

有機合成デザイン支援システムの試作

鳥飼 孝幸*・神 和則**・三浦 一男***
 金岡 泰保****・福永 公寿*****・富田 真吾*****

An Approach to the Design of Computer-Assisted Organic Synthesis

Takayuki TORIKAI, Kazunori KOH, Kazuo MIURA, Taiho KANAOKA
 Kimitoshi FUKUNAGA and Shingo TOMITA

Abstract

Recently, researches concerning the computer-assisted organic synthesis design are attracted. Constructing the system is very interesting problem as the application of knowledge engineering. Since chemist Corey has been developed the OCSS and LHASA as such system, several systems have been developed by many researchers. However, the details of the systems can not be opened to the public.

The authors try to development of the original organic synthesis design supporting system based on LHASA and present the fundamental framework of the system. This system runs on ACOS-6 through the use of color graphic display(NEC N6940). Each module is written in proper language, LISP and FORTRAN.

1. まえがき

近年、計算機科学のハード・ソフト両面に互る急速な進展に伴い、多数の専門的な知識または経験則を組織的に蓄えたコンピュータと対話しながら、専門家の発想を刺激しようとする、いわゆる知識ベースシステム¹⁾が新しいコンピュータ利用法として医療、化学^{2)~6)}など種々の分野で注目されている。その応用分野の一つとして、豊かな経験と発想、および莫大な知識が要求される有機合成デザインがある。化学者E. J. Corey等はこの分野にコンピュータ利用を積極的に推進し、知識ベースシステム的手法を利用してプロスタグランジンの合成経路探索に画期的な成果を収めている。有機合成デザイン支援システムとしてOCSS(1969年)³⁾や、LHASA(1972年)^{4~6)}をCorey等が発表して以来、それらの改良あるいは応用的な種々のシステム⁷⁾が発表されているが、化学情報や経験則のコンピュー

タ内部表現形式、パターン認識および化学構造式の出力部分などについてはほとんど述べられていない。

筆者等は独自の有機合成デザイン支援システムの開発を試みており、本論文はその報告を行うものである。実用化に至るには多くの問題を残しているが、システムの基本的な枠組を与えることが本論文の主な目的である。本システムは図形データ変換、識別、戦略制御、対話および変換の各モジュールで構成されているが、入出力モジュールおよび対話モジュールの作成には、コンピュータ操作に不慣れた化学者にとっても容易に操作できるようにすることを心がけた。又、化学構造式のコンピュータ内部表現形式として戦略制御モジュール等の後の処理に都合のよい、従来とは異なる結合表形式を考案した。更に、目標分子内における環や官能基の認識および戦略箇所の探索には部分マッチング法を提案している。なお、本システムはTSSカラーグラフィックディスプレイ端末(NECN6940)を介してACOS-6上で稼働する。又、使用言語は、図形モジュールおよびディスプレイ表示では、FORTRAN、その他のモジュールではLISPである。

*九州電気製造(株)

**九州日本電気ソフトウェア(株)

***松下電器産業(株)

****電子工学科 *****化学工学科

2. システムの概要

本システムは、実際の有機化合物の合成とは逆方向の解析を行っている。その解析方向を合成と対応させて逆合成と呼び、又、逆合成で導出された有機化合物を前駆物質という。目標分子をシステムに入力すると前駆物質が出力され、更に、その前駆物質を目標分子とすることによって新たな前駆物質が出力される。このような操作を繰り返すことによって、Fig.1に示すような合成の木を構成することができる。本論文では、合成の木を構成するための基礎となる1ステップの合成操作について論じている。

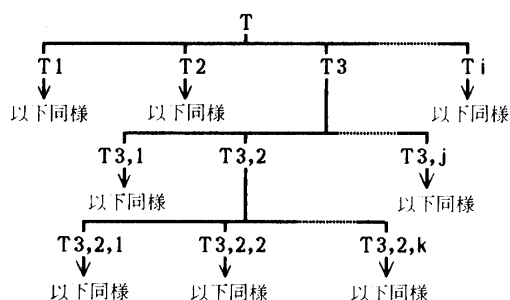


Fig. 1 Tree of synthesis.
(T: Target molecule)

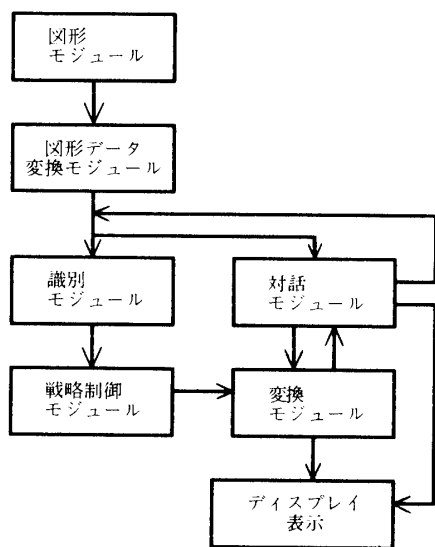


Fig. 2 Flow chart of system.

