

# 计算机自动结构解析专家系统研究\*

## —— ESESOC-II 结构产生器

胡昌玉 许 禄 \*\*

(中国科学院长春应用化学研究所应用谱学开放实验室, 长春130022)

### 摘要

ESESOC-II是一个用于含O,N,P,S及卤素等多种杂原子的有机化合物结构自动解析专家系统。它由3个部分组成：数据分析，结构产生器，结构验证。结构产生器是系统的核心部分，它从分子出发，穷举生成结构基元集和结构片断集，最后对接成穷举的、非冗余候选结构。文中详细介绍了结构产生器的结构生成算法。

**关键词** 结构解析、专家系统、结构产生器

有机化合物结构自动解析是集计算机科学、光谱学、化学等诸种手段于一体的复杂过程。计算机自动结构解析的研究主要包括4个方面<sup>[1-3]</sup>：数据库，模式识别，波谱模拟，人工智能。有关结构解析方面的人工智能，如专家系统的研究是计算机化学的前沿课题。目前，已有的结构解析专家系统有：DENDRAL<sup>[4]</sup>，CASE<sup>[5]</sup>，CHEMICS<sup>[6,7]</sup>以及作者实验室中ESESOC(Expert System for the Elucidation of the Structures of Organic Compounds)等。ESESOC-I系统只能用于含C,H,O有机化合物结构解析<sup>[8,9]</sup>，现在一个能用于含多种杂原子的有机化合物结构自动解析系统ESESOC-II已组建成功。

ESESOC系统包括3个部分：光谱知识库及其解析系统<sup>[10,11]</sup>，结构产生器和结构验证系统。其结构解析过程为：从分子式或光谱数据出发推导出与之相一致的子结构约束条件，在其约束下穷举生成所有的结构异构体，即候选化合物，最后进行结构的认证并进行结构输出。

结构产生器是结构解析专家系统的核心部分。一个结构产生器应满足4个要求：(1)穷举性；(2)非冗余性；(3)有效性；(4)能接受各种约束条件(包括人机接口所提供的)的约束。一些著名的结构解析专家系统的结构产生器都有一个不断发展的过程。DENDRAL系统是一个最典型的例子，早期的结构产生算法“DENDRAL ALgorithm”<sup>[12]</sup>只能生成链式结构；引入结点图(vetex graphs<sup>[12]</sup>)的概念后，结构产生器STrGEen<sup>[12]</sup>能生成链式和环结构；为了处理

1993-08-25收稿，1993-12-11收修改稿。

\* 国家自然科学基金资助项目。

\*\* 联系人。

子结构约束条件,加入了超原子(superatom),则得到了约束生成的结构产生器 CONGEN<sup>[12]</sup>;为了处理<sup>13</sup>CNMR光谱数据解析所得到的原子间相互有重叠的子结构,又开发了一个新的结构产生器——GENOA<sup>[13]</sup>. CASE 系统也有两个结构产生器: ASSEMBLE<sup>[14]</sup>, COCOA<sup>[15]</sup>, 其中 COCOA 采用了和 ASSEMBLE 完全不同的结构对接策略 - 消去法(reduction). Sasaki 的 CHEMICS 系统的结构产生器也在不断发展<sup>[16,17]</sup>. 另外 Zupan<sup>[17]</sup>, Contreras<sup>[18]</sup> 等人都对结构对接算法作了大量的研究工作.

在作者的实验室中,王立等曾对链式结构的生成算法做了研究<sup>[19]</sup>, ESESOC-I 的结构产生器采用深度优先的结构生成树策略<sup>[9]</sup>. 而 ESESOC-II 的结构生成系统采用结合算法,由分子式穷举出结构基元集和结构片断集,然后采用深度优先的策略填充压缩的键属性连接矩阵的方法将结构片断对接生成整体结构,同时提出了新的冗余性消除策略.

