

А. В. Козодоев, Е. М. Козодоева

*Институт оптики атмосферы им. В. Е. Зуева СО РАН
пл. Академика Зуева, 1, Томск, 634055, Россия*

kav@iao.ru, klen@iao.ru

БИНАРНЫЕ ОПЕРАЦИИ В ИНФОРМАЦИОННОЙ СИСТЕМЕ «МОЛЕКУЛЯРНАЯ СПЕКТРОСКОПИЯ»

Представлен подход, использованный при разработке и реализации модуля, выполняющего бинарные операции над данными в ИС «Молекулярная спектроскопия». Приводится формализация бинарных операций над наборами спектроскопических данных с учетом особенностей предметной области. Описываются алгоритм действий и интерфейс пользователя для проведения бинарных операций – единые для имеющихся баз данных по различным веществам и нескольким типам спектроскопических данных.

Ключевые слова: структуры данных, количественная спектроскопия, бинарные операции, базы данных.

Введение

Основным способом представления широкой общественности результатов научной деятельности являются публикации в научных изданиях. При небольшом объеме численных данных, полученных в ходе исследования, они публикуются непосредственно в статье в виде таблиц или графиков. В случае если объем численных данных велик для публикации в статье, такие данные публикуют в сети Интернет, размещая файлы, например, на FTP-серверах.

Современные исследования по молекулярной спектроскопии высокого разрешения в области состояний молекул и характеристик спектральных переходов дают постоянно растущие объемы численных данных благодаря совершенствованию как измерительного оборудования, так и вычислительной техники. Таким образом, результаты экспериментальных работ содержат значения параметров от десятков и сотен до десятков тысяч спектральных переходов или состояний молекулы [1; 2]. Теоретические (расчетные) работы могут содержать характеристики нескольких миллиардов переходов [3; 4].

В информационной системе (ИС) «Молекулярная спектроскопия» ведется накопление численных значений данных, публикуемых исследователями в статьях [5–8]. Совокупность извлеченных из опубликованных материалов численных значений спектроскопических данных будем называть набором данных. В отдельных статьях, как правило, представляются данные по свойствам лишь одной молекулы или небольшой группы молекул. Зачастую приводятся только часть набора параметров спектральных линий или исследования в узком спектральном диапазоне. Однако исследователям из прикладных областей, использующим необходимые спектральные параметры, например, в расчетах радиационных или климатических моделей, требуются данные в достаточно широком спектральном диапазоне. Это приводит к необходимости формирования составных наборов данных путем комбинирования данных из различных источников. Среди составных наборов можно выделить экспертные

наборы, предоставляющие достоверные, согласованные и наиболее полные данные как по перечню спектральных параметров, так и по спектральному диапазону [9].

Процесс создания таких наборов можно разделить на два вида действий: отбор данных с помощью условий в пределах одного набора данных и объединение отобранных данных в новый составной набор данных. Формирование составных наборов данных – процесс трудоемкий, но часть действий можно автоматизировать с помощью информационных технологий. Действия по выборке части данных из конкретного источника, удовлетворяющих наложенным ограничениям, были автоматизированы созданным в рамках ИС «Молекулярная спектроскопия» модулем «унарные операции» [10]. Для создания составного экспертного источника данных из отобранных частей исходных источников необходим механизм их объединения, позволяющий учитывать как формализуемые ограничения предметной области, так и принимаемые экспертами решения, не поддающиеся алгоритмизации.

Определение бинарных операций в молекулярной спектроскопии

Подробно структура данных молекулярной спектроскопии, над наборами которых производятся манипуляции, была рассмотрена нами в работе [10]. Здесь лишь коротко напомним, что набор данных представим в виде таблицы независимо от способа хранения, и имеются *ограничения на операции*, продиктованные особенностями предметной области.

1. Во всех наборах данных есть две обязательные части: набор квантовых чисел и набор некоторых физических величин. В зависимости от спектроскопической задачи обязательной физической величиной является уровень энергии или частота спектрального перехода. Остальные физические величины могут отсутствовать совсем либо иметь «пустоты» в отдельных строках.

