

**МЕТОДЫ И СИСТЕМЫ ОБРАБОТКИ
ИЗМЕРИТЕЛЬНОЙ ИНФОРМАЦИИ НА ЭВМ**

УДК 547.02 : 518.5

М. И. ГОРФИНКЕЛЬ, С. А. НЕХОРОШЕВ
(Новосибирск)

**АЛГОРИТМ И ПРОГРАММА
(НА ВХОДНОМ ЯЗЫКЕ ЭВМ «НАИРИ»)
ДЛЯ ПОДБОРА ВОЗМОЖНЫХ ЭМПИРИЧЕСКИХ ФОРМУЛ
ПО ЗАДАННОМУ ТОЧНОМУ ЗНАЧЕНИЮ
МОЛЕКУЛЯРНОГО ВЕСА СОЕДИНЕНИЯ**

Как известно, точные значения атомных весов изотопов различных элементов несколько отличаются от целочисленных значений, и это обстоятельство послужило основой для разработки метода определения элементного состава молекулярных и осколочных ионов в масс-спектрах по точным значениям соответствующих массовых чисел [1].

Опубликовано несколько типов табличных систем для подбора эмпирических формул по точному значению массового числа [2, 3], причем иногда таблицы вводят с этой целью в память ЭВМ (см., например, [3]). Однако табличная система подбора брутто-формул сама по себе (в любом варианте — ручном или машинном) обладает рядом существенных недостатков: 1) круг рассматриваемых элементов (обычно С, Н, N и О) является жестко заданным и не может быть изменен без замены всей таблицы; 2) доступный интервал массовых чисел ограничен сверху; 3) при использовании ЭВМ (даже при условии, что соединение может содержать только элементы С, Н, N и О) табличная система требует очень большого объема памяти (порядка нескольких десятков тысяч слов).

Недавно Бёрлинггейм предложил алгоритм для генерирования элементных формул по заданному точному значению массового числа ионов, содержащих элементы С, Н, N и О [4]. Однако качество работы этого алгоритма существенно зависит от разумности выбранных при каждом расчете максимально возможных количеств атомов С, Н, N и О, а распространение его на большее число элементов вызывает резкое усложнение расчетов за счет увеличения числа подлежащих рассмотрению различных сочетаний; реализация этого алгоритма на ЭВМ предполагает наличие весьма значительной по объему оперативной памяти. В [5] упоминается «итеративная методика», но сущность ее не описана. Сообщалось также о существовании программы (для ЭВМ типа PDP-8), предназначенной для приписывания элементных составов ионам, содержащим С, Н, N, О, S^{13} и еще 3 других (произвольных) элемента, но без описания самой программы или алгоритма, положенного в ее основу [6].

Нами разработан простой алгоритм для подбора возможных эмпи-

рических формул по точному значению массового числа, который затем был реализован в программе, написанной на входном языке малой цифровой универсальной вычислительной машины «Наири». Основным вариантом программы предназначен для подбора эмпирических формул молекулярных ионов, которые могут содержать элементы С, Н, N, О, F, Cl и Br. Путем внесения небольших изменений в основную программу круг рассматриваемых элементов легко может быть расширен или сужен, а сама программа приспособлена к подбору эмпирических формул осколочных ионов и радикалов наряду с эмпирическими формулами молекулярных ионов.

В качестве исходных выбраны следующие данные:

- 1) ближайшее к точному значению целочисленного молекулярного веса (массового числа молекулярных ионов);
- 2) разность точного и целочисленного значений молекулярного веса;
- 3) числа атомов хлора и брома, определяемые по интенсивностям изотопных линий в масс-спектре;
- 4) погрешность определения массового числа (принята равной ± 10 миллионным долям);
- 5) максимально возможное число атомов фтора и минимально возможное число атомов углерода (задаются формулой насыщенного перфторуглеводорода с молекулярным весом, равным целочисленному молекулярному весу соединения после замены атомов хлора и брома на атомы водорода); предельное число атомов азота принято равным 8, кислорода — 10. Кроме того, принято, что число атомов азота не должно превышать числа атомов углерода более чем на единицу, а число атомов кислорода не должно превышать числа атомов углерода более чем на две единицы.