システムの概略をFig.2に示す。目標分子の化学構造式の入力は、図形モジュールを用いて、ライトペンでカラーグラフィックディスプレイ (CRT) 上を操作することでなされる。図形データ変換モジュールは、入力された目標分子の化学構造式をコンピュータ内部表現である結合表に変換する。識別モジュールは、結合表から官能基、環などの後の処理で必要となる箇所を認識する。それに基づいて、戦略制御モジュールは逆合成反応を適用し戦略箇所を決定する。変換モジュールはその戦略に従い結合表を操作する。ディスプレイ表示は、操作された結合表の化学構造式をCRT上に表示する。対話モジュールは、化学者自身が戦略を行いたい場合や、目標分子を修正したい場合に用いるモジュールである。以下、各モジュールについて詳述しよう。

3. 図形モジュール

このモジュールは、有機化合物を逆合成する際に標的となる目標分子の化学構造式を入力する部分である (Fig.3参照)。コンピュータの操作を容易にするために、図形モジュールにおける全ての入力はライトペンで行うようにした。ライトペンのペン先をCRT上の任意の点に当てることにより、化学構造式における原子の位置の指定や、特定のコマンドの選択によって、種々の操作を行うことができる。本モジュールにおいて、化学構造式を描写する際に用いる記号としては、まず、原子を表示する“C”, “H”, “N”, “O”, “S”, “P” および、ハロゲン元素を示す“X”がある。又、電荷を表示する“+”, “-”およびラジカル (不対電子) の表示は“!”を用いている。

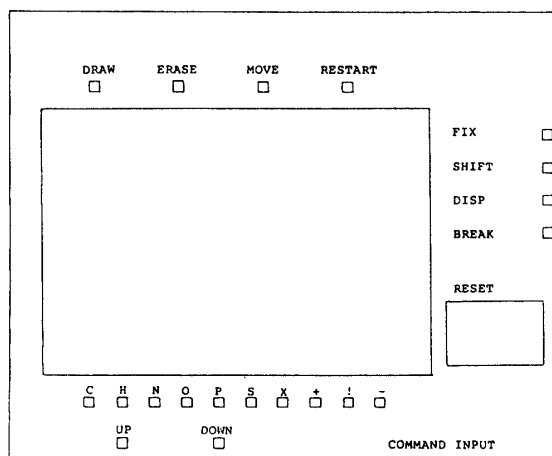


Fig. 3 Display

化学結合として単結合，二重結合，三重結合および立体化学的配置（破線）が表示可能である。

化学構造式はグラフとして表示され，節は原子を，又，節と節を結ぶ枝は化学結合を示す．無表示の節は炭素原子を表し，その他の原子は適当な記号を節上を書くことにより表される．節の末端が無表示のときは水素原子を表すことにした．目標分子の入力部分の起動方法，各コマンド等の詳細は文献(8)を参照されたし．

4. 図形データ変換モジュール

識別モジュール，戦略制御モジュール等の後の操作を可能にするためには，化学構造式の適当な内部表現への変換が必要となる．その変換を行うのが図形データ変換モジュールである．

本論文では，内部表現形式として以下に述べるような結合表形式を考案した．

4.1 結合表

結合表とは，化学構造式の構成原子とその結合関係を数学的手法で表現するものである．

化学構造式の表現形式としてMORGANの結合表⁹⁾がよく知られているが，化学構造式との対応を一意にすることを意図しているために，計算機処理時間や記憶容量に難点があり，更に，原子の付加や削除などに伴う結合表の変更手順が複雑になるという問題がある．

本システムでは，化学構造式と結合表との一対一対応の必要性はない．従って，次に述べるような作成および変更が容易な結合表形式を考案した．まず結合表を定義する．

〔定義〕

原子数 m の化学構造式 CS (Chemical structure) が与えられて，各原子の位置（以後節という）に無作為に 1 から m まで番号付けがなされているものとする．このとき CS の結合表 CT (Connection Table) は次の 3 項組で与えられる．

$$CT = (A, B, C)$$

$$A = (a_1, a_2, \dots, a_m)$$

但し， a_i ($1 \leq i \leq m$) は節番号 i の元素記号である．

$$B = (B_1, B_2, \dots, B_m)$$

但し，各 i ($1 \leq i \leq m$) に対して，

$$B_i = \{i_x | i_x \text{ は節番号 } i \text{ と直接結合している節番号}\}$$

$$C = (C_1, C_2, \dots, C_m)$$

但し， $B_i = \{i_1, i_2, \dots, i_{k_i}\}$ のとき，

$$C_i = (C_{i1}1, C_{i2}2, \dots, C_{ik_i}k_i)$$

C_{ij} ($1 \leq j \leq K$) は節番号 i_j と j との結合多重度である．

すなわち， A は節番号の小さい順に，その節の元素記号を列挙した組である． B の i 番目の要素は節番号 i から出ている枝と結合している節の節番号の集合である． C の i 番目の要素は節番号 i から出ている各枝の結合の多重度を表している．例題を Fig. 4 に示す．

