**扩展的32结构向量SOL表达式**

**摘要**：在原有SOL及陈辉扩展后的29个结构向量的基础上，又增加了3个结构，用以进一步区分一些原有SOL无法表达的分子。本文对增加的SOL结构及其与原有结构的关系进行说明。

**关键词**：SOL；扩展；32结构向量；

一、扩充前的基础SOL

原有的24个结构向量如下：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| A6 | A4 | A2 | N6 | N5 | N4 | N3 | N2 | N1 | R | br | me | H | AA | S | RS | AN | NN | RN | O | RO | O= | Ni | V |

各结构向量所表达的含义可参见以下文献：

*——Quann R J, Jaffe S B. Structure-oriented lumping: describing the chemistry of complex hydrocarbon mixtures[J]. Industrial & Engineering Chemistry Research, 1992, 31(11):2483-2497.*

*——邱彤, 陈金财, 方舟. 基于结构导向集总的石油馏分分子重构模型[J]. 清华大学学报(自然科学版), 2016(4):424-429.*

由于24 SOL中能够描述支链的只有br和me，无法描述分子中存在的乙基丙基，也无法清晰说明这些支链是位于链上还是环上。陈辉在此基础上增加了5个结构：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| index | et-sc | et-ring | pr-ring | bu-ring |

index：用以区分芳环上取代基的位置，1表示邻位，2表示间位，3表示对位；

et-sc：碳链上的乙基数量；

et-ring：环上的乙基数量，包括环烷烃及芳环；

pr-ring：环上的丙基数量，包括环烷烃及芳环；

bu-ring：环上的丁基数量，包括环烷烃及芳环；

以上向量仅是对分子结构的补充描述，对分子的元素计量数不会产生影响，各取代基的C数均已包含在原有的R结构向量里。

二、多核分子

上述的29 SOL表达式很好地描述出分子中存在的乙基丙基等基团的数量及位置，但是后续发现，对分子中双键的位置、环丙烷、环丁烷等29 SOL仍然无法描述。尤其当分子的C碳较大时，会出现一些之前未考虑的新的结构，例如1,4-二苯基丁烷等，见图1。

C:\Users\Administrator\Desktop\1083-56-3.png

图1 1,4-二苯基丁烷

在原有的24 SOL体系中，存在着多核分子的概念，用来表达复杂分子。简而言之，就是先将分子划分为多个核心组成部分，每个核心均用一行24 SOL表示，若分子可以划分为n个核心，即可以得到一个n行24列的矩阵来表示该分子，并调整R与AA的值来描述这些核心之间的具体连接方式。接下来经过一定的变换，并将同一列的n行元素全部陈列到一起，得到一个新的1行24列的矩阵，其形式与原有24 SOL一致。新的矩阵为一个混合矩阵，每一列中的元素均为一个向量，向量则由多核分子相应列中的n个数值共同构成。

例如，图1中的1,4-二苯基丁烷可以如图2划分。

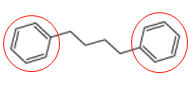


图2 多核分子划分

即该分子可以通过2行24列的矩阵来表示，其中，省略的结构向量均为0：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| A6 | …… | R | …… | AA | …… |
| 1 | …… | 0 | …… | 200020000 | …… |
| 1 | …… | 40004 | …… | 100020000 | …… |

进一步转换成1行24列SOL：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| A6 | …… | R | …… | AA | …… |
| 1 1 | …… | 0 40004 | …… | 200020000 100020000 | …… |

R及AA的具体含义可参见以下文献：

*——Jaffe S B, And H F, Olmstead W N. Extension of Structure-Oriented Lumping to Vacuum Residua[J]. Industrial & Engineering Chemistry Research, 2005, 44(26):págs. 9840-9852.*

通过多核分子的表达方法，可以写出绝大多数分子的SOL表达式。但是，文献中的多核分子旨在对复杂渣油分子进行描述，像图2这一类相对简单的分子，若都用多核分子进行表示，会增加后续数据处理及建模的复杂性。因此，优先考虑对已有的29 SOL进行扩展，暂不引入多核分子的方法。

三、扩充的32 SOL

在29 SOL的基础上，再增加3个结构向量：

|  |  |  |
| --- | --- | --- |
| 碳链上的双键数量 | 构成脂肪环的碳数 | 两环之间的碳数 |
| db-sc | c-ring | bt-ring |

各结构向量的相关说明如下：

**db-sc：碳链上的双键数量**

双键在分子中的位置可能在脂肪环上，也可能在碳链上：

当R-c\_ring=0时，即不存在碳链，若存在双键则会位于脂肪环上，故db-sc=0；

当N6+N5+c\_ring=0时，即不存在脂肪环，若存在双键则肯定会位于碳链上，故也无需用db-sc表示，db-sc=0；

当R-c\_ring≠0且N6+N5+c\_ring≠0，db-sc才有意义；

【对分子化学计量数的影响】

分子的双键数（缺氢度）由H结构向量决定，bt-sc仅用于描述双键的位置，对分子的化学计量数没有影响；

**c-ring：构成脂肪环的碳数**

原SOL体系可以用N6=1、N5=1来表示环己烷、环戊烷，现增加c-ring来表示环丙烷、环丁烷等。例如环丙烷完整的32 SOL表达式为，其中省略的结构向量均为0：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| …… | R | …… | c-ring | …… |
| …… | 3 | …… | 3 | …… |

规定c-ring≠5,6且c-ring≥3；

只用于描述分子除了五元环、六元环之外，另外形成的单个脂肪环，若分子中除了N6、N5还存在多个脂肪环，建议用多核分子的表达形式；

【对分子化学计量数的影响】

c-ring的碳数已包含在R结构向量中，对原有的C含量没有影响，对其他元素的含量也没有影响；

**bt-ring：两环之间的碳数**

bt-ring表示两环之间的所有碳数，环可以是脂肪环或芳环，碳数包括直接连接两环的碳链及其上的支链的所有C原子。例如图3的分子，其bt-ring=6。

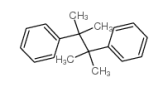


图3 2,3-二甲基-2,3-二苯基丁烷

因为bt-ring=1时，原SOL体系可以用N1=1来表示，故此处规定bt-ring≠1；

只能用于描述两个环之间的碳数，若有三个以上的环通过碳链连接在一起，建议采用多核分子的表达形式；

【对分子化学计量数的影响】

bt-ring的碳数已包含在R结构向量中，对原有的C含量没有影响；

当bt-ring≠0时，H数量-2；

**其他调整解释【尚未完成】**

为了利用已有结构向量描述更多的分子，做出如下的规定与调整：

1.参考原SOL体系中，当芳环上的R=1时，me忽略不计数的做法，规定

……

若环上或链上的支链太长，则用现有的最长的支链向量来代替。例如，当环上存在戊基以上的侧链需要表示时，用bu-ring代替；当链上存在丙基以上的侧链需要表示时，用et-sc代替。

五元环、萘类分子只有邻间异构，故其index值只能为1或2。

r-ring有S时也跟R一致

et也可以是有双键

当bt-ring看直接连C数量