2. Значения квантовых чисел в каждой строке таблицы является уникальным идентификатором, который однозначно определяет, к какому уровню энергии или спектральному переходу относятся значения набора физических величин, представленных в этой строке таблицы, согласно используемой модели молекулы. Сравнение строк производится только по набору значений квантовых чисел, который мы рассматриваем как единый элемент (идентификатор). Два идентификатора равны, если равны между собой все соответствующие квантовые числа. Остальные физические характеристики имеют приближенные значения в силу неточности самих моделей или измерений в эксперименте. Таким образом, все строки в одном наборе данных должны быть с различными (уникальными) идентификаторами как в исходных наборах, так и в результирующем наборе.

3. Операндами в бинарных операциях являются только канонические наборы данных, в которых квантовые числа удовлетворяют ограничениям на состояние и правилам отбора (определяют допустимые сочетания квантовых чисел верхнего и нижнего состояния). Операции производятся над данными, относящимися к одному типу спектроскопических задач (прямая / обратная) и одному веществу [11].

4. В исходных и результирующем наборах данных может быть только по одной колонке с данными по конкретной спектральной характеристике (физической величине).

Рассмотрим три операции над парой наборов данных по аналогии с теорией множеств: объединение, пересечение и разность. Эти операции являются манипуляциями с данными, т. е. действиями, не изменяющими сами значения данных. *Операция разности* наборов данных дает множество строк одного набора данных, не имеющих пары по набору значений квантовых чисел во втором наборе данных. *Операция объединения* наборов данных, позволяет соединять в единый набор строки из двух разных наборов независимо от наличия совпадений по набору значений квантовых чисел. С помощью *операции пересечения* наборов данных, можно выбрать строки из двух наборов с совпадающими наборами значений квантовых чисел.

Специфика предметной области приводит к особому процессу выполнения некоторых операций. Так, в операции разности результат однозначен и может быть получен автоматически, так как среди отобранных строк набора данных не может быть строк с одинаковыми наборами значений квантовых чисел. В то время как в операциях объединения и пересечения

имеет место *промежуточный результат операции*, где могут быть строки с совпадающими наборами значений квантовых чисел, но из разных наборов данных. Согласно принятым нами правилам результирующий набор данных должен содержать только канонические данные [11], т. е. среди прочих условий не иметь строк с совпадающими наборами значений квантовых чисел. Выбор строк, которые попадут в результирующий набор данных, может произвести только пользователь системы (эксперт) на основании имеющихся у него неформализуемых знаний. Это и есть неавтоматизированная часть операций манипулирования наборами данных. Таким образом, в операциях объединения и пересечения автоматически выполняется формирование промежуточного результата с выбранными из разных источников строками и создание нового набора данных с сохранением в него строк, отобранных экспертом.

Формализация бинарных операций

Для описания сути бинарных операций нам удобно использовать термины теории множеств. Обозначим набор данных по конкретному веществу, как $ND^j = \{K_i\}$, где $j \in \{\text{спектральные переходы, профили спектральных линий, уровни энергии молекулы}\}$ – тип спектроскопической задачи, а $\{K_i\}$ – это множество строк набора данных. Согласно теории множеств при выполнении операций (пересечения, объединения и разности) производится сравнение элементов по их значениям, однако в нашей предметной области строка в наборе данных не является элементарным объектом с одним значением. Каждая строка K_i набора данных имеет сложную структуру, для описания которой можно применить кортежи из алгебры кортежей [12].

Тогда элемент множества (набора данных) K_i можно представить кортежем, элементы которого разделены по смыслу на две части, как упоминалось ранее:

$$K_i = (i, S_i^j),$$

где $i = (qn_1, \dots, qn_M)$ – кортеж, содержащий значения квантовых чисел из предметной области, уникальная комбинация которых (идентификатор) в пределах набора данных однозначно идентифицирует K_i ; $S_i^j = (s_{i,1}^j, \dots, s_{i,N}^j)$ – кортеж, содержащий численные значения физических величин, набор которых зависит от спектроскопической задачи j , описывающих свойства i -го спектрального перехода или состояния молекулы, N – число физических характеристик.

Пересечение

Операция «пересечение» двух наборов данных производится построчно на основе сравнения идентификаторов. Выбираются строки с одинаковыми идентификаторами из обоих наборов данных, прочие строки в результат операции не попадают. Промежуточный результат может содержать по два значения одной и той же физической величины, соответствующих одному идентификатору. Так как в нашей информационной системе принято, что одному идентификатору должно соответствовать по одному значению каждой физической величины, то далее пользователь должен выбрать, какое значение физической величины включать в результат.