<例1>

$$\begin{pmatrix} (C) & (C) & (N) & (N) & (O) & (O) & (C) & (C) \\ \{(2\ 3\ 4)\} & \{(1\ 5\ 6)\} & \{(1\ 7)\} & \{(1\ 8)\} & \{(2)\} & \{(2)\} & \{(3\ 8)\} & \{(4\ 7)\} \\ \{(1\ 1\ 2)\} & \{(1\ 1\ 2)\} & \{(1\ 1)\} & \{(2\ 1)\} & \{(1)\} & \{(2)\} & \{(1\ 1)\} & \{(1\ 1)\} \end{pmatrix}$$

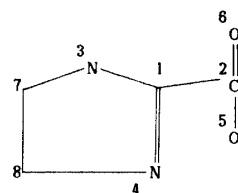


Fig. 4 Chemical formula and connection table.

ここで提案した結合表は原子記号を単に記号に置き換えて得られる， $R_1-C-N=R_2$ のような構造式に対しても有効であり，以後このような構造式も化学構造式と呼ぶことにする．

5. 識別モジュール

識別モジュールは，いわばこのシステムの眼に相当する部分であって，結合表が作成されると，これに基づいて戦略制御モジュールの作動に必要な構造概念を導き出す．現段階では環と Fig. 5 に示す官能基の認識が可能である．官能基は認識に加えて，電子的性質に基づく分類も行われ，その結果としてそれぞれの基について“一般電子的性質記述子”が付与される．例えば，電子吸収性の基は W で電子供与性の基は D である．

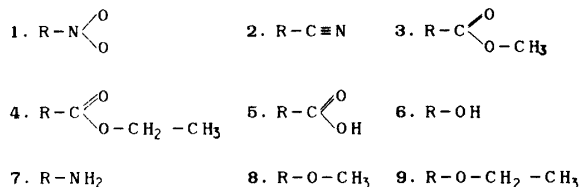


Fig. 5 Chemical formula of functional groups in WD-FILE.

```
(W ((N O O R) ((2 3 4) (1) (1) (1)) ((1 1 1) (1) (1) (1))) NO2 )
((C N R) ((2 3) (1) (1)) ((3 1) (3) (1))) CN )
((C O O C R) ((2 3 5) (1) (1) (1) (3) (1)) ((2 1 1) (2) (1 1) (1) (1)))
COOCH3 )
((C O O C R) ((2 3 6) (1) (1) (1) (3 5) (4) (1))
((2 1 1) (2) (1 1) (1) (1) (1))) COOCH2CH3 )
((C O O R) ((2 3 4) (1) (1) (1)) ((2 1 1) (2) (1) (1))) COOH )
)
(D ((O R) ((2) (1)) ((1) (1))) OH )
((N R) ((2) (1)) ((1) (1))) NH2 )
((O C R) ((2 3) (1) (1)) ((1 1) (1) (1) (1))) OCH3 )
((O C R) ((2 4) (1 3) (2) (1)) ((1 1) (1 1) (1) (1))) OCH2CH3 )
)
```

Fig. 6 WD-FILE (file of functional group)

官能基の認識は、Fig.6に示すようなファイル (WD-FILE) に蓄えられている官能基の結合表と、目標分子の結合表との部分マッチングで行われる。1, 2, …, 8 および 9 は、それぞれ NO₂, CN, …OCH₃ および OC₂H₅ である。

目標分子内のある部分が官能基であると認識されると、その基は電子的性質により、W又はDと置き換えられる。基が複数個存在する場合は、変換される順に W₁, W₂, W₃, … (又、D₁, D₂, D₃…) が官能基の代わりに付加される。変換した官能基を、PUTPROP でW又はDに性質付けする。例題をFig.7にしめす。

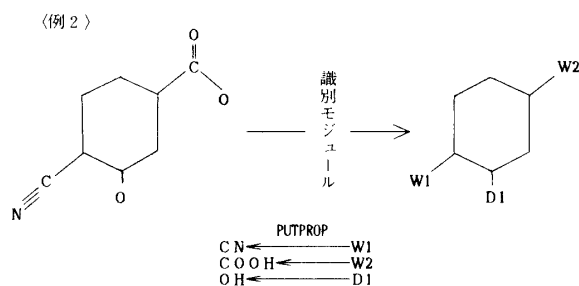


Fig. 7 Perception module.

