

МЕТОДЫ И ТЕХНИЧЕСКИЕ СРЕДСТВА АВТОМАТИЗАЦИИ ЭКСПЕРИМЕНТА

УДК 543.51 : 681.32

Б. Г. ДЕРЕНДЯЕВ, В. А. КОПТЮГ, К. С. ЛЕБЕДЕВ,
О. Н. ШАРАПОВА

(Новосибирск)

МАШИННАЯ ИНФОРМАЦИОННО-ПОИСКОВАЯ СИСТЕМА НА БАЗЕ КАТАЛОГА ПОЛНЫХ МАСС-СПЕКТРОВ

Масс-спектрометрия наряду с другими физическими методами занимает важное место при структурном анализе органических соединений. Однако повседневное использование этого метода для установления строения исследуемых веществ сопряжено с определенными трудностями, так как спектроструктурные корреляции в масс-спектрометрии выражены далеко не в таком явном виде, как, например, в спектроскопии ядерного магнитного резонанса. К настоящему времени накоплены большие массивы масс-спектрометрических данных (имеются, например, каталоги, содержащие до 40 000 спектров [1]), эффективное использование которых с целью идентификации исследуемого соединения по его масс-спектру или установления структурных особенностей, по-видимому, становится невозможным без широкого использования ЭВМ.

Изложенное выше привело к созданию (в НИОХ СО АН СССР) информационно-поисковой системы (ИПС) [2] на базе каталога 8-пиковых масс-спектров, опытная эксплуатация которой дала достаточно хорошие результаты. Однако восемь наиболее интенсивных пиков далеко не всегда характеризуют спектральные и соответственно структурные признаки исследуемого соединения, что особенно сильно сказывается при использовании ИПС для решения задач распознавания структурных блоков сложных органических соединений. Обращение к массивам, содержащим достаточно полное описание масс-спектров, как показано в работе [3], позволяет получать более широкую информацию о структуре неизвестного соединения, если в используемой машинной библиотеке имеются его структурные аналоги. По этой причине нами разработана новая ИПС на базе каталога полных масс-спектров [4]* различных органических соединений.

Организация поискового массива. Работа ИПС непосредственно с массивами, содержащими полную спектральную информацию, требует слишком больших затрат машинной памяти и времени. Поэтому при организации поискового массива желательно использовать различные методы сокращения вводимых в ЭВМ спектральных данных, которые должны, с одной стороны, сохранять спектральные признаки, наиболее ответственные за структурные особенности соединения, а с другой —

* Авторы выражают благодарность Ф. Г. Унгеру за предоставление каталога на машинно-читаемых носителях (~5000 спектров).

по возможности исключать малоинформативную часть спектральной информации.

Применительно к ИПС для масс-спектрометрии в качестве сокращенной выборки принято использовать одну [5] или две [6—8] наиболее интенсивных линии в каждом интервале по 14 массовых единиц (м/е) на всем протяжении спектра. В то же время статистический анализ, проведенный нами на массиве полных масс-спектров, показал, что для произвольно взятого спектра средняя вероятность появления любых линий, находящихся в интервале массовых чисел 20÷117 (низкая массовая область) и имеющих интенсивность $\geq 1\%$ * от основной линии спектра, составляет $\sim (1/2)^2$, а для интервалов 118÷187 (средняя массовая область) и 188÷705 (высокая массовая область) $\simeq (1/2)^4$ и $(1/2)^8$ соответственно. Отсюда следует, что при сокращении спектра, по-видимому, целесообразно использовать одинаковое количество линий на различных участках масс-спектра. Если появление отдельных линий в массиве спектров при любых значениях м/е рассматривать как независимые случайные события [10—12], то вероятность появления четырех линий в низкой массовой области $(1/2)^2 \cdot (1/2)^2 \cdot (1/2)^2 \cdot (1/2)^2$ будет равна вероятности появления двух любых линий в средней массовой области $(1/2)^4 \cdot (1/2)^4$ и соответственно одной в высокой массовой области $(1/2)^8$, т. е. целесообразное количество кодируемых линий на соответствующих участках масс-спектра можно представить последовательностью 4:2:1. Аналогичная оценка, проведенная нами по данным работы [11] для массива, содержащего значительно большее число спектров ($\sim 18\,800$), дала близкое соотношение — 3:2:1. Учитывая полученные статистические характеристики, мы решили при сокращении спектров использовать в каждом из интервалов длиной в 14 м/е три наиболее интенсивных линии в низкой массовой области, две — в средней и одну — в высшей.

Отбор пиков для сокращенного спектра проводился следующим образом. На первом этапе в каждом из интервалов в 14 м/е полного спектра выбиралось число интенсивных пиков на единицу больше, чем требовалось для соответствующей массовой области сокращенного спектра. Отобранные пики располагались в порядке убывания интенсивностей. Затем по отношению интенсивностей двух последних линий из этого списка решался вопрос о том, какая из этих линий остается в сокращенном спектре. В том случае, когда отношение интенсивностей этих линий было $\leq 1,25$ [5], предпочтение отдавалось линии с большим массовым числом. Такой способ сокращения числа используемых в поисковом массиве спектральных линий уменьшает их количество на один спектр в среднем с 88 (в полном массиве) до 20; при этом, как видно из рис. 1, основная спектральная информация сохраняется.

Рассмотрим некоторые статистические характеристики каталога сокращенных таким образом масс-спектров. На рис. 2 представлена зависимость частоты появления пиков (N) в каталоге сокращенных масс-спектров при различных значениях массовых чисел. Легко видеть, что пики при различных значениях массовых чисел должны по-разному учитываться при оценке подобия спектров. Например, если в спектре неизвестного соединения присутствуют пики с м/е 29, 41, 43, 111, 154 и 271, то при сопоставлении его со спектрами каталога с целью отыскания наиболее похожего спектра следует придать большее значение пикам с м/е 111, 154 и 271, частоты появления которых в массиве спектров значительно ниже, чем для пиков с м/е 29, 41 и 43. В качестве величины, характеризующей значимость совпадения отдельного пика, удобно использовать фактор M (см., например, [11]), равный отрицательному значению двойного логарифма от вероятности появления пика с дан-

* Здесь и далее пики, имеющие интенсивности $< 1\%$, не рассматривались (ср. [9]).

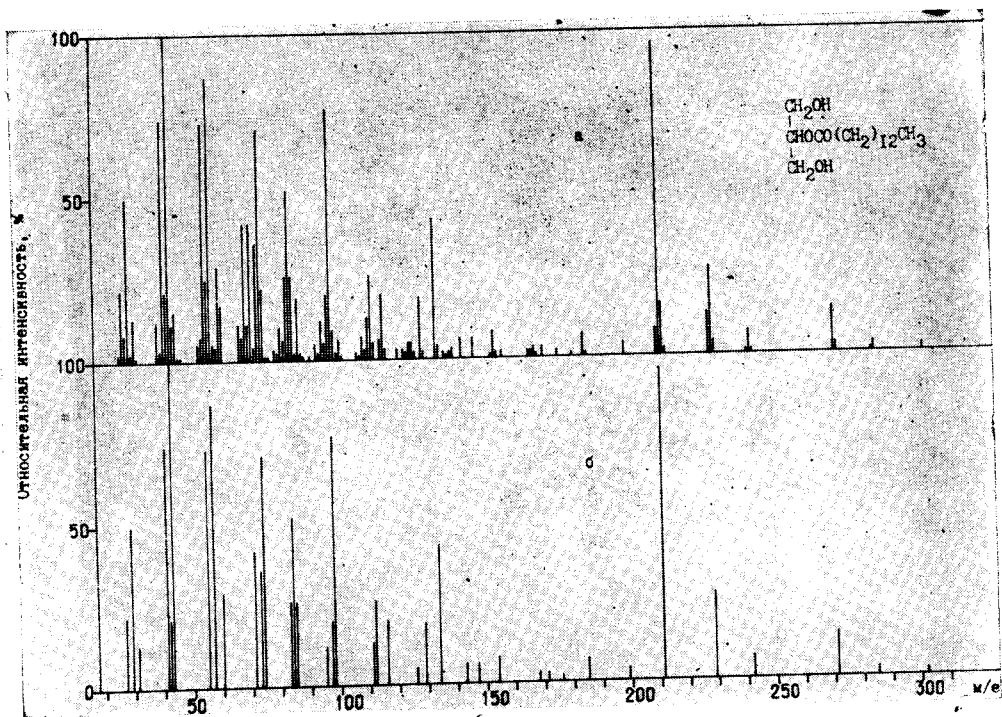


Рис. 1. Полный (а) [4] и сокращенный (б) масс-спектры 2-мономирстилового эфира глицерина.

