

CICSJ Bulletin

Published Bimonthly by Division of
Chemical Information and Computer Sciences
The Chemical Society of Japan

コンピュータを利用したドラッグデザイン—研究の現場から 特集号

日本化学会
情報化学部会

Volume 11, Number 1
January 1993

目 次

特集：コンピュータを利用したドラッグデザイン—研究の現場から	
特集にあたって.....	編集 幹 事 2
構造化学とドラッグデザイン.....	秋 元 利 夫 3
Pharmacophore の探索に基づく薬物設計.....	小 貫 慶 昭 6
コンピュータを利用したドラッグデザイン—研究の現場から	片 倉 晋 一 10
コンピュータを利用したドラッグデザイン—研究の現場から	多 田 幸 雄 13
3D データベースの状況.....	時 実 象 一 17
コンピュータを利用したドラッグデザイン—研究の現場からの雑感	張ヶ谷 泰 二 20
ドラッグデザインのためのコンピュータ利用—研究の現場からの雑感	藤 原 巖 24
部 会 行 事	27
第15回情報化学討論会 報告	
第3回ケモメトリックスワークショップ 報告	
第2回情報化学部会講習会—化学者のためのパーソナルコンピュータ利用講習会 報告	
第3回情報化学部会講習会—パーソナルコンピュータを利用した文献整理講習会 案内	
関 連 行 事	43
実践パソコン分子科学計算講習会	
記 事	
UNIXは本当に難しいものだろうか.....	小 出 昭 夫 44
海 外 動 向	48
Meetings	
文 献 紹 介	49
コンテンツ・リスト：	
Journal of Molecular Graphics (Vol.10, No.3~4)	
Journal of Chemical Education (Vol.69, No.9)	
Quantitative Structure Activity Relationships (Vol.11, No.1~4)	
Journal of Computational Chemistry (Vol.13, No.8~10, Vol.14, No.1)	
Journal of Chemical Information and Computer Sciences (Vol.32, Nov./Dec.)	
Journal of Mathematical Chemistry (Vol.11, No.4)	
編 集 後 記	59

特集にあたって

今回の特集では「コンピュータを利用したドラッグデザイン—研究の現場から」というタイトルで、実際に製薬企業の研究現場に籍をおく6人の方から原稿をいただく共に、最近注目されている3Dデータベースの紹介を時実氏にお願いした。

本特集のきっかけは大学に籍を置くある部会員の方から「企業は高価なCADDシステムを導入し、かなりのお金と人手を投入している様子だが、実際どのように利用されているのか、どのように評価しているのか、ほとんど知る機会がない。大変興味あるところなので是非部会誌でとりあげてほしい」との提案をいただいたことによる。製薬企業にとっては最も機密性の高い部分であり、引き受けていただけるとか危ぶまれたが、幸い協力者を得て特集としてお届け出来ることになった。

記事を拝見すると、CADDを単なるコンピュータ技術の一つとして捉えるのではなく研究そのものとして捉え、それを推進するのはコンピュータ技術者ではなく研究者であるという認識は大方定着しているように伺える。その上で研究者一人一人にまで展開するにはシステムとしてもまだまだ改善されるべき点が多いこと、又、CADDの利用は病気のメカニズムや、ターゲットとなる生体高分子物質の原子レベルでの解明と一体となってその有用性を増していくものであることを複数の方が指摘している。現状には満足していないが自分達の掌中のものにしようとする研究者の方々の苦勞と希望を感じることが出来る。

7人の著者の方に心から御礼申し上げます。

(編集幹事)

構造化学とドラッグデザイン

中外製薬富士御殿場研究所

秋元 利夫

序

上記の様な標題としたが、コンピュータが役に立っていないという訳ではない。だが、学会などを聴講した際、この部分は実験的裏付けが必要だと思われたり、コンピュータのみではあまりに不十分な'データ'だと感じる事が多いからである。実験が必要と思われる部分はやはり実験を行なうべきであろう。まだまだ、コンピュータが一人歩きするのは早すぎる。その上で、コンピュータの長所および不十分さを充分認識して研究を行なうべきであると考え。コンピュータ作業に行きずまりを感じたら、文献整理や実験を行なってみるべきである。

薬業界を取り巻く情勢は段々と厳しくなり、Me Too 医薬品開発にたいする国際的な評価は益々厳しくなるであろう。そこで従来のようにレセプターマッピングを行ない、特許やぶりの少しばかり modify した'薬'を上市することは難しくなるであろう。国際的に薬に新規性が要求される時代になりつつある。

我が社でも多くの検体のデザインにコンピュータが利用されている。だが、合成の後のまとめ研究であることも多い。これでは創造的な医薬品開発に寄与したことにはならない場合が多い。

取り組まねばならない事柄、運営体制

コンピュータは創薬の初期の段階から使われねばならない。実際に使うヒトはコンピュータの専門家である必要はない。リード化合物は合成的観点、代謝的観点など多くの分野から検討されねばならない。また、コンピュータを用いて得た構造には疑問点も多い(特にoptimizeした構造)から、X線やNMRの構造専門家からの情報も必要である。また、物質の荷電状態などを調べながら進めて行かねばならない。

コンピュータは創薬に対して過大評価を受けながら登場してきた面も多いだろう。筆者もそのため良い思いもした。だが、その反動として現在コンピュータの役割を否定している人も多いだろう。中には、コンピュータを利用すればレセプターの構造が分かると期待した上司もいたからである。コンピュータ専門家の中にまだこの部類のヒトがいるのも事実である。我々の使っている構造は一つのモデルであり、多くの仮定がなされて出てきたもので、実際の実験的結果に合わなければ修正されねばならないモデルだということを謙虚に、科学的に考えていなければならない。コンピュータは速く結果がでるが、仮定が沢山入っていて、その仮定の一つ一つを実験をしている人と討議しながら、よりよいモデルへと高めて行かねばならない。

運営体制は従って、CADD従事者は研究者であるということであり、コンピュータの他の従事者と一緒にされるべきではないと思われる。本来、CADD従事者は構造を扱う研究者であるから、構造や物性を研究する人達とグループを組むべきである。他いろいろな運営体制も考えられようが、コンピュータグループなどと独立しているのはどう考

えてもおかしい。CADDはchemistryである。一般的コンピュータ取り扱い者はChemistではない。だが、一方化学構造の知識に乏しいCADD研究者も多い。この中からでは新しい化学構造的アイデア、特にlead generationの発想は生まれにくい。従って、CADD従事者は合成研究者などと密接な協力関係になければならない。

逆に合成研究者などにCADDを普及するのはどうであろうか。この方が研究の能率は上がるのではないだろうか。合成研究者が自分で考え、合成して行く方が合成研究者にとってやりやすいだろうし、やる気も起こるだろう。但し、合成研究者は構造的な知識や計算機の出力結果の理解に不足なところもあるから、その点は構造化学研究者やCADDの専門家が補ってやるのが大いに必要だろう。

システム評価

会社組織の中で研究していくにはGeneralなシステムが良い。日本で創業のためのソフト開発を行なっている研究組織は数があまりに少ない。一般的には物理化学的な研究があまりに多い。それも実験を伴わないコンピューター一人歩きの研究が多い。または、全ての合成を終え、生物活性を得た上での結果整理の論文が多い。従って、我々会社の研究所の研究者としては、高価な外国のソフトを購入することとなる。

日本でのシステムを考える時、外国のそれらとは大きく異なるgeneralなソフトの開発を望みたい。会社の研究者が必要だと考えている事を米国、英国では行なっている様に見える。インターフェイスも取りやすく、使いやすい。だが、ソースコードがないので使っていて心配になる事もある。JCPEが設立され日本でも良いソフトが必要に応じて入手できるようにはなってきたが、創業においてはまだ一部分のみにしか使用できないのが現状である。日本の研究組織が小人数で、また大学では各自が学位論文を出さねばならないので、研究のスケールがどうしても小さくなってしまう。数年かまえて、複数の一人前の研究者でoriginalityのあることをやって欲しい。外国のソフトを使って、論文の数を増やす（これでは会社の研究者が内部で時々やっていることと余り変わらない。会社では論文にしない場合が多い）より、地道に自分の道を確立して行って欲しい。

大学の薬学部の教室でも創業を頭に置いて研究している教授は小数派であろう。これでは日本の製薬企業が欧米の大会社と5分に渡り合うことは困難である。創業にAcademic interestは重要であるが、それを実際に使えるまで高めるのは会社だけの作業ではない。だが、会社でも基礎的な研究が必要になって来るであろう。大学など研究機関でコンピュータを使っている研究者には、創業はもっと泥臭いと考えて頂き、他の分野の研究者には創業はエレガントな部分もある事は頭に入れて置いて頂こう。

今後の方向

コンピュータだけで薬を作ることは出来ないことははっきりしている。幸いにこの10年の蛋白質化学（特に遺伝子工学）の発展による標的蛋白の大量発現が可能となりやX線やNMRなどの構造化学の発展には目を見張るものがある。CADDはこれらの構造化学的手段と同居しながら将来的には伸びて行くべきであろう。CADD自身も発展するだろうし、この専門家も重要である事は勿論であろう。だが、若い研究者の間に10年

前に見られたコンピュータアレルギーはもう見られない。教育さえしっかりしていれば CADD はこれらの研究者の間にスムーズに入って行くであろう。

レセプター（広義の）の構造が分かった時、その阻害剤の構造をデザインすることが可能になりつつある。まだまだ不十分な点はこれらのソフトにはあるが、10年後の製薬企業の研究のやりかたが大きく変わって来る可能性もある。Agonist の設計の研究は Antagonist の設計の場合よりX線やNMR の構造により強く依存するであろう。

また、毒性や代謝の問題も分子のレベルに降りて来るであろう。分子の化学構造や立体構造からものが言える時代が近づくであろう。これらの問題についても、ある程度有望な化合物が見いだされた段階から、早期に実験的な裏付けが必要になり、ドラッグエンジニアリングに生かされねばならない。

以上

Pharmacophoreの探索に基づく薬物設計

持田製薬（株）富士中央研究所

小貫 慶昭

1. はじめに

現在、コンピュータ支援によるドラッグデザインという名のもとで、多くの手法が紹介され普及しているが、ほとんどが理論化学計算や解析についてである。これらの理論化学計算や解析を徹底的に実施しても、その結果から新しい化合物は生まれていないのが現状である。化合物の設計は、計算結果を使用しながら研究者の頭の中で行われている。したがって、コンピュータは研究者の頭脳の中で行われている創造のプロセスを手助けする道具にすぎないと言える。そこで、要求されるのは道具であるコンピュータを使いこなす技術の確立またはシステム化である。

最近、理論化学計算や解析のソフトウェアはQCPEやJCPEより普及しており、市販のソフトウェアも高価ではあるが普及してきている。しかし、ソフトウェアは入手すれば済むが、ドラッグデザインを指向した使いこなす技術は市販されていない。したがって、自社において使いこなすノウハウを体系化して、理論化学計算や解析を道具として用いて新しい化合物を構築する手順を確立する必要がある。

弊社では、アプローチを体系化してドラッグデザインを実施する手順の確立を進行中である。今回はその体系づけられたシステムの中でコンピュータがどう活用され、化合物の設計にどう生かされているかについて述べる。

2. コンピュータ支援によるアプローチ

最近、コンピュータ支援によるドラッグデザインについて簡潔明瞭にアプローチを紹介している総説が出ている。¹⁾ このアプローチの仕方は我々が目指しているのと方向性が似ている。まず、コンピュータ支援によるアプローチを考えると、以下の2つのステージによって異なってくる。

- ・リード化合物の発見または創出 (Lead-molecule development)
- ・リード化合物の最適化 (Lead-molecule optimization)

従来においてコンピュータを利用したドラッグデザインの成功例の多くは後者のリード化合物の最適化においてである。この場合は、Hansch-Fujita法などのQSARを中心に、構造と活性との相関性を解析して、目的とする薬理活性を上げ、不都合な副作用（毒性）を下げるまたは無くするような化合物を探索している。このアプローチでは多くの場合、リード化合物が母核化合物となり、最適の置換基を探索している。また、このアプローチは確立していて総説²⁾にも詳しく紹介されている。

前者のリード化合物の発見または創出は偶然によるのが大部分である。過去の新薬を概観すると天然物よりの抽出物やランダムスクリーニングなどの偶然によって得られているものである。この段階ではコンピュータが関与する余地はまったくないが、一般に偶然見つけられた化合物は薬理活性が弱い。そこで、薬理活性を上げるために種々の合成が試みられることになる。この段階で初めてコンピュータが関与してくる。薬理活性を上げるために単なる置換基変換でなく、分子に関する構造や物理化学的性質などの情報を理論化学

計算や解析より得て、分子の構造を構築していくのである。このアプローチについては、今回の主旨になりますので詳細については後述する。

別の視点よりコンピュータ支援のアプローチを整理すると、次の2つに分けられる。

・ C A D A (Computer Aided Drug Analysis) ・ ・ コンピュータ支援による解析

・ C A D D (Computer Aided Drug Design) ・ ・ ・ コンピュータ支援によるデザイン

現在、学会や論文で紹介されているアプローチの多くは C A D A である。また、実際の作業でも C A D A が中心になっている。C A D A により、後付け的な側面はあるが、化合物に活性が何故あるのかの説明が付けられる。しかし、設計を指向した適用を考えると C A D A には限界があるので、C A D D の技術の確立が必要となってくる。

また、コンピュータを利用したドラッグデザインを進める場合、レセプターの構造が判明しているかどうかでアプローチが異なってくる。現在、進めているものについてはそのほとんどがレセプターの構造が不明のものである。まさに暗闇の中でドラッグデザインの作業を行っているのである。その場合のアプローチとしていくつかある。

アプローチの一つとして Q S A R があり、他のアプローチとして、A C E 阻害薬である Ro31-2201, Ro31-2848 の薬物設計のエlegantな実例⁹⁾のような方法がある。弊社では、後者のアプローチの根底に流れている発想を踏襲している。このアプローチで根幹となるのは、既知の化合物の構造より、活性発現部分の原子の空間的配置すなわち薬理作用団 (Pharmacophore) のジオメトリー (geometry) を見だし、レセプターの発現機構の作業仮説を立てることにある。この薬理作用団を有し、レセプターに対する作業仮説に合致するように、既知構造を修正または改変して新しい骨格を構築する。この作業は研究者の創造性に依存するところが大きい、コンピュータグラフィックス上で3次元的に表示しながら実施することにより、設計が創造的になる。また、薬理作用団を部分構造として有する化合物を自社のデータベースより検索する。このようにして設計した化合物を合成し、データベースの検索の結果得られた化合物をサンプルストックより持ち出して、薬理活性を測定する。測定の結果を上記の作業仮説にフィードバックして作業仮説をより充実させ、さらに化合物の設計を試みる。弊社ではこのアプローチにより、目的とする薬理活性を有する新規骨格の化合物がいくつか見いだされている。

この一連の作業の中でコンピュータが有効的に使用されるのが、薬理作用団を見いだすところである。手法としては以下のとおりである。

- ・ 配座解析
- ・ 形状解析
- ・ アクティブ アナログ アプローチ (Active Analogue Approach)⁴⁾
- ・ レセプターマッピング⁵⁾
- ・ ディスタンス マッピング⁶⁾
- ・ 重ね合わせ
- ・ C o M F A (3 D Q S A R)⁷⁾

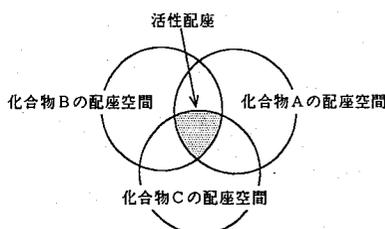
以上の手法は、コンピュータがなければできない作業であり、コンピュータの威力が発揮される場所である。

3. 薬理作用団の探索

薬理作用団のジオメトリーを探索する際に、ドラッグとレセプターの相互作用にどのような種類があるかについて考慮しておく必要がある。その相互作用として4つが考えられている。⁸⁾ まず、Steric repulsion (立体的反発)、さらにHydrophobic interaction (疎水性相互作用)、Electrostatic interaction (静電相互作用)、Dispersive interaction (分散力相互作用) などがある。これらの相互作用を考慮しながら、薬理作用団を探索していく。探索の手法はターゲットに応じて考える必要があるが、ここでは紙面の関係で一例を示す。

探索するにあたって、最安定エネルギーまたはそれに近いエネルギーの配座を使用するので、最初に配座解析を実施する。配座解析には種々の方法があるが、目的および計算時間を考慮して使用する方法を決めている。自由に回転する結合の数が1~3個なら、各結合の二面角を $10^\circ \sim 30^\circ$ 間隔で変えそれぞれについて分子力学法を用いてエネルギー極小化計算を実施する。各配座のエネルギーの値と2次元のエネルギーマップなどを用いて配座を求めていく。3個以上の結合なら、フレキシビリティが高いだけでなく計算時間を要するので、分子動力学法を用いる。その際に安定な配座を有効に見いだすために、Simulated-Annealing法を用いる。この場合多くの配座については求められないが、最安定エネルギーの配座を中心として20~30個の安定配座を求める。各配座を求めた後、用いる配座に重み付けを行う目的で、各配座間が熱力学的に平衡であるとの仮定のもとで、Boltzmann分布を求めて各配座の存在確率を算出する。

つぎに、薬理活性のあった化合物を概観してポイントとなる原子または官能基に着目する。窒素や酸素などのヘテロ原子、ベンゼンなどの芳香族の官能基などは一般に着目すべきものである。着目した原子あるいは官能基の距離を上記で求めた配座について算出する。これらの距離を座標軸に取り各化合物毎に2次元または3次元マップ(ディスタンスマップ)を作成する。これは一種の化合物の配座空間を表現している。これらの配座空間の積集合をとれば、その配座が活性を有するものであると言える。



求められた配座空間と着目した原子または官能基の位置関係を考慮して、薬理作用団の仮説を立てる。その場合に疎水性相互作用や静電相互作用について、可能性を考慮しながら実施する。

4. 今後の方向

現在は、適用できる手法でコンピュータを利用したドラッグデザインを実施しているが、道具として見た場合、まだまだ不十分である。しかし、ここ数年の間に新しい手法が開発され普及しつつある。⁹⁾ それはデータベースの3D-SearchとDe Novo設計である。3D-searchは探索された薬理作用団を3次元構造を用いてデータベースを検索する手法である。

現行の2次元構造処理では限界があるところを見事に解消している。De Novo設計はレセプターの構造が既知の時に、レセプターの形に対応して水素結合などを考慮しながらそこに当てはまる適切な官能基を選んできて構造を自動的に構築する手法である。また、藤田稔夫先生が開発を進めている、構造変換の知識ベースを用いるリード化合物構造展開支援システム (Emil) も新しい手法¹⁰⁾として興味がある。これらの手法については、コンピュータを利用したドラッグデザインを変革する可能性を秘めているので今後とも着目していきたいものである。

研究者の頭脳を刺激する観点より、バーチャルリアリティ (仮想現実: Virtual Reality) の技術に着目している。¹¹⁾ 現行のソフトウェアはいずれもマウスを使用しているが、どうしても人間にとって違和感がある。バーチャルリアリティにより研究者自身がレセプターの世界に入り込んで分子を当てはめようとした時に感覚的に実感できる状態になれば、研究者の創造性を刺激して新しい構造を思いつく可能性が高まると思われる。

5. おわりに

コンピュータ利用のドラッグデザインの手法はまだまだ発展途上にあり、今後もコンピュータの発達に伴い理論化学計算や解析もますます充実してくるだろうと考えられる。現状では、コンピュータから新薬は生まれていないが、研究者の頭脳を刺激して創造性を活発化させる道具であると考え、コンピュータは非常に素晴らしく無限の可能性を与えてくれるものであると言える。21世紀には現在問題になっている点も解決されて、コンピュータ利用のドラッグデザインがルーティン作業になることが期待される。

参考文献

- 1) J. Scott Dixon, (1992) Trends Biotechnol., 10, 357-363.
- 2) 藤田稔夫編: 「構造活性相関とドラッグデザイン」、化学同人、1986年。
構造活性相関懇話会編: 「薬物の構造活性相関II」、南江堂、1982年。
- 3) Hassall, C. H., (1985), Chemistry in Britain, 21, 39-46.
- 4) Marshall G. R., Barry C. D., Bosshard H. E., Dammkoehler R. A., Dunn D. A., (1979),
Computer-Assisted Drug Design, ACS Symposium Series 112, 205-226.
- 5) J. Sufrin et al., (1981), Mol. Pharmacol., 19, 307-313.
- 6) D. Mayer, C. B. Naylor, I. Motoc, G. R. Marshall, (1987), 1, 3-16.
- 7) R. D. Cramer III, D. E. Patterson, and J. D. Bunce, (1988), J. Am. Chem. Soc., 110,
5959-5967.
- 8) Arup K. Ghose, Avis Pritchett, and Gordon M. Crippen, (1988), J. Comput. Chem.,
9, 80-90.
- 9) Stu Borman, (1992), Chem. Eng. News, 70, 18-26.
- 10) 藤田稔夫等、(1991), 第19回構造活性相関シンポジウム講演要旨集、268-279.
- 11) Rod Hubbard, (1992), CDA News, 7, 1, 34-37.