6. 戦略制御モジュール

前節の識別の結果に基づいて、戦略制御モジュールではこのシステムに組み込まれているいくつかの有機合成の基本的な経験則の適応が試みられる。

経験則は、逆合成反応ごとに、逆合成Tableにまとめられている。例として、Knoevenagal Reactionの逆合成TableをFig.8に示す。図中使用する“ \Rightarrow ”は逆合成を示す記号である。

Fig.9中の①はKnoevenagal Reactionの逆合成式であり②, ③, ④は①の右辺の構造式の官能基を決定する条件である。

識別モジュールで処理された目標分子が戦略制御モジュールに入力されると、目標分子と全ての逆合成Tableの①の左辺の化学構造式との部分マッチングがなされる。マッチングされた反応は目標分子に対して可能な反応として蓄積される。その可能な反応群の中から化学者がひとつの反応を指定すると、その反応に関して逆合成が実行される。すなわち、化学者の指示を受けた反応の逆合成Tableが取り出され②以下の評価に移る。②以下で①の反応式の右辺の化学構造式の官能基が決定される。

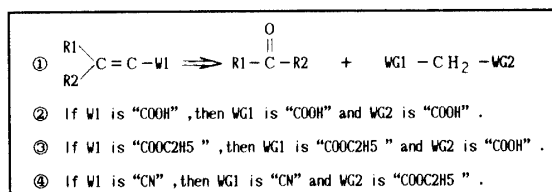


Fig. 8 Antithesis table of Knoevenagal reaction.

```
(
  ( ( C C W1 R1 R2 )
    ( ( 2 3 ) ( 1 4 5 ) ( 1 ) ( 2 ) ( 2 ) )
    ( ( 2 1 ) ( 2 1 1 ) ( 1 ) ( 1 ) ( 1 ) )
    ( ( C O R1 R2 ) ( ( 2 3 4 ) ( 1 ) ( 1 ) ( 1 ) ) ( ( 2 1 1 ) ( 2 ) ( 1 ) ( 1 ) ) )
    ( ( C W1 W2 ) ( ( 2 3 ) ( 1 ) ( 1 ) ) ( ( 1 1 ) ( 1 ) ( 1 ) ) )
    ( ( PRINT $$$R1 OS )
      ( PRINT $$$ I "$" )
      ( PRINT1 $$$ C=C-W1 => R1-C-R2$ )
      ( PRINT1 $$$ + W1-CH2-W2$ ) ( TERPRI )
      ( PRINT $$$ I$ )
      ( PRINT $$$R2$ )
      ( TERPRI ) ( TERPRI ) ( TERPRI ) )
    ( R-REAC1 )
  )
)
DEFINE( (
  ( R-REAC1 ( LAMBDA (
    ( PROG ( NCT-1 FG1 AD )
      ( SETQ NCT-1 ( CAR TARGET ) )
      ( SETQ AD ( PICK ( PICK 3 E-CRD ) NCT-1 ) )
      ( SETQ FG1 ( GET AD ' F-GROUP ) )
      ( REMPROP AD ' F-GROUP )
      ( COND ( ( NOT ( EQUAL ' COOH FG1 ) ) ( GO G1 ) ) )
      ( PUTPROP ' W1 ' F-GROUP ' COOH )
      ( PUTPROP ' W2 ' F-GROUP ' COOH )
      ( GO G3 )
      ( COND ( ( NOT ( EQUAL ' COOC2H5 FG1 ) ) ( GO G2 ) ) )
      ( PUTPROP ' W1 ' F-GROUP ' COOC2H5 )
      ( PUTPROP ' W2 ' F-GROUP ' COOH )
      ( GO G3 )
      ( COND ( ( NOT ( EQUAL ' CN FG1 ) ) ( GO G3 ) ) )
      ( PUTPROP ' W1 ' F-GROUP ' CN )
      ( PUTPROP ' W2 ' F-GROUP ' COOC2H5 )
      ( RETURN NIL )
    )
  )
)

```

Fig. 9 Inner representation of Fig. 8.

逆合成Tableの内部表現は、Fig.9に示されるように①は結合表のS表現¹⁰⁾で、②以下はLISP関数として定義している。Fig.8の① (1 ≤ i ≤ 4) とFig.9の①は対応している。

目標分子内に、①の左辺の構造をいかにして探索するかは目標分子内の結合表とFig.8の①の左辺の結合表との部分マッチングの問題に帰着する。なお、この部分マッチングは識別モジュールで官能基を目標分子内で探索する際にも使用される。

6.1 部分マッチングの手法

目標分子の結合表をTCT (Target Connection Table), それとマッチングの対象となる化学構造式の結合表をDCT (Data Base Connection Table) とする。原子数をそれぞれm, n (m ≥ n) とし、TCTおよびDCTを次のようにおく。

$$TCT = (A^T, B^T, C^T)$$

$$\text{但し, } A^T = (a_1^T, a_2^T, \dots, a_m^T)$$

$$B^T = (B_1^T, B_2^T, \dots, B_m^T)$$

$$C^T = (C_1^T, C_2^T, \dots, C_m^T)$$

$$DCT = (A^D, B^D, C^D)$$

$$\text{但し, } A^D = (a_1^D, a_2^D, \dots, a_n^D)$$

$$B^D = (B_1^D, B_2^D, \dots, B_n^D)$$

$$C^D = (C_1^D, C_2^D, \dots, C_n^D)$$

部分マッチング法は大きく次の I, II の過程より成る.