ным массовым числом в массиве масс-спектров, т. е. $M = -\log_2 (N/4903)$. На правой оси ординат рис. 2 представлена шкала факторов M для массовых чисел в интервале от 20 до 400 м/е. Для практического использования факторы M округлялись до ближайшего целого значения. Например, для пиков с м/е 29, 32, 43, 197 и 303 (на рис. 2 помечены звездочками) факторы M принимались равными 1, 4, 1, 6 и 9 соответственно. Таблица значений M хранится в памяти ЭВМ.

Интенсивности спектральных пиков — вторая важная характеристика описания масс-спектров. Известно, однако, что интенсивности пиков могут существенно зависеть от используемой системы ввода образца в источник и ее температуры, типа прибора, метода развертки и т. д. Поэтому в массиве сокращенных спектров, по-видимому, нет необходимости представлять точные значения интенсивностей пиков. В связи с этим для их описания мы воспользовались средними значениями факторов значимости интенсивностей (I), представляющих собой отрицательные значения двоичных логарифмов вероятностей того, что интенсивности линий попадут в определенные интервалы значений. Границы интервалов выбирались таким образом, чтобы внутри каждого из них указанные факторы имели целочисленные значения. Расчет производился следующим образом. В сокращенном массиве имеется 98 101 пиков с интенсивностью $\geq 1\%$. Половина этих пиков имеет интенсивность $\geq 8,3\%$, т. е. вероятность появления пиков с интенсивностью от 1 до 8,3% равна $1/2$ и, следовательно, $I = 1$; половина оставшихся пиков — интенсивность $\geq 24,2\%$, т. е. соответствующая вероятность появления пиков с интенсивностью от 8,4 до 24,2% равна $(1/2)^2$ и $I = 2$. Продолжив эту процедуру, мы нашли, что величинам I , равным 3, 4 и 5 соответствуют следующие интервалы значений интенсивностей: (24,3—50,2), (50,3—87,8) и (87,9—100) %.

С целью форматного представления сокращенных спектров в поисковом массиве и экономии памяти ЭВМ при записи информации на МЛ

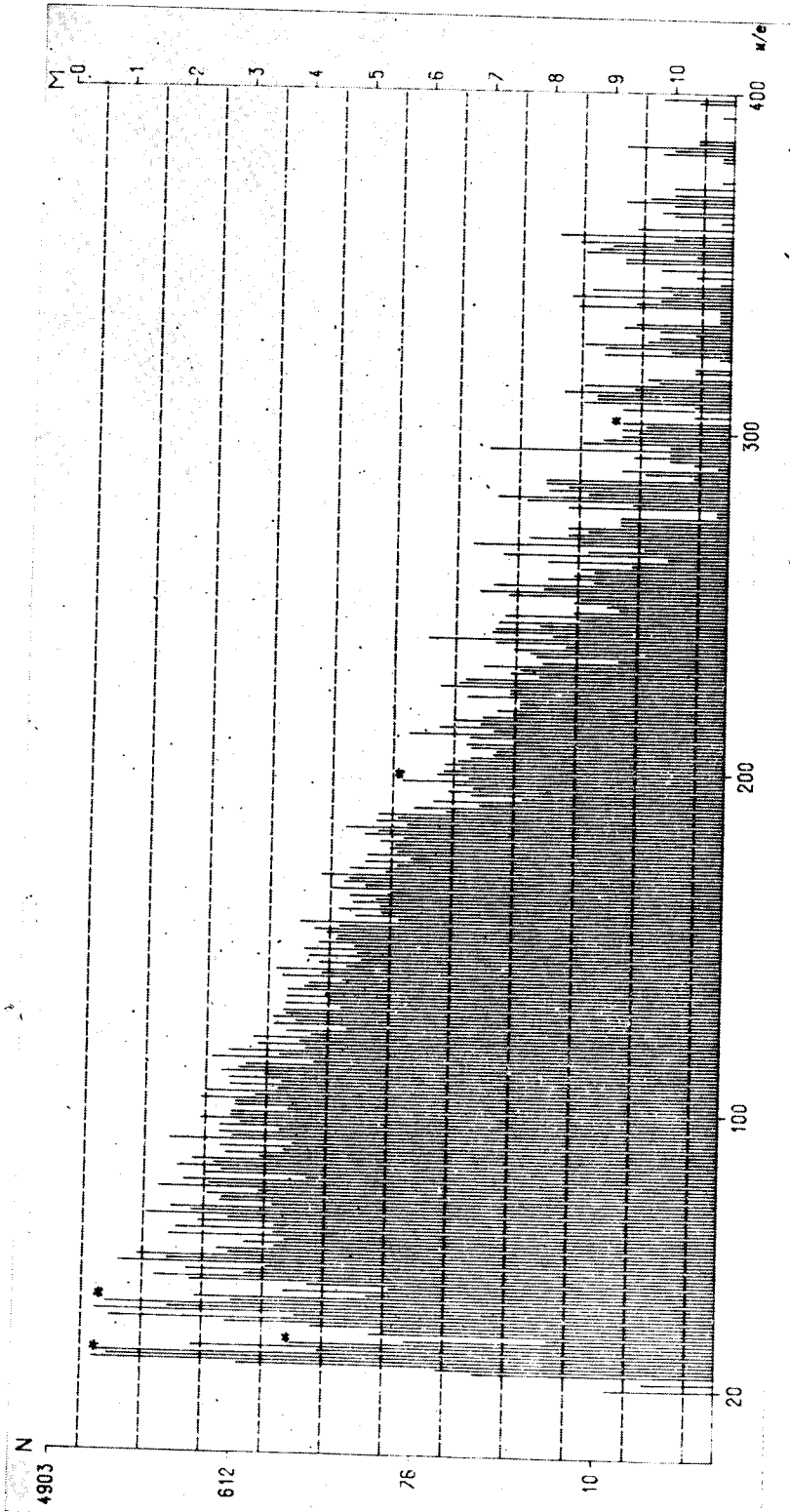


Рис. 2. Частоты появления пиков (N) при различных значениях m/e (шкала по N логарифмическая) в массиве, содержащем 4903 сокращенных масс-спектра.