Пусть $ND1^j = \{(i, S_i^j)\}$ и $ND2^j = \{(p, S_p^j)\}$ – два исходных набора данных. Их пересечение можно расписать следующим образом:

$$ND1^j \cap ND2^j = ND3^j = \{(r, S_r^j)\},$$

где $r \in \{i\}$ и $r \in \{p\}$, а $S_r^j = ([s_{i,1}^j | s_{p,1}^j], \dots, [s_{i,N}^j | s_{p,N}^j])$.

С помощью записи $\left[s_{i,n}^j | s_{p,n}^j \right]$ мы обозначили, что эксперт, производящий операцию, делает выбор между $s_{i,n}^j$ и $s_{p,n}^j$, основываясь на имеющихся у него неформализуемых знаниях.

Объединение

Операция «объединение» двух наборов данных проводится, как и операция пересечения, на основе сопоставления идентификаторов строк. Выбираются все строки из первого набора данных, а затем к ним добавляются строки из второго набора данных. Из получившегося промежуточного результата в итоговый набор данных попадают строки с несовпадающими идентификаторами. Значения физических величин в строках с совпадающими идентификаторами могут быть включены в результирующий набор данных в зависимости от выбора пользователя либо все из первого набора данных, либо все из второго набора данных. Допускается объединение наборов данных независимо от заполненности строк, т. е. с возможностью образования пропусков в данных (отсутствие значений физических величин).

Пусть имеется два набора данных:

$$ND1^j = \{(i, S_i^j)\}, \quad ND2^j = \{(p, S_p^j)\}.$$

Тогда их объединение можно записать так:

$$ND1^j \cup ND2^j = ND3^j = \{(r, S_r^j)\},$$

где S_r^j содержит строки, удовлетворяющие следующим условиям:

$$\begin{aligned} S_r^j &= \left(\left[s_{i,1}^j | s_{p,1}^j \right], \dots, \left[s_{i,N}^j | s_{p,N}^j \right] \right) \text{ при } r \in \{i\} \text{ и } r \in \{p\}, \\ S_r^j &= (s_{i,1}^j, \dots, s_{i,N}^j) \text{ при } r \in \{i\} \text{ и } r \notin \{p\}, \\ S_r^j &= (s_{p,1}^j, \dots, s_{p,N}^j) \text{ при } r \notin \{i\} \text{ и } r \in \{p\}. \end{aligned}$$

Разность

Операция «разность» наборов данных $ND1^j$ и $ND2^j$ аналогична разности (дополнению) в теории множеств. Так же как и операция пересечения, она выполняется на основе сравнения идентификаторов. Разность множества $ND1^j$ и множества $ND2^j$ – это множество, содержащее в себе элементы множества $ND1^j$, но не входящие в $ND2^j$. Обозначается: $ND1^j$ без $ND2^j$, $ND1^j \setminus ND2^j$ или $ND1^j - ND2^j$. Мы будем использовать последний вариант обозначения для записи операции разности наборов данных.

В результате может получиться пустой набор данных, часть $ND1^j$ или полностью $ND1^j$.

$$\begin{aligned} ND1^j &= \{(i, S_i^j)\}, \quad ND2^j = \{(p, S_p^j)\}, \\ ND1^j - ND2^j &= ND3^j = \{(r, S_r^j)\}, \end{aligned}$$

где $r \in \{i\}$ и $r \notin \{p\}$.

Обзор интерфейса пользователя модуля «бинарные операции»

Обзор интерфейсов в информационной системе «Молекулярная спектроскопия» для проведения автоматических операций (не требующих участия пользователя) был представлен

в работах [10; 13]. Здесь рассмотрим интерфейс той части ИС, которая относится к выполнению неавтоматизированного этапа операций «объединение» и «пересечение».

После выбора двух источников данных, над которыми будут производиться манипуляции, пользователь делает выбор операции и параметров ее выполнения (рис. 1). В качестве дополнительной информации при принятии решения эксперт может воспользоваться имеющейся в системе характеристикой разупорядочения между выбранными наборами данных.