I. 集合 M を作成する. 集合 M は,

$$M = \{M_i \mid 1 \leq i \leq n\}$$

$$\text{但し, } M_i = \{q \mid a_q^D = a_i^D \text{ かつ } \widetilde{C}_q^D = \widetilde{C}_i^D\}$$

$$C_q^D = (C_{q1}^D, C_{q2}^D, \dots, C_{qk}^D)$$

$$\widetilde{C}_q^D = \{C_{ik}^D \mid 1 \leq q \leq k\} \text{ である.}$$

すなわち, M_i は DCT の節番号 i に対応する TCT の節番号の候補の集合であり, 候補の条件は, “原素記号” と “結合の多重度の集合” が等しいことである.

II. DCT の各節番号とマッチングした TCT の節番号の一連の組 Δ を求める.

$$\Delta = (\delta_1, \delta_2, \dots, \delta_n)$$

$$(\delta_i \in M_i)$$

S_i は DCT の節番号 i とマッチングする TCT の節番号である.

Δ は δ_1 から順に決定してゆき, δ_n が決定されて初めて TCT と DCT は部分マッチングがなされたという.

δ_p (DCT の節番号 p とマッチングする TCT の節番号) の決定方法を示す.

$\delta_p = f (\in M_p)$ と仮定する, つまり DCT の節番号 p と TCT の節番号 f がマッチングすると仮定する.

ここで仮定の正当化をみるために次の集合 Y を考える.

$$Y = \{Y_i \mid 1 \leq i \leq |B_p^D|\}$$

$$\text{但し, } B_p^D = \{g_q \mid 1 \leq q \leq |B_p^D|\}$$

$$\text{とすると, } Y_i = B_i^D \cap M_{g_i}$$

$$\text{かつ, } C_{p g_i}^D = C_{i f}^D, j \in Y_i$$

但し, $|U|$ は集合 U の要素数を示す.

$\delta_p = f$ の仮定の正当性を満たす条件は次の条件である.

$$\text{i) } \forall Y_\alpha \in Y \text{ について } \alpha \neq \beta$$

$$\text{ii) } \forall Y_\alpha \in Y, \forall Y_\beta \in Y \text{ について}$$

$$|\alpha| = |\beta| = 1 \text{ のとき, } \alpha \neq \beta$$

$\delta_p = f$ の仮定は DCT の節番号 p と結合している節の候補を限定する. ここで Y_i は節番号 p と結合している節番号 g の限定された新たな候補である. 条件 i) は候補を持たない節があってはならないこと, 条件 ii) は節の候補が重複してはならないことを示している.

i), ii) の条件を満たさない場合, $\delta_p = f$ は正当でない. 従って TCT と DCT はマッチングしない. $\delta_p = f$ の仮定が正当であると,

$$M = \{M_i \mid 1 \leq i \leq p-1 \text{ のとき } M_i = M_i$$

$$i = p \text{ のとき } M_i = f$$

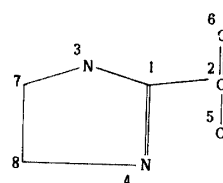
$$i \in B_p^D \text{ のとき } i = g_j \text{ とすると}$$

$$M_i = Y$$

$$i \in B_p^D \text{ のとき } M_i = M_i\}$$

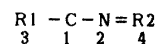
とする. すなわち集合 M に関して, 節 $g_i \in B_p^D (1 \leq i \leq |B_p^D|)$ の候補は $Y_i \in Y (1 \leq i \leq |B_p^D|)$ として, $\delta_{i+1} = f' (\in M_{i+1})$ の仮定の正当性をみる. そのとき集合 M' を新たに集合 M とおき, 同様の操作を繰り返す. 具体的に例を示す.