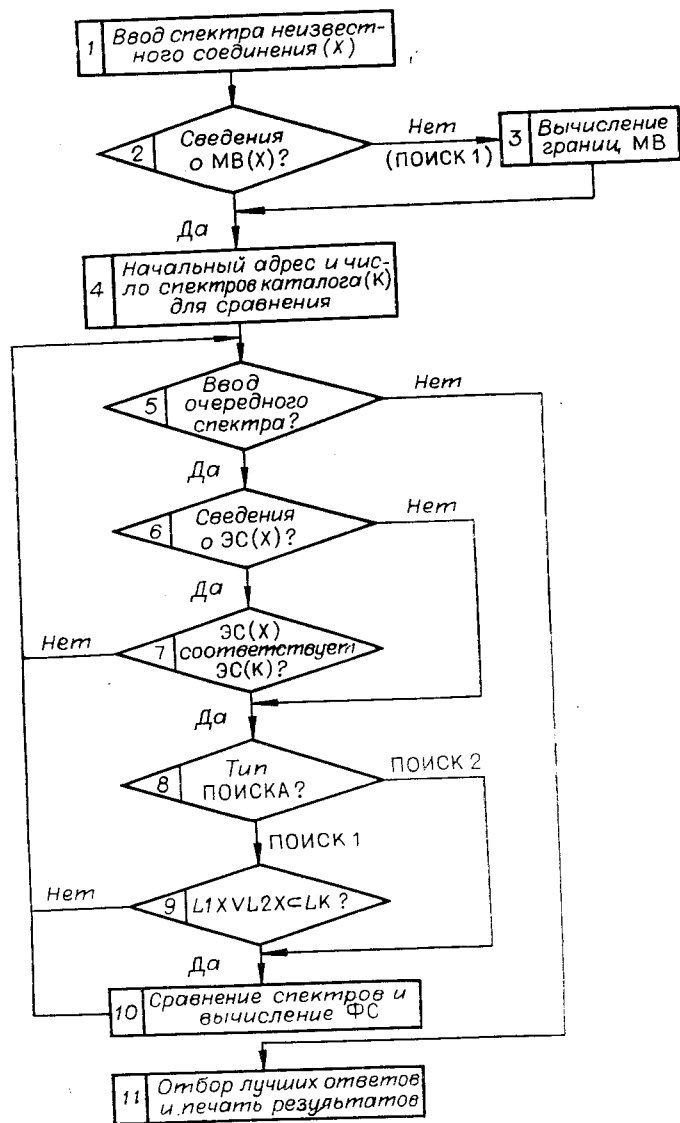


Рис. 3. Блок-схема программы ПОИСК.

использовались не сами значения m/e , а их 4-разрядные коды, определяющие относительные положения линий в соответствующих интервалах по 14 m/e [5]. Например, пик с $m/e=41$ (интервал 34—47) и интенсивностью, равной 65%, кодируется в виде 7-разрядного символа «0111100», где первые четыре разряда — код m/e ($47-41+1=7$), а оставшиеся — код интенсивности (фактор $I=4$). Весь спектр в диапазоне массовых чисел от $m/e=20$ до $m/e=705$ занимает 68 символов или 14 неполных 37-разрядных машинных слов. Кроме спектральной информации, массив содержит порядковый номер спектра, брутто-формулу, название соединения и таблицу факторов M . Полная информация о спектре составляет отдельную запись длиной в 38 ячеек. Поисковый массив представлен на МЛ зонами по 50 записей, причем записи идут по возрастанию молекулярного веса. Подобное расположение записей позволяет сократить время поиска соединений с заданными молекулярными весами. Это достигается за счет таблицы-указателя, в которой хранятся

данные о начальных адресах (номер зоны; номер записи в массиве) спектров соединений с различными молекулярными весами.

Работа поисковой системы. Поиск в машинном каталоге спектра, который наилучшим образом соответствует спектру исследуемого соединения, осуществляется с помощью программы ПОИСК, блок-схема которой приведена на рис. 3. Информация об исследуемом соединении (массовые числа и интенсивности пиков, сведения о молекулярном весе (МВ) и элементном составе (ЭС)) вводится в ЭВМ с внешнего терминального устройства («Consul-254») или с магнитной ленты. В зависимости от имеющейся у исследователя информации сведения о молекулярном весе могут быть заданы различными способами: а) одно значение МВ, б) верхняя и нижняя границы МВ, в) несколько значений МВ, г) молекулярный вес неизвестен. Аналогичным образом могут быть заданы сведения об элементном составе: а) известная брутто-формула, б) обязательное наличие в составе неизвестного соединения определенных элементов с указанием минимального числа их атомов, в) присутствие произвольных элементов и указание минимального числа атомов обязательных элементов, г) сведения об ЭС отсутствуют.

Программа системы может работать в двух режимах: ПОИСК 1 и ПОИСК 2. В обоих режимах данные о молекулярном весе исследуемого соединения и его элементном составе используются системой для сокращения числа спектров библиотеки (машинного каталога), которые необходимо далее сопоставлять со спектром исследуемого соединения.

Если исследователь не располагает указанными данными, то в режиме ПОИСК 1 используются два условия для предварительного отбора спектров библиотеки.

1. Молекулярный вес (Q) соединения из каталога должен удовлетворять неравенству (1):

$$0,6S \leq Q \leq 1,9S, \quad (1)$$

где S — максимальное массовое число линии, имеющей интенсивность $\geq 0,5\%$ в спектре неизвестного соединения. Целесообразность такого выбора объясняется тем обстоятельством, что, во-первых, не всегда в спектрах исследуемых соединений наблюдаются молекулярные ионы ($MV > S$) и, во-вторых, в спектрах изучаемых соединений возможно появление линий, обусловленных примесями ($MV < S$). Статистический анализ при сопоставлении линий S в спектрах машинного каталога с молекулярными весами соединений показал, что 99% молекулярных весов соединений удовлетворяют этому неравенству.

2. Одна из двух наиболее интенсивных линий спектра неизвестного соединения ($L1X$ или $L2X$) должна присутствовать среди трех наиболее интенсивных линий спектра машинного каталога (LK), т. е. должно выполняться логическое выражение вида

$$L1X \vee L2X \subset LK. \quad (2)$$

Если в число двух наиболее интенсивных линий попадают линии с m/e 27, 29, 41, 43, 55 и 57, которые очень часто присутствуют в масс-спектрах каталога, программа заменяет их следующими по интенсивности линиями, но интенсивность последних должна быть при этом $\geq 10\%$ от основной линии спектра. При отсутствии таковых программа использует указанные нежелательные линии, но при этом вторая линия должна иметь интенсивность $\geq 30\%$ от основной; если это условие невыполнимо, то для предварительного отбора используется всего одна (основная) линия.

Предварительно отобранные масс-спектры подвергаются далее детальному сопоставлению с исследуемым спектром для вычисления фак-

тора совпадения (ФС), характеризующего меру близости неизвестного спектра и спектров библиотеки.

В режиме ПОИСК 2 указанные условия предварительного отбора не применяются.

Сравнение спектров (блок 10, рис. 3) осуществляется следующим образом. По начальному адресу, определяющему положение первого ненулевого символа в формате неизвестного спектра, читаются коды массовых чисел спектра неизвестного соединения и библиотеки. В том случае, когда коды m/e совпадают, вычисляется величина W :

$$W = \sum_{j=1}^{m_1} M_j + \sum_{\gamma=1}^{m_2} I_{\gamma}, \quad (3)$$

где M_j — значение фактора M для совпадающих по коду значений m/e ; I_{γ} — значение фактора I для совпавших по коду градаций интенсивностей. Суммирование I_{γ} производится только в случае совпадения кодов значений m/e . Обычно $m_1 > m_2$, а при полном совпадении спектров их значения равны. Если коды m/e не совпадают, то рассматриваются следующие символы. Сравнение спектров заканчивается на последнем ненулевом символе спектра неизвестного соединения. Если в результате сравнения оказалось, что $W \geq 17^*$, то вычисляется фактор совпадения спектров:

$$\text{ФС} = [W / (W + W1)] 100\%, \quad (4)$$

где W и $W1$ — сумма весовых факторов совпавших и несовпавших характеристик спектра неизвестного соединения; легко видеть, что знаменатель выражения (4) представляет собой сумму факторов всех спектральных характеристик масс-спектра исследуемого соединения, т. е.

$$W + W1 = \sum_{i=1}^m (M + I)_i. \quad (5)$$

Из (4) следует, что ФС стремится к 100% при уменьшении значения $W1$. Чем ближе ФС к 100%, т. е. чем меньше интенсивности несопадающих спектральных линий и чем выше частоты появления массовых чисел этих линий в каталоге сокращенных спектров, тем больше уверенность в том, что совпадение спектров не случайно.