コンピュータを利用したドラッグデザイン

— 研究の現場から —

第一製薬（株）探索第二研究所 片倉晋一

私は、入社以来ずっと合成研究に従事してきましたが、数年前から、コンピュータを利用したドラッグデザインに関わってきました。今回は、主に合成研究者から見たコンピュータの利用について当社の現状を以下に述べさせていただきます。

「活用事例」

三次元グラフィックス上で行うドラッグデザインでは、まず最初に各化合物の安定な三次元構造を分子力場法により算出し、それらを重ね合わせることから、活性発現に必要な原子及び官能基の位置、性質（静電ポテンシャルなどを利用）を推定することに使用しています。また分子軌道法を利用することで、合成反応での反応性の予測に使用したり、化合物の生物活性値との相関関係...例えば、一連の化合物のスカベンジャーとしての活性と軌道（HOMO、LUMO）のエネルギー値との相関...をドラッグデザインに活かすことを行っています。

最近になって、ターゲットとする蛋白質の三次元構造が推定可能なケースが幾つか見られるようになってきました。当社においてもターゲットの三次元構造を推定できる研究テーマが幾つかあります。そのような場合には、まず蛋白質の三次元モデルを構築した後、デザインした化合物をドッキングし、蛋白質と化合物との相互作用を検討することによって化合物の活性を定性的に予測しています。

但し、当然のことですが、薬物のターゲットの三次元構造に関して情報がどの程度明らかになっているかによってその精度は大きく変わってきます。しかしながら、重ねあわせ等、化合物側だけの情報で行うデザインに比べ、ターゲットの三次元構造が見える場合のデザインは、明かにデザインの効率化（効かない化合物を排除する点）において差があるようです。実例としては、抗菌化合物（ケトカルボン酸誘導体）とDNAとのドッキング、あるいは酵素阻害剤と酵素とのドッキングなどに適用しています。

以上述べてきたようにしてコンピュータをドラッグデザインに活用していますが、ドッキングスタディでも、あるいは重ねあわせでも、得られた情報から新しいドラッグをデザインするのは最終的にはあくまで各研究者のアイデアであって、化合物の形は各人の過去の経験に基づくものです（例えば、デザインした化合物の合成のしやすさなど）。従って、現時点においてドラッグデザインにおけるコンピュータの役割はデザインのヒントを与えてくれるものであり、無駄なデザインを排除してくれるものと私は考えています。

また、このような三次元グラフィックスをもとにした試みとは別に、従来の統計学的手法による定量的構造活性相関の検討も随時行っています。

「取り組み、運営体制」

最近のソフトウェアは以前に比べるとだいぶ使いやすくなってきていますので、基本的な項目、例えば化合物の安定構造の算出、重ねあわせなどは、一般の研究者自身が行えるようになってきました。しかしながら、逆に見えない部分が多くなり、一般の研究者にとってはどうしてもブラックボックス化の傾向が出てきています。例えば、アトムタイプのアサインなどをとって、ソフトウェアが自動で行いますので誤ってアサインしてもチェックしなければ、そのまま誤った構造を生成し、それを使うことになってしまいます。従って、ソフトウェアがいくら使いやすくなって、一般の研究者に対するアトムタイプの意味の説明などのきちっとした教育は欠かすことができません。

ハード及びソフトの管理に関しては、従来、全てシステム部門が行ってきましたが、前述したような背景から、これとは別にCADD運営委員会（合成研究者を中心にシステム部門、あるいは構造解析部門等の総勢8名からなる委員会）を3年前に組織しました。運営委員の主な活動としては、一般研究者に対する分子力場計算などについての教育、啓蒙、並びに実際の操作時に発生する種々の問題に対するサポートなどがあります。

また、一般研究者に対する支援とは別に、各委員は自らが属する各研究所部門内のテーマに即した形で、より複雑なアプローチ、あるいはコンピュータを利用した新しい解析手法の適用を実施しています。例えば、蛋白質のモデリングそれに続くドッキング、結合様式の解明などの研究がそれに該当します。この結果から得られた情報を研究者にフィードバックし、次のドラッグデザインに反映させるよう努力しています。

以上のように当社においては、コンピュータを利用したドラッグデザインとして1)一般の研究者に対する基本的な解析手法の徹底2)新しいアプローチの実際の研究への適用を二つの柱として運営しています。

「システムの評価」

システムとしては、ワークステーション（IRIS）をプラットフォームとしモデリングソフトを稼働させています。ワークステーションは、ETHERNETを経由してパソコン端末（PC-9800並びにマッキントシュ）及びホストコンピュータ（VP-1100）に接続されネットワークを形成しています。従って、ほとんどの仕事はワークステーションの上で行い、GAUSSIANのような重たいプログラムを走らせる場合のみ、ホストコンピュータを利用するようになってきました。単独のソフトをワークステーション上で使用する場合は問題はないのですが、複数のプログラムを使用するには、ファイルの変換などが必要となってきます。ホストコンピュータでのプログラムを使用する場合は、更に面倒になります。しかしながら、近年、ホストコンピュータで作動するプログラムもワークステーションに移植される傾向があり、このような問題もしばらくすればかなり改善されるものと思います。

前にも述べましたようにソフトウェアが、毎年改良され使いやすくなってきていることは喜ばしいことなのですが、未だ、他のソフトウェアとの関係がうまくいかないことが多く、作業の効率化を妨げています。プラットフォームがワークステーション（UNIXマ

シーン)に、ほぼ統一されていることを考えると、データファイルのフォーマットの統一も、そろそろあっていいのではないのでしょうか。

「今後の方向」

コンピュータを利用したドラッグデザインに関して、今ある数多くの手法の細かいところまで議論できるほど、私自身知識も経験ありませんが、今までの数少ない経験から言えることは、ターゲットとする相手側の三次元構造がわかるとドラッグデザインが非常に効率的になる点です。確かに、ターゲットの三次元構造が明らかな、或は推測可能な例は極僅かであるとは思いますが、三次元構造を解析する努力を最大限行っていくことが、今後ますます必要になると思います。現在、いろいろと行われているコンピュータを利用したドラッグデザインに関する手法は、当然フォローしていかなければならないとは思いますが、それと同じくらいに蛋白質の三次元構造を決定する技術(X線結晶解析、NMR解析など)をいかに駆使できるかが大切な気がします。

また、三次元的定量的構造活性相関の手法として有名なCOMFA法は、最近いろいろな領域のドラッグデザインに適用されています。この手法は、特にターゲットの三次元構造がわからない場合には、有効であるようですし、将来において適用例が更に増えることによってより発展していくことと思います。

以上、思うままに述べてきましたが、合成する者の立場からいわせてもらえれば、効率化の面でも、またアイデアの糸口を提供するツールとしてもコンピュータは頼もしい存在になってきたというのが実感です。

今後、コンピュータを利用したドラッグデザインがどのように変わっていくのか期待をもって見ていきたいと思ひますし、その変化にあわせて実際の研究へ、どのように取り入れていくかを常に考え実践していきたいと思ひています。

〈大鵬薬品工業（株）創薬センター・多田幸雄〉

1 はじめに

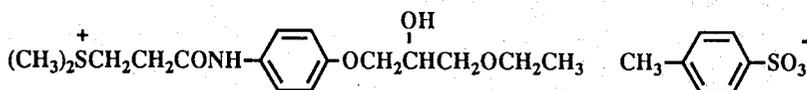
新薬の開発が成功するか否かは薬剤コンセプトと評価系に掛かっていると言っても過言ではない。「開発しようとする薬剤のコンセプトとそれを具現化する評価系」があって初めてドラッグデザインが可能になる。まず薬の開発のスタート時点において勝算のある薬剤コンセプトを設定できる創造的企画力が要求される。10年近い将来の医療ニーズを押し量ってコンセプトを決める訳で、リスクを背負っての決定になる。次に適切な評価系を構える必要がある。できる限りヒトの臨床に近い病態モデル、もしくはこれに準じる *in vivo* と *in vitro* の系を組み合わせて化合物が越えるべきバーを設定する。ここからドラッグデザインが始まる。評価系に関係する化合物などあらゆる情報と知識を基にリード化合物を決定し、毒性も考慮してリード最適マイゼーションを実行する。この過程のエキスパートシステムもいくつか出て来てはいるが、まだ新薬は手作りに近いと言える。しかしながらコンピュータを利用して論理的に新薬をデザインする事は可能と考えている。コンピュータの具体的な利用としては種々のデータベースの検索と構築、物性パラメータの算出、生体分子との相互作用シミュレーション、構造活性相関解析などが挙げられる。ここではコンピュータを利用するドラッグデザインの考え方と、できるだけ論理的に新薬を開発しようとした例を示したい。

2 構造活性相関を用いるドラッグデザインの考え方とその実例

まずドラッグデザインという言葉の意味としては、用いる方法に依らず「デザインした化合物から得られたデータが次の化合物のデザインに論理的につながれば良い」と考えている。ドラッグデザインのやり方として大きく二つの方法に分けることができる。一つはあたかも有機合成化学の知識と同じようにメディシナルケミストリーに関する化合物の構造変換知識を持っており、この知識を経験に照らしながら活用し、望ましい活性を有する化合物を作って行く方法である。この方法ではそれほどコンピュータを必要としない。もう一つはドラッグデザインに構造活性相関を利用する方法である。この方法では常に物性データと生物データとの相関を解析しながら化合物をデザインするため、その解析と予測過程でコンピュータの利用が必須になる。構造活性相関を利用した方がより論理的にドラッグデザインができるとの考えから、必然的にコンピュータを利用している訳である。

ドラッグデザインで構造活性相関を利用する目的は生物データを解析し活性を予測して化合物の構造展開を計ることである。生物データとしては病態モデルを含む *in vivo*、薬物動態（血中濃度、代謝、分布、排泄、タンパク結合など）、毒性と *in vitro* データなどがある。この中で *in vitro* データはモデル化された系に於ける生物活性であって薬物としてのポテンシャルを表す一種のパラメータである。したがってこの *in vitro* や薬物動態のデータをパラメータとして *in vivo* や毒性を説明できる場合も多い。以上のような化合物と生体との相互作用の結果生じる生物データと化合物の構造との相関を解明する手段として種々の物性パラメータを用いる。しかし、化合物のどのようなパラメータが何の生物データを解析できるかは個々に検討しなければ分からない。したがって可能性の高いと判断したパラメータから順次、生物データとの相関を検討する必要がある。決して易しくはないけれども、パラメータを検討することで生物データが少しづつ見えて来るのが

実感できる過程でもある。また、たとえ相関がなくても、あるパラメータに関しては化合物が非常に片寄っていることが分かる場合がある。化合物の姿をいろんな角度から理解することができ、構造展開の際に有用な情報になる。したがってできる限り多くのパラメータを検討しておく事は決して無駄にはならない。このような方法の例として現在申請中の抗アレルギー剤 IPD-1151T の開発当時に検討したパラメータを示す¹⁾。



IPD-1151T

1) 分配係数:

化合物が有機塩であるスルホニウム化合物であるため極めて水に解け易く、分配係数は非常に小さく、化合物間に大きな差はみられなかった。したがって効力との相関もあまりなかった。

2) Hammett σ

ベンゼン環の置換基はアルコキシ基のような電子供与基が効力を強めた。

3) 電荷

導入した置換基のベンゼン環の電荷を CNDO/2 で計算し毒性との相関を検討したが明確な解析はできなかった。

4) 分子の形と大きさ

スルホニオ基に続くメチレン鎖は2個が望ましく、枝分かればのぞましくなかった。

5) メチルプロトンの酸性度

アリル基のような不飽和結合を持つ置換基は顕著にメチルプロトンの酸性度を上昇させ、これに比例して急性毒性が強くなった。

6) メチル基の反応性

大部分の化合物がメチルプロトンの酸性度が大きい程メチル基の反応性は高かった。急性毒性の面からも反応性の大きくないジメチル基が望ましいと判断した。

7) カチオンの強さと広がり (電気泳動による移動度の測定)

分子内にカルボキシル基やアミノ酸残基があると分子全体のスルホニウムのカチオン性は低下し極めて低毒性となった。また CNDO/2 の計算で硫黄原子のチャージをメチル基の反応性に差がある化合物間で比較したがほとんど差はなかった。しかしスルホニオ基をスルホキシド基やスルホン基に変換すると全く効果を失った。従って広く置換基のプロトンに分布しているプラスチャージが抗アレルギー作用発現には必要と判断した。

8) 酸素原子の水素結合力

エーテル酸素を持つと抗アレルギー作用の他に抗腫瘍効果を示す化合物や、望ましくない副作用を発現する化合物が見られた。一方、ヒドロキシ基の存在は毒性の軽減に寄与した。従って一個のヒドロキシ基は残した。

9) イオンペアー形成能

スルホニウム化合物の塩の種類を検討し、イオンペアーを形成し易い程、吸収面と効力面で有利であった。

10) 液膜を用いた水-有機層の移動速度

イオンペアーとして移動していく速度が吸収や分布に関連すると考え移動速度の大きな塩として p-トルエンスルホン酸塩を選択した。

以上の物理化学的パラメータに基づいてできるだけ論理的に IPD-1151T の構造展開を計ったつもりです。これらの検討を実施したのは10年近くも前の事ですので量子化学計算と Hansch-Fujita 解析のための重回帰分析にコンピュータを使っただけです。現在ではコンピュータを身近に使い種々の計算化学の方法が発達し、今まで以上にドラッグデザインに利用できるパラメータが増え、メカニズムにまで迫ることができるようになりドラッグデザインに幅が出てきたと言える。さらにグラフィックスによる視覚化、相関解析の多様化、データベースの検索、エキスパートシステムなどをコンピュータは可能にし、ドラッグデザイン技術は年々向上している。これらの技術をいかに薬の開発の現場で応用して行けるかが今後の課題である。ここでコンピュータを使っただけのドラッグデザインの進め方のポイントと [] 内に利用している主なソフトウェアおよびデータベースを示す。

- 1) リード化合物の設定にはデータベースの検索を軸に関連すると考えた化合物群のエッセンスを基に化合物を合成し、テストスクリーニングを実施する。

[RIS (治験薬情報支援システム)、EMIL、CSD、PDB]

- 2) はじめは相関があると考えられる数種のパラメータに関して、一連の化合物で活性がなくなるまで幅広く化合物を合成する。いかに少しの構造の変化で活性が消失する化合物を見つけるかがポイントである。また活性のあった化合物にすぐには集中しないで同等の活性を持つ、できるだけ骨格の違った化合物を見つける努力をして、時間と人の許す限り複数のサブリード化合物を構える。

[QUANTA、Molgraph、CLOGP]

- 3) 生体高分子 (レセプター、標的酵素、標的ペプチドなど) の構造が利用できる場合は実験データ (NMR、X線結晶解析) との比較から化合物の活性構造を推定する。

[AMBER、CHARMm、NMR2]

- 4) 病態モデルや in vivo データを評価系として生物学的パラメータ (in vitro、吸収、代謝、排泄等) と物理化学的パラメータ (分配係数、電荷等) を検討し生物活性と相関のあるパラメータを見出す。特にひとつのパラメータを通して見ると構造式だけでは気付かない例外化合物を発見できる事があり、この情報を旨く生かす工夫をする。また理想的には発ガン性も考慮し化合物の最適化を目指す。

[代謝と発ガンのデータベースは検討中]

- 5) 化合物の多くの物性パラメータに起因して生体反応が惹起しているのでどのパラメータがその活性の主な原因となっているのか見出すのは難しい。そこで化合物の数がある程度揃ってくれば、グループ分けのできるパラメータを探す。そして各グループごとに生物データを説明できる新たなパラメータを検討してみる。旨くグルー

分けができていれば生物データを定量的に解析できる場合が多い。また種々の相関解析法も検討する。

〔Hansch-Fujita法、ALS、CoMFA〕

3 ドラッグデザインに構造活性相関を利用する問題点

ドラッグデザインに構造活性相関を利用する利点としては、論理的な構造展開ができる、例外化合物が発見できる、化合物の活性を客観的に評価できる、根拠をもって前臨床試験薬の選定できる、作用メカニズムの解明の一助となる、経験した事が客観的な知識データとして残り、参照できる等、数多く挙げる事はできる。しかし現状ではそれほど一般化していない。そこで構造活性相関を利用しない理由を列記してみる。

- 1) 構造活性相関を利用しなくてもドラッグデザインは十分可能である。
- 2) 構造活性相関は後付けの知識であって予測性は低く役に立たない。
- 3) 構造活性相関の解析に時間を取られるより少しでも多くの化合物を合成して生物データを見た方が効率が良い。
- 4) 相関のあるパラメータ見つける事のひとつを取っても結局、試行錯誤で効率が良いとは言えない。
- 5) 数式、物理化学的パラメータ、計算化学、多変量解析など出てくるので構造活性相関は難しそうで敷居が高い。