Сущность разработанного нами алгоритма состоит в следующем. Вначале формируется базовая эмпирическая формула, в состав которой входят только атомы углерода и водорода* (при четном целочисленном молекулярном весе) или только атомы углерода, водорода и один атом азота (при нечетном целочисленном молекулярном весе), причем число атомов углерода выбирается максимально возможным. Затем с помощью ряда циклов в базовой формуле осуществляют все возможные замены $2\text{CH}_2 \rightarrow \text{N}_2$, $\text{CH}_4 \rightarrow \text{O}$ и $\text{CH}_7 \rightarrow \text{F}$, вырабатывают новую базовую формулу путем замены $\text{C} \rightarrow 12\text{H}$ в предыдущей базовой формуле, вновь производят указанные выше циклы замен и т. д. Каждая полученная формула проверяется на соответствие требуемому значению точного молекулярного веса и правилам валентности. В тех формулах, которые удовлетворяют обоим условиям, соответствующее число атомов водорода заменяют первоначальным числом атомов хлора и брома и получают искомые эмпирические формулы. Приводимая ниже программа, реализующая этот алгоритм, написана на входном языке ЭВМ «Наири», который весьма близок к обычному математическому языку и поэтому позволяет легко проследить отдельные детали алгоритма непосредственно при рассмотрении программы.

Программа состоит из 71 оператора *an* (автопрограммирующей программы) и написанной в машинном коде программы печати.

Оператор 1 осуществляет ввод исходных данных — целочисленного молекулярного веса *m*, разности между точным и целочисленным значениями молекулярного веса *l*, числа атомов хлора *ц* и числа атомов брома *щ*. Операторы 2 и 36 осуществляют переход к программе печати.

* Атомы хлора и брома предварительно заменяют атомами водорода (один атом галоида одним атомом водорода) и вычисляют новые значения точного и целочисленного молекулярного веса.

Операторы 3—12 осуществляют вычисление некоторых вспомогательных величин — в частности, значения целочисленного молекулярного веса m и пределов (верхнего p и нижнего c) разности между точным и целочисленным значениями молекулярного веса после замены атомов хлора и брома атомами водорода, максимально возможного числа атомов фтора l и минимально возможного числа атомов углерода u .

Операторы 13—27 и 64—69 формируют первую базовую формулу (a — число атомов углерода, $ч$ — число атомов водорода, $в$ — число атомов азота, $г$ — число атомов кислорода, $д$ — число атомов фтора).

Операторы 29—41 осуществляют замены $2\text{CH}_2 \rightarrow \text{N}_2$ («азотный» цикл), проверку полученной формулы на соответствие заданному точному значению молекулярного веса и правилам валентности, а также обратную замену атомов водорода атомами хлора и брома перед печатью; кроме того, здесь же находится оператор (34), обеспечивающий вычисление интенсивности линии $M+1$, которая наблюдалась бы в масс-спектре для ионов с соответствующим составом. С целью проверки формулы на соответствие правилам валентности вычисляется величина формальной ненасыщенности $ш$, которая должна быть больше или равна нулю [7]. Переход ко второму («кислородному») циклу осуществляется по одному из следующих условий: 1) если в последней из полученных в цикле формул $в \geq 7$; 2) если $а \leq в$; 3) если $ч < ж = ч + ш + 4$; 4) если в последней из полученных в цикле формул точный молекулярный вес оказался ниже нижнего предела.

«Кислородный» цикл (операторы 42—53) осуществляет замены $\text{СН}_4 \rightarrow \text{O}$. После каждой такой замены оператор 53 обеспечивает переход к «азотному» циклу, который является, таким образом, внутренним циклом по отношению к «кислородному». Переход к третьему («фторному») циклу осуществляется по одному из следующих условий: 1) если $г = 10$; 2) если $а \leq г - 2$; 3) если в базовой формуле число атомов водорода $ч$ меньше, чем $ж + 4г + 7д$; 4) если для последней из полученных в цикле формул точный молекулярный вес оказался ниже нижнего предела.

Третий («фторный») цикл (операторы 54—63) осуществляет замены $\text{СН}_7 \rightarrow \text{F}$. После каждой такой замены оператор 63 обеспечивает переход к «азотному» циклу; «фторный» цикл является, таким образом, внешним по отношению к «кислородному» и «азотному». Переход к последнему («водородному») циклу осуществляется по одному из следующих условий: 1) если в базовой формуле число атомов водорода $ч$ меньше, чем $ч + ш + 7д + 7$; 2) если для последней из полученных в цикле формул точный молекулярный вес оказался ниже нижнего предела.

Четвертый («водородный») цикл (операторы 64—69) осуществляет замены $\text{C} \rightarrow 12\text{H}$, причем последний оператор этого цикла осуществляет переход к формированию новой базовой формулы. Цикл является, таким образом, самым внешним (по отношению к «азотному», «кислородному» и «фторному»); конец работы программы определяют условия $а < и$ или $д > l$.

Вся программа (включая программу печати) занимает 700 ячеек оперативного запоминающего устройства (ОЗУ). Расчет (среднее быстрое действие ЭВМ «Наири» при работе с числами, представленными в виде с плавающей запятой, составляет 100—300 операций в секунду) и печать результатов для молекулярного веса порядка 400 требует около 10 мин. Сокращенный вариант программы (для элементов С, Н, N и O) занимает 470 ячеек ОЗУ и требует для расчета и печати результатов (молекулярный вес порядка 400) около 1,5 мин.