## 1 由分子式穷举生成结构基元集及结构片断集

在 ESESOC 系统中,把有机化合物结构的最基本单位的元素组成称为结构基元. 如表 1 所示,含有 C,H,O,N,S,P,F,Cl,Br,I 等 10 种元素的有机化合物,总共有 18 种结构元素

本系统的结构片断是带有键属性的基本结构单位及其组合. ESESOC-II 选择了 99 个结构片断(见表 2),由此可以组成含上述 10 种元素的所有的有机化合物的结构异构体.

表 1 18 种结构基元

1	CH <sub>3</sub>	7	SH	13	PH
2	CH <sub>2</sub>	8	S	14	P
3	CH	9	NH <sub>2</sub>	15	F
4	C	10	NH	16	Cl
5	OH	11	N	17	Br
6	O	12	PH <sub>2</sub>	18	I

表 2 99 个结构片断

1 — CH <sub>2</sub> —	2 CH <sub>3</sub> CH <sub>2</sub> —	3 CH <sub>2</sub> =
4 >CH—	5 CH <sub>3</sub> CH<	6 (CH <sub>3</sub> ) <sub>2</sub> CH—
7 —CH=	8 CH <sub>3</sub> CH=	9 ArCH <sup>a)</sup>
10 —CH=O	11 —CH=S	12 —CH=NH
13 N <sup>⊕</sup> =N <sup>⊖</sup> =CH—	14 >C<	15 CH <sub>3</sub> C—   
16 (CH <sub>3</sub> ) <sub>2</sub> C<	17 (CH <sub>3</sub> ) <sub>3</sub> C—	18 >C=
19 CH <sub>3</sub> C=	20 (CH <sub>3</sub> ) <sub>2</sub> C=	21 >C=O
22 >C=S	23 >C=NH	24 >C(=N <sup>⊕</sup> =N <sup>⊖</sup> )
25 CH <sub>3</sub> C(=O)—	26 CH <sub>3</sub> C(=S)—	27 CH <sub>3</sub> C(=NH)—
28 CH <sub>3</sub> C(=N <sup>⊕</sup> =N <sup>⊖</sup> )—	29 ArC—	30 CH <sub>3</sub> ArC
31 ArC	32 —C≡	33 CH <sub>3</sub> —C≡
34 CH≡C—	35 N≡C—	36 N(→O)≡C—
37 =C=	38 O=N=	39 S=C=
40 NH=C=	41 N <sup>⊕</sup> =N <sup>⊖</sup> =C=	42 C <sup>⊕</sup> ≡N <sup>⊖</sup> =
43 HO—	44 —O—	45 CH <sub>3</sub> O—
46 ArO	47 NH <sub>2</sub> —	48 —NH—
49 CH <sub>3</sub> NH—	50 ArNH	51 >N—
52 CH <sub>3</sub> N<	53 (CH <sub>3</sub> ) <sub>2</sub> N—	54 —N=

a) Aromatic.

表 2(续)

55	CH <sub>3</sub> N=	56	O=N—	57	S=N—
58	(O $\leftarrow$ )S=N—	59	NH=N—	60	N $^{\ominus}$ =N $^{\oplus}$ =N—
61	ArN	62	>N( $\rightarrow$ O)—	63	CH <sub>3</sub> N( $\rightarrow$ O)<
64	(CH <sub>3</sub> ) <sub>2</sub> N( $\rightarrow$ O)—	65	—N( $\rightarrow$ O)=	66	CH <sub>3</sub> N( $\rightarrow$ O)=
67	O=N( $\rightarrow$ O)—	68	S=N( $\rightarrow$ O)—	69	NH=N( $\rightarrow$ O)—
70	N $^{\ominus}$ =N $^{\oplus}$ =N( $\rightarrow$ O)—	71	ArNO	72	HS—
73	—S—	74	CH <sub>3</sub> S—	75	ArS
76	—S( $\rightarrow$ O)—	77	CH <sub>3</sub> S( $\rightarrow$ O)—	78	ArS( $\rightarrow$ O)
79	—S( $\rightarrow$ O) <sub>2</sub> —	80	CH <sub>3</sub> S( $\rightarrow$ O) <sub>2</sub> —	81	PH <sub>2</sub> —
82	PH <sub>2</sub> ( $\leftrightarrow$ O)—	83	—PH—	84	CH <sub>3</sub> —PH—
85	—PH( $\rightarrow$ O)—	86	CH <sub>3</sub> PH( $\rightarrow$ O)—	87	>P—
88	CH <sub>3</sub> P<	89	(CH <sub>3</sub> ) <sub>2</sub> P—	90	>P( $\rightarrow$ O)—
91	CH <sub>3</sub> P( $\rightarrow$ O)<	92	(CH <sub>3</sub> ) <sub>2</sub> P( $\rightarrow$ O)—	93	>P( $\rightarrow$ S)—
94	CH <sub>3</sub> P( $\rightarrow$ S)<	95	(CH <sub>3</sub> ) <sub>2</sub> P( $\rightarrow$ S)—	96	F—
97	Cl—	98	Br—	99	I—

