



Title	一般式化学構造のコンピュータによる蓄積と検索
Author(s)	時實, 象一
Citation	大阪大学, 1987, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/35538
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、大阪大学の博士論文についてをご参照ください。

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

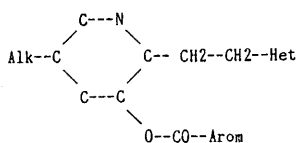
The University of Osaka

氏名・(本籍)	とき	さね	そう	いち
学位の種類	理	学	博	士
学位記番号	第	7590	号	
学位授与の日付	昭和62年3月18日			
学位授与の要件	学位規則第5条第2項該当			
学位論文題目	一般式化学構造のコンピュータによる蓄積と検索			
論文審査委員	(主査)			
	教授 千原 秀昭			
	(副査)			
	教授 京極 好正	教授 桑田 敬治	教授 佐々木慎一	

論文内容の要旨

コンピュータによる化学構造の蓄積と検索は既に実用化され、数百万件の化学構造がオンラインで検索できるようになっている。しかし、特許文献等で多用される一般式を含む化学構造型式(図1)については、フラグメント・コードを用いる古典的な方法を除いては、いまだにコンピュータによる取扱い手法が見出されていない。

一般式構造を含め、化学構造をコンピュータ上で正確に取扱うには結合表が最適である。一般式を表現するため結合表の各原子(ノード)にアトリビュートを付与することとした。アトリビュートは環、鎖、原子アトリビュートにわかれ、それぞれ化学的特徴を表わす複数の要素(0または1の値を持つ)からなるビットマップベクトルである。一般式が範囲を持つ(C3-C5など)場合は要素を「以上」



Alk = Alkyl groups with C1-C5 chain. {2, 5}
 Arom = Aromatic rings (i. e., benzene or naphthalene) {4}
 Het = Heterocyclic (i. e., rings containing one or more non-carbon atoms) groups with ring size of 5 or 6 {3, 6}

図1 一般式構造の例

FILE STRUCTURE		QUERY STRUCTURE	
STRUCTURE	1 2 3 Ph-(CH ₂) _n -OH n = 3-5	a b c Ph-(C-C) _n -O n = 2-4	
ATTRIBUTE	2 = 3-5 carbon chain	b = 2-4 carbon chain	
Bit Position	Description	C2 C3 C4 C5	File Str. (OR)
1	Chain Length		
2	1 or more	1 1 1 1	1 1
3	2 or more	0 1 1 1	1 1
4	3 or more	0 0 1 1	1 0
5	4 or more	0 0 0 1	1 0
6	1 or less	0 0 0 0	0 0
7	2 or less	1 0 0 0	0 0
8	3 or less	1 1 0 0	1 0
9	4 or less	1 1 1 0	1 1
10	5 or less	1 1 1 1	1 1

図2 アトリビュートの例

と「以下」に分割し、かつファイル中の構造式についてはベクトルのOR結合、質問構造式についてはAND結合を取ることで、部分的（オーバーラップ）一致が検出できることがわかった。（図2）

一般式ノードは原子グループであるから、複数のノードに対応する可能性がある。このような一対多の対応関係を含む原子結合検索法は知られていない。そこで結合スタック法を拡張することで、このような検索を可能とした。この方法では、質問構造式の一般式ノードを含む各ノードにファイル構造のノードを順次対応させていく。各対応において元素、結合関係、およびアトリビュートの一致が検査される。一般式ノードの場合は複数のノードに対応する。このようにして一般式を含む化学構造の検索（図3のような）が可能となった。

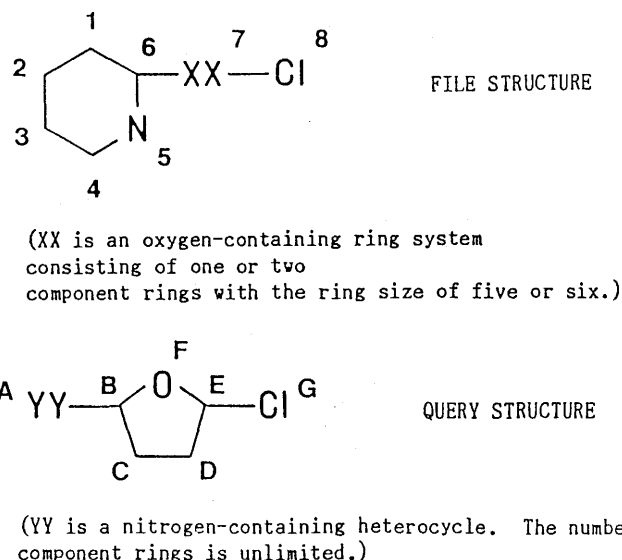


図3 一般式構造検索の例

論文の審査結果の要旨

時實象一君の論文は、一般式を用いて表現される有機化合物の構造式をコンピュータの記憶部に蓄積する方法と、そのようにして蓄積されたファイルから、特定の構造または一般式表現の構造による質問式に対して真となるレコードを検索する新しい方法に関するものである。

特定構造式のファイル化とこれを特定構造式で検索する方法は既にいろいろな方式が提案され、そのうちいくつかは実用化も行われている。しかし、化学の研究論文や特許明細書等で多用されている一般式表現のファイル化と検索についてはこれまで確立された方法がない。時實君はこの問題を構造式上のノードにアトリビュートを付与し、それをビットマップベクトルで表わすことにより、結合表をいたずらに拡大することなく効率よくファイル化できることを示した。このファイルから検索するとき、ある特定の構造ないし、部分構造を質問としても、それと厳密にヒットする構造はもちろん、その質問を含む一般式レコードをもヒットとすること、また反対に一般式表現の質問構造の場合ファイル中から特定構造とともにその一般式を包含する一般式構造をも検索すること、この両者を満足する検索方式として、総合スタック法の拡張を考案した。この方法では、質問構造式の各ノード（原子または原子群）に、ファ

イル中の一つずつの構造のノードとの対応をしらべ、その対応関係が元素の種類、結合数と次数、およびアトリビュートが満足されるレコードを選択していく。

この蓄積・検索方法は、これまで困難とされていた一般式（Markush式を含む）のコンピュータ処理に画期的な進歩をもたらすものであり、化学情報学の発展ひいては化学の研究そのものにも大きな貢献であり、理学博士の学位論文として十分価値あるものと認める。