Выбор бинарной операции и способа обработки данных

Источники информации выбранные для проведения бинарной операции

1	1981_TaDaGo_PH3	G. Tarrago, M. Dang-Ithu, and A. Goldman, Analysis of Phosphine Absorption in the Region 9-10 μm and High Resolution Line-by-Line Simulation of the ν_2 and ν_4 Bands, Journal of Molecular Spectroscopy, 1981, Volume 88, Issue 2, Pages 311-322, DOI: 10.1016/0022-2852(81)90182-X. Annotation
2	2011_JaCrArBo_PH3	N. Jacquinet-Husson, L. Crepeau, R. Armante, C. Boutammine, A. Chédin, N.A. Scott, C. Crevoisier, V. Capelle, C. Boone, N. Poulet-Crovisier, A. Barbe, A. Campargue, D. Chris Benner, Y. Benilan, B. Bézard, V. Boudon, L.R. Brown, L.H. Coudert, A. Coustenis, V. Dana, V.M. Devi, S. Fally, A. Fayt, J.-M. Flaud, A. Goldman, M. Herman, G.J. Harris, D. Jacquemart, A. Jolly, I. Kleiner, A. Kleinböhl, F. Kwabia-Tchana, N. Lavrentieva, N. Lacome, Li-Hong Xu, O.M. Lyulin, J.-Y. Mandin, A. Maki, S. Mikhailenko, C.E. Miller, T. Mishina, N. Moazzen-Ahmadi, H.S.P. Müller, A. Nikitin, J. Orphal, V. Perevalov, A. Perrin, D.T. Petkie, A. Predoi-Cross, C.P. Rinsland, J.J. Remdios, M. Rotger, M.A.H. Smith, K. Sung, J. Tennyson, R.A. Toth, A.-C. Vandaele, J. Vander Auwera, The 2009 edition of the GEISA spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 2011, Volume 112, Issue 15, Pages 2395-2445, DOI: 10.1016/j.jqsrt.2011.06.004. Annotation

Число строк с совпадающей идентификацией: 927

Показать характеристики разупорядочения

Выберите бинарную операцию и ее параметры

Разность Пересечение Объединение

Пересекающуюся по квантовым числам часть обработать следующим образом:

Взять все строки из 1-го набора данных
 Взять все строки из 2-го набора данных
 Выбрать по колебательным плосам
 Выбрать по строкам

Выполнить бинарную операцию

Рис. 1. Интерфейс выбора операции и способа обработки пересекающейся части данных

В операции *объединения* выбор эксперта сводится к указанию, что сделать со всеми строками с совпадающими наборами значений квантовых чисел. Их можно все взять из одного или другого набора данных либо не брать их в результирующий набор данных вообще. Поэтому промежуточный результат операции не отображается, а сразу выдается окончательный результат.

Выберите полосы

	ν_1^1	ν_2^1	ν_3^1	$ \nu_3^1 $	ν_4^1	$ \nu_4^1 $	l^1	k^1	Γ^1	ν_1^f	ν_2^f	ν_3^f	$ \nu_3^f $	ν_4^f	$ \nu_4^f $	l^f	j^f	k^f	Γ^f	Br	Brk	Вакуумные волновые числа (ω)
<input checked="" type="radio"/> 1981_TaDaGo_PH3	0	0	0	0	0	0	2	2	E	0	0	0	0	1	1	-1	1	1	E	P	P	1098.585
<input type="radio"/> 2011_JaCrArBo_PH3	0	0	0	0	0	0	2	2	E	0	0	0	0	1	1	-1	2	1	E	Q	P	1098.58682
	0	0	0	0	0	0	2	2	E	0	0	0	0	1	1	-1	3	1	E	R	P	1116.852
	0	0	0	0	0	0	2	2	E	0	0	0	0	1	1	-1	3	1	E	R	P	1116.80232
	0	0	0	0	0	0	2	2	E	0	0	0	0	1	1	-1	3	1	E	R	P	1144.053
	0	0	0	0	0	0	2	2	E	0	0	0	0	1	1	-1	3	1	E	R	P	1144.05915
	0	0	0	0	0	0	2	2	E	0	0	0	0	1	1	-1	3	1	E	R	P	1099.106
	0	0	0	0	0	0	17	7	E	0	0	0	0	1	1	-1	17	6	E	Q	P	1117.072
	0	0	0	0	0	0	18	2	E	0	0	0	0	1	1	-1	18	1	E	Q	P	1116.77451
	0	0	0	0	0	0	18	2	E	0	0	0	0	1	1	-1	18	1	E	Q	P	1124.95
	0	0	0	0	0	0	18	2	E	0	0	0	0	1	1	-1	18	1	E	Q	P	1124.4985
<input type="radio"/> 1981_TaDaGo_PH3	0	0	0	0	0	0	0	0	A1	0	0	0	0	1	1	1	1	1	A2	R	R	1130.522
<input type="radio"/> 2011_JaCrArBo_PH3	0	0	0	0	0	0	0	0	A1	0	0	0	0	1	1	1	1	1	A2	R	R	1130.52313
	0	0	0	0	0	0	1	0	A2	0	0	0	0	1	1	1	1	1	A1	Q	R	1121.274
	0	0	0	0	0	0	1	0	A2	0	0	0	0	1	1	1	2	1	A1	R	R	1121.26863
	0	0	0	0	0	0	1	0	A2	0	0	0	0	1	1	1	2	1	A1	R	R	1140.174
	0	0	0	0	0	0	1	0	A2	0	0	0	0	1	1	1	2	1	A1	R	R	1140.17514