結局のところ化合物のデザインに携わる人が構造活性相関はドラッグデザインに役に立つと実感しない限りはこれを利用しないという結論になると思う。ただ私自身を含めてドラッグデザインに経験が浅い人程、構造活性相関は利用価値が高いと考えていますがいかがでしょうか。

4 今後のコンピュータ利用によるドラッグデザインの方向性

現在ではほとんどの研究所で何らかの分子設計システムを持っているのが普通に成ってきている。しかし、まだ現在の分子設計システムではドラッグデザインに利用するには機能が十分とは言えない。つまり今のシステムは薬物分子の作用のシミュレーションや各種パラメータの算出には使えるが、その後の生物活性との相関を解析する機能が不十分で、直接ドラッグデザインにつながっていない。しかし CoMFA や Apex などの様にこのあたりの機能が少しずつ付加されつつある。またドラッグデザインのための3次元化合物情報データの構築、代謝、発ガン、毒性のデータベースの利用も試行されている。^{2), 3)} したがってコンピュータ利用による論理的なドラッグデザインシステムとして実用化されるのもそう遠くはないと思う。

- 1) 多田幸雄、山脇一朗、松浦直資、安本三治、堀幹夫、江田昭英、第16回構造活性相関シンポジウム講演要旨集、p272 (1988)。
- 2) H. Waterbeemd, Quant. Struct. - Act. Relat. 11, 200 (1992)。
- 3) S. Borman, C&EN. AUGUST 10, 18 (1992)。

3 D データベースの現状

社団法人 化学情報協会
時実 象一

近年分子設計の有用な道具としての三次元座標データベースの需要が高まっている。従来は特定化学構造を三次元グラフィック表示するための計算の出発点として使われるのが主であったが、最近では三次元構造検索用のデータベースとしても広く使われるようになった。三次元構造検索システムについては別稿にゆずることにして、ここでは三次元データベースの種類について解説したい。総合的な解説としては Chem. & Eng. News の記事 1) が優れているので一読をおすすめする。

三次元座標データは、実測値のものと計算値のものにわけられる。実測値として有名なものは有機物質分野では Cambridge Crystallographic Data Centre (CCDC) の Cambridge Structure Database (CSD) と Brookhaven National Laboratory の Protein Data Bank (PDB)である。

CSD は有機低分子に関する世界最大の結晶構造データベースで、現在 100,000 件弱の有機化合物および有機金属化合物の座標データを収録している。データは学術文献から収録するほか、協力団体から直接入手する場合もある。CSD のデータで三次元構造検索をおこなうために CCDC では GEOM ソフトウェア²⁾ が開発されている。最近 CCDC が発表した提供ソフトウェアの Version 5 によれば、分子だけでなく結晶しての三次元構造検索が可能となっている。CCDC では MDL の MACCS-3D や Tripos の Sybyl/3DB Unity 用の形式も提供する予定である。対象化合物が必ずしも分子設計上有用な物質とは限らないが、部分構造的なデータ、すなわち最適な配座、原子間・分子間距離、といったデータも有用である。

PDB はタンパク質・核酸の事実上唯一の三次元構造データベースである。収録は学術文献からのほか研究者からの寄託データも多い。最近では X 線構造解析データのほか溶液中の NMR スペクトルからの解析データも収録されるようになった。対象が巨大分子であるため、CSD に比べて大幅にデータ数が少ないのはやむを得ない。しかし実際には全構造だけではなく部分構造をもよく利用することを考えれば、やはり大量のデータが収録されていると考えることができる。データ数の増加率が上昇しつつあるため、評価前データのオンライン提供、CD-ROM による提供も始まっている³⁾。

わが国でも日本結晶学会により X 線結晶構造データの収集がおこなわれており、データベースが日本科学技術情報センターの JOIS-F システムからオンライン提供されている⁴⁾。

なお無機結晶データとしてはドイツの GMELIN Institute と FIZ Karlsruhe が作成している Inorganic Crystal Structure Database (29,000 件) とカナダの NRCC Metals

Crystallographic Data File (CRYSTMET) (21,000 件)がある。これらは最近超伝導セラミックなどの出現により注目されている。両者とも磁気テープとオンライン (STN) で提供されているが、最近 CD-ROM での提供も始まった。

上記実測値データは信頼性は高いが、対象化合物は分子設計を念頭において選択されている訳ではない。分子設計担当者としては、医薬・農薬などの候補化合物の構造が必要である。そうした要求を満たすために計算値のデータベースが登場した。多くの場合、この座標計算は Tripos 社の CONCORD プログラムを用いていることが多い。もちろんこのような計算は利用者が自分でおこなうことが可能であるが、数万の化合物のデータを自分で生成することは実際上困難であると思われる。

これまでに使われてきた計算値の三次元座標データベースは MACCS-3D 用のものが多い。例としては Fine Chemicals Directory-3D (60,000 件)、Comprehensive Medicinal Chemistry-3D (6,000 件)、MACCS-II Drug Data Report (16,000 件) などがある。これらは既存の MACCS 用データベースの二次元結合表から CONCORD で計算したものであり、計算値の信頼性はばらついていると考えられる。

Chemical Abstracts Service (CAS) は従来 CONCORD で生成した 4,500,000 件以上の座標データを CAS ONLINE の Registry ファイルで提供してきた。このデータは STN Express や Kermit を用いることにより、ダウンロード可能である。今回この 4,500,000 件のうち選択された座標データ 370,000 件を CAST-3D の名前で磁気テープ形式で提供することとなった^{5,6)}。対象化合物は Registry ファイル中の構造が堅い (rigid) 化合物、すなわち配座の柔軟性がすくない物質であり、CONCORD の計算値の信頼性も高いと考えられる。配座の柔軟性の決定は CAS が開発した柔軟性指標 (CAS Topological Flexibility Index) によった。収録されている化合物の種類はエキソ結合を持つ環系 (72%)、その他の環系 (26%)、非環系 (2%) となっている。したがってこのデータベースは分子設計の基本骨格を提供するものと考えられる。実際 CAST とは CAS Template file の意味で、ユーザがこのデータを分子設計の鋳型として使ってほしいという意味である。使い方としては単純な基本骨格としての利用のほか、これを用いて活性点に結合する鋳型を検索することなどが提案されている。370,000 件の化学構造のうち、各利用者が必要なものだけを二次元構造検索などで抽出してライセンス利用することが可能である。

このように実測・計算を含め、分子設計のためのデータベースが大いに入手しやすくなった。今後はこれらの活用方法が発展していくものと思われる。

[参考文献]

- 1) Borman, Stu, "New 3-D Search and De Novo Design Techniques Aid Drug Development", Chem. & Eng. News, August 10, 18-26 (1992).
- 2) Allen, Frank H.; Davies, John E.; Galloy, Jean J.; Johnson, Owen; Kennard,

Olga; Macrae, Clare F.; Mitchell, Eleanor M.; Mitchell, Gary F.; Smith, J. Mechael; Watson, David G., "The Development of Version 3 and 4 of the Cambridge Structural Database System", J. Chem. Inf. Comput. Sci., 31, 187-204 (1991).

- 3) Koetzle, Thomas F., 「タンパク質データベース-生体高分子の立体構造の総合データベース」、第 8 回化学情報講演会、1992. 10. 16., 東京 (化学情報協会主催)。
- 4) 早瀬修一、岡野恵子「JICST ファクトデータベース JICST 結晶構造データベース」、情報管理 33(4), 319-331 (1990).
- 5) "CAST-3D - Background Material", Chem. Abstr. Serv., 1992. (化学情報協会から入手可能)
- 6) Fisanick, William; Cross, Kevin P.; Rusinko, Andrew, III, "Characteristics of computer-generated 3D and related molecular property data for Chemical Abstracts Service registry substances", Tetrahedron Comput. Methodol., 3(6c), 635-652 (1992).

1. はじめに

企業での研究内容や活用事例を具体的には話しにくい。それを技術力のなさの隠れ蓑としていては、技術と製品の入超の典型である日本の製薬メーカーはいつまで経っても欧米には追い付けない、と個人的には思っている。だからと言って、「ハードやソフトにこのようなものがあります。」との、通り一辺の紹介だけでは、部会員や編集委員の皆様には申し訳ない。そこで、執筆を受けたからには所属とは関係なく、CADDについて個人的な意見を雑多に述べたい。

老人性痴呆、癌、エイズのようにリードとなる化合物がない領域での新薬作り（創薬研究）は、製薬メーカーで研究する者にとって大きな夢であろう。その中で、コンピュータによる医薬品設計（CADD）を担当する研究者は、その基となる化合物が設計できれば最高の栄誉であろう。そこまでできなくとも、すでにある薬とは全く構造を異にする化合物が設計できれば、それだけで大成功といえる。最近は特に合成以外の現場組織が、CADDに対してこのような熱い期待をかけてきている。

私がコンピュータの世界に足を踏み入れた動機は、新しい構造の化合物を設計するには新しい手法を取り入れることがベストである、と思ったからである。薬の設計では、ちょっとした構造変化で活性や安全性等が大幅に向上したりするので、良く使われる回帰式での解析は必須である。だが、この方法でどこまで新しい構造の化合物に変化できるであろうか。

2. 医薬品設計研究の現状

化合物の構造パラメータについて、例えば分子の長さや運良く縦軸である活性との相関が見い出されたとする。そのデータをもとに基本骨格まで変えた斬新な薬物の設計は難しいであろう。特定原子の電荷とか置換基の脂溶性（ π ）とかいった部分的な構造情報では勿論、logP（分子の脂溶性を表す実験値）などの化合物構造全体を表すパラメータでも、斬新な構造に展開できる可能性は小さい。それでも、苦勞して得た回帰式を用いて論理的に構造を変換させ、設計した化合物に期待した活性や物性が得られれば、CADD研究としては成功である。まして、それが製品となろうものならその喜びは大きい。勿論、化合物の設計とまでできなくとも構造相関の論理的説明ができただけでも研究としては成功と言える。しかし、その成功を担当者として本当に満足できるだろうか。

当然であるが、この回帰法では合成された化合物の構造式とその活性データなどがなければCADD研究は始まらない。言い換えれば、CADD研究は既存化合物のマイナーな変換が主体となり、皆が期待している斬新な構造への変換や、全く新規なリード化合物の設計は非常に難しい。だが、それをCADDの問題とか力のなさや片付けでは、担当者にはあまりに酷であろう。薬物設計でもっとも重要な情報である薬が作用するメカニズムについて、人類はどこまで解明できているであろうか？

消化性潰瘍は分子レベルで作用メカニズムが解明され、酸分泌をおさえる選択的なヒスタミン-IIレセプター (H2) 拮抗作用を持った化合物がCADDで研究されている。しかし、これらの化合物の活性コンフォメーションですらいまでも論議的である¹⁾。H2レセプターの構造を推定した報告があるが²⁾、そのレセプターの構造がアミノ酸配列から側鎖の原子レベルまでわかった時、薬の本当の活性コンフォメーションが推定できる。

一般論として、病気のメカニズムの解明なくして論理的な薬物設計は難しい。このような中で新しい薬物を発見しようとする、合成研究の力に頼る方法に重点を置かざるを得ない。もし、薬物が発見できたとしてもそれは運がよかったと言うべきである。生体の仕組みは難解であり、それを解明するのは科学技術の総力をあげた戦いでもある。その結果として斬新な薬が生まれてくると思う。

コンピュータを導入すればすぐにいい薬ができると思っている人はもういないであろうが、訳のわからないところ (病気のメカニズム) で訳のわからないもの (化合物) をコンピュータに入れても、いい答え (治療薬) はでてこない。だが、これをCADDの言い訳とはとって欲しくはないし、担当者はそれに甘えてはいけない。

3. 期待されるCADD

本当にCADDで斬新な薬物設計ができるのであろうか。1991年、薬物設計のすばらしい成功例がAgouron社から報告された³⁾。病気の作用メカニズムからターゲットとなる酵素を決め、その活性部位 (酵素のレセプター部位) の立体構造から既存の化合物とは全く異なる構造の化合物を設計し (図-1)、それが期待できる活性を持っていた。「構造が大きく違うからいい薬」とは臨床試験の結果がでるまで言えないが、非常に大きな期待が持てる。このことがCADD研究にとって重要だったのは、レセプターの構造さえわかれば、既存の化合物がなくとも、薬の設計が精度良くできる (de novo design) ことが見事に実証されたことである。逆に、コンピュータでなければこのような化合物を設計できなかったのである。

この成功はCADDの進歩だけがもたらしたのではない。基盤技術 (バイオ、分析機器、合成、評価等) の発達のおかげで、病気のメカニズムからターゲットとなる生体物質を原子レベルまで解明することが現実となってきた。これでコンピュータによる薬物設計が論理的により効率的に進めることができるようになったと言えよう。

それでは、レセプターの立体構造から新しい構造の化合物をコンピュータで作れるのであろうか。化合物の三次元構造データベースからレセプター部位に当てはまる化合物の検索法⁴⁾ や、レセプターの立体構造からの化合物自動設計法⁵⁾ など、まさに、コンピュータならではの新しい設計法が報告されている。コンピュータから出力された化合物は恐らくいろいろな問題 (合成、安定性、特許性等) があると思われる。そのときは、実験データがほとんどないなかでの化合物のさらなる構造変換が、CADD担当者の知恵のだしどころである。この時こそ、CADD技術が本当に頼られる最高のときであらうが、そのためにも常にCADD技術に磨きをかけセンスを培っておきたい。

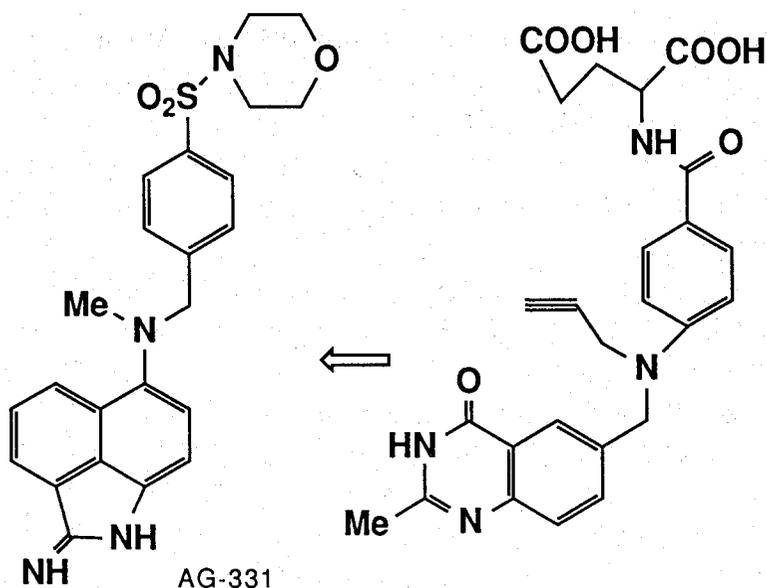


図-1、斬新な薬物設計 (de novo design) の成功例
 (レセプターの立体構造から設計したもので、
 右から左への化合物の構造発展には繋がりはない)

4. ソフト会社をお願いしたいこと

最近のハードの進歩、価格の低下はユーザーにとって非常に好ましいことであるが、保守料はついにハード/ソフトが逆転してしまった。ワークステーションはソフトの互換性がいいことが売り物であったが、市販ソフトを使っているとそれは幻想でしかすぎない。グラフィック関係ソフトであれば誰かが使用中であれば、すぐに使えなくとも我慢はできる。しかし、バッチ計算は時間がかかるので、できる限り空いているコンピュータを使って効率良く仕事を進めたいと思うが、ライセンスの問題で、隣に同じ機種のコピー機があっても計算には使えない。ソフトライセンスをすべて一律に扱うのではなく、特にバッチ計算部分は2台目以後でもライセンス料は格安になるよう契約内容を是非見直して欲しい。計算が勝負のCADD研究の発展はソフト会社の発展でもある。

5. CADD研究における法的保護の問題

最後に、将来的にCADD研究での活性予測精度が向上したときに考えられる、法的な問題について述べたい。現在の物質特許制度では、合成された化合物もしくはそれに準ずる化合物だけが、法的な保護を受けられるようになっている。薬の歴史を考えれば、生薬からの抽出物や生理活性物質、偶然発見された活性物質の化学修飾が薬作りで、その中での物質特許制度である。従って、現在の制度の元では、CADDの技術を向上し活性予測精度の良い化合物が設計できるようになっても、合成されなければそれらは日の目を見ることなく葬り去られる運命にある。もし、設計化合

物の構造が学会等どこかで報告されたら、物質特許が取りにくい化合物としてのよくない烙印を押されてしまう。

生体成分のもっとも上流に位置する遺伝子構造は、その機能がはっきりしたものであれば、特許として認められる方向にある。遺伝子から作られるレセプターや蛋白などの難しい立体構造を解明し、その構造を基にして薬物を設計しても法的保護はない。そのため、薬作りに重要なデータの公開を遅らせたり、他社での特許取得を妨害するために設計した化合物を多数発表することがありうるであろう。このことは人類にとって果してプラスなのだろうか。遺伝子から病気のメカニズムの研究が始まり、薬の候補として化合物が決まるまでのどこかの段階、例えば設計化合物特許のような法的保護が受けられるようになると、企業の研究投資の姿勢に大きな変化をもたらすであろう。素人的な発想で問題点は多々あり、これについてはいろいろな面から検討すべきであると思う。しかし、生薬成分のものまねの'もの作り'から、病気のメカニズムからのより科学的な精度の高い薬作りが可能となってくると、物質特許だけでない法的保護制度を考えることは、人類にとって有益なことである、と思うのは私の独断であろうか。

[参考文献]