Легко видеть, что как алгоритм, так и программа могут быть легко

распространены на другое число и другие сочетания элементов (путем изменения числа и содержания внутренних циклов), а также приспособлены к вычислению элементных составов осколочных ионов и радикалов наряду с эмпирическими формулами для молекулярных ионов (в этом случае базовая формула должна содержать только атомы углерода и водорода, а «азотный» цикл — осуществлять замены $\text{CH}_2 \rightarrow \text{N}$).

Ниже в качестве примера приведены результаты, полученные для соединения с точным молекулярным весом 356,150 (в углеродной шкале, $\text{C}^{12} = 12,000000$), не содержащего атомов хлора и брома.

ОПЕРАТОРЫ АВТОПРОГРАММИРУЮЩЕЙ ПРОГРАММЫ

1 введем м л ц ш	38 если $v > 6,9$ иди к 42
2 программа 250	39 если $c < ж$ иди к 42
3 допустим $a=0 \ v=0 \ g=0 \ d=0 \ й=1$	40 вставим $n=n-0,025154 \ a=a-2 \ c=c-4$ $v=v+2$
4 вычислим $ь=n+ш$	41 иди к 30
5 вычислим $з=ь+2$	42 если $г > 9,9$ иди к 54
6 вычислим $ж=ь+4$	43 вставим $г=г+1$
7 вычислим $m=m-34n-78ш$	44 вычислим $a=y-g-d$
8 вычислим $s=0,038973n+0,089435ш$	45 вычислим $л=г-2$
9 вычислим $p=л+s+m/100000$	46 если $a \leq л$ иди к 54
10 вычислим $c=л+s-m/100000$	47 вычислим $e=ь+4г+7d$
11 вычислим $l=0,04m+0,48$	48 если $z < e$ иди к 54
12 вычислим $u=0,02m-0,76$	49 вычислим $ч=z-e+ь$
13 вставим $a=a+1$	50 вычислим $n=t-0,036386г-0,056374d$
14 вычислим $ч=m-12a$	51 если $n < 0$ иди к 54
15 если $ч > 12,1$ иди к 13	52 допустим $v=ф$
16 вычислим $x=2й-ч$	53 иди к 30
17 вставим $й=й+1$	54 если $d > 1$ иди к 70
18 если $x < -0,1$ иди к 16	55 вставим $d=d+1$
19 если $ч \geq ь$ иди к 22	56 вычислим $й=ь+7d$
20 допустим $y=a \ z=ч \ ф=0$	57 если $z < й$ иди к 64
21 иди к 64	58 допустим $v=ф \ g=0$
22 если $x < 0,9$ иди к 24	59 вычислим $a=y-g$
23 если $ч < з$ иди к 20	60 вычислим $ч=z-7d$
24 вычислим $n=0,007825ч-0,012577x$	61 вычислим $n=t-0,056374d$
25 вычислим $a=a-x$	62 если $n < 0$ иди к 64
26 вычислим $ч=ч-2x$	63 иди к 30
27 вычислим $v=v+x$	64 вставим $y=y-1 \ z=z+12 \ t=t+0,093900$
28 допустим $y=a \ z=ч \ t=n \ ф=v$	65 допустим $v=ф \ g=0 \ d=0 \ a=y \ ч=z \ n=t$
29 если $n < 0$ иди к 42	66 если $a < и$ иди к 70
30 если $n > p$ иди к 37	67 если $ч < ь$ иди к 64
31 вычислим $ш=a+0,5v-0,5(ч+d)+1$	68 если $ф=1$ иди к 28
32 если $ш < -0,1$ иди к 37	69 иди к 22
33 вычислим $б=ч-ь$	70 интервал 3
34 вычислим $m=1,12a+0,016ч+0,38v$	71 иди к 1
35 вычислим $ы=m+n-s$	исполним 1
36 программа 150	
37 если $a \leq v$ иди к 42	

ПРОГРАММА ПЕЧАТИ

250ко2274н	о2226н	о2258н
о2274н	о2229н	о2177н
о2270н	о2270н	о2274н
о2214н	о2270н	о2274н
о2223н	о2220н	н66н ш150кпп42н1
о2216н	о2216н	о2270н
о2270н	о2226н	о2270н
о2270н	о2229н	пп43н1
о2270н	о2270н	о2270н
о2270н	о2270н	о2270н
о2213н	о2222н	пп34н1
о2226н	о2229н	о2270н
о2215н	о2226н	о2270н
о2270н	о2217н	пп44н1