При значениях $W < 17$ и $\text{ФС} < 30\%$ результат сравнения признается отрицательным и программа переходит к сравнению спектра неизвестного соединения со следующим спектром библиотеки. После того как исчерпан весь список спектров, предназначенных для детального сравнения, осуществляется отбор соединений, спектры которых имеют наивысшие значения ФС. Результаты поиска печатаются на терминальном устройстве. Максимально возможное число соединений, выдаваемых исследователю, ограничено пятнадцатью.

На рис. 4 приведен пример диалога оператора с ЭВМ при работе информационно-поисковой системы. В качестве «неизвестного» соединения выбран этилдихлорацетат [4]. В первоначальном задании на поиск указаны массовые числа и соответствующие интенсивности пиков масс-спектра «неизвестного» соединения; предполагаемые границы молекулярного веса (от 100 до 300 единиц); сведения об элементном составе «неизвестного» (не указаны); число соединений машинного ответа (7 соединений); режим поиска (ПОИСК 1). На этом же рисунке представлен вид печати результатов поиска и один из возможных вариантов редакции первоначального задания.

* В этом случае теоретическая надежность идентификации неизвестного соединения по одному из его масс-спектров, имеющемуся в массиве, содержащем ~ 5000 спектров, должна быть более 97% [12].

```

ПРОГРАММА ПОИСКА К ВАШИМ УСЛУГАМ
ИНСТРУКЦИЯ К ПРОГРАММЕ (ДА ИЛИ НЕТ)
- НЕТ
ЗАДАНИЕ НА ПОИСК (ДА ИЛИ НЕТ)
- ДА
НОСИТЕЛЬ СПЕКТРА:
- ПМ
МАССА/ИНТЕНСИВНОСТЬ:
001- 25/41 27/174 28/101 29/1000 30/32 35/16
002- 41/35 42/35 43/18 45/17 47/32 48/73
003- 49/22 50/24 76/45 77/46 78/18 79/15
004- 83/106 85/68 87/11 111/20 113/18
005- *
МОЛЕКУЛЯРНЫЙ ВЕС:
020- 100-300
ЭЛЕМЕНТНЫЙ СОСТАВ:
030-
ЧИСЛО ОТВЕТОВ:
040- 7
РЕЖИМ ПОИСКА:
050- 1
ПОИСК 1: 209 СПЕКТРОВа

```

	ФС%	N ^б	ИСТОЧНИК ^в	МВ ^г	БФ ^д
1	100	2871	MOR-130	156	C4H6O2CL2
			ETHYL DICHLOROACETATE		
2	93	2867	API-789	156	C4H6O2CL2
			ETHYL DICHLOROACETATE (GAS)		
3	93	2864	DOW-1328	156	C4H6O2CL2
			ETHYL DICHLOROACETATE		
4	39	1914	DOW-901	162	C2H1O2CL3
			TRICHLOROACETIC ACID		
5	39	3562	DOT-191	162	C1H1CL3BR1
			BROMODICHLOROMETHANE		
6	36	3156	DOW-489	168	C2H1CL3F2
			1,1,2-TRICHLORO-2,2-DIFLUOROETHANE		
7	36	2438	DOW-1851	142	C3H4O2CL2
			METHYL DICHLOROACETATE		

```

РЕДАКЦИЯ ЗАДАНИЯ (ДА ИЛИ НЕТ)
- ДА
НОМЕР СТРОКИ- ДАННЫЕ:
020- 156
030- CL2, } е
050- 2
ПОИСК 2: 8 СПЕКТРОВ

```

	ФС%	N	ИСТОЧНИК	МВ	БФ
1	100	2871	MOR-130	156	C4H6O2CL2
			ETHYL DICHLOROACETATE		
2	93	2867	API-789	156	C4H6O2CL2
			ETHYL DICHLOROACETATE (GAS)		
3	93	2864	DOW-1328	156	C4H6O2CL2
			ETHYL DICHLOROACETATE		

```

РЕДАКЦИЯ ЗАДАНИЯ (ДА ИЛИ НЕТ)
- НЕТ
ДО СВЕДАНИЯ

```

Рис. 4. Пример диалога оператора с ЭВМ при работе информационно-поисковой системы:

a — число спектров, отобранных для детального сравнения; *б* — порядковый номер спектра в поисковом массиве; *в* — литературный источник полного масс-спектра (см. [4]); *г*, *д* — молекулярный вес и брутто-формула соединений из машинного каталога; *е* — редакция задания (указаны молекулярный вес, минимальное число атомов обязательного элемента и изменен режим поиска).

яти первых соединений, соответственно выдаваемых в машинном ответе. Введение в запрос сведений о молекулярном весе «исследуемого» соединения заметно улучшает результаты. Обращает на себя внимание тот факт, что вероятность появления идентифицируемого соединения на первом месте в машинном ответе для соединений группы *B* несколько ниже, чем для групп *A* и *C*. По-видимому, это объясняется тем, что в группе *B* оказалось большое число изомерных соединений, масс-спектры которых очень мало отличаются друг от друга (например, 1, 2, 4-триметилбензол и 1, 2, 3-триметилбензол, терефталевая кислота и изофталевая кислота и т. п.).

Обсуждение результатов.

Работа поисковой системы была проверена на примере 217 различных соединений, отобранных в машинном каталоге таким образом, чтобы они были представлены в нем несколькими спектрами. Отобранные соединения были разбиты на три группы: соединения с молекулярными весами от 25 до 117 (группа *A*, 80 соединений), 118—187 (группа *B*, 80 соединений) и 188—500 (группа *C*, 57 соединений). По результатам работы системы в режиме ПОИСК 1 проведена оценка способности ИПС решать задачи двух типов: а) идентификации «исследуемых» соединений по их масс-спектрам, б) получения информации о структурном типе соединения, если его масс-спектр не представлен в машинном каталоге. Критериями способности системы идентифицировать «исследуемое» соединение по его масс-спектру могут служить средние вероятности появления (*P*) этого соединения (найденного по одному из его масс-спектров, присутствующих в машинном каталоге спектров) в пределах определенного числа соединений, упорядоченных в машинном ответе по величинам факторов совпадений спектров. Соответствующие результаты поисков по спектрам 217 различных соединений сведены в табл. 1.

Как видно из таблицы, вероятности *P* для различных групп соединений (*A*, *B* и *C*) несколько колеблются, составляя в среднем 78,0; 95,7 и 98,5% для одного, пяти и де-

Таблица 1

Средняя вероятность появления (P) «исследуемого» соединения
среди первых K соединений машинного ответа

K	P, %				P, %*			
	A	B	C	Среднее значение	A	B	C	Среднее значение
1	78,8	73,8	82,5	78,0	87,5	82,5	89,5	86,1
5	96,3	96,3	94,8	95,7	99,0	100	98,5	99,2
10	97,5	99,0	98,5	98,5	99,0	100	98,5	99,2

* Результаты получены при указании в запросе молекулярного веса «исследуемого» соединения.