- 1) H. Höltje, and A. Batzenschlager; *J. Computer-Aided Mol. Design*, 4, 391-402 (1990)
- 2) I. Gantz, J. DelValle, L. Wang, T. Tashiro, G. Munzert, Y. Guo, Y. Konda, and T. Yamada; *J. Biol. Chem.*, 267(29)20840-20843 (1992)
- 3) (1) K.A. Applet, R.J. Bacquet, C.A. Bartlett, C.L.J. Carol, S.T. Freer, M.A.M. Fuhry, M.R. Gehring, S.M. Herrman, E.F. Howland, C.A. Janson, T.R. Jones, C. Kan, V. Kathardekar, K.K. Lewis, G.P. Marzoni, D.A. Matthews, C. Mohr, E.W. Moomaw, C.A. Morse, S.J. Oatley, R.C. Ogden, M.R. Reddy, S.H. Reich, W.S. Schoettlin, W.W. Smith, M.D. Varney, J.E. Villafranca, R.W. Ward, S. Webber, K.M. Welsh, and J. White; *J. Med. Chem.*, 34(7)1925-1934 (1991).
(2) M.D. Varney, G.P. Marzoni, C.L. Palmer, J.G. Deal, S. Webber, K.M. Welsh, R.J. Bacquet, C.A. Bartlett, C.A. Morse, C.L.J. Booth, S.M. Herrmann, E.F. Howland, R.W. Ward, and J. White; *J. Med. Chem.*, 35(4)663-676 (1992)
(3) S.H. Reich, M.A.M. Fuhry, D. Nguyen, M.J. Pino, K.M. Welsh, S. Webber, C.A. Janson, S.R. Jordan, D.A. Matthews, W.W. Smith, C.A. Bartlett, C.L.J. Booth, S.M. Herrmann, E.F. Howland, C.A. Morse, R.W. Ward, and J. White; *J. Med. Chem.*, 35(5)847-858 (1992)
(4) 化学と工業、とびくす、45(9)1721-1722 (1992)
- 4) E.C. Meng, B.K. Shoichet, I.D. Kuntz; *J. Computational Chem.*, 13(4)505-524 (1992)
具体例：(1) R.L. DesJarlais, G.L. Seibel, I.D. Kuntz, P.S. Furth, J.C. Alvarez, P.R. Ortiz de Montellano, D.L. DeCamp, L.M. Babe, and C.S. Craik; *Proc. Natl. Acad. Sci. U.S.A.*, 87, 6644-6648 (1990)
(2) Y.C. Martin; *J. Med. Chem.*, 35(12)2146-2154 (1992)
- 5) (1) Y. Nishibata, and A. Itai; *Tetrahedron*, 47, 8986-8990 (1991)
(2) A. Inoue, T. Kanazawa, and A. Itai; 第19回構造活性相関シンポジウム講演要旨集、316-319 (1991) 川口
(3) T. Kanazawa, A. Inoue, and A. Itai; 第20回構造活性相関シンポジウム講演要旨集、258-261 (1992) 京都

1. はじめに. 周知のように医薬品の研究開発は長期間のプロセスをたどる。製品の社会的責任や使命から当然かもしれないが、俗に「研究開発に十年」と言われるのは決して誇張した表現ではない。CADD（コンピュータ支援ドラッグデザイン）の手法が適用されるのは、その長い研究開発期間のなかで前臨床のさらに基礎研究の段階であり、製品のライフサイクルを考えるといかにも短い。にもかかわらず、研究開発の効率化に対するこの分野への期待が非常に大きいのは各社に導入された高額なモデリングシステムの数を見れば一目瞭然である。ともあれ、殆どの製薬企業が多かれ少なかれコンピュータによる分子設計技術の応用を試みている筈であり、現在はその成果が現われ始めているか、ないしは成果をあげるべく奮闘している時期とってよいだろう。標題のテーマについての適任者は他の部会員の中に数多くおられると思うが、僅かな経験ながら以下にコンピュータを利用した分子設計に携わる者の一人として当社の現状と感想を簡単に述べてみたい。

2. 分子設計のためのコンピュータ環境. 人的資源を含めたコンピュータ環境の整備と組織づくりは管理職ならずとも興味あるところであろうし、その運営体制次第では当該業務の成否に関わるといっても過言ではないであろう。

1) ハードウェア・ソフトウェアの環境. 下図に分子設計支援用に筆者らが用いているハードウェアおよびソフトウェアの環境を示した。端末としてのパーソナルコンピュータの類やQ C P Eのプログラム群は頻繁に利用しているが省略してある。特に取り立てて目新しい構成でもなく既にシステム構築をされている方が見ても参考になるところはないと思う。唯、データベースマシン（VAX）と計算化学用（ワークステーション）とを分散してネットワーク上に配置し必要に応じてコミュニケーションをとる方式を基本としている。ワークステーション同志はNFSによりディスクの共有を計り、分子軌道法や分子動力学計算等のCPU負荷の高いプロセスはいずれのワークステーションでも実行可能な状態になっている。

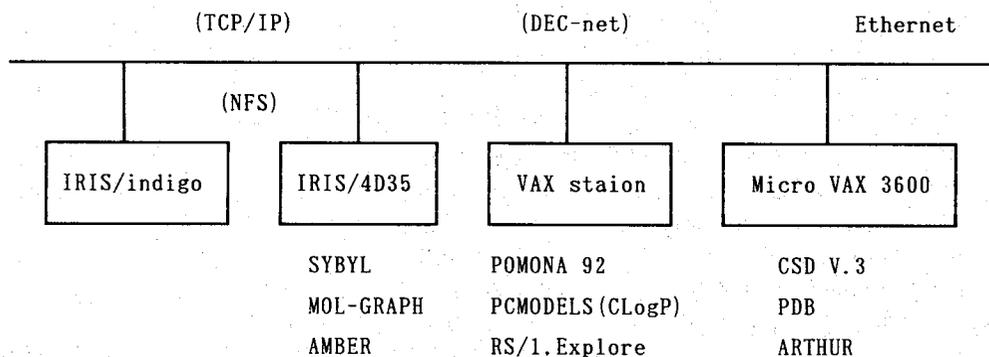


図. 分子設計のためのコンピュータ環境

ソフトウェアの構成は米国製のモデリングシステムを中心に据え、その回りにClogPのような物性推算や統計解析のプログラムを用意している。ケンブリッジの結晶構造データベース(CSD)やProtein Data Bank(PDB)もモデリングシステムを補完する重要な役割を果たしている。

この形態は殆どの企業の研究所が採用しているように思われる。もちろん企業規模の大きさに比例してハードウェアの規模や数及び導入しているCCSパッケージの種類が多いと推察できるが。

2) 運営体制。 上記システムの運用は筆者が所属する研究所のコンピュータグループと合成部門に配属している専任の担当者と共同で行なっており、組織的には暫定的な処置と考えている。コンピュータの操作上、Hansch-藤田法や簡単な分子のモデリングは実験者にも容易であるから「解析ができる合成化学者」を増やしていくのも必要であろうが現時点では専任者が行なう体制をとっている。CADDグループのあり方は各社悩みの種かもしれない。個人的には、ある程度コンピュータ環境が整ったら計算機化学者を中核として実験者を含めたプロジェクト的な組織に移行するのがベストと思っている。いずれにせよ実験化学者との緻密な協力関係が保てる配置が必要なのは言うまでもないだろう。

3. システムの評価。 当社ではCADDを用いてデザインされた薬剤が臨床試験に進んだというような成功例は未だ無いが、様々な分子計算がルーティン的に行なわれるようになってきた。構造式と活性値のデータセットからQSARを行なうための分子モデリングや物性計算の他に、有機反応機構の解析や特定部位のアミノ酸を置換した蛋白質のコンフォメーションの計算などもそのなかに含まれる。QSARモデルから活性を予測した新規化合物が合成されるケースも増えてきており、これも日常的な風景になるのであろう。このような状況は徐々にではなく最近数年で起きてきた変化であり、市販の分子設計支援システム(CCSパッケージまたはモデリングシステムと同義)と高速のグラフィックワークステーションの導入が契機となっているのは間違いない。

80年代の前半から国内で使われ始めた欧米のモデリングシステムは最近、益々機能が追加・拡張されかつ洗練されてきた。ほんの数年前までのQCPEプログラムと統計パッケージだけを組み合わせ使用していた頃を思うとまさに隔世の感がある。以前はQCPEから取り寄せたプログラムを動かすのですら研究所のコンピュータ業務を統括する部門の協力が必須であり、まして多少とも統合的なシステムを意図するならそれはコンピュータグループそのものの仕事であった。85年、米国の製薬会社G.D.SearleのA.J.Hopfinger(現在はイリノイ大)はJ. Med. Chem.の総説¹⁾の中で将来「製薬会社の中でCADD用ソフトウェアを開発しているグループは市販パッケージの管理ユーザー(caretaker user)になるだろう」と予測した。たまたま当社ではエンドユーザーを兼ねているに過ぎず、そのとおりになっていると云わざるを得ない。ここで市販のプログラムを使う是非を論じるつもりはないが、読者の中には化学の理論がブラックボックス化することを懸念する向きもあるかもしれない。特に多少なりとも化学ソフトウェアを開発ないし改良の経験がある年配の方はそう思うのではなからうか。斯言う筆者もCCSを導入した頃はそれを嬉々として使う若手を見ながら「これでは本当に理解していることにはならない」などと独り言をつぶやいたものだった。しかし、今では使い込む毎にその出来栄と欧米の

モデリング業界の実力に驚き、資本が投入されるということはこういうことかと感心するばかりである。

市販モデリングソフトは各社それぞれ特徴があり、どの機能を重視するかによって選択の条件が決まる。例えば、力場の信頼性やPLSに代表されるChemometricsの機能、グラフィックスや総合的な操作性など。筆者らは立体配座解析と3D-QSARを最終的な判断材料とした。自明のことだが分子設計とは分子設計用のソフトウェアを開発することではない。目的に叶った、信頼の置けるソフトウェアを選び、背景となっている技術や理論を使いこなすことに時間を費やすべきと考える。今はそれができる時代なのだから。

4. 今後の方向。 お叱りを覚悟で言えば、CADDの要素技術あるいはもっと狭義に（定量的）構造活性相関のための個々の基礎的な方法論はほぼ出尽くし、今後は得られた3D-QSARモデルやPharmacophoreモデルに適合する新規な化学構造式をデータベースから発見する試みと立体化学的形狀や物理化学（物性）的な拘束条件のもとでの構造創出の自動化に興味は移るものと思われる。

前者は3-D Searchと呼ばれる一種の構造検索手法であり、分子の三次元的形状が質問条件ということになる。相応の数の三次元構造が格納されたデータベースが必要なことと、柔らかい分子の取り扱いに難があるが実用の域に達したとっていいだろう。当社でも導入の検討を開始したが、CSD以外に既に3つのシステムが国内の代理店を通じて入手可能である。これにたいして後者はDe Novo Designと呼ばれる。受容体の形状が既知かどうかは問題ではなく、あくまでリガンドの候補構造の自動発生を目的にしているシステムである。データベースを必要としないことや極めて特異的な構造が得られる可能性があり魅力があるが、容易に想像できるように候補構造の数が膨大になるのをいかに絞るかが鍵になるであろう。こちらは最近になってやっと市販品が現れた。できれば3-D SearchとDe Novo Designを相補的に使ってみたいと思う。

今までのCADD技術は「化学構造と生物活性の相関を解析してそのcriteriaを確立する」ためのコンピュータ手法を意味していたのに対し、これらの新しい技術はリード化合物を含め新規化合物をコンピュータが発見的に呈示する可能性を秘めている。国内の製薬企業では未だいずれも導入されていないようだがそれも時間の問題だろう。

最近、C & ENにこれら2つの新しいCADD技術の詳細な特集記事²⁾が掲載された。興味ある方は参照されたい。

5. おわりに。 完成度の高いモデリングシステムと高速のワークステーションが各製薬会社に行き渡り、所有しているコンピュータシステムでは差別化ができない時代になっていると言って差し支えあるまい。CADDの手法を適用するための技術を磨くことと同時に、CADDを通じて合成部門や生物検定部門とのより一層の協力関係が求められているのであろう。

1) A. J. Hopfinger, J. Med. Chem., vol. 28, 1133 (1985)

2) S. Borman, Chemical and Engineering News, Aug., No. 10, p18-26 (1992)

第 15 回情報化学討論会報告

11月25日(水)から11月27日(金)までの3日間、京都市左京区の京都大学薬学部および京都教育文化センターにおいて、第15回情報化学討論会が第20回構造活性相関シンポジウムと合同で開催された。両会場間の距離は徒歩にて約4分である。準備全般、特に教育文化センター会場および懇親会場の設営には、今回の構造活性相関シンポジウム世話人京都大学化学研究所西岡孝明助教授の全面的なご協力を頂いた。情報化学討論会の一般講演(25件)には京都大学薬学部記念講堂、ポスターセッション(19件)には京都教育文化センター302号室、特別講演(1件)および構造活性相関シンポジウムの一般講演(30件)には京都教育文化センターホールを使用した。また、昨年度と同様、パソコン、ワークステーションの展示および関連ソフトの実演(8社協賛)が会期中を通じ、京都教育文化センター202、203号室にて行われた。

構造活性相関シンポジウムの講演申込件数は昨年度と同様であったが、情報化学討論会の講演件数は昨年度より10件少なく、両者合同の参加者数も予約、当日を合わせ360名と大幅に減少した。その原因は不況による企業からの参加者の減少および同系統の学会が同時期に集中したためであり、したがって一過性の現象と考えたい。毎回参加される熱心な支持者の数が減っていないことは、懇親会参加申込者数が120名とほぼ予想通りであったことから明らかである。講演件数が少なかったため、両討論会とも、講演時間18分、討論時間7分とし、2件ないし3件ごとに休憩を入れる余裕のあるプログラムを組むことができた。一方、ポスターセッション開催中は両討論会とも一般講演なしとしたので、展示会場とともに大盛況となった。結果的に会場が狭すぎる感じを与えたのは否めない。

特別講演は今回は一件のみとし、京都開催の機会を活かして、財団法人基礎化学研究所長福井謙一先生にお願いしたところ、幸いにもご快諾を得、「化学反応の一般化」と題して量子力学による化学反応の捉え方のポイントを解説していただくことができた。ご講演の締めくくりとして、情報化学部会長大沢映二教授によるサッカーボール分子 C_{60} の存在予言に言及され、先生みずから高く評価されたことは、聴衆の部会員諸氏にとって大きな刺激となったと思われる。企画した者として喜びに堪えない。

第3日全講演終了後、大沢部会長の司会で、「情報化学部会の今後」と題するパネル討論が行われた。部会員数の増加、部会活動の活性化など、かねてからの懸案にとって、現在の部会名が適切か否かについて、それぞれ賛否両論の立場からの討論(トップは部会の生みの親である米田幸夫東海大学教授による現部会名肯定論)があった。フロアからの発言も多数あり、この問題はさらに時間をかけて検討されることになった。

今回は徳島大学工学部月原富武教授のお世話により、同大学薬学部長寺田弘教授の世話される第21回構造活性相関シンポジウムとの合同で、徳島で開催されることとなった。参加会員数が再び上昇することを切望する次第である。

(京都大学薬学部 町田勝之輔)

25I01 "Partial-heat molecular dynamics"

立体配座解析の効率的な手法の開発

(北里大薬) 辻下英樹、広野修一、森口郁生

立体配座解析を効率的におこなえるPHMD、PAMPER法をあらたに開発し、その解析例を報告している。あらかじめ、エネルギー極小点がわかっているコンホメーションに対して計算すれば、それが極小点と判断できるが、そうでないときに、それがグローバルミニマムであると判断できるだろうか。ノーマルMD法でも、適当に初期値を設定すれば、その「極小点」を計算できる可能性があるのではないか。パラメータが多いときは、いづれにしても、初期値に依存した結果が得られると思います。

25I02 分子動力学シミュレータのためのワークベンチ: MASPHYC

(富士通) 酒井広太他

分子動力学計算をするのに、その系にとってに適切なポテンシャル関数とパラメータを選択する必要がある、そのためのインタラクティブなエディタを開発したことを報告している。原子間相互作用において、ペアポテンシャル、ペア汎関数、クラスターポテンシャル、クラスター汎関数とパラメータをライブラリー化することにより、シミュレータを使いやすくしている。

(分子) 藤田 忠男

25I03 Ab initio 分子軌道法による不斉フッソ原子をもつ強誘電性液晶の

モデリングおよび静電ポテンシャルと物性との関係

(分子) 藤田 忠男

ジアステレオマーA、BのAb initio 計算についての報告である。Ab initio 計算はGAMESSで、基底関数 6-31G* を用いて行い、双極子モーメントと静電ポテンシャルが求められた。その結果を使い分子構造と分子物性の関係が調べられた。液晶材料の物性(屈折率、自発分極)と分子物性との間の関係などについての質問・討議がなされた。

25I04 ベクトルプロセッサとパラレルIOを利用したオキシヘムの大規模

CASSCF計算

(中京大教養) 山本 茂義, (九工大情報工) 柏木 浩

分子軌道法の一方法であるCASSCF法の大規模計算についての報告である。前半で、CASSCFの方法とオキシヘムの計算例が述べられた。後半では、パラレルIOと拡張記憶を用いてCASSCFが高速化できることが報告された。パラレルIOと拡張記憶の効果についての質問・討議がなされた。

(富士通・三上益弘)

25I05 メリチンの高次構造形成機構 (東京農工大工) ○園山正史、安川民男

25I06 球内残基数を用いた蛋白質立体構造予測の試み(2)

(旭化成) ○鎌田朋子、戸潤一孔

25I07 類縁蛋白質のモデリング法の進展

(北里大・薬) 山下憲文、梅山秀明

セリンプロテアーゼ類における側鎖の2面角の保存性に関する報告である。 x_1 および x_2 は約50%の高い保存性を示し、保存されている残基や芳香族・複素環アミノ酸残基ではさらに高い保存性があるということである。Ramachandranプロットのように x_1 、 x_2 をプロットしてみるとどのような結果が得られるのかなと思った。

25I08 タンパク質受容体部位の自動推定: 静電解析に基づく候補部位の評価

(豊橋技科大) Carlos A. Del Carpio, 高橋由雅, 佐々木慎一

タンパク質受容体における、仮想リガンド分子の位置を自動推定するためのアルゴリズムの開発とその応用例の報告である。幾何学的なフィッティングにより得られるタンパク質ポケット領域の候補から、仮想リガンド分子との間の静電的な相補性の解析に基づき最も妥当な受容領域を推定する方法を取っている。会場からは、非常に野心的で、興味深いとの発言があった。

(東京農工大工 園山 正史)

DNAを基礎にした蛋白質一次シーケンス解析

【25I09】(北里大薬) ○孤岡仁志、梅山秀明

コドン在前面に出したアミノ酸配列解析手法に関する報告。3個の内、任意の2個が一致するという単純なルールでアミノ酸での取扱と大体同じ結果が得られるとのこと。計算上は問題が無いということだったが、個人的には、塩基配列とアミノ酸配列を複合した形式に冗長な印象を受けた。会場からのコメントもあったが、おおもとのデータベースに含まれる誤植等はいつも悩ましい問題である。

アンチセンス分子設計のための核酸三重らせんのモデリング法

【25I10】(東大薬) ○富岡伸夫、板井昭子

核酸トリプレックスのモデリングに関する報告。糖部分について8つの立体異性体を検討し、既知のもの以外にも安定構造を作る可能性があるものを見出したとのこと。会場からの質問は、安定化エネルギーを分子単独で見積るのは問題では無いかという点であったが、講演者も触れていた様に、水分子やイオンの影響まで考慮した計算は時間が掛かる為、その前の検討手段としてある程度の単純化は仕方の無いところであろう。

(旭化成・CS室 戸潤一孔)