Рис. 2. Интерфейс обработки промежуточного результата по полосам

В операции *пересечения* эксперт имеет возможность детальной обработки промежуточно-го результата. Он может указать, из какого источника брать данные по конкретной спектральной линии или спектральной полосе в целом (рис. 2). Название источника данных и значения конкретных параметров спектральной линии, взятых из него, подсвечиваются одним цветом. Полученный результат сохраняется как новый набор данных (рис. 3).

Результат операции: Пересечение
Пересекающуюся по квантовым числам часть обработать следующим образом: Выбрать по колебательным плоскам

Показать		10	строк от		0	Всего строк 927										> >>		Настройки				
v1 ⁱ	v2 ⁱ	v3 ⁱ	l3 ⁱ	v4 ⁱ	l4 ⁱ	l ⁱ	j ⁱ	k ⁱ	l ⁱ	v1 ^f	v2 ^f	v3 ^f	l3 ^f	v4 ^f	l4 ^f	l ^f	j ^f	k ^f	l ^f	Br	BrK	Вакуумные волновые числа (ω)
0	0	0	0	0	0	0	13	6	A1	0	0	0	0	1	1	-1	13	5	A2	Q	P	1115.565
0	0	0	0	0	0	0	8	3	A1	0	0	0	0	1	1	-1	7	2	A2	P	P	1046.723
0	0	0	0	0	0	0	9	3	A1	0	0	0	0	1	1	-1	8	2	A2	P	P	1038.565
0	0	0	0	0	0	0	9	3	A1	0	0	0	0	1	1	-1	9	2	A2	Q	P	1119.283
0	0	0	0	0	0	0	8	3	A1	0	0	0	0	1	1	-1	8	2	A2	Q	P	1118.374
0	0	0	0	0	0	0	11	3	A1	0	0	0	0	1	1	-1	11	2	A2	Q	P	1120.544
0	0	0	0	0	0	0	10	3	A1	0	0	0	0	1	1	-1	10	2	A2	Q	P	1119.664
0	0	0	0	0	0	0	12	3	A1	0	0	0	0	1	1	-1	12	2	A2	Q	P	1120.745
0	0	0	0	0	0	0	14	3	A1	0	0	0	0	1	1	-1	14	2	A2	Q	P	1121.843
0	0	0	0	0	0	0	13	3	A1	0	0	0	0	1	1	-1	13	2	A2	Q	P	1121.817
Показать		10	строк от		0	Всего строк 927										> >>		Настройки				

Для сохранения результатов операции задайте имя нового набора данных
"2017-Mar-27_12h58m10s"

Рис. 3. Интерфейс сохранения результата операции в новый набор данных

Такая реализация позволяет сохранять результат детальной обработки в отдельном наборе данных и использовать его потом в других операциях, тем самым уменьшается количество потенциальных ошибок, когда пользователь в при первой обработке промежуточного результата выбрал одно, а при повторной выбрал другое. Детальная обработка пересекающейся части (по строкам и по спектральным полосам) имеется только в операции пересечения.

Заключение

В рамках информационной системы «Молекулярная спектроскопия» создан модуль позволяющий формировать составные наборы данных, путем их комбинирования из различных источников. Это позволяет создавать составные экспертные наборы данных, учитывая как формализуемые ограничения предметной области, так и принимаемые экспертами решения, не поддающиеся алгоритмизации.

Одно из направлений дальнейших исследований – определение критериев для автоматического формирования рекомендаций для эксперта по выбору окончательного результата полуавтоматической части операций.