Отметим также, что вероятность идентификации в значительной степени зависит от величин факторов совпадения сравниваемых спектров. Так, если в режиме ПОИСК 1 оказывается, что для соединения библиотеки выдаваемого на первом месте машинного ответа фактор совпадения масс-спектра со спектром «исследуемого» соединения лежит в диапазоне от 100 до 81% или от 80 до 41%, то средняя вероятность то-

<p>"НЕИЗВЕСТНОЕ": 1,1-DIETHOXYHEPTANE (C11H24O2, MB= 188)</p> <p>ПОИСК 1: 183 СПЕКТРА</p> <table border="1"> <thead> <tr> <th>№</th> <th>ФС%</th> <th>N</th> <th>ИСТОЧНИК</th> <th>MB</th> <th>ВФ</th> </tr> </thead> <tbody> <tr><td>1</td><td>64</td><td>4229</td><td>SIK-2297</td><td>230</td><td>C14H3002</td></tr> <tr><td></td><td></td><td></td><td>1,1-DIETHOXYDECANE</td><td></td><td></td></tr> <tr><td>2</td><td>56</td><td>3364</td><td>SIK-2293</td><td>174</td><td>C10H2202</td></tr> <tr><td></td><td></td><td></td><td>1,1-DIETHOXYHEXANE</td><td></td><td></td></tr> <tr><td>3</td><td>57</td><td>3878</td><td>SIK-2295</td><td>202</td><td>C12H2602</td></tr> <tr><td></td><td></td><td></td><td>1,1-DIETHOXYOCTANE</td><td></td><td></td></tr> <tr><td>4</td><td>47</td><td>2492</td><td>WUR-56</td><td>174</td><td>C10H2202</td></tr> <tr><td></td><td></td><td></td><td>1,1-DIETHOXYHEXANE</td><td></td><td></td></tr> <tr><td>5</td><td>42</td><td>4070</td><td>WUR-66</td><td>216</td><td>C13H2802</td></tr> <tr><td></td><td></td><td></td><td>1,1-DIETHOXYNONANE</td><td></td><td></td></tr> <tr><td>6</td><td>41</td><td>3876</td><td>WUR-65</td><td>202</td><td>C12H2602</td></tr> <tr><td></td><td></td><td></td><td>1,1-DIETHOXYOCTANE</td><td></td><td></td></tr> <tr><td>7</td><td>36</td><td>4230</td><td>WUR-67</td><td>230</td><td>C14H3002</td></tr> <tr><td></td><td></td><td></td><td>1,1-DIETHOXYDECANE</td><td></td><td></td></tr> <tr><td>8</td><td>32</td><td>2627</td><td>WUR-13</td><td>146</td><td>C8H1802</td></tr> <tr><td></td><td></td><td></td><td>1,1-DIETHOXYBUTANE</td><td></td><td></td></tr> <tr><td>9</td><td>32</td><td>2625</td><td>SIK-2291</td><td>146</td><td>C8H1802</td></tr> <tr><td></td><td></td><td></td><td>1,1-DIETHOXYBUTANE</td><td></td><td></td></tr> <tr><td>10</td><td>31</td><td>4324</td><td>SIK-2324</td><td>242</td><td>C15H3002</td></tr> <tr><td></td><td></td><td></td><td>1,1-DIETHOXY UNDECENE-10</td><td></td><td></td></tr> </tbody> </table>					№	ФС%	N	ИСТОЧНИК	MB	ВФ	1	64	4229	SIK-2297	230	C14H3002				1,1-DIETHOXYDECANE			2	56	3364	SIK-2293	174	C10H2202				1,1-DIETHOXYHEXANE			3	57	3878	SIK-2295	202	C12H2602				1,1-DIETHOXYOCTANE			4	47	2492	WUR-56	174	C10H2202				1,1-DIETHOXYHEXANE			5	42	4070	WUR-66	216	C13H2802				1,1-DIETHOXYNONANE			6	41	3876	WUR-65	202	C12H2602				1,1-DIETHOXYOCTANE			7	36	4230	WUR-67	230	C14H3002				1,1-DIETHOXYDECANE			8	32	2627	WUR-13	146	C8H1802				1,1-DIETHOXYBUTANE			9	32	2625	SIK-2291	146	C8H1802				1,1-DIETHOXYBUTANE			10	31	4324	SIK-2324	242	C15H3002				1,1-DIETHOXY UNDECENE-10			<p>"НЕИЗВЕСТНОЕ": METHYL HEPTANOATE (C8H16O2, MB= 144)</p> <p>ПОИСК 1: 166 СПЕКТРОВ</p> <table border="1"> <thead> <tr> <th>№</th> <th>ФС%</th> <th>N</th> <th>ИСТОЧНИК</th> <th>MB</th> <th>ВФ</th> </tr> </thead> <tbody> <tr><td>1</td><td>52</td><td>3298</td><td>RYS-5</td><td>172</td><td>C10H2202</td></tr> <tr><td></td><td></td><td></td><td>METHYL N-NONANOATE</td><td></td><td></td></tr> <tr><td>2</td><td>50</td><td>3593</td><td>RYS-6</td><td>186</td><td>C11H2202</td></tr> <tr><td></td><td></td><td></td><td>METHYL N-DECANOATE</td><td></td><td></td></tr> <tr><td>3</td><td>50</td><td>3594</td><td>WUR-1834</td><td>186</td><td>C11H2202</td></tr> <tr><td></td><td></td><td></td><td>METHYL DECANOATE</td><td></td><td></td></tr> <tr><td>4</td><td>49</td><td>1762</td><td>RYS-4</td><td>158</td><td>C9H1802</td></tr> <tr><td></td><td></td><td></td><td>METHYL N-OCTANOATE</td><td></td><td></td></tr> <tr><td>5</td><td>48</td><td>2857</td><td>AST-1834</td><td>186</td><td>C11H2202</td></tr> <tr><td></td><td></td><td></td><td>METHYL DECANOATE</td><td></td><td></td></tr> <tr><td>6</td><td>47</td><td>4326</td><td>RYS-10</td><td>242</td><td>C15H3002</td></tr> <tr><td></td><td></td><td></td><td>METHYL N-TETRADECANOATE</td><td></td><td></td></tr> <tr><td>7</td><td>47</td><td>1969</td><td>DOW-2390</td><td>130</td><td>C7H1402</td></tr> <tr><td></td><td></td><td></td><td>METHYL N-HEXANOATE</td><td></td><td></td></tr> <tr><td>8</td><td>45</td><td>2942</td><td>MOR-132</td><td>158</td><td>C9H1802</td></tr> <tr><td></td><td></td><td></td><td>METHYL CAPRYLATE</td><td></td><td></td></tr> <tr><td>9</td><td>42</td><td>4490</td><td>USI-1907</td><td>270</td><td>C17H3402</td></tr> <tr><td></td><td></td><td></td><td>METHYL PALMITATE</td><td></td><td></td></tr> <tr><td>10</td><td>42</td><td>2945</td><td>WUR-1820</td><td>158</td><td>C9H1802</td></tr> <tr><td></td><td></td><td></td><td>METHYL OCTANOATE</td><td></td><td></td></tr> </tbody> </table>					№	ФС%	N	ИСТОЧНИК	MB	ВФ	1	52	3298	RYS-5	172	C10H2202				METHYL N-NONANOATE			2	50	3593	RYS-6	186	C11H2202				METHYL N-DECANOATE			3	50	3594	WUR-1834	186	C11H2202				METHYL DECANOATE			4	49	1762	RYS-4	158	C9H1802				METHYL N-OCTANOATE			5	48	2857	AST-1834	186	C11H2202				METHYL DECANOATE			6	47	4326	RYS-10	242	C15H3002				METHYL N-TETRADECANOATE			7	47	1969	DOW-2390	130	C7H1402				METHYL N-HEXANOATE			8	45	2942	MOR-132	158	C9H1802				METHYL CAPRYLATE			9	42	4490	USI-1907	270	C17H3402				METHYL PALMITATE			10	42	2945	WUR-1820	158	C9H1802				METHYL OCTANOATE		
№	ФС%	N	ИСТОЧНИК	MB	ВФ																																																																																																																																																																																																																																																																
1	64	4229	SIK-2297	230	C14H3002																																																																																																																																																																																																																																																																
			1,1-DIETHOXYDECANE																																																																																																																																																																																																																																																																		
2	56	3364	SIK-2293	174	C10H2202																																																																																																																																																																																																																																																																
			1,1-DIETHOXYHEXANE																																																																																																																																																																																																																																																																		
3	57	3878	SIK-2295	202	C12H2602																																																																																																																																																																																																																																																																
			1,1-DIETHOXYOCTANE																																																																																																																																																																																																																																																																		
4	47	2492	WUR-56	174	C10H2202																																																																																																																																																																																																																																																																
			1,1-DIETHOXYHEXANE																																																																																																																																																																																																																																																																		
5	42	4070	WUR-66	216	C13H2802																																																																																																																																																																																																																																																																
			1,1-DIETHOXYNONANE																																																																																																																																																																																																																																																																		
6	41	3876	WUR-65	202	C12H2602																																																																																																																																																																																																																																																																
			1,1-DIETHOXYOCTANE																																																																																																																																																																																																																																																																		
7	36	4230	WUR-67	230	C14H3002																																																																																																																																																																																																																																																																
			1,1-DIETHOXYDECANE																																																																																																																																																																																																																																																																		
8	32	2627	WUR-13	146	C8H1802																																																																																																																																																																																																																																																																
			1,1-DIETHOXYBUTANE																																																																																																																																																																																																																																																																		
9	32	2625	SIK-2291	146	C8H1802																																																																																																																																																																																																																																																																
			1,1-DIETHOXYBUTANE																																																																																																																																																																																																																																																																		
10	31	4324	SIK-2324	242	C15H3002																																																																																																																																																																																																																																																																
			1,1-DIETHOXY UNDECENE-10																																																																																																																																																																																																																																																																		