25I11 DNAモデリングのエキスパートシステムの進展

(北里大薬) ○四宮和子、梅山秀明

25I12 大腸菌とバクテリオファージのゲノムコード領域における σ^{70} -プロモータ類似配列の抑制性の相違 (山形大工) ○金谷重彦、工藤喜弘

26I13 延伸による結晶性高分子の挙動と弾性率

(呉羽化学) ○栗原あづさ、中馬 寛、榎本 聡

26I14 木構造間の距離と化学への応用 (神大工) ○田中栄一、増田澄男

26I15 有機構造解析の考え方とその実践 (山形大工) ○工藤喜弘

26I16 化学名(体系的名称)から平面構造式を生成するシステムの構築

(その2) 立体化学の認知と表現

(中外製薬) 松浦育敏

IUPACおよびCAの命名法による化学名から平面構造式を得るためのシステム開発についての昨年に続く報告であり、立体化学表現を正しく平面構造式に表現するためのアルゴリズムが考案されている。本研究に関連した報告は本討論会では比較的多く(例えば、藤原議ら(第7回)、鈴木隆弘ら(第9回)、阿部英次ら(第12回)、米田幸夫(第13回))、"不斉"の定義をめぐってこれらの研究に関わっている研究者らから活発な討論がなされた。このように多くの研究者が関わっている問題をCAではどのように対処しているのでしょうか。

26I17 創薬と安全性研究のための分子計算環境MOSAIIの開発

(国立衛試・群馬大・ア・イ研) 神沼二真、中野達也、中田琴子、五十嵐貴子、中田吉郎、浜谷和彦

本講演は、パーソナルコンピュータとワークステーションをネットワーク化し、その上にEMBL、GENBANK、PDBなどの各種データベースおよび分子計算ツールを統合的に利用できる環境を構築したということであるらしい。レクチャーに関する知識を蓄えることの重要性が示唆されているが、その実現は構想の段階にあるとのことである。質問およびコメントがこの点に集中してなされた。

26P01 分子動力学シミュレータのためのワークベンチ：MASPHYC

(富士通) 並木武文他

成膜シミュレーションのためのヒューマンインタフェースを開発したことを報告している。このような、ヒューマンインタフェースの評価は、多くのユーザの意見を聞かないとできないでしょう。実際のユーザの意見を取り入れれば、さらに優れたGUIができると思います。実際の材料開発に役立つという実績が必要になります。

26P02 長距離相互作用を考慮した2段階cutoff-MD計算法

(北里大薬) 山乙教之、広野修一、森口郁生

分子動力学計算における長距離相互作用を計算する際のcutoff半径を中心残基が電荷をもつときもたないときで、その値を第一、第二のふたつのcutoff半径を使う計算法を発表している。一段階cutoffを使う計算での自由エネルギーのデータとの比較において、その理論的説明にまだ曖昧さが残っているように思われます。

(分子) 藤田 忠男

26P03 分子動力学を用いた分子内環化反応の位置選択性に関する検討

(旭化成) 田頭瑞加, 戸潤一孔

アセチルグリシンをオリゴペプチドを介してアントラキノンに結合させた分子(AGAQ)の分子内環化反応の位置選択性に関する報告である。オリゴペプチドの長さとそれを構成するアミノ酸の種類異なる六種類のAGAQの分子動力学計算を行い、各AGAQそれぞれの100個の最適化構造について反応部位間の距離を求め、分子構造と反応選択性の関係が調べられた。

26P04 MNDO計算によるコラーゲン関連ペプチドの安定構造と

そのX線解析およびNMR解析

(近大理工) 高橋紘一, 山口仁宏, 茨木雅人

コラーゲンの繰り返し単位であるBoc-L-Pro-L-Pro-GlyのX線構造解析と分子力学/分子軌道法計算についての報告である。X線構造解析から得られた安定構造が、半経験的分子軌道法でも再現されたことが述べられた。このことから計算化学の手法による生体高分子の構造予測の可能が議論された。

(富士通・三上益弘)

ANALOGSのニューラルネットワーク・システムによるレセプター・マッピング
(第4報) (東京農工大工・大正製薬・阪薬大) ○山野辺進、安川民男、

【26P05】 角谷重幸、北村一泰、石田寿昌

ニューラルネットワークの構造活性相関解析への応用に関する報告。以前に比べて、パラメーターの数を減らしたり、学習の手順を変更する等して、随分とスッキリした判りやすい内容になっていた。しかし、その分だけ古典的な構造活性相関解析研究との差はなくなってきたのではないかと、いう気もした。

生体高分子の基準振動解析法と酵素反応系への応用

【26P06】 (北里大薬) ○神谷健秀、梅山秀明

部分的に自由度を調節出来る基準振動解析のプログラムに関する報告。定式化は京都大学の郷教授達のものに従っているが、部分構造をバラバラの原子と見て多体系への拡張をうまく利用している。個人的には、基準振動解析では蛋白質全体の動きしか見えないという印象が強いが、このプログラムで応用面での有効性が示されることを期待したい。

(旭化成・CS室 戸潤一孔)

26P07 キメラ参照蛋白質を利用した蛋白質モデリング法の開発
(北里大) 梶原明朗, 菰岡仁志, 神谷健秀, 梅山秀明

キメラ参照タンパク質および類縁参照タンパク質を利用したタンパク質モデリングを比較し、前者の方法の精度、妥当性の考察に関する報告である。タンパク質のモデリングでは、参照タンパク質の選択が重要となるが、キメラタンパク質を参照タンパク質とする方法はなかなか興味深いと感じた。キメラタンパク質を用いる方が良好な結果が得られることが示されており、今後の展開が期待される。

26P08 類縁蛋白質の構造のずれと疎水コア距離の相関
(持田製薬) 鮫嶋理恵 (北里大薬) 梅山秀明

疎水コア距離と類縁タンパク質の構造のずれとの相関を検討し、疎水コア距離が構造保存部分を予測する有用な指標となるという報告である。5種類のタンパク質ファミリーを取り上げ、3残基平均の疎水コア距離と類縁タンパク質の構造のずれの傾向はよく一致しているとのことである。いくつかまだ問題点があるようだが、タンパク質のモデリングや進化を考えるときに、有益な情報をもたらし得る確実な方法に発展させて頂きたい。

(東京農工大工 園山 正史)

- 26P09 フレキシブル分子を含む多分子間の三次元共通部分構造の認識
(豊橋技科大) ○ヴォルフ・イーレンフェルト、高橋由雅、阿部英次、佐々木慎一
- 26P10 構造-活性相関を意識した合成支援システム (富士通) ○湯田浩太郎
- 26P11 合成設計支援システムSPEKの開発 (神奈川大理) ○中山 堯、伊藤尚樹、川村健史、菅野 崇、北原清隆
- 26P12 有機化合物のIUPAC命名支援システムの開発名称-構造変換システム II
(豊橋技科大) ○阿部英次、曾田健一、佐々木慎一
- 26P13 パソコン用分子特性推算システムの開発
(豊橋技科大) ○水元英詔、高橋由雅、助川正之、佐々木慎一
- 26P14 ニューラルネットワークによるスペクトルパターン認識(第16報)赤外スペクトルからの各種官能基の推定

(化技研) 田辺和俊、田村禎夫、(富士通) 上坂博亨

18種の官能基について、約100個の赤外スペクトルを基に入力層250、中間層100、出力層1からなるニューラルネットを構築し、約100個のテストデータによりこれら構築されたニューラルネットの信頼性を検討したところ平均88%の的中率が得られたという研究報告である。的中率には、(1)ある官能基をもつ化合物についてその官能基が存在することを推定する場合および、(2)ある官能基を持たない化合物についてその官能基を持たないと推定する場合の両方の正解率を含んでいるとのことであり、これらの詳細な考察がなされることを今後に期待したい。また、本システムを構造推定に用いるならば、再現率100%を保証するニューラルネットを構築することは重要なことの一つかもしれない。

- 26P15 並列カルマンフィルターによるスペクトルデータ解析

(豊橋技科大) 吉田浩士、李志良、宮下芳勝、佐々木慎一

PhenolとResorcinolの混合系におけるUVスペクトルデータを例に用い、並列カルマンフィルターにより成分組成および濃度を推定することを試み、(1)第1成分に対する第2成分の濃度が0.1%であっても成分組成および濃度を推定することができること、および(2)不純物の混入をも検出することが可能であることが示されている。さらに複雑な系においても理論的には成分組成および濃度を推定することが可能であるとのことであるから、この点においてさらに発展することを期待したい。

26P16 TUTORs : 部分構造情報に基づく候補構造精選知識ベースシステムの開発

豊橋技科大 石山 雄一, 高橋 由雅, 佐々木 慎一

分子設計支援システムTUTORsは基本部分構造をもとに与えられた一定条件下での可能な候補化合物を構造式を組み立て表示するが, 実際には候補化合物も数が多く更にそれらを絞って行く必要がある. その為のキーの一つとして部分構造と各種活性に関連する経験的な知識を システムの中に知識エディタを介して付加し, それにより精選プログラムが動作する方式を UNIX WS 上に FORTRAN,C により X-window による表示ユーティリティを用いて実現した.

今の処, その知識の数が実用化するには充分とは云えないようだが, とにかく今後の進展を待つことにしよう.

26P17 化学構造式とNMRスペクトルデータ処理統合システムCIMSの開発 (V)

CIMSのワークステーションへの実装と並列計算機能の実現

豊橋技科大 奥山 徹, 五月女 優, 豊田 学, Joe Lian Min,
橋詰 裕行, 阿部 英次, 佐々木 慎一

有機化合物の構造とNMRスペクトルのための統合データベースをUNIX-WSによるネットワーク上で実現し, ユーザインタフェースとして X-Window を利用しようとする本報告の構想は既に第五回を迎えている.

Ethernet上でTCP/IPプロトコルにより75,000+10,000件の C,H-NMR スペクトルをデータの制御と分散化を図り乍ら, しかもこれらのファクトデータをリレーショナル型にもって行くと云う, この難しい課題をどう実現させるのか, さらに一方では電子メールと云う形も云々されている状況もあり乍ら, これが完成した暁には果してユーザにはどう云うメリットが期待出来るのか, どう云うユーザが想定されるのか 今の処一向に見えて来ない.

もっと解り易い形で, 素人の吾々にも先が見える形で提示載けると有難い.

(東京家政学院筑波短期大学 前田 浩五郎)

- 26P18 3次元類似性検索の応用の試み (石原産業) ○赤木俊夫
- 26P19 パーソナルコンピュータによる学会講演要旨集のデータベース化とその利用。情報化学討論会および構造活性相関シンポジウム要旨集(続報)
(電通大・DAKO Japan・北里大薬) 山崎 昶、○梶川博司、岡地優司、平賀やよい、木下俊夫
- 27I18 90MHZプロトンNMRフルスペクトルデータベースの構築
(豊橋技科大) ○阿部英次、石井孝幸、佐々木慎一
- 27I19 電子メールを用いたNMRスペクトルデータ検索システムの開発
(豊橋技科大) ○奥山 徹、阿部英次、佐々木慎一
- 27I20 研究室所蔵各種文献集のデータベース化とその利用
(電通大・明治大理工) ○山崎 昶、梶川博司、尾崎 剛、渡邊耕司、鈴木康雄
- 27I21 IARCモノグラフからの発癌物質データベースの作成
(国立衛試) 神沼二真、五十嵐貴子、○中野達也
- 27I22 類推による有機合成研究用情報システム(OS-IBS)の構築
筑波大 電子・情報工学系 王 忠清、鄭 四清、藤原 謙

有機合成反応に関する情報にはエキスパートシステム型とデータベース型の二つがあるが、ともに充分でないこと、そしてこれから新しい知識を獲得する学習、類推等の機能が必須であるとする立場から SUN-WS上の X-Window環境上で OS-IBSを CORESを情報源として利用することで実現した。本システムでは類推機能を可能にするため化合物概念構造、反応概念構造等を導入し、まず類似化合物、類似反応を検索し、次ぎに反応ルールによる推論、その精度評価と云う手続きで進む形になっている。

時間が充分でなく成果の例示が多くなかったのが惜まれる。今後の進展による実用化に期待しよう。

- 27I23 シフト検定法の開発と化学・薬学への応用(4)
ニューラルネットワークへの適用

阪大薬 高木 達也、中井 宏紀、和田 千佐、野田 明宏、小野澤 学寿、
松村 謙一、藤原 英明

従来迄の線形判別分析に代わって、Oswaldのシフト検定法をBootstrap法の考え方を導入した拡張シフト検定法を考案し、それをバックプロパゲーション法によるニューラルネットワークの手法を適用することにより、対象の認識に必要な記述子の有意性の判定を例として検討してみた。応用例としてアヤメ科植物の分類とトリアジン誘導体の坑コクシジウム症活性的の判別問題を取り上げ良好な結果を得たとしている。富士通のS-4/IPとEPSONのPC386GSと云う小型システムによる試みだけに今後の、例えば野外での応用なども期待出来るだけに収束の時間短縮も含めて一層の改良、進展を待つことにしよう。

(東京家政学院筑波短期大学 前田 浩五郎)

27124 反応データベースの構築とその利用システムの開発

(チッソ石化・情報大・ミドリ十字・三井東圧化・藤沢薬品・大正製薬・出光石油化学・吉富製薬・NTT・クラレ・神奈川大・萬有製薬・筑波大・中外製薬・田辺製薬・三共) 潮田 誠、石塚英弘、井上佳久、小田研悟、加藤一孝、川島 豊、○清水恵次郎、白木政巳、菅原駿吾、高島裕幸、中山 堯、西岡 亮、藤原 譲、松浦育敏、宮城島利一、松井孝司

27125 N I S T - M S D S における C A S 化合物の側鎖部分 (その5) 該当化合物高速抽出の試み (筑波研究学園・DBS・東京家政学院筑波短大) ○仲久保正人、仲久保公子、和知孝宜、前田浩五郎

27118 パーソナルコンピュータによる¹H-NMRフルスペクトルデータベース (豊橋技科大) ○阿部英次、石井孝幸、河野英二、佐々木慎一

NMRは有機化合物の同定に不可欠の汎用的で強力な分析手段となっているが、最終的な同定には標準スペクトルとの比較が不可欠である。演者らは、かなり以前から、このためのデータ収集に着手しており、今回の発表もその一環である。急速に発展するNMR装置とコンピューターに追従すべく90MHzで測定した全領域のスペクトルを登録する方式を採用しているが、MHz数が年々向上されるのが、この種の(特に¹Hの)データを収集する際の問題となっている。当分の間はユーザーが可能な範囲で(即ち、複雑な高次分裂の部分を除いて)MHzの差を考慮して使用するしかないだろうと言うのが座長の所感である。

27119 電子メールを用いたNMRスペクトルデータ検索システムの開発 (豊橋技科大) ○奥山徹、阿部英次、佐々木慎一

本講演は上の講演と関連しているが、中心となるセンターに標準スペクトルデータを整備しておき、ユーザーが電子メールでそのデータを取り寄せ検索する方式を検討している。問題点としては、データの機密性の保護が確保できるかや、通信回線のオンライン接続が常時可能かなどが指摘されている。膨大な標準データの収集の苦労がうかがい知られる。これと全く対照的に、NMRユーザーが自分で測定したスペクトルをデータベースとして保存しておき、将来的に自分の検索に利用できるようなシステムも有用ではないかと言うのが座長の感じたことである。一連の類似化合物を扱うユーザーが多いことから、このようなパーソナルなデータベースを作るシステムも一考に値するのではないだろうか?

(阪大薬 藤原英明)

第3回ケモメトリックスワークショップ

報告

主催 日本化学会情報化学部会
日時 1992年2月10日(月) 9時55分~16時45分。
会場 大阪大学付属図書館 吹田分館、視聴覚ホール

ケモメトリックスワークショップも第3回を迎え、今回は、塩野義製薬の山川眞透先生と私(高木)がお世話をさせていただくこととなりましたが、講演して頂いた先生方を含めると、60名という参加者を数えるに至る盛況ぶりとなり、感謝の念に耐えません。又、情報化学講習会との併催となり、会場に持ち込まれたパーソナルコンピューターで、実際にデモンストレーションが行われました。

特に今回、筑波大の藤原譲先生、キッコーマンの相島鐵郎先生、お茶の水女子大学の藤枝修子先生のご尽力で、中国から Dr. Leming Shi をお迎えすることができ、中国のケモメトリックスの現状について、大変貴重で、かつ、興味深いお話を伺うことができました。ご尽力いただいた先生方に心から深謝いたします。今回のワークショップ開催に関しましては、文責者(高木)の経験の浅さから、講師の先生方をはじめ、日本化学会情報化学部会事務局の皆様にお世話になりっぱなしで、特に、前回、前々会の世話人をされた豊橋技術科学大学の宮下芳勝先生、お茶の水女子大学の藤枝修子先生には、種々ご指導頂き、お礼の言葉もございません。今回、成功のうちに本ワークショップ、講習会を終えることができましたのは、ひとえに、皆様方のおかげだと思っております。講師の先生方、ご指導頂きました皆様、ご参加頂きました皆様方に、改めて、お礼申し上げます。

以下に、当日の内容を略記させて頂きました。

1. ケモメトリックスの分光光度計への応用
(大塚電子) ○ Davar Pishva、大嶋浩正
2. Distance Geometry 法によるキモトリブシン阻害剤の構造活性相関
(豊橋技科大) ○長谷川 清・宮下芳勝・佐々木慎一
3. 一般化加法モデルの考え方と適用事例：ACEプログラムの提示
(塩野義製薬) ○田崎武信・財前政美・後藤昌司 (広島大 原医研) 大瀧 慈
4. ニューラルネットワークの化学に対する適用時の特徴及び留意
(富士通) 湯田浩太郎

5. 食品機器分析データパターン認識～ニューラルネットワークと多変量解析
(キッコーマン) ○相島鐵郎・(富士通) 湯田浩太郎

6. ソフトウェア・デモンストレーション

7. (特別講演) Simultaneous Multicomponent Analysis and Structure-Activity Relationship Study Using Chemometric Methods (Institute of Chemical Metallurgy, Chinese Academy of Science) Leming Shi

Dr. Shi の講演を伺って驚いたのは、何よりもまず、中国におけるケモトリックスの普及、発展ぶりでした。既に海外においては、普及、発展の著しいケモトリックス分野ですが、日本国内においてはなぜか、普及が遅れているのが現状です。私の勤務先(薬学部)では、中国からも多くの留学生を迎えています。こと、ケモトリックス分野に関しては、日本から中国へ留学生を送らなければならない状況と言って、差し支えないでしょう。

又、先述のように、今回は情報化学講習会との併催であったため、関係各位の協力をえて、一般化加法モデル・ACEプログラムと、PLSやクラスター分析法を組み込んだ分光光度データ処理プログラムのデモンストレーションを行うことができました。日本でケモトリックスが普及できない原因の一つは、対応するソフトウェアの不足にあると考えられます。こうした試みが重なり、ソフトウェアが多く蓄積されていくことによって、ケモトリックスの普及の手がかりができれば、望外の喜びと考えています。