о2270н	о2270н	о2270н
о2270н	о2270н	о2270н
о2221н	о2270н	пл36н1
о2236н	о2270н	о2270н
о2226н	о2270н	о2270н
о2230н	о2270н	пл61н1
о2270н	о2217н	о2270н
о2270н	о2213н	о2270н
о2248н	о2270н	пл62н1
о2225н	о2270н	о2270н
о2212н	о2270н	о2270н
о2216н	о2270н	пл55н4
о2270н	о2270н	о2270н
о2270н	о2270н	о2270н
о2239н	о2270н	пл51н2
о2230н	о2217н	о2274н
		н66н ш

**ПРИМЕР РАСЧЕТА ДЛЯ СОЕДИНЕНИЯ
С ЭКСПЕРИМЕНТАЛЬНО ОПРЕДЕЛЕННЫМ МОЛЕКУЛЯРНЫМ ВЕСОМ 356,150,
НЕ СОДЕРЖАЩЕГО АТОМОВ ХЛОРА И БРОМА**

356								
0,150								
0								
0								
угл	вод	азот	кисл	фтор	хлор	бром	мв	м+1
19,0	16,0	8,0	0,0	0,0	0,0	0,0	356,1497	24,57
23,0	20,0	2,0	2,0	0,0	0,0	0,0	356,1524	26,83
18,0	20,0	4,0	4,0	0,0	0,0	0,0	356,1484	21,99
16,0	17,0	8,0	1,0	1,0	0,0	0,0	356,1509	21,23
21,0	19,0	2,0	0,0	3,0	0,0	0,0	356,1500	24,58
17,0	24,0	0,0	8,0	0,0	0,0	0,0	356,1471	19,42
15,0	21,0	4,0	5,0	1,0	0,0	0,0	356,1495	18,65
13,0	18,0	8,0	2,0	2,0	0,0	0,0	356,1520	17,88
18,0	20,0	2,0	1,0	4,0	0,0	0,0	356,1511	21,23
14,0	25,0	0,0	9,0	1,0	0,0	0,0	356,1482	16,07
12,0	22,0	4,0	6,0	2,0	0,0	0,0	356,1507	15,31
10,0	19,0	8,0	3,0	3,0	0,0	0,0	356,1532	14,54
13,0	20,0	4,0	3,0	4,0	0,0	0,0	356,1471	16,39
11,0	17,0	8,0	0,0	5,0	0,0	0,0	356,1496	15,63
15,0	21,0	2,0	2,0	5,0	0,0	0,0	356,1523	17,89
10,0	21,0	4,0	4,0	5,0	0,0	0,0	356,1482	13,05
8,0	18,0	8,0	1,0	6,0	0,0	0,0	356,1507	12,28
12,0	22,0	2,0	3,0	6,0	0,0	0,0	356,1534	14,55
13,0	20,0	2,0	0,0	8,0	0,0	0,0	356,1498	15,63

ЛИТЕРАТУРА

1. Дж. Бейнон. Масс-спектрометрия органических соединений при высокой разрешающей силе.— В кн. «Успехи масс-спектрометрии». М., «Мир», 1964, стр. 326.
2. J. H. Beynon, A. E. Williams.— Mass and Abundance Tables for Use in Mass Spectrometry. Amsterdam, Elsevier, 1963. D. Henneberg, K. Casper. Ein einfaches Tabellensystem zur Zuordnung von Bruttoformeln zu Massenwerten aus hochaufgelösten Massenspektren organischer Verbindungen.— Z. Anal. Chem., 1967, v. 227, № 4, p. 241. W. Benz.— Massenspektrometrie Organischer Verbindungen. Leipzig, Akademische Verlagsgesellschaft, 1969, p. 356. G. P. Moss. A New Method for the Calculation of Molecular Formulae from Accurate Mass Data and vice versa.— OMS, 1971, v. 5, № 3, p. 353.
3. L. Lederberg.— Computation of Molecular Formulas for Mass Spectrometry. San Francisco, Holden Day, 1964.

4. A. L. Burlingame. Data Acquisition, Processing and Interpretation via Coupled High — Speed Real — Time Digital Computer and High Resolution Mass Spectrometer Systems.— In "Advances in Mass Spectrometry". London, Petroleum Institute, 1968, v. 4, p. 15.
5. R. Venkataraghavan, F. W. Mc Lafferty, J. M. Amy. Automatic Reduction of High Resolution Mass Spectral Data. Computer Techniques for Improved Mass-Measuring Accuracy and Resolution.— *Analyt. Chem.*, 1967, v. 39, № 2, p. 178.
6. Mass Spectrometer Data Acquisition and Analysis System. Publication T. P. 22. Manchester, GEC — AEL (Electronics) Ltd.
7. М. И. Горфинкель.— Метод масс-спектрометрии в структурных исследованиях. Новосибирск, НГУ, 1969, стр. 47.

*Поступила в редакцию
19 июля 1971 г.*
