индексом.
4. Если номера столбцов выбранных элементов строки отсутствуют в массиве, записываем их в конец массива.
 5. Помеченный индекс увеличиваем на 1.
 6. Если помеченный индекс не превышает размера массива, то переходим к пункту 3.
 7. Полученные элементы массива упорядочиваем по возрастанию.
 8. Вызываем процедуру построения матрицы смежности из матрицы большего размера по указанным номерам.

В пунктах с 1 по 6 выделяются номера атомов, входящих в выбранное изображение химического соединения. Алгоритм процедуры, вызываемой в пункте 8, описан в [4]. Использование данной процедуры позволяет изображать в одном окне несколько структурных формул.

Заключение. Вышеописанные алгоритмы использовались при создании графического интерфейса, послужившего связующим звеном, дающим пользователю простую возможность ввода информации об органических соединениях в виде структурных формул для ее дальнейшего использования в программах, реализующих расчетные методы предсказания термодинамических характеристик. Создание данного интерфейса является одним из важнейших шагов в нашей работе по созданию баз термодинамических данных нового поколения с элементами искусственного интеллекта. Все разработки проводятся в среде визуального программирования Delphi. Данные алгоритмы реализованы с использованием объектно-ориентированного программирования в нескольких модулях. Это позволяет использовать их при создании других приложений данного направления.

Литература

1. Репин А.А., Осин В.Н. Концепция создания автоматизированных баз термодинамических данных нового поколения // VII Региональные Каргинские чтения: «Химия и химическая технология». Тез. Докл. Тверь, 2001. С. 35.
2. Репин А.А., Осин В.Н. Разработка автоматизированной базы термодинамических данных с элементами искусственного интеллекта // 8-я Всеросс. межвузовская науч.-техн. конф. «Микроэлектроника и информатика – 2001»: Тез. Докл. Москва, 2001.
3. Репин А.А., Орлов Ю.Д. Тополого-групповой алгоритм фрагментации органических соединений. // Учен. записки ТвГУ. 1999. Т. 5. С. 102-105.
4. Репин А.А., Орлов Ю.Д. Алгоритм автоматизации расчёта энергий диссоциации связей // Учен. записки ТвГУ. 2001. Т. 6. С. 78-81.