采用一些简单的组合数学方法,如整数的分解(Composition)及划分(partition)算法,即可由分子式穷举生成结构基元集,然后由结构基元集穷举生成结构片断集。如对于分子式 $C_6H_{10}N_2O_2$ ,穷举得到的结构基元集(见表3)。

表 3

结构基元集 1			结构基元集 2			结构基元集 3		
ID#	分子式	序数	ID#	分子式	序数	ID#	分子式	序数
1	CH <sub>3</sub>	1	1	CH <sub>3</sub>	1	1	CH <sub>3</sub>	1
2	CH <sub>2</sub>	1	2	CH <sub>2</sub>	1	3	CH	2
4	C	1	4	C	1	5	HO	1
5	HO	1	5	HO	1	9	NH <sub>2</sub>	1
9	NH <sub>2</sub>	1	10	NH	2	11	N	1
11	N	1						
	...			...			...	

得到的结构片断集(见表 4).

表 4

结构片断集 1			结构片断集 2			结构片断集 3		
ID#	分子式	序数	ID#	分子式	序数	ID#	分子式	序数
1	$-\text{CH}_2-$	1	1	$-\text{CH}_2-$	1	1	$-\text{CH}_2-$	1
14	$>\text{C}<$	1	14	$>\text{C}<$	1	15	$\text{CH}_3\text{C}<$	1
43	$\text{HO}-$	1	43	$\text{HO}-$	1	43	$\text{HO}-$	1
47	$\text{NH}_2-$	1	47	$\text{NH}_2-$	1	47	$\text{NH}_2-$	1
52	$\text{CH}_3\text{N}<$	1	55	$\text{CH}_3\text{N}=$	1	51	$- \text{N}<$	1
...			...			...		