Список литературы

1. Сердюков В. И., Синуца Л. Н., Быков А. Д., Щербаков А. П. Уширение и сдвиг спектральных линий метана в области 11000–11400 см⁻¹ // Оптика атмосферы и океана. 2017. Т. 30, № 12. С. 1023–1026.
2. Лукашевская А. А., Люлин О. М., Perrin A., Перевалов В. И. Глобальное моделирование центров спектральных линий молекулы NO₂ // Оптика атмосферы и океана. 2015. Т. 28, № 01. С. 12–27.
3. Yurchenko S. N., Barber R. J., Tennyson J. A variationally computed line list for hot NH₃ // Monthly Notices of the Royal Astronomical Society. 2011. Vol. 413. Iss. 3. P. 1828–1834. <https://doi.org/10.1111/j.1365-2966.2011.18261.x>

4. *Yurchenko S. N., Tennyson J.* ExoMol line lists IV: The rotation-vibration spectrum of methane up to 1500K. arXiv:1401.4852v1 [astro-ph.EP] 20 Jan 2014
5. *Козодоев А. В.* Система загрузки данных в распределенной информационной системе «Молекулярная спектроскопия» // Материаловедение, технологии и экология в 3-м тысячелетии: Материалы IV Всерос. конф. молодых ученых. Томск: Изд-во Ин-та оптики атмосферы СО РАН, 2009. С. 587–591.
6. *Ахлестин А. Ю., Воронина С. С., Лаврентьев Н. А., Фазлиев А. З.* Информационные ресурсы по спектроскопии в ИОА СО РАН // Оптика атмосферы и океана. 2015. Т. 28, № 05. С. 480–488.
7. *Ахлестин А. Ю., Воронина С. С., Науменко О. В., Половцева Е. Р., Фазлиев А. З.* Информационная система для решения задач молекулярной спектроскопии. 6. Систематизация спектроскопических данных по дейтерозамещенным изотопологам молекулы сероводорода // Оптика атмосферы и океана. 2016. Т. 29, № 05. С. 386–396.
8. *Ахлестин А. Ю., Воронина С. С., Привезенцев А. И., Родимова О. Б., Фазлиев А. З.* Информационная система для решения задач молекулярной спектроскопии. 7. Систематизация информационных ресурсов по поглощению для основного изотополога молекулы метанола // Оптика атмосферы и океана. 2016. Т. 29, № 10. С. 876–887.
9. *Лаврентьев Н. А., Макогон М. М., Фазлиев А. З.* Сравнение спектральных массивов данных HITRAN и GEISA с учетом ограничения на опубликование спектральных данных // Оптика атмосферы и океана. 2011. Т. 24, № 4. С. 279–292.
10. *Козодоев А. В., Козодоева Е. М.* Универсальный модуль «унарные операции» в ИС «Молекулярная спектроскопия» // Вестн. НГУ. Серия: Информационные технологии. 2015. Т. 13, № 1. С. 46–54.
11. *Быков А. Д., Науменко О. В., Родимова О. Б., Сеница Л. Н., Творогов С. Д., Тонков М. В., Фазлиев А. З., Филиппов Н. Н.* Информационные аспекты молекулярной спектроскопии. Томск: Изд-во Ин-та оптики атмосферы СО РАН, 2008. 360 с.
12. *Кулик Б. А., Зуенко А. А., Фридман А. Я.* Алгебраический подход к интеллектуальной обработке данных и знаний. СПб.: Изд-во Политехн. ун-та, 2010. 235 с.
13. *Козодоев А. В., Козодоева Е. М.* Формирование наборов данных в ИС «Молекулярная спектроскопия» с использованием бинарных операций // Оптика атмосферы и океана. Физика атмосферы: Материалы XXII Междунар. симп. Томск: Изд-во ИОА СО РАН, 2016. А209-2012.

Материал поступил в редколлегию 03.04.2018

A. V. Kozodoev, E. M. Kozodoeva

*V. E. Zuev Institute of Atmospheric Optics SB RAS
1 Academician Zuev Square, Tomsk, 634055, Russian Federation*

kav@iao.ru, klen@iao.ru

THE BINARY OPERATIONS IN THE INFORMATION SYSTEM «MOLECULAR SPECTROSCOPY»

The paper describes an approach to the development and implementation of functions for performing binary operations on data in the IS «Molecular Spectroscopy». The formalization of binary operations over sets of spectroscopic data is made, taking into account the features of the subject area. Describes an action algorithm and a user interface for performing binary operations, one for databases for various substances and several types of spectroscopic data.