№	ФС%	N	ИСТОЧНИК	MB	ВФ																																																																																																																																																																																																																																																																
1	52	3298	RYS-5	172	C10H2202																																																																																																																																																																																																																																																																
			METHYL N-NONANOATE																																																																																																																																																																																																																																																																		
2	50	3593	RYS-6	186	C11H2202																																																																																																																																																																																																																																																																
			METHYL N-DECANOATE																																																																																																																																																																																																																																																																		
3	50	3594	WUR-1834	186	C11H2202																																																																																																																																																																																																																																																																
			METHYL DECANOATE																																																																																																																																																																																																																																																																		
4	49	1762	RYS-4	158	C9H1802																																																																																																																																																																																																																																																																
			METHYL N-OCTANOATE																																																																																																																																																																																																																																																																		
5	48	2857	AST-1834	186	C11H2202																																																																																																																																																																																																																																																																
			METHYL DECANOATE																																																																																																																																																																																																																																																																		
6	47	4326	RYS-10	242	C15H3002																																																																																																																																																																																																																																																																
			METHYL N-TETRADECANOATE																																																																																																																																																																																																																																																																		
7	47	1969	DOW-2390	130	C7H1402																																																																																																																																																																																																																																																																
			METHYL N-HEXANOATE																																																																																																																																																																																																																																																																		
8	45	2942	MOR-132	158	C9H1802																																																																																																																																																																																																																																																																
			METHYL CAPRYLATE																																																																																																																																																																																																																																																																		
9	42	4490	USI-1907	270	C17H3402																																																																																																																																																																																																																																																																
			METHYL PALMITATE																																																																																																																																																																																																																																																																		
10	42	2945	WUR-1820	158	C9H1802																																																																																																																																																																																																																																																																
			METHYL OCTANOATE																																																																																																																																																																																																																																																																		
<p>"НЕИЗВЕСТНОЕ": ALLYL O-PHTHALATE (C14H14O4, MB= 246)</p> <p>ПОИСК 1: 62 СПЕКТРА</p> <table border="1"> <thead> <tr> <th>№</th> <th>ФС%</th> <th>N</th> <th>ИСТОЧНИК</th> <th>MB</th> <th>ВФ</th> </tr> </thead> <tbody> <tr><td>1</td><td>52</td><td>4344</td><td>SIK-2368</td><td>222</td><td>C12H1404</td></tr> <tr><td></td><td></td><td></td><td>ETHYL O-PHTHALATE</td><td></td><td></td></tr> <tr><td>2</td><td>32</td><td>4116</td><td>DOW-2855</td><td>222</td><td>C12H1404</td></tr> <tr><td></td><td></td><td></td><td>DIETHYL TEREPHTHALATE</td><td></td><td></td></tr> <tr><td>3</td><td>52</td><td>4596</td><td>SIK-2376</td><td>306</td><td>C18H2604</td></tr> <tr><td></td><td></td><td></td><td>N-AMYL O-PHTHALATE</td><td></td><td></td></tr> <tr><td>4</td><td>51</td><td>2585</td><td>SIK-2371</td><td>250</td><td>C14H1804</td></tr> <tr><td></td><td></td><td></td><td>N-PROPYL O-PHTHALATE</td><td></td><td></td></tr> <tr><td>5</td><td>49</td><td>4510</td><td>SIK-2375</td><td>278</td><td>C16H2204</td></tr> <tr><td></td><td></td><td></td><td>ISOBUTYL P-PHTHALATE</td><td></td><td></td></tr> <tr><td>6</td><td>49</td><td>3589</td><td>DOW-2285</td><td>250</td><td>C14H1804</td></tr> <tr><td></td><td></td><td></td><td>DI-N-PROPYL PHTHALATE</td><td></td><td></td></tr> <tr><td>7</td><td>48</td><td>4503</td><td>DOW-2373</td><td>278</td><td>C16H2204</td></tr> <tr><td></td><td></td><td></td><td>DI-ISOBUTYL PHTHALATE</td><td></td><td></td></tr> <tr><td>8</td><td>47</td><td>4511</td><td>DOW-2608</td><td>278</td><td>C16H2204</td></tr> <tr><td></td><td></td><td></td><td>DIISOBUTYLTEREPHTHALATE</td><td></td><td></td></tr> <tr><td>9</td><td>44</td><td>4508</td><td>SIK-2373</td><td>278</td><td>C16H2204</td></tr> <tr><td></td><td></td><td></td><td>N-BUTYL O-PHTHALATE</td><td></td><td></td></tr> <tr><td>10</td><td>41</td><td>3590</td><td>SIK-2372</td><td>250</td><td>C14H1804</td></tr> <tr><td></td><td></td><td></td><td>ISOPROPYL O-PHTHALATE</td><td></td><td></td></tr> </tbody> </table>					№	ФС%	N	ИСТОЧНИК	MB	ВФ	1	52	4344	SIK-2368	222	C12H1404				ETHYL O-PHTHALATE			2	32	4116	DOW-2855	222	C12H1404				DIETHYL TEREPHTHALATE			3	52	4596	SIK-2376	306	C18H2604				N-AMYL O-PHTHALATE			4	51	2585	SIK-2371	250	C14H1804				N-PROPYL O-PHTHALATE			5	49	4510	SIK-2375	278	C16H2204				ISOBUTYL P-PHTHALATE			6	49	3589	DOW-2285	250	C14H1804				DI-N-PROPYL PHTHALATE			7	48	4503	DOW-2373	278	C16H2204				DI-ISOBUTYL PHTHALATE			8	47	4511	DOW-2608	278	C16H2204				DIISOBUTYLTEREPHTHALATE			9	44	4508	SIK-2373	278	C16H2204				N-BUTYL O-PHTHALATE			10	41	3590	SIK-2372	250	C14H1804				ISOPROPYL O-PHTHALATE			<p>"НЕИЗВЕСТНОЕ": 1-ARACHIDYL-2,3-DIACETIN (C27H50O6, MB= 470)</p> <p>ПОИСК 1: 48 СПЕКТРОВ</p> <table border="1"> <thead> <tr> <th>№</th> <th>ФС%</th> <th>N</th> <th>ИСТОЧНИК</th> <th>MB</th> <th>ВФ</th> </tr> </thead> <tbody> <tr><td>1</td><td>76</td><td>4829</td><td>HOL-273</td><td>386</td><td>C21H3806</td></tr> <tr><td></td><td></td><td></td><td>1-MYRISTO-2,3-DIACETIN</td><td></td><td></td></tr> <tr><td>2</td><td>64</td><td>4914</td><td>HOL-443</td><td>470</td><td>C27H5006</td></tr> <tr><td></td><td></td><td></td><td>2-ARACHIDYL-1,3-DIACETIN</td><td></td><td></td></tr> <tr><td>3</td><td>57</td><td>4874</td><td>HOL-437</td><td>414</td><td>C23H4206</td></tr> <tr><td></td><td></td><td></td><td>1-PALMITYL-2,3-DIACETIN</td><td></td><td></td></tr> <tr><td>4</td><td>51</td><td>4902</td><td>HOL-266</td><td>442</td><td>C25H4606</td></tr> <tr><td></td><td></td><td></td><td>1-STEARO-2,3-DIACETIN</td><td></td><td></td></tr> <tr><td>5</td><td>49</td><td>4872</td><td>HOL-438</td><td>414</td><td>C23H4206</td></tr> <tr><td></td><td></td><td></td><td>2-PALMITYL-1,3-DIACETIN</td><td></td><td></td></tr> <tr><td>6</td><td>44</td><td>4623</td><td>HOL-238</td><td>386</td><td>C23H4604</td></tr> <tr><td></td><td></td><td></td><td>1-MONOARACHIDIN</td><td></td><td></td></tr> <tr><td>7</td><td>36</td><td>4644</td><td>HOL-233</td><td>330</td><td>C19H3804</td></tr> <tr><td></td><td></td><td></td><td>1-MONOPALMITIN</td><td></td><td></td></tr> <tr><td>8</td><td>35</td><td>4897</td><td>HOL-9405</td><td>438</td><td>C25H4206</td></tr> <tr><td></td><td></td><td></td><td>1-LINOLEYL-2,3-DIACETIN</td><td></td><td></td></tr> <tr><td>9</td><td>33</td><td>4873</td><td>HOL-9438</td><td>414</td><td>C23H4206</td></tr> <tr><td></td><td></td><td></td><td>2-PALMITYL-1,3-DIACETIN</td><td></td><td></td></tr> <tr><td>10</td><td>30</td><td>4645</td><td>HOL-412</td><td>330</td><td>C19H3804</td></tr> <tr><td></td><td></td><td></td><td>2-MONOPALMITIN</td><td></td><td></td></tr> </tbody> </table>					№	ФС%	N	ИСТОЧНИК	MB	ВФ	1	76	4829	HOL-273	386	C21H3806				1-MYRISTO-2,3-DIACETIN			2	64	4914	HOL-443	470	C27H5006				2-ARACHIDYL-1,3-DIACETIN			3	57	4874	HOL-437	414	C23H4206				1-PALMITYL-2,3-DIACETIN			4	51	4902	HOL-266	442	C25H4606				1-STEARO-2,3-DIACETIN			5	49	4872	HOL-438	414	C23H4206				2-PALMITYL-1,3-DIACETIN			6	44	4623	HOL-238	386	C23H4604				1-MONOARACHIDIN			7	36	4644	HOL-233	330	C19H3804				1-MONOPALMITIN			8	35	4897	HOL-9405	438	C25H4206				1-LINOLEYL-2,3-DIACETIN			9	33	4873	HOL-9438	414	C23H4206				2-PALMITYL-1,3-DIACETIN			10	30	4645	HOL-412	330	C19H3804				2-MONOPALMITIN		
№	ФС%	N	ИСТОЧНИК	MB	ВФ																																																																																																																																																																																																																																																																
1	52	4344	SIK-2368	222	C12H1404																																																																																																																																																																																																																																																																