## 2 由结构片断集穷举生成整体结构

### 2.1 键性联接矩阵

在 ESESOC-II 结构生成系统中, 使用键性联接矩阵来表示化合物的结构。结构 I 及其对应的联接矩阵见图 1。

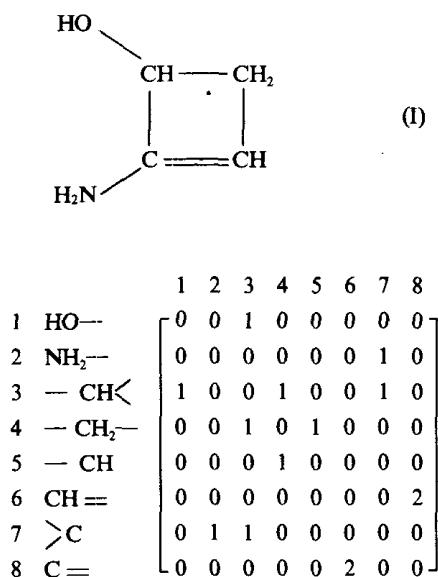


图 1 结构 I 及其联接矩阵

其中各矩阵元的意义为, 0: 不成键, 1: 单键, 2: 双键, 3: 三键, 9: 芳香键。

对键属性编号的规则为: (1)从不同原子上引出的键赋以不同的序号, 如 HO—, NH<sub>2</sub>— 上的单键分别赋以序号 1, 2; (2) 从同一原子引出的具有相同键属性的多个键, 赋予一个序号, 如—CH<上的三个单键赋以序号 3, —CH<sub>2</sub>—上的两个单键赋以序号 4; (3) 从同一个原子上引出的具有不同键属性的几个键, 则要分别编号, 如—CH= 有两个自由键, 一个单键, 一个双键, 属于不同的键属性, 则分别赋以序号为 5 CH—, 6 CH=, 同样 >C=的两个单键赋以序号 7 >C, 双键赋以序号 8 C=。

我们的键性联接矩阵同时具有 Munk<sup>[15]</sup> 和 Zupan<sup>[17]</sup> 的键性联接矩阵的优点, 即(1)由于只考虑结构片断的原子上的自由键, 因而可以任意加入其它结构片断, 如本系统没有选入的 >Si<, >B— 等和含有多个有自由键的原子的结构片断(随着分析手段的提高, 现代谱学方法已能提供较大的含较多原子的结构片断, 且这些结构片断上不止一个原子含自由键), 因而增加本系统的适用范围; (2)在结构对接中不用进行键属性匹配检测, 从而加快结构对接的速度, 而且大大地减小矩阵的维数, 以利于计算机的处理, 从下面可知, 它还可以大大缩小结构生成树的规模, 从而提高了结构对接的效率。

### 2.2 超结构的键性联接矩阵

由结构片断集进行结构对接时, 首先把结构片断集中的各结构片断及其自由键按上面的编号规则赋给一任意序号, 然后构造一个键性联接矩阵。开始时, 先把所有的矩阵元都填上

$b$ (bond, 任意键属性, 其值可以是 0, 1, 2, 3, 9 中的任一个), 后对其进行如下的处理: (1) 删去同一结构片断间的键; (2) 删去键属性不同的结构片断间的键; (3) 删去形成小分子的结构片断间的键; (4) 删去与子结构约束条件不一致的键, 这样得到的联接矩阵即为超结构联接矩阵, 它包含了该结构片断集可以产生的所有的结构异构体。

如由分子式  $C_3H_8N_2O$  穷举得到的某一结构片断集为

$$\begin{array}{l} \text{(结构片断: } -CH_2- = CH- HO- -N= -NH_2\text{)} \\ \text{(个数: } 2 \quad 1 \quad 1 \quad 1 \quad 1\text{)} \end{array}$$