Keywords: data structure, quantitative spectroscopy, binary operations, database.

References

1. Serdyukov V. I., Sinitsa L. N., Bykov A. D., Shcherbakov A. P. Broadening and shift of the methane absorption lines in the 11000–11400 cm^{-1} region. *Optika Atmosfery i Okeana*, 2017, vol. 30, no. 12, p. 1023–1026. (in Russ.)
2. Lukashetskaya A. A., Lyulin O. M., Perrin A., Perevalov V. I. Global modeling of NO_2 central line positions. *Optika Atmosfery i Okeana*, 2015, vol. 28, no. 01, p. 12–27. (in Russ.)
3. Yurchenko S. N., Barber R. J., Tennyson J. A variationally computed line list for hot NH_3 . *Monthly Notices of the Royal Astronomical Society*, 2011, vol. 413, iss. 3, 21 p. 1828–1834. <https://doi.org/10.1111/j.1365-2966.2011.18261.x>
4. Yurchenko S. N., Tennyson J. ExoMol line lists IV: The rotation-vibration spectrum of methane up to 1500K. arXiv:1401.4852v1 [astro-ph.EP] 20 Jan 2014
5. Kozodoev A. V. Data loading system in the distributed information system «Molecular spectroscopy». *Materials of the IV all-Russian conference of young scientists materials. Science, technology and ecology in the 3rd Millennium*. Tomsk, Publishing House of the Institute of Atmospheric Optics SB RAS, 2009, p. 587–591. (in Russ.)
6. Akhlyostin A. Yu., Voronina S. S., Lavrent'ev N. A., Fazliev A. Z. Spectroscopic information resources at Institute of Atmospheric Optics SB RAS. *Optika Atmosfery i Okeana*, 2015, vol. 28, no. 05, p. 480–488. (in Russ.)
7. Akhlyostin A. Yu., Voronina S. S., Naumenko O. V., Polovtseva E. R., Fazliev A. Z. Information system for molecular spectroscopy. 6. Systematization of spectral data on deuterio-substituted isotopologues of hydrogen sulfide molecule. *Optika Atmosfery i Okeana*, 2016, vol. 29, no. 05, p. 386–396. (in Russ.)
8. Akhlyostin A. Yu., Voronina S. S., Privezentsev A. I., Rodimova O. B., Fazliev A. Z. Information system for molecular spectroscopy. 7. Systematization of information resources on the main isotopologue of methanol molecule. *Optika Atmosfery i Okeana*, 2016, vol. 29, no. 10, p. 876–887. (in Russ.)
9. Lavrent'ev N. A., Makogon M. M., Fazliev A. Z. Comparison of HITRAN and GEISA spectral data, based on taking into account the existent constraints. *Optika Atmosfery i Okeana*, 2011, vol. 24, no. 04, p. 279–292. (in Russ.)
10. Kozodoyev A. V., Kozodoyeva E. M. Extensible module «Unary operations» in the Information system «Molecular spectroscopy». *Vestnik NSU. Series: Information Technologies*, 2015, vol. 13, no. 1, p. 46–54. ISSN 1818-7900. (in Russ.)
11. Bykov A. D., Naumenko O. V., Rodimova O. B., Sinitsa L. N., Tvorogov S. D., Tonkov M. V., Fazliev A. Z., Filippov N. N. Information aspects of molecular spectroscopy. Tomsk, Publishing House of the Institute of Atmospheric Optics SB RAS, 2008, 360 p. (in Russ.)
12. Kulik B. A., Zuenko A. A., Fridman A. Ya. Algebraic approach to intelligent processing of data and knowledge. St. Petersburg Polytechnic University Publishing House, 2010, 235 p. (in Russ.)
13. Kozodoev A. V., Kozodoeva E. M. Building of data sets in is «Molecular spectroscopy» using binary operations. *Atmospheric and ocean Optics. Atmospheric physics: Proceedings of the XXII International Symposium*. Tomsk, Publishing House of the Institute of Atmospheric Optics SB RAS, 2016, A209-2012. (in Russ.)

For citation:

Kozodoev A. V., Kozodoeva E. M. The Binary Operations in the Information System «Molecular Spectroscopy». *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 2, p. 70–77. (in Russ.)

DOI 10.25205/1818-7900-2018-16-2-70-77