			ETHYL O-PHTHALATE																																																																																																																																																																																																																																																																		
2	32	4116	DOW-2855	222	C12H1404																																																																																																																																																																																																																																																																
			DIETHYL TEREPHTHALATE																																																																																																																																																																																																																																																																		
3	52	4596	SIK-2376	306	C18H2604																																																																																																																																																																																																																																																																
			N-AMYL O-PHTHALATE																																																																																																																																																																																																																																																																		
4	51	2585	SIK-2371	250	C14H1804																																																																																																																																																																																																																																																																
			N-PROPYL O-PHTHALATE																																																																																																																																																																																																																																																																		
5	49	4510	SIK-2375	278	C16H2204																																																																																																																																																																																																																																																																
			ISOBUTYL P-PHTHALATE																																																																																																																																																																																																																																																																		
6	49	3589	DOW-2285	250	C14H1804																																																																																																																																																																																																																																																																
			DI-N-PROPYL PHTHALATE																																																																																																																																																																																																																																																																		
7	48	4503	DOW-2373	278	C16H2204																																																																																																																																																																																																																																																																
			DI-ISOBUTYL PHTHALATE																																																																																																																																																																																																																																																																		
8	47	4511	DOW-2608	278	C16H2204																																																																																																																																																																																																																																																																
			DIISOBUTYLTEREPHTHALATE																																																																																																																																																																																																																																																																		
9	44	4508	SIK-2373	278	C16H2204																																																																																																																																																																																																																																																																
			N-BUTYL O-PHTHALATE																																																																																																																																																																																																																																																																		
10	41	3590	SIK-2372	250	C14H1804																																																																																																																																																																																																																																																																
			ISOPROPYL O-PHTHALATE																																																																																																																																																																																																																																																																		
№	ФС%	N	ИСТОЧНИК	MB	ВФ																																																																																																																																																																																																																																																																
1	76	4829	HOL-273	386	C21H3806																																																																																																																																																																																																																																																																
			1-MYRISTO-2,3-DIACETIN																																																																																																																																																																																																																																																																		
2	64	4914	HOL-443	470	C27H5006																																																																																																																																																																																																																																																																
			2-ARACHIDYL-1,3-DIACETIN																																																																																																																																																																																																																																																																		
3	57	4874	HOL-437	414	C23H4206																																																																																																																																																																																																																																																																
			1-PALMITYL-2,3-DIACETIN																																																																																																																																																																																																																																																																		
4	51	4902	HOL-266	442	C25H4606																																																																																																																																																																																																																																																																
			1-STEARO-2,3-DIACETIN																																																																																																																																																																																																																																																																		
5	49	4872	HOL-438	414	C23H4206																																																																																																																																																																																																																																																																
			2-PALMITYL-1,3-DIACETIN																																																																																																																																																																																																																																																																		
6	44	4623	HOL-238	386	C23H4604																																																																																																																																																																																																																																																																
			1-MONOARACHIDIN																																																																																																																																																																																																																																																																		
7	36	4644	HOL-233	330	C19H3804																																																																																																																																																																																																																																																																
			1-MONOPALMITIN																																																																																																																																																																																																																																																																		
8	35	4897	HOL-9405	438	C25H4206																																																																																																																																																																																																																																																																
			1-LINOLEYL-2,3-DIACETIN																																																																																																																																																																																																																																																																		
9	33	4873	HOL-9438	414	C23H4206																																																																																																																																																																																																																																																																
			2-PALMITYL-1,3-DIACETIN																																																																																																																																																																																																																																																																		
10	30	4645	HOL-412	330	C19H3804																																																																																																																																																																																																																																																																
			2-MONOPALMITIN																																																																																																																																																																																																																																																																		