〈例 3〉



$$TCT = (A^T, B^T, C^T) = \left(\begin{array}{cccccc} \{ \begin{array}{c} 2 \\ 3 \\ 4 \end{array} \} & \{ \begin{array}{c} 1 \\ 5 \\ 6 \end{array} \} & \{ \begin{array}{c} 1 \\ 7 \end{array} \} & \{ \begin{array}{c} 1 \\ 8 \end{array} \} & \{ 2 \} & \{ 2 \} \\ \{ \begin{array}{c} 1 \\ 2 \end{array} \} & \{ \begin{array}{c} 1 \\ 2 \end{array} \} & \{ \begin{array}{c} 1 \\ 1 \end{array} \} & \{ \begin{array}{c} 2 \\ 1 \end{array} \} & \{ 1 \} & \{ \begin{array}{c} 3 \\ 8 \end{array} \} \\ & & & & & \{ \begin{array}{c} 4 \\ 7 \end{array} \} \end{array} \right)$$

Fig.10 Chemical formula and TCT of target molecule.



$$DCT = (A^D, B^D, C^D) = \left(\begin{array}{ccc} \{ \begin{array}{c} 2 \\ 3 \end{array} \} & \{ \begin{array}{c} 1 \\ 4 \end{array} \} & \{ \begin{array}{c} 1 \\ 2 \end{array} \} \\ \{ \begin{array}{c} 1 \\ 1 \end{array} \} & \{ \begin{array}{c} 1 \\ 2 \end{array} \} & \{ \begin{array}{c} 1 \\ 2 \end{array} \} \end{array} \right)$$

Fig.11 Chemical formula and DCT of object for matching.

ここで $DCT(i)$, $TCT(i)$ は各々結合表 DCT, TCT の節番号 i を表す. $R1$ と $R2$ は任意の分子を表しているのでマッチングは考えない. すなわち, $DCT(1)$ および $DCT(2)$ のみをマッチングの対象とする.

(I) $DCT(1)$ は元素記号 “C”, 結合多重度の集合は “{1, 1}” である. この条件を満たす TCT の節は, $TCT(7)$ と $TCT(8)$ である. $DCT(2)$ は, 元素記号が “N”, 結合多重度の集合が “{1, 2}” で, 同じ条件をもつ TCT の節は $TCT(4)$, つまり

$$M = (\{7, 8\}, \{4\}), M = \{7, 8\}$$

$$M = \{4\}$$

(II) DCT の節番号とマッチングした TCT の節番号の一連の組 $\Delta = (\delta_1, \delta_2) (\delta_1 \in M_1, \delta_2 \in M_2)$ を求める.

$$\delta_1 = 7, \text{ つまり } DCT(1) \text{ と } TCT(7) \text{ がマッチング}$$

すると仮定する。そのとき、

$$B_1^p = \{2, 3\}$$

$$B_7^r = \{3, 8\}$$

となる。DCT(3)は考慮しないから、 B_1^p における節番号2の候補となるTCTの節番号の集合は M_2^p であり、従って、

$$Y_1 = M_2 \cap B_1^r = \phi$$

となる。 $Y_1 = \phi$ なので条件 i) を満足しない、故に $\delta_1 = 7$ は正当ではない。

$\delta_1 = 8$ 、つまりDCT(1)とTCT(8)がマッチングすると仮定する。そのとき、

$$B_1^p = \{2, 3\}$$

$$B_8^r = \{4, 7\}$$

であるから、

$$Y_1 = M_2 \cap B_8^r = \{4\}$$

Y_2 は考えない。この Y は条件 i) ii) を満足する。

故に $\delta_1 = 8$ は正当であり、 $M^r = (\{8\}, \{4\})$ である。 M^r を M とする。 $M = (\{8\}, \{4\})$ 、 $\delta_2 = 4$ 、つまりDCT(2)とTCT(4)がマッチングすると仮定する。そのとき、

$$B_2^p = \{1, 4\}$$

$$B_4^r = \{1, 8\}$$

であるから、 $Y_1 = M_1 \cap B_4^r = \{8\}$

Y_2 は考慮しない。この Y は条件 i), ii) を満足する。

故に $\Delta = (8, 4)$ となる。

7. 戦略対話モジュール

対話モジュールは、化学者自身が戦略を行いたい場合や目標分子を修正したい場合に用いるモジュールであり、目標分子の原子の削除、結合の切断および官能基の置換を対話的に行う機能を有する。

Fig.12にこのモジュールの流れを示す。

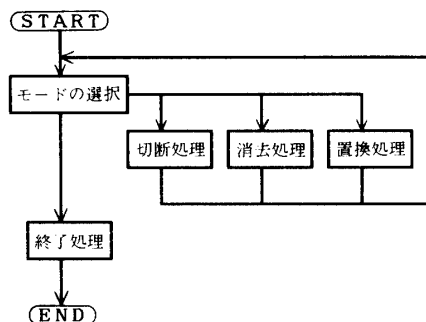


Fig.12 Dialogue module.