会の最後に残されていた予備の時間を利用して、参加して頂いた方々、特に大阪府立大学名誉教授、宗森 信先生をはじめ、多くの先輩方と、ケモトリックスの今後について語り合うことができたのも、恐らく参加者の皆様方にとって、大きな収穫だったことだと思います。これを機会に、より多くの方々に、ケモトリックスの研究、発展に参加して頂きますよう、お願い申し上げます。

文責 大阪大学薬学部 高木達也

第2回 情報化学部会講習会報告 文責 エーザイ(株)張ヶ谷 泰二 —化学者のためのパーソナルコンピュータ利用講習会—

主催 日本化学会情報化学部会
協賛 キヤノン販売(株)

日時：平成4年12月9日(水) 10:00~16:00

会場：キヤノン販売(株)セミナールーム 東京 大手町

内容：[午前の部] マッキントッシュの基本的な使い方 キヤノン販売(株)
[午後の部] 化学関係ソフトの使い方 京都大学 渡辺芳人先生

テキスト 「Macintoshによる科学論文の仕上げ方」

参加人数： 予約36名(定員40名で募集したが、都合で36名締切り)

事前に渡辺先生、キヤノン販売の人と打ち合わせを持ったところ、先生の意欲的な構想、豊富な内容を40名に短時間で講習することは難しいことがわかり、途中で募集を打ち切った。事前アンケートでは、レベルに大きな差があることがわかり、このような中でも参加者全員が満足できる中身の濃い講習会にすることで望んだ。

さて、午前の基礎的な講習で初心者にはマウスをこわごわ動かしながらも、比較的順調に進んだ。通常2日の講習をたった2時間に圧縮したので、昼食時間を短縮せざるを得なかった。午後の部で渡辺先生の熱心な講習に食後の眠気を忘れ、午後4時をすぎても2/3の人は残り、会場都合の6時になって、仕方なく余韻を残し閉会した。

マッキントッシュ(MAC)は簡単に使えるため、アプリケーションをすぐ使い出す人が、マニュアルを熱心に読んでいる人は少ない。このようなことから、午前中の初心者向けのワープロ講習でもタブの使い方、脚注の使い方など、いつも使っている人でもいろいろ参考となる場所があったようである。午後はChemDrawではpreferencesの使い方、簡単な化合物の構造修正、グループ化、alignの方法と、通常知らなければ使わない機能を中心に講習された。そして、ChemDrawからChem3Dへの化合物のカットandペーストによる転送、Chem3Dでの化合物の構築、preferencesの変更と進んだ(図-1)。4時以後もほとんど休憩なしで、DeltaGraphによるグラフのスムージング、EndNoteでの文献整理と引用文献としてのコピー機能と大急ぎで進行された。それでも渡辺先生は用意された内容すべてを紹介できなかったようである。最後まで忙しく、内容濃く進んだので後半はフォローしきれなかった人もいたようだ。しかし、事後アンケートの結果(表-1)では大半の人は満足されたようで、幹事としては「ほっと」した次第である。

反省点

1. 広報活動

CICSJ、化学と工業、化学と教育、ぶんせき、それぞれ8、9月号に案内が掲載されたが、この企画を知らない人が多く、広報活動に問題を残した。

2. レベルの差

今回のアンケートで「午前中の講習はなくてもよかった」との意見や、「よくわかった」との回答が多かったが、これには、「もっと詳しく知りたい」との期待が込められていると思われる。参加者にレベルに差がある中で講習の内容を欲張ったため、特に午後は時間が足りなくなり、講師の先生も参加者にも欲求不満を残す結果になってしまった。このような講習会を企画すれば、いろいろなレベルの人が集まり、多様な期待で参加する。初心者向けの講習会はすでにいろいろあるので、部会主催とすれば、テーマを絞り技術をさらに引き上げる講習会にした方がより充実したものになったかもしれない。

3. アシスタント

先生一人で説明やデモを行い、また、フロアーに回ってのトラブル対応をしたりで、非常に大変そうであった。パソコン講習会にはやはり、フロアーを回るアシスタントが必須であろう。

ともあれ、第2回の講習会は盛況のうちに無事終わることができた。是非関西で開催して欲しいとか、ネットワークでのMACの使い方をとか、アンケート以外にもMAC講習会の要望が来ている。部会としても今後さらに充実した講習会を年1~2回は企画したいので、皆様の忌憚のないご意見をお願いします。

最後に、講師を引受けていただいた京都大学の渡辺先生、セミナー室とマシンを提供していただいたキヤノン販売(株)の関係者に厚くお礼申し上げます。

以上

表-1、第2回講習会事後アンケート結果

(参加35名、欠席1名、回答24名)

1. 午前中の講習会 (ワープロソフト/表計算ソフト)

よくわかった	15名	難しかった	0名
もの足りなかった	9名	その他	2名

2. ChemDraw, Chem3D, その他

よくわかった	18名	難しかった	5名
もの足りなかった	3名	その他	1名

3. 講習会全体

よくわかった	17名	難しかった	3名
もの足りなかった	1名	その他	2名

4. マッキントシュの講習会企画希望

内容を絞って詳しく	7	ワープロ関係	5	ソフト比較	2
MAC入門	1	DeltaGraph, CricketGraph	1	ネットワーク	1
他のパソコンとの比較	1				

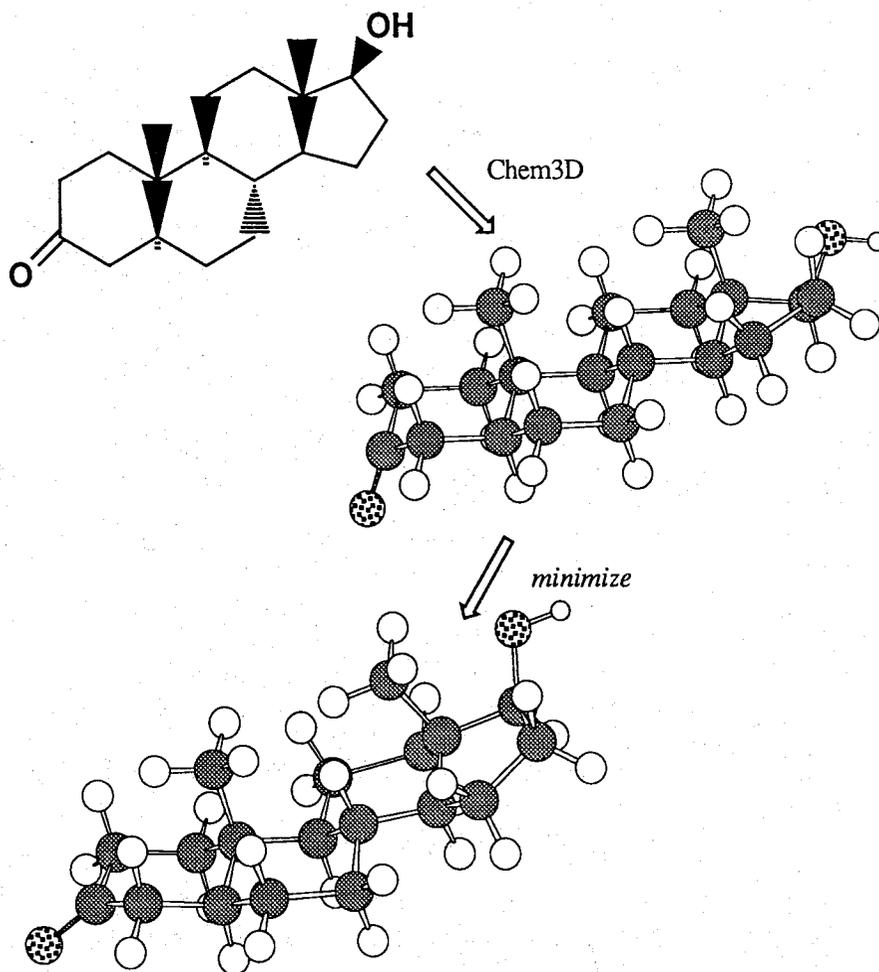
5. 他の講習会希望

UNIX	1	分子軌道法	1
文献データベース	1	LA	1

6. その他

時間不足	7	一人1台	3
各ソフトの概略	2	レベル別	2

図-1、第2回情報化学部会講習会の一例
 (ChemDrawで平面構造式を描き、Chem3Dへコピーandペーストし、
 その構造を最適化するまでの流れをChemDrawで作成した図)



第3回情報化学部会講習会の御案内

パーソナルコンピュータを利用した文献整理講習会

主催 日本化学会情報化学部会

協賛 ユサコ(株)・キャノン販売(株)

研究において文献整理は重要な仕事ですが、その量が多くなると必要なものがすぐに見つからない、と言ったことがよくあります。最近、抄録文献がフロッピーで配付され、ネットワークによる抄録文献データ利用や、個人の文献整理が非常に便利にできるようになってきました。またそれ以外にも、文献の引用のためのコピーandペーストなど、パソコンだからできるいろいろな活用が考えられます。そこで、この新しい方法についてのパソコン(マッキントッシュ)講習会を企画しましたので、お気軽にご参加下さい。

日時：平成5年5月26日(水) 午前10時～午後5時

会場：キャノン販売(株) セミナールーム 東京 大手町(詳細は後日案内)

参加申し込み：20名(定員になり次第締切)

参加資格：パソコン基礎知識(マッキントッシュのワープロが使える程度)のある方

参加費：10,000円(但し、情報化学部会非会員は15,000円)

内容：マッキントッシュを1台を2人で触りながら講習する。

10:00～12:00 CCOD (Current Contents On Diskette) の説明および使い方
ユサコ(株) 畠中正己

13:00～15:30 検索データ変換・処理の方法、文献整理ソフト(Pro-Cite)の使い方
ユサコ(株) 畠中正己

15:30～17:00 応用事例— 個人文献データベースCCLDBとC60LDB
豊橋技術科学大学 工学部知識情報工学系教授 大澤映二

参加申し込み方法：FAXまたは葉書で「第3回情報化学部会講習会申し込み」と必ず標記し、参加者氏名、年齢、所属、連絡先(住所、電話番号、FAX番号)を明記して、下記お申し込み下さい。

(注) キャンセルは整理の都合で受け付けません。参加できなくなった時はできるだけ代理の方を立ててください。

申し込み、問い合わせ：101 東京都千代田区神田駿河台1-5

日本化学会 情報化学部会 講習会係

電話03-3292-6163、FAX 03-3292-6318

支払方法：1.郵便振替 口座/東京7-6058 社団法人 日本化学会、

2.銀行振込 第一勧銀神田支店 普通1073505 (社)日本化学会(振込日を連絡下さい)、3.現金書留

(いずれも第3回情報化学部会講習会と明記して下さい。)

なお、情報化学部会の入会(日本化学会員、非会員共 年会費2,000円)は常時受け付けておりますので、非会員の方は是非加入されることをお勧めします。問い合わせは上記「情報化学部会」までお願いします。

以上

実践パソコン分子科学計算講習会

主 催 化学ソフトウェア学会

協 賛 日本化学会情報化学部会，日本化学プログラム交換機構，
神戸学院女子短期大学

日 時 4月24日(土)，25日(日) 10時 - 17時

会 場 神戸学院女子短期大学 (神戸市長田区林山町27-1)

[交通] JR神戸駅より市バス11系統板宿行乗車25分、宮川町9丁目下車
徒歩 10分

趣 旨 化学ソフトウェア学会編「実践パソコン分子科学計算」(サイ
エンスハウス刊)をテキストとし、これから分子科学計算を始めようとす
る人を対象に、一人一台のパソコンを使用した講習会

定 員 30名

第1日(4月24日)

1. 分子モデリング (MOLDA&GRIMM) (広島大理) 吉田 弘
2. 分子力学計算 (MM2) (広島大理) 深澤 義正
3. 分子構造表示 (MODRAST) (姫工大工) 中野 英彦

第2日(4月25日)

4. 分子モデリング (GUMMOS) (東洋紡) 相川 泰
5. 分子軌道計算 (GAUSSIAN70) (大阪芸大) 荒川 透
6. 分子軌道表示 (PSI/77) (兵庫教育大) 木原 寛

参加費 (テキスト代を含む) 主催学会正会員 (賛助会員および維持会
員構成員は、それぞれ2名および5名以内を会員扱いとする) 18,000
円，学生会員 10,000 円，非会員 30,000 円

申込期限 3月19日(金) (但し定員に達し次第締め切ります)

参加申込方法 官製葉書に氏名，勤務先(学校)，会員資格の有無(会
員番号)，連絡先(郵便番号，所在地，電話番号)を明記のうえ下記あて
お申し込みください。申込者には受理通知票と共に，参加費の振込先，会
場地図等を記載した詳細案内をお送りします。

申込先 671-22 姫路市書写 2167 姫路工業大学工学部応用化学科 中野
英彦 (電話 (0792) 66-1661 内線 325, Fax (0792) 66-8868)

UNIXは本当に難しいものだろうか

日本アイ・ビー・エム株式会社

小出昭夫

1. はじめに

ワークステーションの普及とともに、この数年間日本でUNIX、UNIXとやたらに騒がしくなっている。IBMでも、1985年以来、AIXの名前で『UNIX』路線を進めてきた。では、UNIXとはなんだろうか。難しいものなのか。難しいとしたらどのような難しさなのだろうか。既にUNIXについての各種の解説本や実用書が多数世に出ているので、ここでは、文化評論のような切り口でこれらの問題を論じてみたいと思う。

2. UNIXとはなにか

UNIXは、他のDOS/V、MVS、VM/CMSと同様にオペレーティング・システムの一つである。オペレーティング・システムとはコンピューター資源を管理する基本ソフトのことである。同時に、UNIXはその生まれて既に20年近い歴史の中で一つの独特なコンピューター文化を築いている。また、この間にハードウェアや通信やユーザ層などのコンピューター環境が大きく変わり、『純正』UNIXは『AT&Tの商標登録』という冗談の中でしか今や存在しない。従って、UNIXについて論ずることは、オペレーティング・システムのみならず、文化、歴史の側面からも語ることである。

UNIXは茶面気のある私的研究（お宅族）の中で生まれた。UNIXの名前自体、ベル研究所のB. Kernighanが冗談で名付けたものである。ベル研究所が1969年に撤退した企業何社かと大学との失敗した共同研究プロジェクトMulticsからの語呂合わせである。さて、当時、ベル研究所はコンピューターで商売することが禁じられていた。これが幸いして、当面の性能を無視し、具体的ハードウェアから独立し、理想を求めてオペレーティング・システムを設計開発できた。また、これをソースの形で公開したので、他の研究グループによる追体験や拡張などの科学としての必要な手続きが可能となった。よく述べられる事柄、UNIXが安い値段で使用できたので広まり多数のソフトウェア資産がその上に築かれたということは、単なる余話である。

UNIXがコンピューター環境の大変化に柔軟に対応し生き延びてきたのは、オペレーティング・システムが階層構造をもっていることにある。いわゆる古典的UNIXに即していうと次の二層に大きく分けることができる。

まず、仮想コンピューターとしての層でカーネルと呼ばれる。木構造のファイル管理機構やマルチプロセス機構がその代表例である。次がユーザとのインタフェースとしての層で、シェルやC言語がその例である。

階層構造があるがために、部分的手直して環境の変化に対応してこられた。これらについては次章以降具体的に論ずるとして、UNIXのもつ二つの大きな問題点についてこの章で述べておこう。第一はユーザ層の変化に関連する問題で、第二は柔軟性が生む多様性の問題である。

現在のユーザの多くからみれば、オペレーティング・システムとはコンピューターを使うための必要だが退屈な手続きのことである。例えば、地下鉄の乗り方やマクドナルドでの注文の仕方のようなものである。コンピューターの利用がユーザにとって難しいとすればソフトウェアを作る側が悪いと非難される時代である。一方、UNIXが生まれた頃のユーザ層は趣味としてプログラミングする人々であり、従って、前述のユーザ・インタフェース層はコンピューターの原理の理解を前提にしていた。

次に、UNIXが簡単で柔軟にできているため、非常に多様な文化がUNIXの上に花開くことができ、多様性のための混乱が生じている問題である。例えば、シェルにはBシェル、Cシェル、Kシェルなど色々あり、エディターのようにユーザの好みで選択できる。シェルとは、キーボードから入力した文字列をコマンドとして解釈するソフトのことである。また、おのおのシェルも個人専用に修正（カスタマイズ）できる幅が大きい。更にコマンドの追加が容易なため、機能的に重複したコマンドの洪水も起きている。陰口を叩く人は『バベルの塔』現象がUNIXとともに始まっていると嘆く。

これらの問題については後でもう一度取り上げる。

3. UNIXのファイル・システムは簡単明瞭である

MVSやVM/CMSを知っている人なら、UNIXのファイル・システムが良くできていることに納得していただけたらと思う。実際、パソコンのDOSもUNIXのファイル・システムを真似ている。ワークステーションを表層で利用するのなら、ファイル・システムとエディターと利用する応用ソフトを知っていれば充分である。シェル・コマンドで言えば、pwd、ls、cd、mv、cp、mkdir、rm、rmdir、chmod、chownが使いこなせればよい。

UNIXでは、ハードディスク上のファイルを郵便の住所のように階層構造で管理できる。今、鈴木さんという人がいたとする。鈴木という名前の人は一杯いるから、名前だけではわからないかもしれない。ある局面では鈴木さんで充分だが、ある局面では何々会社の何々部門の何々課の鈴木さんという必要がある。UNIXではこれを木構造のファイル管理でこれを可能にしている。

この仕組みはディレクトリーという特別のファイルをもつことと現行のディレクトリーの概念で実現されている。ここでディレクトリー・ファイルとは子供のファイル名とそれが置かれている場所などの情報をもったファイルのことである。例として、

／IBM／東京基礎研究所／SPI／教育システムズ応用／鈴木

を考えてみよう。一番上のディレクトリーは会社名の一覧を与える。IBMというディレクトリーを選ぶと大組織名の一覧が得られる。次にSPIを選ぶと次の下位の組織名の一覧がという仕掛けで、最終的に普通のファイル鈴木に到達する。局面を設定するとは、現行のディレクトリーを一つ選ぶこと（cdコマンド）で、東京基礎研究所が現行なら、

でファイル鈴木が特定できる。

このディレクトリーはmkdirコマンドで自由に作ることができる。階層の深さも自由である。名前の付け方も自由である。すなわち、UNIXのカーネルは基本ルールを与えるだけで、後はユーザ任せである。ところが、実際の使いやすさはこの後の部分に依存する。従って、図書の整理学のように、ディレクトリー・システムの実際の運用の仕方が真剣に議論される日も近いと思う。

ワークステーションはパソコンと違い、複数の人によって共有されることが多い。このため、UNIXにはホーム・ディレクトリーという概念がある。UNIXでは、ユーザがシステムにログインすると各個人専用のディレクトリー（ホームと呼ぶ）が最初の現行ディレクトリーになる。また、自分のファイルを他のユーザが読んでも良いか書き直しても良いかどうかなどを個別に設定できる。ファイルにlkコマンドで別名を付けられる。すなわち、DOSよりかなり強力である。