Рис. 5. Примеры машинных ответов при поиске соединений, масс-спектры которых отсутствовали в поисковом массиве.

Таблица 2
Средняя вероятность появления в машинном ответе структурно-подобного соединения в зависимости от фактора совпадения спектров (ФС)

ФС, %	100 n_k/n , %			Среднее значение
	A	B	C	
100—81	91,0	100	100	96,1
80—71	90,0	96,0	100	94,1
70—61	76,0	82,7	96,0	84,0
60—51	68,5	57,5	87,5	66,5
50—41	57,5	62,5	70,1	62,0
40—31	23,5	43,3	34,4	34,0

го, что «исследуемое» соединение опознано, составляет 87,5 и 58,0% соответственно. При задании (в запросе на поиск) молекулярных весов «исследуемых» соединений соответствующие вероятности увеличиваются и составляют для указанных выше диапазонов факторов совпадения 93,5 и 80,0% соответственно.

Важной характеристикой поисковой системы является и ее способность отыскивать все спектры исследуемого соединения, имеющиеся в машинном каталоге и, возможно, записанные в условиях, отличных от использованных при регистрации спектра исследуемого

соединения. Ранее неоднократно отмечалось (см., например, [7]), что при изменении условий регистрации масс-спектры одного и того же соединения могут значительно отличаться друг от друга. Проведенная оценка показала, что средняя вероятность нахождения всех спектров «исследуемого» соединения в числе первых пяти и десяти результатов поиска составляет 87,5 и 94,5% соответственно.

При практическом использовании ИПС с достаточно большой вероятностью можно столкнуться с ситуацией, когда масс-спектр исследуемого соединения отсутствует в машинном каталоге, и поэтому оно не может быть идентифицировано по масс-спектру. По этой причине, наряду со способностью системы осуществлять идентификацию исследуемых соединений, мы попытались оценить ее способность отбирать масс-спектры структурно-подобных соединений. С этой целью при обработке результатов работы системы в режиме ПОИСК 1 отбрасывалась та часть машинного ответа, которая относилась к спектрам соединения, фигурирующего в запросе, а в качестве корректных ответов ЭВМ рассматривались те соединения, которые содержали достаточно большие структурные блоки или функциональные группы, присутствующие в «исследуемом» соединении.

В том, что система обеспечивает поиск структурных аналогов, можно убедиться, просмотрев результаты нескольких машинных ответов, приведенных на рис. 5.

Чтобы охарактеризовать способность системы отыскивать структурные аналоги, можно воспользоваться величиной вероятности присутствия структурно-подобного соединения среди той части отобранных машиной соединений, для которой значения факторов совпадения спектров попадают в определенный заданный интервал, т. е. величиной $100 n_k/n$, где n — общее число отобранных соединений для выбранного диапазона ФС, а n_k — число структурных аналогов исследуемых соединений в этой группе. Значения указанных вероятностей, вычисленные (на основании результатов тех же экспериментов, которые легли в основу табл. 1) для шести различных интервалов ФС, представлены в табл. 2. Из таблицы видно, что достоверность выявления структурных аналогов при значениях ФС в диапазоне от 100 до 71% для всех групп соединений превышает 90%, а для группы соединений с молекулярными весами 188—500 равна 100%. В среднем по данным для всего массива при $ФС \geq 60\%$ (достаточно низкий фактор совпадения спектров) вероятность нахождения структурных аналогов превышает 84%.

Заметим, что отсутствие строгости в определении «структурный аналог» (в описываемых экспериментах решение о наличии или отсутствии структурного подобия принималось химиком) делает приведенные выше

результаты ориентировочными. Однако увеличение объема машинного каталога спектров и введение в него кодовых записей структурных формул позволит делать количественные оценки подобного рода для любых конкретных структурных фрагментов.

В целом результаты экспериментальной работы с описываемой ИПС позволяют заключить, что предложенные способ сокращения полных масс-спектров и метод определения меры близости спектров могут оказаться весьма полезными.

Поисковая система реализована на языке символического кодирования и ФОРТРАНе для ЭВМ «Минск-32». Время работы системы в режиме ПОИСК 1 не превышает одной минуты.

ЛИТЕРАТУРА

1. Heller R. S., Milne G. W. A., Feldmann R. J., Heller S. R. An international mass spectral search system (MSSS).—*J. Chem. Inf. Comput. Sci.*, 1976, vol. 16, N 3, p. 176.
2. Дерендяев Б. Г., Покровский Л. М., Нехорошев С. А., Смирнов В. И., Коптюг В. А. Машинная информационно-поисковая система для масс-спектрометрии.—*Изв. СО АН СССР. Сер. хим. наук*, 1977, вып. 2, с. 109.
3. Kwok K.-S., Venkataraghavan R., McLafferty F. W. Computer-aided interpretation of mass spectra. A self-training interpretive and retrieval system.—*J. Am. Chem. Soc.*, 1973, vol. 95, N 13, p. 4185.
4. Stenhagen E., Abrahamson S. A., McLafferty F. W. Atlas of mass spectral data. Ed., John Wiley and Sons. N. Y., 1969.
5. Grotch S. L. Computer identification of mass spectra using highly compressed spectral codes.—*Anal. Chem.*, 1973, vol. 45, N 1, p. 1.
6. Hertz H. S., Hites R. A., Biemann K. Identification of mass spectra by computer-searching a file of known spectra.—*Anal. Chem.*, 1971, vol. 43, N 6, p. 681.
7. Knock B. A., Smith I. C., Wright D. E., Ridley R. G. Compound identification by computer matching of low resolution mass spectra.—*Anal. Chem.*, 1970, vol. 42, N 13, p. 1516.
8. Heller S. R. Conversational mass spectral retrieval system and its use as an aid in structure determination.—*Anal. Chem.*, 1972, vol. 44, N 12, p. 1951.
9. Wangen L. E., Woodward W. S., Isenhour T. L. Small computer, magnetic tape oriented, rapid search system applied to mass spectrometry.—*Anal. Chem.*, 1971, vol. 43, N 12, p. 1605.
10. McLafferty F. W., Hertel R. H., Villwock R. D. Probability based matching of mass spectra, rapid identification of specific compounds in mixtures.—*Org. Mass. Spectrom.*, 1974, vol. 9, N 7, p. 690.
11. Pesyna G. M., McLafferty F. W., Venkataraghavan R., Dayringer H. E. Statistical occurrence of mass and abundance values in mass spectra.—*Anal. Chem.*, 1975, vol. 47, N 7, p. 1161.
12. Pesyna G. M., Venkataraghavan R., Dayringer H. E., McLafferty F. W. Probability based matching system using a large collection of reference mass spectra.—*Anal. Chem.*, 1976, vol. 48, N 9, p. 1362.

Поступила в редакцию 28 августа 1978 г.

УДК 65.011.56 : 681.32 : 543.51

Б. Г. ДЕРЕНДЯЕВ, С. А. НЕХОРОШЕВ, Л. М. ПОКРОВСКИЙ

(Новосибирск)

АВТОМАТИЗИРОВАННАЯ ОБРАБОТКА ДАНЫХ МАСС-СПЕКТРОМЕТРИЧЕСКИХ ЭКСПЕРИМЕНТОВ НА БАЗЕ ЭВМ «МИНСК-32» В РЕЖИМЕ ВЫСОКОГО РАЗРЕШЕНИЯ

Развитие систем автоматического опознавания структур неизвестных соединений по их спектральным характеристикам [1, 2] в значительной мере зависит от степени автоматизации процедуры обработки данных