7.1 各モードの概要

i) 切断処理モード

このモードは目標分子内で指定された2つの原子間の結合を消去する。

ii) 消去処理モード

このモードは目標分子内で指定された原子を消去する。又、その原子から出ている結合も削除する。

iii) 置換処理モード

このモードは目標分子内で指定された官能基を他の

```

WHAT DO YOU WANT TO DO ?
SEVERANCE
ERASURE
SUBSTITUTION
DONE
PLEASE SELECT ORDER.
=ERASURE
  
```

```

*ERASURE MODE*
WHICH ATOM DO YOU ERASE ?
INPUT ATOM NUMBER YOU WANT TO ERASE.
=(3)
  
```

```

YOU JUST INPUTTED 3 .
ALL RIGHT ? = YES OR NO =
=YES
  
```

ERASURE IS COMPLETE.

```

WHAT DO YOU WANT TO DO ?
SEVERANCE
ERASURE
SUBSTITUTION
DONE
PLEASE SELECT ORDER.
=SEVERANCE
  
```

```

*SEVERANCE MODE*
WHICH COMBINATION DO YOU SEVER ?
INPUT ATOM NUMBER AT BOTH ENDS.
  
```

```

ONE END IS..?
=(4)
  
```

```

YOU JUST INPUTTED 4 .
ALL RIGHT ? = YES OR NO =
=YES
  
```

```

ANOTHER END IS..?
=(5)
  
```

```

YOU INPUTTED 4 AND 5 .
ALL RIGHT ? = YES OR NO OR INITIATE =
=YES
  
```

SEVERANCE IS COMPLETE.

```

WHAT DO YOU WANT TO DO ?
SEVERANCE
ERASURE
SUBSTITUTION
DONE
PLEASE SELECT ORDER.
=DONE
  
```

Fig.13 Example of dialogue.

官能基と置換するものである。

IV) 終了処理モード

切断処理や消去処理によって構造が複数に分割される場合には、1つの結合表内に複数の構造の結合表が存在するので、このモードは、1つの結合表を複数個の結合表に分割する。又、処理で不規則になった節番号を再び規則的にする。

戦略対話モジュールの対話例を次に示す。

〈例4〉 Fig.13では目標分子における節番号3の原子を消去し、節番号4および5の結合を切断する例である。

8. 変換モジュール

変換モジュールは化合物を記号的に化学変換処理して前駆物質を作り出す。このモジュールはいくつかの

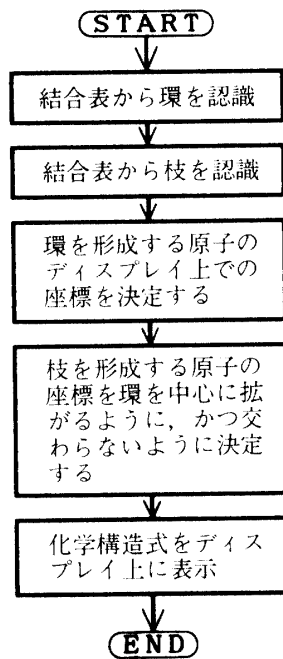
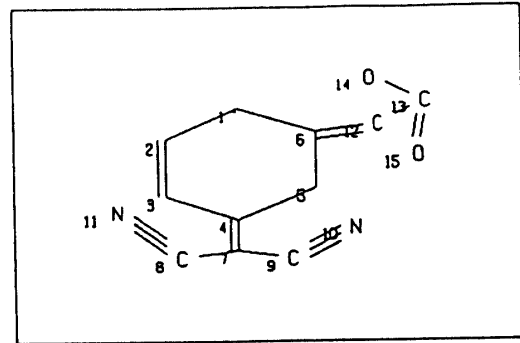


Fig.14 Flow chart of representation on display.

サブルーチンから成っている。これらのサブルーチンを駆動するとその結果として結合の切断や原子の削除および官能基の置換がなされる。

対話モジュールは化学者の指定でそのサブルーチンを作動させたが、変換モジュールは対話によって化学者から与えられた情報を逆合成 Table から抽出してルーチンを作動させる。そして逆合成 Table ①の左辺の結合表を作成する。結合表はディスプレイ表示に転送され、CRT上に化学構造式として表示される。

TARGET MOLECULE



I MAKE THE CONNECTION TABLE FROM THE INPUT DATA.

DO YOU HAVE A STRATEGY ?
-----YES OR NO-----
=NO

I RECOGNIZE SOME SIGNIFICANT STRUCTURAL FEATURE
IN TARGET MOLECULES.

PERCEPTION FINISHED.

I MATCH TARGET MY RULES.

TEKIO DEKIRU RULES.

TEKIO NUMBER =1

```

R1          O
I          ||
C=C-W1 => R1-C-R2 + WG1-CH2-WG2
I
R2
  
```

TEKIO NUMBER =2

```

R1 W1          O
I I          ||
C=C => R1-C-R2 + WG1-CH2-WG2
I I
R2 W2
  
```

SELECT THE MATCHING NUMBER.
-----NUMERAL-----
=(2)

I ACT DATA FUNCTION.
FINISHED.

I MUNIPULATE THE CONNECTION TABLE.