若假定子结构约束条件为, 分子中没有  $HO-N=$  和  $HO-CH=$  这两个子结构, 则其超结构联接矩阵的生成过程如图 2 所示。

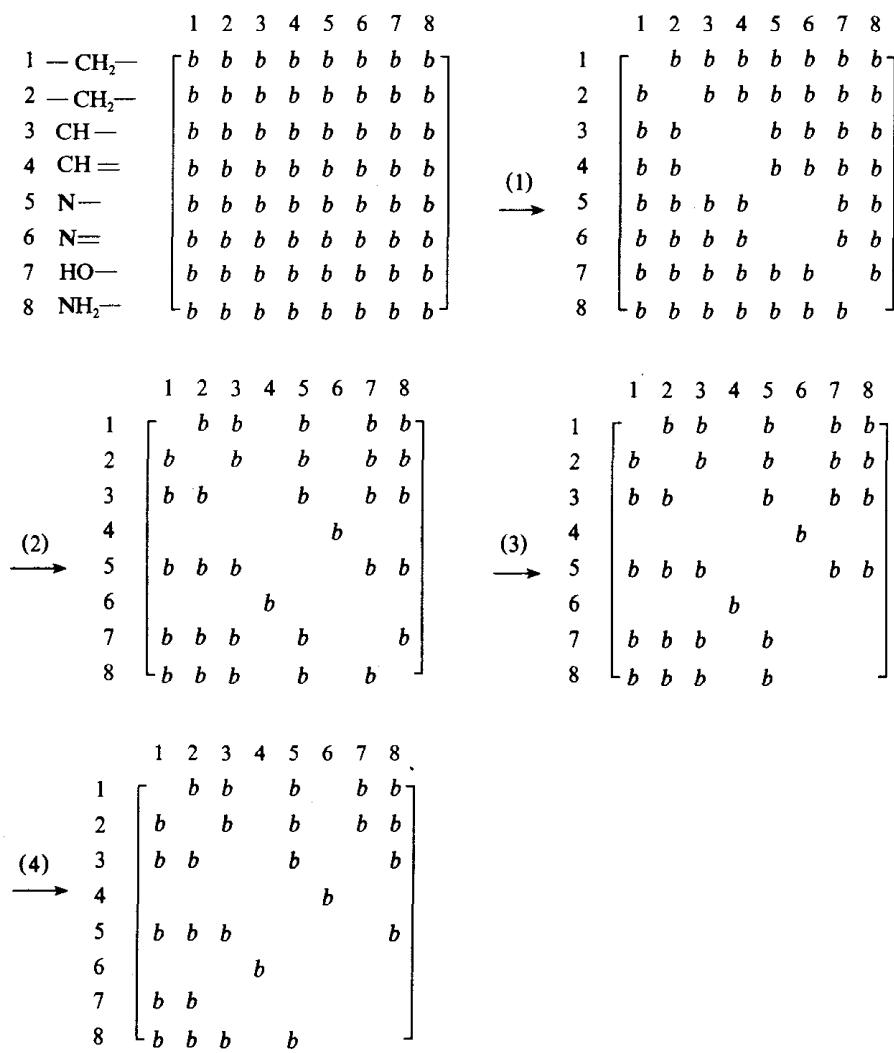


图 2 超结构联接矩阵的构造过程示意图

(1) 删去同一结构片断间的键, (2) 删去键属性不同的结构片断间的键,

(3) 删去形成小分子的结构片断间的键, (4) 删去与子结构约束条件

不一致的键

### 2.3 结构对接

在超结构键属性联接矩阵中, 其矩阵元  $b$  的值可以是它所对应的行或列上的结构片断键属性所对应的键, 也可以是 0。把超结构键属性联接矩阵中的矩阵元赋以一组确定的适当的数值, 然后补上 0(因为超结构键属性联接矩阵中已被删去的矩阵元都应是 0, 即不能成键), 即可得到一个键属性联接矩阵及其对应的结构。如对上述得到的超结构键属性联接矩阵的矩阵元赋以下面的一组确定的数值,

	1	2	3	4	5	6	7	8
1 —CH <sub>2</sub> —	1	0	0	1	0			
2 —CH <sub>2</sub> —	1	1	0	0	0			
3 CH—	0	1	0	0	0			
4 CH=				2				
5 N—	0	0	0			1		
6 N=				2				
7 HO—	1	0						
8 NH <sub>2</sub> —	0	0	0	1				

补上 0 后得到了如下的键属性联接矩阵, 其对应的结构为结构 II

	1	2	3	4	5	6	7	8
1 —CH <sub>2</sub> —	0	1	0	0	0	0	1	0
2 —CH <sub>2</sub> —	1	0	1	0	0	0	0	0
3 CH—	0	1	0	0	0	0	0	0
4 CH=	0	0	0	0	0	2	0	0
5 N—	0	0	0	0	0	0	0	1
6 N=	0	0	0	0	2	0	0	0
7 HO—	1	0	0	0	0	0	0	0
8 NH <sub>2</sub> —	0	0	0	0	1	0	0	0

HO—CH<sub>2</sub>—CH<sub>2</sub>—CH=N—NH<sub>2</sub>  
(II)

结构对接就是要将包含在超结构键属性联接矩阵中的所有异构体都穷举出来, 即, 要求出所有可能的矩阵元的能够表示不同结构的各组数值, 如对图 2 中的超结构键属性联接矩阵, 上面的一组数值即是其中之一。