さて、いままで普通のファイルと言ってきたが、実はファイルの実体はかなり広い。コンピューターは閉じた系でなく、色々な入出力、例えば、サウンド、グラフィックスをもつ。これらの入出力をファイル・システムが統一的に管理する。すなわち、入出力装置を見かけ上ファイルとして扱う。また、ディスクに限っても、ネットワークで結ばれている他のワークステーションのディスクを自分のファイル・システムの一部と見ることができる。

4. マルチプロセスがUNIXを柔軟にしている

前章でファイルの階層について述べたが、UNIXの実行での階層的な管理を可能にしているのは、実はマルチプロセス機構である。マルチプロセスは子供のプロセスを起こし、見かけ上、それを親と同時に走らせる機構である。ここで、同時に走らせるプロセスの数は原則上制限がない。

例えば、シェルはloginプロセスの子プロセスである。ユーザがログインしたとき、loginプロセスはファイル/etc/passwdを見て、そのユーザの指定するプログラム（プロセス）を走らせる。このファイルは通常のテキスト・ファイルであり、書換えれば違ったシェルが走る。シェルだけでなく、通常のプログラムも指定できる。ここで、shutdownを指定すれば、システム自動終了だけのユーザができる。

また、どのシェルも実行可能なファイル（プログラム）があれば、そのファイル名の入力でそのプログラムを子プロセスとして動かす。従って、自分専用のコマンドを既存のシェルに追加することもできる。

このように、単なるマルチタスクではなく、親子関係の設定できるマルチプロセス機構や、テキスト・ファイルを媒介にしてカスタマイズの幅を広げる文化風土がUNIXの柔軟性を広げている。子プロセスの変更や追加で、UNIXをどんどん拡張できる。

マルチプロセスでもう一つ重要なのは、プロセス間でどのようにコミュニケーションするかである。この点に関してはネットワークの出現などの新しい要請のためかなり変化してきた。マルチメディアの出現は、現在、音とアニメーションの同期の問題を引き起こしており、プロセス間コミュニケーションの更なる強化を必要としている。

5. ユーザ・インタフェースは日々進化しつつある

前章でUNIXはログインを通じて各ユーザ専用インタフェースのプロセスを走らせることができると述べた。さて、シェルはそれ自体として制御コマンドをもったプログラミング言語であるので、本業としてワークステーションを使う人（例えばシステム開発者）にとっては非常に便利なものである。しかし、そうでない多くのユーザにとっては表層のレベルでのシェル利用で充分こと足りている。それならば、UNIXでは多様なインタフェースが同居できるので、そのようなユーザは全く別のインタフェースを使用してもいいのではないだろうか。

実際、UNIXの世界でもグラフィックス・ユーザ・インタフェース（GUI）化が進行中である。X-Windowsはグラフィックスやマウス操作の基本ルーチンとマルチウィンドウ環境とを与える。Motifは更に高度な画面構築やユーザ入力処理のためのライブラリーを与える。画面上のものをマウスで直接操作するプログラムは簡単に書ける時代になった。スーパー・ユーザの仕事（システム管理）もメニュー形式でできるようになった。従って、今後は、プログラムを作成する人だけでなくプログラムを利用する人のためにやさ（優、易）しいインタフェースになるだろう。

このような活動は、現在、OSFのように外に開かれた協力団体の下になされている。しかし、このUNIXの標準化は、UNIX文化を強制的に統一することを目指しているのではない。強権で文化の標準化を進めることは中央統制や談合のようなものである。UNIX文化を構成する要素を部品と見なし、部品の規格化を進めているのである。部品の組合せで、デザイナーがユーザ・インタフェースを自由に設計できる余地を残している。率直にいうと、UNIX上のソフトにはグラフィックス・ユーザ・インタフェースの設計が悪いものが多いが、これはそのソフトを作るデザイナーに責任がある。時間と共に悪いものは自然淘汰されるだろう。

6. おわりに

初めて知らない町に訪れて安くて美味しく食事できる店を探すことは難しい。しかし、だれもこの難しさを知的なものとは思わないだろう。極限すれば、UNIXの難しさも知的な難しさでなく、慣れれば実的には何でもない。基本ルールは簡単である。求められるのは、初心者（旅行者）のための判りやすいインタフェースで、これはUNIX文化の問題である。

UNIXの講義を昨年化学系の大学院生を対象に行ったが、使うことより教えることがずっと難しいことを味わった。反省してみると、教えようとした内容が多過ぎたためだ。古いUNIX文化はあまり教えないほうが良い。

UNIX文化の現状は、たとえば、中世のギルト的迷路のような都市から判りやすい近代的都市への過度期の感である。旧市街のまわりに新市街ができつつある。都市の場合は、都市人口の急増、大量生産による工業製品の規格化、交通網の発展がその要因であったが、ユーザの増大、大量販売のソフトウェアの出現、ネットワークの普及と同様のことが起きている。雑多な多様性を保ちながら、今後4、5年で、UNIXはやさしいと印象を与えるようになるだろうと楽観している。

MEETINGS

28 March-2 April 1993.

205th ACS National Meeting.
Denver, CO.

Contact: American Chemical Society, Meetings Dept., 1155 16th St., NW,
Washington, DC 20036 (202-872-4396).

6-10 June 1993.

3rd International Conference on Chemical Structures.
Leeuwenhorst Congress Center, Noorwijkerhout, The Netherlands.

国内問合先: 113 東京都文京区弥生2-4-16学会センタービル 社団法人 化学情報協会
千原秀昭 (電話03-3816-3389 FAX 03-3816-7826)

13-17 June 1993.

3rd Scandinavian Symposium on Chemometrics.
Arhus, Denmark.

Contact: SSC3 Secretariat, Department of Chemical Technology, Danish Technological
Institute, Teknologiparken, DK-8000 Arhus C, Denmark. Tel. : +45 86 142400; Fax:
+45 86 147445.

August, 1993.

Gordon Conference on QSAR.
Tilton, NH.

Contact: Alexander M. Cruikshank, Gordon Research Conferences, Gordon Research
Center, University of Rhode Island, Kingston, RI 02881-0801 (401-783-4011; FAX: 401-
783-7644)

22-27 August 1993.

206th ACS National Meeting.
Chicago, IL.

Contact: American Chemical Society, Meetings Dept., 1155 16th St., NW,
Washington, DC 20036 (202-872-4396).

5-11 Sept. 1993.

Euroanalysis VIII. European Conference on Analytical Chemistry.
Univ. of Edinburgh, Edinburgh, UK.

Contact: Miss P.E. Hutchinson, Analytical Div., The Royal Society of Chemistry,
Burlington House, Piccadilly, London W1V 0BN, UK. Tel. : +44 71 437 8656; Fax:
+44 71 734 1227.

文 献 紹 介

[コンテンツリスト]

JOURNAL OF MOLECULAR GRAPHICS

Sept. 1992, Vol.10, No.3

- 131 **Automated site-directed drug-design: a method for the generation of general three-dimensional molecular graphs**
R. A. Lewis
- 144 **The calculation of molecular similarity: alternative formulas, data manipulation and graphical display**
A. C. Good
- 152 **Automation of conformational analysis and other molecular modeling calculations**
R. Taylor, G. W. Mullier and G. J. Sexton
- 161 **Color illustrations**
- 166 **Object Command Language: a formalism to build molecular models and to analyze structural parameters in macromolecules, with applications to nucleic acids**
J. Gabarro-Arpa, J. A. H. Cognet and M. Le Bret
- 174 **Finding and filling protein cavities using cellular logic operations**
J. S. Delaney
- 178 **FOCUS: a program for analyzing molecular dynamics simulations, featuring digital signal-processing techniques**
D. J. Osguthorpe and P. Dauber-Osguthorpe

Dec. 1992, Vol.10, No.4

- 194 **Pharmacophoric pattern matching in files of three-dimensional chemical structures: use of bounded distance matrices for the representation and searching of conformationally flexible molecules**
D. E. Clark, P. Willett and P. W. Kenny
- 205 **Searching for geometric molecular shape complementarity using bidimensional surface profiles**
A. Badel, J. P. Mornon and S. Hazout
- 212 **KEY, LOCK, and LOCKSMITH: Complementary hydrophobic map predictions of drug structure from a known receptor-receptor structure from known drugs**
G. E. Kellogg and D. J. Abraham
- 218 **Photorealistic image generation of molecular structure on PC screen using the ray-tracing technique**
S. Kim, C. W. Yoon, B. J. Mhin, H. S. Kim and K. S. Kim
- 222 **An easily fixed error in the Nyburg algorithm for discovering the best fit between molecules**
D. P. Dolata and J. Arnold
- 225 **Color illustrations**
- 229 **POCKET: A computer graphics method for identifying and displaying protein cavities and their surrounding amino acids**
D. G. Levitt and L. J. Banaszak
- 235 **Molecular illustration in black and white**
D. S. Goodsell and A. J. Olson
- 241 **MOL3D—A modular and interactive program for molecular modeling and conformational analysis: II. Extended modules**
D. Pattou, L. H. Bach, J. P. Kocher and B. Maigret
- 247 **SQUID: A program for the analysis and display of data from crystallography and molecular dynamics**
T. J. Oldfield
- 253 **Using a RIBBON program to illustrate lipid bilayer packing**
A. Hybl and B. P. Gaber
- 257 **A unique structural feature of a phospholipase A₂ is probed by molecular dynamics**
J.-P. Demaret and S. Brunie

JOURNAL OF CHEMICAL EDUCATION

Sept. 1992, VOL.69, NO.9

Computer Series:141

LIMSport:Spreadsheet Laboratory Information Management System for Chemistry

Ed Vitz, p.744.

QUANTITATIVE STRUCTURE

ACTIVITY RELATIONSHIPS

Vol.11, No.1, 1992

- | | | |
|---|----|--|
| Pratesi, P., Caliendo, G.,
Silipo, C. and Vittoria, A.: | 1 | A QSAR approach to the study of structural requirements of muscarinic receptor ligands. Part I: Agonists |
| Miyashita, Y., Ohsako, H.
Takayama, C. and Sasaki, S.: | 17 | Multivariate structure-activity relationships analysis of fungicidal and herbicidal thiolcarbamates using partial least squares method |
| Simon, Z. and Bohl, M.: | 23 | Structure-activity relations in gestagenic steroids by the MTD method. The case of hard molecules and soft receptors |
| Paberit, N. and Paberit, M.: | 28 | Inhibition of thermolysin and <i>Bacillus brevis</i> neutral proteinase by organic solvents and some related compounds |
| Reynolds, C.A. Burt, C.
and Richards, W.G.: | 34 | A linear molecular similarity index |
| Dross, K.P., Mannhold, R.
and Rekker, R.F. | 36 | Drug lipophilicity in QSAR practice: II. Aspects of R_M -determinations; critics of R_M -corrections; interrelations with partition coefficients |
| Balaz, S., Wiese, M.
and Seydel, J.K. | 45 | A simple model-based function for description of non-linear bioactivity-hydrophobicity profiles |
| Raevsky, O.A.,
Grigor'ev, V.Y., Kireev, D.B.
and Zefirov, N.S.: | 49 | Complete thermodynamic description of H-bonding in the framework of multiplicative approach |

Kim, K.H.:	127	3D-quantitative structure-activity relationships: description of electronic effects directly from 3D structures using a GRID-comparative molecular field analysis (CoMFA) approach
Liu, Q., Hirono, S. and Moriguchi, I.:	135	Application of functional-link net in QSAR. 1. QSAR for activity data given by continuous variate
Haaksma, E.E.J., Voss, H.-P., Donné-Op den Kelder, G.M. and Timmerman, H.:	142	An interaction model for tiotidine and analogues with the histamine H ₂ -receptor
Pratesi, P., Caliendo, G., Silipo, C. and Vittoria, A.:	151	A QSAR approach to the study of structural requirements of muscarinic receptor ligands. Part II: Antagonists
Famini, G.R., Ashman, W.P., Mickiewicz, A.P. and Wilson, L.Y.:	162	Using theoretical descriptors in quantitative structure-activity relationships: Oplate receptor activity by fentanyl-like compounds
Garrone, A., Marengo, E., Fornatto, E. and Gasco, A.:	171	A study on pK _a ^{app} and partition coefficient of substituted benzoic acids in SDS anionic micellar system
Klopman, G.:	176	MULTICASE 1. A hierarchical computer automated structure evaluation program
	185	Future events/New Books
Special issue:	187	Contributions of the Symposium on Biological Activity and Chemical Structure: Current aspects and future perspectives in Medicinal Chemistry, Heinrich-Heine-Universität, Düsseldorf, December, 6th 1991:
Ariens, E.J.:	190	QSAR: Conceptions and misconceptions
Rekker, R.R.:	195	The history of drug research: From Overton to Hansch
van de Waterbeemd, H.:	200	The history of drug research: From Hansch to the present
Seydel, J.K., Albores Velasco, M., Coats, E.A., Cordes, H.P., Kunz, B. and Wiese, M.:	205	The importance of drug-membrane interaction in drug research and development
Lemoine, H.:	211	β-Adrenoceptor ligands: characterization and quantification of drug effects
Timmerman, H.:	219	Ways to new H ₂ -agonists
Höltje, H.-D.:	224	Molecular modelling studies on 1,4-dihydropyridines acting at the Ca-channel
Edgar, B.:	228	Industrial drug research: Development of Ca-antagonistic vasodilators
Mannhold, R.:	232	University drug research: current state and future perspectives

□ Vol.11, No.3 ,1992

Kim, K.H.:	309	3D-Quantitative structure-activity relationships: nonlinear dependence described directly from 3D structures using a GRID-comparative molecular field analysis (CoMFA) approach
Liu, Q., Hirono, S. and Moriguchi, I.:	318	Application of functional-link net in QSAR. 2. QSAR for activity data given by ratings
Moriguchi, I., Hirono, S., Liu, Q. and Nakagome, I.:	325	Fuzzy adaptive least-square and its application to structure-activity studies
Hemken, H.G. and Lehmann F., P.A.:	332	The use of computerized molecular structure scanning and principal component analysis to calculate molecular descriptors for QSAR.
Hecht, P., Vyplel, H., Nussbaumer, P. and Berner, H.:	339	A combined use of quantum chemical parameters, hydrophobic and geometrical descriptors to establish QSAR's of allylamine antimycotics
ter Laak, A.M., van Drooge, M.J., Timmerman, H., and Donné-op den Kelder:	348	QSAR and molecular modelling studies on histamine H ₁ -receptor agonists
E.A. Coats, M. Wiese, H.-L. Chi, H.P. Cordes and J.K. Seydel:	364	Computer modelling and NMR studies of interactions between phospholipids and benzylamines

□ Vol.11, No.4 ,1992

Kim, K.H.:	453	3D-Quantitative structure-activity relationships: Investigation of steric effects with descriptors directly from 3D structures using a comparative molecular field analysis (CoMFA) approach
Greco, G., Novellino, E. Silipo, C. and Vittoria, A.:	461	Study of benzodiazepine receptor sites using a combined QSAR-CoMFA approach
Nakayama, A.:	478	Quantitative structure-activity relationship of histamine H ₂ antagonists. Electrostatic similarity of cimetidine derivatives
Kamphuis, J., Leij, F., Tancredi, T., Toniolo, C., and Temussi, P.A.:	486	SAR of sweet molecules: Conformational analysis of two hypersweet and two conformationally restricted aspartame analogues
Rose, V.S., Wood, J. and MacFie, H.J.H.:	492	Generalised single class discrimination. A new method for the analysis of embedded structure-activity relationships
Barnett, S.P., Hill, A.P., Livingstone, D.J. and Wood, J.:	505	A new method for the calculation of partition coefficients from experimental data for both mixtures and pure compounds
Avdeef, A.:	510	pH-Metric log P. Part 1. Difference plots for determining ion-pair octanol-water partition coefficients of multiprotic substances
		Letters to the editors:
Mager, H. and Mager, P.P.:	518	Validation of QSAR's: Some reflections
Wold, S.:	521	Answer to Mager and Mager
	521	New Books
	523	Abstracts

□ Oct. 1992, Vol.13, No.8

Theoretical Studies of Organometallic Compounds. I. All Electron and Pseudopotential Calculations of $Ti(CH_3)_nCl_{4-n}$ ($n = 0-4$) <i>Volker Jonas, Gernot Frenking, and Manfred T. Reetz</i>	919
Theoretical Studies of Organometallic Compounds. II. All Electron and Pseudopotential Calculations of $M(CH_3)_nCl_{4-n}$ ($M = C, Si, Ge, Sn, Pb; n = 0-4$) <i>Volker Jonas, Gernot Frenking, and Manfred T. Reetz</i>	935
Correlated Molecular and Multicenter Multipole Moments in Ground and Excited States from Multiple Reference Double-Excitation Configuration Interaction Calculations <i>Szczepan Roszak, W. Andrzej Sokalski, and Joyce J. Kaufman</i>	944
SETTLE: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models <i>Shuichi Miyamoto and Peter A. Kollman</i>	952
Derivation of Fluorine and Hydrogen Atom Parameters Using Liquid Simulations <i>Craig A. Gough, Stephen E. DeBolt, and Peter A. Kollman</i>	963
How Transferable Are Hydrogen Parameters in Molecular Mechanics Calculations? <i>David L. Veenstra, David M. Ferguson, and Peter A. Kollman</i>	971
Molecular Electrostatic Potentials and Partial Atomic Charges from Correlated Wave Functions: Applications to the Electronic Ground and Excited States of 3-Methylindole <i>John D. Westbrook, Ronald M. Levy, and Karsten Krogh-Jespersen</i>	979
Rapid Estimation of Charge Distribution. II. Heteroatoms <i>John J. Houser and Gilles Klopman</i>	990
Does CH_5^+ Prefer a C_{2v} Rather Than a C_s Structure? <i>Paul von Ragué Schleyer and José Walkimar de M. Carneiro</i>	997
Extended Koopmans' Theorem: Approximate Ionization Energies from MCSCF Wave Functions <i>Robert C. Morrison and Guanghua Liu</i>	1004
The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method <i>Shankar Kumar, Djamal Bouzida, Robert H. Swendsen, Peter A. Kollman, and John M. Rosenberg</i>	1011
Molecular Dynamics on a Distributed-Memory Multiprocessor <i>S.L. Lin, J. Mellor-Crummey, B.M. Pettitt, and G.N. Phillips, Jr.</i>	1022