SAVED MOLECULE

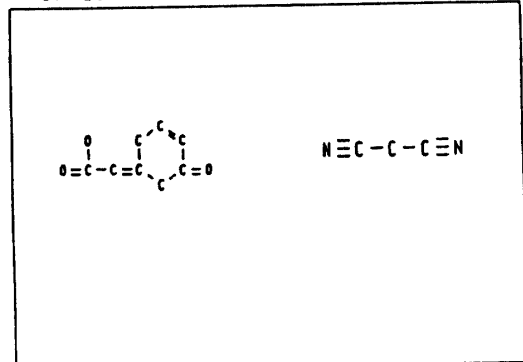


Fig.15 Example of system opration.

9. ディスプレイ表示

ディスプレイ表示は、戦略後、変換モジュールで結合の切断や原子の付加、削除の操作を受けた結合表から図形的に結合の交わりがなく、環に歪みがないような自然な化学構造式をCRT上に表示する。

ディスプレイ表示のフローチャートをFig.14に示す。

10. 入出力例

システムの駆動例をFig.15に示す。

まず化学者によって入力された目標分子 (Target Molecule) に節の番号付けがなされてCRT上に表示される。次に化学者自身に戦略がない場合、目標分子に適用可能なルールが表示される。

ルールの指定を化学者が行うと、戦略が行われ、結果を前駆物質 (Saved Molecule) としてCRT上に表示する。目標分子の入力から、前駆物質がCRT上に表示されるまでに約15分 (目標分子の入力に約8分) を要した。

11. むすび

以上、有機合成デザイン支援システムの基本的枠組みについて述べた。

図形モジュールおよび図形データ変換モジュールはかなり実用的に近いものと思われる。システム全体として実用化に至るには多くの問題点を残しているが、本論文に直接関係するものとして、

1. 識別モジュールは、現段階では識別の対象が環と官能基に限られているが、原子鎖、測鎖基 (環あるいは鎖状分子における)、および分子の骨格構造中の冗長性あるいは対称性などの識別を可能にすること。

2. 戦略制御モジュールで用いた経験則 (合成テーブル) に収率、温度、試薬などの情報を系統的に組み込むこと。

3. 知識獲得モジュールの作成。
等が挙げられるが、これらについては検討中である。いずれにせよ本システムを知識ベースシステムとして更に充実したものへと発展させるためには、情報学者と化学者との強力な共同研究体制の確立が必要であろう。

謝 辞

日頃ご助言頂く本学高浪五男教授、井上克司助教授ならびに故岡田敏彦助教授に深謝すると同時にLISPシステムに関して有益なご意見を頂いた石原好宏助教授に深謝する。又、本研究に際してご指導ご助言頂いた臨床研 (東京都臨床医学総合研究所) 神沼二真博士ならびに臨床研の諸氏に感謝の意を表する。

参 考 文 献

- 1) R. Davis, D. B. Lent, "Knowledge-Based Systems in Artificial Intelligence", McGraw-Hill Inc., (1982)
- 2) 溝口, 北沢, "知識工学入門", 講談社, Blue Backs (1982)
- 3) E. J. Corey, W. T. Wipke, Science, 166, PP178-192 (1969); [訳] 大野, 額田, 坂本, 鈴木, 化学の領域, 26-7, PP18-25, 26-8, PP41-50, 26-9, PP29-37
- 4) E. J. Corey, W. T. Wipke, R. D. Cramer III, W. J. Howe, "Organic and Biological Chemistry", J. Amer. Chem. Soc., 94, PP421-430 (1972)
- 5) E. J. Corey, W. T. Wipke, R. D. Cramer III, W. J. Howe, "Techniques for Perception by a Computer of Synthetically Significant Structural Features in Complex Molecules", J. Amer. Chem. Soc., 94, PP431-439 (1972)
- 6) E. J. Corey, R. D. Cramer III, W. J. Howe, "Computer-Assisted Synthetic Analysis for Complex Molecules. Methods and Procedures for Machine Generation of Synthetic Intermediates", J. Amer. Chem. Soc., 94, PP440-459 (1972)
- 7) W. T. Wipke, G. I. Ouchi, S. Krishnan, "Simulation and Evaluation of Chemical Synthesis-SECS: Artificial Intelligence, 11, PP173-193 (1978)
- 8) N. Ezaki, T. Kanaoka, S. Tomita, T. Okada, "Chemical Structure Input System by Computer Crapics", Memoirs of The Faculty of Engineering Yamaguchi University, 35, PP145-151 (1984)
- 9) H. L. Morgan, "The Generation of a Unique Machine Description for Chemical Structures A Technique Developed at Chemical Abstracts Services" J. Chem. Doc., 5, PP107-113 (1965)
- 10) 中西, "Lisp入門 ——システムとプログラミング——", 近代科学社 (1981)

(昭和61年4月15日受理)