因为联接矩阵是对称的, 因而可以略去下三角阵, 仅讨论上三角阵。对上例中的超结构键属性联接矩阵的上三角阵中的矩阵元按先列后行的自然顺序进行编号, 结果如下:

	1	2	3	4	5	6	7	8
1	1	2	4	8	10			
2		3	5	9	11			
3			6		12			
4				7				
5					13			
6								
7								
8								

把上面已编号的各矩阵元按其编号顺序依次填上 0, 或对应的行或列上的结构片断键属性所对应的键的数值, 考虑所有的组合, 即可穷举出该超结构键属性联接矩阵所包含的所有

的结构异构体。由图论可知,矩阵元填充的过程实际上是一个二叉树生成过程。本文采用深度优先的算法进行二叉树的遍历,上述例子的结构生成树见图 3。图中,○的表示不成键,即在超结构键属性联接矩阵中的该矩阵元为 0, ① 的表示成单键,即在超结构键属性联接矩阵中的该矩阵元为 1, ② 的表示成双键,即在超结构键属性联接矩阵中的该矩阵元为 2。

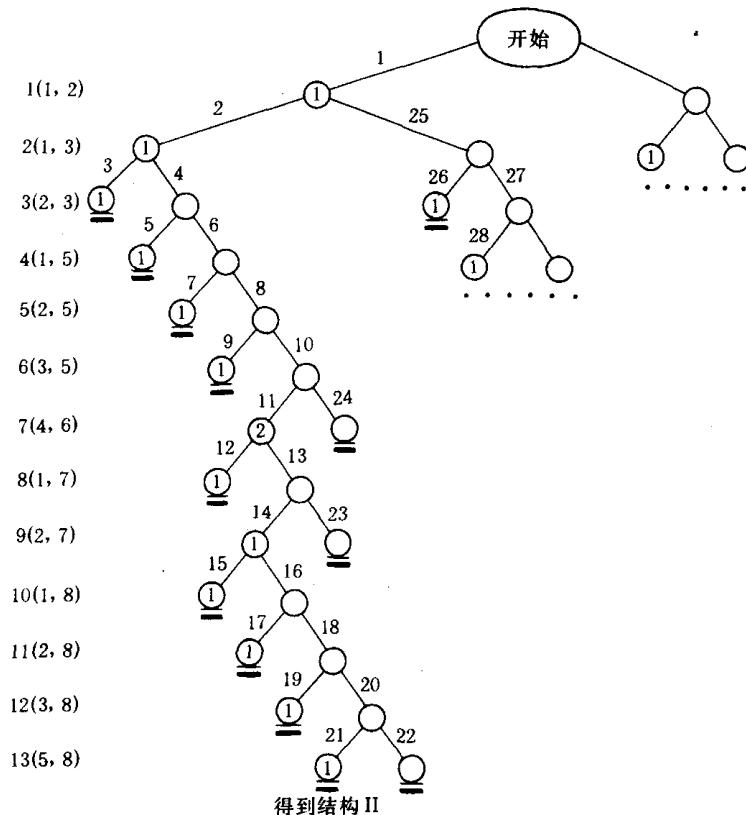


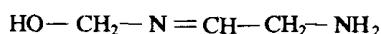
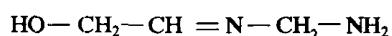
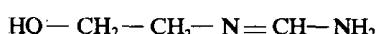
图 3 结构生成树的示意图  
(—表示此路径不能继续向前遍历)

在结构生成树的遍历过程中,要随时进行连接性、冗余性、子结构约束条件等检测,以避免无效对接,即尽早消除结构生树的无效的分枝。

对上例,从 1 号矩阵元开始,该矩阵元对应 1, 2 两个结构片断间的连接关系,这两个结构片断之间可成一个单键,可以不成键,即 1 号矩阵元可以填上 1 或 0, 选填 1 为左子树,按深度优先的规则,结构生成树由此进行下去,填 0 的情况将在以后的过程中讨论。与第 1 步相同,第 2 步的结果是在 2 号矩阵元上填 1; 第 1,2 两步结果得到  $\text{CH}_2 - \text{CH}_2 - \text{CH}_3$ 。现在看 3 号矩阵元,它对应 2, 3 两个结构片断间的连接关系,在第 1,2 步后 3 号结构片断已无自由单键了,因此 2, 3 之间不能再成单键了,由连接性检测结果可知 3 号矩阵元已不能填 1, 即该左子树(第 3 步)为一终结点,到此为止,因而回溯到第 2 步的结果,从右子树(第 4 步)进行下去,即 3 号矩阵元填 0, 以后分析与前面相同,故不在此赘述。进行到第 21 步时得到了本节前面提到的结构 II。对于第 23 步,由于到此为止 7 号结构片断还没有参与成键,而在后面 10, 11, 12, 13 号矩阵元都没有涉及 7 参加成键,所以即使进行下去也不能生成整体结构,因而由联接

性检测结果应舍去结构生成树的此分枝。同样第 24 步也应舍去。现在看第 26 的结果，它将得到子结构— $\text{CH}_2-\text{CH}_2-\text{CH}=\text{}$ 与第 2 步以后得到的子结构— $\text{CH}_2-\text{CH}_2-\text{CH}=\text{}$ 只是在 C 原子的编号上不同，又因为编号是任意的，所以这两个子结构是等价的。即该步对接与第 2 步对接相冗余，应舍去。

由该结构生成树得到的其它结构为：



## 2.4 关于冗余性问题

关于消除冗余性问题，大部分系统都采用子结构检索的方法，Sasaki 采用 canonical stack 的方法<sup>[16]</sup>，但这些方法都比较烦锁。作者发现：当第  $i$  个结构片断和第  $j$  个结构片断对接时，若它们都不存在拓扑等价性，则不会与其它各步对接相冗余；若有等价性存在，则一定有冗余对接产生。避免冗余对接的方法为，若  $i$  和  $j$  均为它们各自等价类中编号最小的一个则保留，否则舍去。如本例中— $\text{CH}_2-\text{CH}_2-$ 的 1 和 2 是等价类，3 没有等价类，取 1,2 中最小的 1 与 3 对接（第 2 步），而舍弃 2 与 3 的对接（第 25 步），由此避免了冗余结构的产生。

我们建议了一个有机化合物结构图的新的不变量，它体现了化合物结构中各节点属性和其在结构图中环境的属性，从而表征了亚结构的拓扑等价性<sup>[20]</sup>。

## 3 结果与讨论

ESESOC-II 的结构生成系统采用许多有效的方法，如组合数学方法，键属性联接矩阵，二叉树遍历，新的消除冗余策略等，且采用分步进行的方法，可随时接受约束条件的约束，这样可大大地提高结构生成的效率。

一个结构产生器必须是穷举的和非冗余的，为了验证本系统的穷举性，我们从分子式出发，在无约束条件下进行异构体的穷举生成，如本文中例子的分子式  $\text{C}_3\text{H}_8\text{N}_2\text{O}$ ，在无约束条件下，可生成 527 个异构体，这与 DENDRAL<sup>[21]</sup> 系统的生成结果相一致（DENDRAL 系统的结构生成算法已经过数学证明是穷举的）。从其它分子式进行异构体的穷举生成的结果列于表 5 中。表中，饱和烷烃的异构体的数目与图论理论计算结果一致，其它分子式的异构体数目与 DENDRAL, CHEMICS<sup>[17]</sup> 等系统的生成结果一致；对含 N, S 等原子的化合物，由于本系统考虑了含五价氮和四价、六价硫的官能团，而 DENDRAL 系统没有考虑这些基团，因而异构体总数不一致，但本系统若扣除含五价氮和四价、六价硫的基团的异构体后与 DENDRAL 系

表 5 ESESOC 系统由分子式穷举所得的异构体数目

分子式 异构体数目	$\text{C}_3\text{H}_8$ 1	$\text{C}_4\text{H}_{10}$ 2	$\text{C}_5\text{H}_{12}$ 3	$\text{C}_8\text{H}_{20}$ 18	$\text{C}_{12}\text{H}_{26}$ 355	$\text{C}_{16}\text{H}_{34}$ 10359	$\text{C}_{20}\text{H}_{42}$ 366319
分子式 异构体数目	$\text{C}_{10}\text{H}_{22}$ 75	$\text{C}_{10}\text{H}_{20}$ 852	$\text{C}_{10}\text{H}_{18}$ 5495	$\text{C}_{10}\text{H}_{22}\text{O}$ 5989	$\text{C}_{10}\text{H}_{20}\text{O}$ 13373	$\text{C}_{10}\text{H}_{18}\text{O}$ 90769	$\text{C}_3\text{H}_{10}\text{N}_2\text{O}$ 102 <sup>a)</sup>
分子式 异构体数目	$\text{C}_6\text{H}_{14}$ 5	$\text{C}_6\text{H}_{12}$ 25	$\text{C}_6\text{H}_{10}$ 77	$\text{C}_6\text{H}_8$ 159	$\text{C}_6\text{H}_6$ 217	$\text{C}_6\text{H}_4$ 185	$\text{C}_6\text{H}_2$ 85

a) 为不含五价氮的异构体数目。

统的结果完全一致。因而本系统的结构生成是穷举和非冗余的。运行结果表明，本系统的结构生成算法效率高，结果良好。

### 参 考 文 献

- [1] 许 禄, 分析化学, 1990, **18**: 180.
- [2] 许 禄、郭传杰, 计算机化学方法及应用, 化学工业出版社, 北京, 1990, 218—240.
- [3] 许 禄, 化学计量学方法, 科学出版社, 北京, 1993.
- [4] Gray, N. A. B., *Chemom. Intell. Lab. Sys.*, 1988, (5): 11.
- [5] Munk, M. E., Christie, B. D., *Anal. Chim. Acta*, 1989, **216**: 57.
- [6] Funatsu, K., Miyabayashi, N., Sasaki, S., *J. Chem. Inf. Comput. Sci.*, 1988, **28**: 18.
- [7] Abe, H., Okuyama, T. et al., *J. Chem. Inf. Comput. Sci.*, 1984, **24**: 220.
- [8] 胡昌玉、许 禄, 分析化学, 1992, **20**: 643.
- [9] 胡昌玉、许 禄, 有机化学, 1992, **12**(6): 129.
- [10] 胡昌玉、许 禄, 波谱学杂志, 1991, (8): 347.
- [11] Zhang, J. G., Xu, L., *Anal. Chim. Acta*, 1988, **210**: 163.
- [12] Lindsay, R. K., Buchanan, B. G., Fengenbaum, E. A. et al., *Applications of Artificial Intelligence for Organic Chemistry, the DENDRAL Project*, McGraw-Hill, New York, 1980.
- [13] Carhart, R. E., Smith, D. H., Gray, N. A. B. et al., *J. Org. Chem.*, 1981, **40**: 1708.
- [14] Shelley, C. A., Hays, T. R., Munk, M. E. et al., *Anal. Chim. Acta*, 1978, **103**: 121.
- [15] Christie, B. D., Munk, M. E., *J. Chem. Inf. Comput. Sci.*, 1988, **28**: 271.
- [16] Kudo, Y., Sasaki, S., *J. Chem. Inf. Comput. Sci.*, 1976, **16**: 43.
- [17] Bohanec, S., Zupan, J., *J. Chem. Inf. Comput. Sci.*, 1991, **31**: 531.
- [18] Contreras, M. L., Valdivia, R., Rozas, R., *J. Chem. Inf. Comput. Sci.*, 1992, **32**: 323.
- [19] 王 立、许 禄, 计算机与应用化学, 1992, (4): 247.
- [20] Hu, C. Y., Xu, L., *Anal. Chim. Acta*, 1994 (in press).
- [21] Smith, D. H., *J. Chem. Inf. Comput. Sci.*, 1975, **15**: 203.