Analysis of the Gas-Phase Addition of Water to Formaldehyde: A Semiempirical and <i>Ab Initio</i> Study of Bifunctional Catalysis by H ₂ O <i>Oscar N. Ventura, Elena L. Coitiño, Agustí Lledós, and Juan Bertran ..</i>	1037
Theoretical Studies of [n]Paracyclophanes and Their Valence Isomers. I. Geometries, Strain Energies, and Enthalpies of the Inter-Conversions of [n]Paracyclophanes and Their Dewar Benzene Isomers <i>F. Bockisch, J.C. Rayez, D. Liotard, and B. Duguay.....</i>	1047
Monte Carlo Simulation of the Diabatic Free Energy Curves for a Dissociative Electron Transfer Reaction in a Polar Solvent <i>Victor Pérez, José M. Lluch, and Juan Bertran.....</i>	1057
<i>Ab Initio</i> Investigation of the Structures and Stabilities of CH ₂ N ₂ , CHFN ₂ , and CF ₂ N ₂ Isomers: Important Consequences of MP2 Optimizations <i>Alexander I. Boldyrev, Paul v.R. Schleyer, D. Higgins, Colin Thomson, and Sofia S. Kramarenko.....</i>	1066
Algorithm to Generate Reaction Pathways for Computer-Assisted Elucidation <i>Raúl E. Valdés-Pérez.....</i>	1079
PM3-SM3: A General Parameterization for Including Aqueous Solvation Effects in the PM3 Molecular Orbital Model <i>Christopher J. Cramer and Donald G. Truhlar.....</i>	1089
Protein Calculations on Parallel Processors. II. Parallel Algorithm for the Forces and Molecular Dynamics <i>J.F. Janak and P.C. Pattnaik.....</i>	1098
Highly Accurate Diatomic Centrifugal Distortion Constants for High Orders and High Levels <i>Mahmoud Korek and Hafez Kobeissi.....</i>	1103
Improvements on the Protein-Dipole Langevin-Dipole Model <i>Y.W. Xu, C.X. Wang, and Y.Y. Shi.....</i>	1109
Solving the Finite-Difference Non-Linear Poisson-Boltzmann Equation <i>Brock A. Luty, Malcolm E. Davis, and J. Andrew McCammon.....</i>	1114
General Expression for the Density Dependence of the Mori Coefficients <i>Juan J. Morales and María J. Nuevo.....</i>	1119
Molecular Mechanics Studies of Ketene Derivatives and Related Structures <i>Eugene L. Stewart and J. Phillip Bowen.....</i>	1125
Intensities of Infrared Bands in Molecular Mechanics (MM3) <i>Jenn-Huei Lii and Norman L. Allinger.....</i>	1138
Treatment of Redundancies among Internal Coordinates in Optimizing Molecular Mechanics Force Constants <i>K. Palmö, L.-O. Pietilä, and S. Krimm.....</i>	1142
Study of Hydrogen Bonding Interactions Relevant to Biomolecular Structure and Function <i>Ya-Jun Zheng and Kenneth M. Merz, Jr.....</i>	1151

<i>Ab Initio</i> Prediction of Possible Crystal Structures for General Organic Molecules <i>H.R. Karfunkel and R.J. Gdanitz</i>	1171
Theoretical Studies of Organometallic Compounds. III. Structures and Bond Energies of FeCH_n and FeCH_n^+ ($n = 1, 2, 3$) <i>Achim Veldkamp and Gernot Frenking</i>	1184
From Linear Combinations to Integrals: A New Approach to the Basis Function Problem <i>Jesús R. Flores</i>	1199
Molecular Dynamics Simulation of a Poly(Oxyethylene) Chain Dissolved in Benzene <i>M. Depner and B.L. Schürmann</i>	1210
Conformational Distribution of Heptaalanine: Analysis Using a New Monte Carlo Chain Growth Method <i>B. Velikson, T. Garel, J.-C. Niel, H. Orland, and J.C. Smith</i>	1216
Critical Analysis of Electric Field Modeling: Formamide <i>François Colonna, Earl Evleth, and János G. Ángyán</i>	1234
Vector Processing Algorithm for Electron Repulsion Integrals in <i>Ab Initio</i> HF Calculation Based Upon the PK Supermatrix <i>Shuichi Yahiro and Yasuhiko Gondo</i>	1246
<i>Ab Initio</i> Computed Molecular Structures and Energies of the Conformers of Glucose <i>P.L. Polavarapu and C.S. Ewig</i>	1255
Conformational Sampling by a General Linearized Embedding Algorithm <i>Gordon M. Crippen, Andrew S. Smellie, and Wendy W. Richardson</i>	1262

Editorial	
<i>Norman L. Allinger and Paul von R. Schleyer</i>	1
Rotational Barrier in Phosphatriafulvene: An MCSCF Study <i>Wolfgang W. Schoeller, Jörg Strutwolf, Wilfried Haug, and Thilo Busch</i>	3
An Electron Repulsion Integral Compression Algorithm <i>Markus P. Fülischer and Per-Olof Widmark</i>	8
Vectorizable Algorithm for Green Function and Many-Body Perturbation Methods <i>Vyacheslav G. Zakrzewski and Wolfgang von Niessen</i>	13
Off-Lattice Monte Carlo Method with Constraints: Long-Time Dynamics of a Protein Model without Nonbonded Interactions <i>E. W. Knapp and A. Irgens-Defregger</i>	19
Optimal Use of the Recurrence Relations for the Evaluation of Molecular Integrals over Cartesian Gaussian Basis Functions <i>Ungsik Ryu, Myeongcheol Kim, and Yoon Sup Lee</i>	30
MNDO Parameters for Helium: Optimization, Tests, and Application to Endohedral Fullerene-Helium Complexes <i>Matthias Kolb and Walter Thiel</i>	37
Force Field for Platinum Binding to Adenine <i>Jiří Kozelka, Roger Savinelli, Gaston Berthier, Jean-Pierre Flament, and Richard Lavery</i>	45
Supplementary <i>d</i> and <i>f</i> Functions in Molecular Wave Functions: Optimum and Nonoptimum Exponents <i>Eric Magnusson</i>	54
Supplementary <i>d</i> and <i>f</i> Functions in Molecular Wave Functions at Large and Small Internuclear Separations <i>Eric Magnusson</i>	67
Comparison of Computational Methods Applied to Oxazole, Thiazole, and Other Heterocyclic Compounds <i>Alan A. Shaffer and Scott G. Wierschke</i>	75
Ability of the PM3 Quantum-Mechanical Method to Model <i>Intermolecular</i> Hydrogen Bonding between Neutral Molecules <i>Marcus W. Jurema and George C. Shields</i>	89
Multigrid Solution of the Poisson-Boltzmann Equation <i>Michael Holst and Faisal Saied</i>	105
Molecular Conformations from Distance Matrices <i>W. Glunt, T. L. Hayden, and M. Raydan</i>	114
Benzene Is Not Very Rigid <i>Kenny B. Lipkowitz and Michael A. Peterson</i>	121

JOURNAL OF CHEMICAL INFORMATION

AND COMPUTER SCIENCES

□ VOL. 32, Nov./ Dec., 1992

- 577 The 1992 Beilstein Workshop on Similarity in Organic Chemistry
Clemens Jochum, Martin G. Hicks, and Josef Sunkel*
- 578 Similarity in Organic Chemistry: A Summary of the Beilstein Institute Conference
Stephen R. Heller
- 580 Definition and Role of Similarity Concepts in the Chemical and Physical Sciences
D. H. Rowray
- 587 A Momentum-Space Approach to Molecular Similarity
Neil L. Allan and David L. Cooper*
- 591 Visualizing Relative Occurrences in Metabolic Transformations of Xenobiotics Using Structure-Activity Maps
Eric M. Gifford, Mark A. Johnson, David G. Kaiser, and Chun-che Tsai**
- 600 Quantum Similarity Measures, Molecular Cloud Description, and Structure-Properties Relationships
Ramon Carbó and Blanca Calabuig*
- 607 Database Searching on the Basis of Three-Dimensional Molecular Similarity Using the SPERM Program
*Nicholas C. Perry and Vincent J. van Geerestein**
- 617 Similarity Searching in Databases of Three-Dimensional Molecules and Macromolecules
Peter J. Artymiuk, Peter A. Bath, Helen M. Grindley, Catherine A. Pepperrell, Andrew R. Poirrette, David W. Rice, David A. Thorner, David J. Wild, Peter Willett, Frank H. Allen, and Robin Taylor*
- 631 Similarity and the Beilstein Information System: Searching for Concepts with Current Facts
Martin G. Hicks
- 639 Automatic Identification of Molecular Similarity Using Reduced-Graph Representation of Chemical Structure
Yoshimasa Takahashi, Masayuki Sukekawa, and Shin-ichi Sasaki*
- 644 Clustering of Chemical Structures on the Basis of Two-Dimensional Similarity Measures
J. M. Barnard and G. M. Downs*
- 650 Shape-Similarity Measures for Molecular Bodies: A Three-Dimensional Topological Approach to Quantitative Shape-Activity Relations
Paul G. Mezey
- 657 Structural Similarity Searching Using Descriptors Developed for Structure-Activity Relationship Studies
Philip N. Judson
- 664 Similarity Searching on CAS Registry Substances. 1. Global Molecular Property and Generic Atom Triangle Geometric Searching
William Fisanick, Kevin P. Cross, and Andrew Rusinko III*
- 675 Organic Reaction Similarity in Information Processing
Alexander J. Lawson

- 680 An Algorithm for the Multiple Common Subgraph Problem
Denis M. Bayada, Richard W. Simpson, A. Peter Johnson, and Claude Laurengo*
- 686 Similarity Based on Extended Basis Descriptors
Milan Randić
- 693 Similarity Ideas in the Theory of Pericyclic Reactivity
Robert Ponec and Martin Strnad*
- 700 Similarity Concepts for the Planning of Organic Reactions and Syntheses
Johann Gasteiger, Wolf-Dietrich Ihlenfeldt, Ralf Fick, and John Royce Rose*
- 713 Reaction Prediction: The Suggestions of the Beppe Program
Guido Sello
- 718 Computer-Assisted Perception of Similarity Using the Chiron Program: A Powerful Tool for the Analysis and Prediction of Biogenetic Patterns
Stephen Hanessian, Maurizio Botta, Benoit Larouche, and Ani Boyaroglu*
- 723 Automatic Interpretation of Infrared Spectra: Recognition of Aromatic Substitution Patterns Using Neural Networks
*U.-M. Weigel and R. Herges**
- 732 Computational Neural Networks as Model-Free Mapping Devices
G. M. Maggiora, D. W. Elrod,* and R. G. Trenary*
- 742 Application of Recurrent Neural Networks in Chemistry. Prediction and Classification of ^{13}C NMR Chemical Shifts in a Series of Monosubstituted Benzenes
Vladimír Kvasnička, Štěpán Sklenák, and Jiří Pospíchal*
- 748 Application of Genetic Algorithms in the Field of Constitutional Similarity
Eric Fontain

JOURNAL OF MATHEMATICAL CHEMISTRY

□ Nov. 1992, VOL. 11, NO. 4.

- Enumeration of perifusenes with one internal vertex: a complete mathematical solution
S.J. Cyvin, F. Zhang and J. Brunvöll 283
- The existence of Kekulé structures in helicenes and enumeration of concealed non-Kekuléan helicenes with $h \leq 13$
F. Zhang and X. Guo 293
- A basis set comparison in a variational scheme for the Yukawa potential
M. Demiralp, N.A. Baykara and H. Taşeli 311
- Formulation of N - and v -representable density functional theory. III: Excited states
E.S. Kryachko, E.V. Ludeña and T. Koga 325
- Summary of S_n mapping properties of some lexical p -tuples, or wordlengths, over $(M_1 - M_n)$ fields of Rota-Cayley algebra in generating Hilbert space $\text{SO}(3) \times S_n$ irreps of higher I_i identical spin clusters, $[A]_n(S_n)$, for $n \leq 7$
F.P. Temme and J.P. Colpa 341
- An existence theorem for molecular graphs determined by a sequence of valence states
V. Kvasnička and J. Pospíchal 353
- A strategy to derive new internal coordinates by partitioning the internal configuration space according to invariance properties
J.P. Leroy, R. Wallace and H. Rabitz 365
- Mathematical foundation of a global strategy for searching reaction paths
L.L. Stachó and M.J. Bán 405

寒い日が続いておりますが、部会員の皆様は如何お過ごしでしょうか。お伺い申し上げます。今年最初の部会誌をお届け致します。今年第11巻ということで10年ひと区切りと考えますとちょうど一巡して新たなスタートの年ということも出来ます。お陰様で欠号もなく順調に発行しておりますが、現状に満足することなく頑張っていきたいと思っております。皆様の御協力をお願い致します。

'89年から4年間、編集幹事としてCICSJの編集にあたってこられた日本アイ・ビー・エム株式会社の小出昭夫氏が今月(2月)をもって退任されることになりました。最近では理論化学計算や超並列コンピュータ等、小出氏が中心となって企画・編集し、好評をいただいた特集も多く、皆様の御記憶にも新しいことと思っております。本号にもUNIXに関して執筆していただくなど最後まで御無理をお願いしました。この場をお借りして小出氏に心から御礼申し上げます。

新しい編集幹事としては3月から三井東圧化学株式会社 三戸邦郎氏に加わっていただくことになりました。新しい視点から、これまでになかった企画・特集等をお届け出来るのではないかと思います。御期待下さい。

(編集代表 四居)

CICSJ Bulletin

Published Bimonthly by Division of
Chemical Information and Computer Sciences
The Chemical Society of Japan

部会員名簿 掲載号

日本化学会
情報化学部会

Volume 11, Number 2
March 1993

目 次

部 会 行 事	2
第16回情報化学討論会	
第3回情報化学部会講習会 —パーソナルコンピュータを利用した文献整理講習会 案内	
計算化学・情報化学サマーセミナー 案内	
部 会 記 事	5
情報化学部会 部会員名簿	
関 連 行 事	38
第21回構造活性相関シンポジウム	
記 事	
データベース作成における機械翻訳の利用	佐藤 雅之 39
日本特許情報機構の特許情報サービスについて	大山 勝弘 43
日本の研究者はCCSで困っている?	黒坂 厚 51
掲 示 板	54
文 献 紹 介	56
コンテンツ・リスト:	
Journal of Computer-Aided Molecular Design (Vol.6, No.6)	
Journal of Computational Chemistry (Vol.14, No.2~3)	
Journal of Chemical Information and Computer Sciences (Vol.33, Jan./Feb.)	

部 会 行 事

第16回 情報化学討論会

主催 日本化学会情報化学部会

共催 日本化学会、日本薬学会、日本農芸化学会、日本分析化学会、

日本プログラム交換機構

日時：11月24日(水)～26日(金)

会場：徳島大学大塚講堂、長井記念ホール、歯学部(徳島市庄町1丁目78-1)。

[交通] JR蔵本駅下車 徒歩5分、JR徳島駅下車 バス15分

なお、本討論会は第21回構造活性相関シンポジウムと併催されます。

討論主題：A. 化学情報, B. 計算化学, C. ケモメトリックス, D. 化学用ソフトウェア
E. ラボラトリーオートメーション

講演申込締切：7月9日(金)(郵送の場合)、7月23日(金)(電子メールの場合)

今回、従来の郵送による申し込みと電子メールによる申し込みを受け付けます。

学会等のデータベース化に伴い、なるべく電子メールでの申し込みをお願い致します。

【講演申込】〒101 東京都千代田区神田駿河台1-5 化学会館内 情報化学・構造活性
委員会 講演申込係 電話(03)3292-6190(担当：石鍋、清水)、FAX(03)3292-6318

電子メール：RGD03172@niftyserve.or.jp

【郵送の場合】1件ごとにA4版(縦)の用紙に横書きで次の事項を記入のうえ、上記
講演申込先宛てお送りください。(a)部会名：情報化学、(b)演題、(c)発表者と所属(講演者
に○印、全員にふりがな、所属が複数の場合は誰がどこの所属かわかるように記入してく
ださい)、(d)キーワード(討論主題を含め全角40字、10個以内)、(e)発表形式：口頭、
ポスター(申し込み件数が多い場合には、ポスター発表に切り換えて下さるようお願いす
ることがあります)、(f)連絡先。

【電子メールの場合】郵送の場合と同様の事項を電子メールにてお送りいただきますが、
詳細資料を上記講演申し込み先または下記問い合わせ先宛てご請求下さい。

講演要旨締切：9月18日(土)

講演1件についてA4版用紙4枚。他に半ページの英文要旨を添付して下記宛てお送りく
ださい。記入要領の詳細は申し込みを受理したのち送付します。

懇親会：11月25日(木)18時から。徳島大学蔵本学生会館にて第21回構造活性相関
シンポジウムと合同で開催します。会費：予約5,000円、当日6,000円

参加登録費(構造活性相関シンポジウムと共通)：

予約(一般)6,000円、(学生)4,000円。当日(一般)7,500円、
(学生)5,500円。講演要旨集(構造活性相関シンポジウム要旨集と合本)1冊を
含みます。講演要旨集の前送を希望される方は送料として1,000円追加。

参加登録、懇親会予約、要旨集前送申し込み締め切り：11月5日(金)

参加申込方法：A4版(縦)用紙に横書きで、氏名、勤務先、住所(TEL、FAX)、学生は
指導教官名、懇親会出欠、要旨集前送の有無をご記入の上、下記宛てお送りください。

銀行口座：銀行名 第一勧業銀行徳島支店

口座番号 普通預金 645-1431070

名義 情報活性委員会

実行委員会：寺田弘(徳島大学薬学部)、月原富武(徳島大学工学部)

講演要旨、参加予約申し込み、問い合わせ先

770 徳島市南常三島町2-1 徳島大学工学部

月原富武、森本幸生(TEL&FAX 0886-55-3160 ダイアルイン)

第3回情報化学部会講習会の御案内

パーソナルコンピュータを利用した文献整理講習会

主催 日本化学会情報化学部会

協賛 ユサコ(株)・キヤノン販売(株)

研究において文献整理は重要な仕事ですが、その量が多くなると必要なものがすぐに見つからない、と言ったことがよくあります。最近、抄録文献がフロッピーで配付され、ネットワークによる抄録文献データ利用や、個人の文献整理が非常に便利にできるようになってきました。またそれ以外にも、文献の引用のためのコピーandペーストなど、パソコンだからできるいろいろな活用法が考えられます。そこで、この新しい方法についてのパソコン(マッキントッシュ)講習会を企画しましたので、お気軽にご参加下さい。

日時：平成5年5月26日(水) 午前10時～午後5時

会場：キヤノン販売(株) セミナールーム 東京 大手町(詳細は後日案内)

参加申し込み：20名(定員になり次第締切)

参加資格：パソコン基礎知識(マッキントッシュのワープロが使える程度)のある方

参加費：10,000円(但し、情報化学部会非会員は15,000円)

内容：マッキントッシュを1台を2人で触りながら講習する。

10:00～12:00 CCOD (Current Contents On Diskette) の説明および使い方
ユサコ(株) 畠中正己

13:00～15:30 検索データ変換・処理の方法、文献整理ソフト(Pro-Cite)
の使い方
ユサコ(株) 畠中正己

15:30～17:00 応用事例一 個人文献データベースCCLDBとC60LDB
豊橋技術科学大学 工学部知識情報工学系教授 大澤映二

参加申し込み方法：FAXまたは葉書で「第3回情報化学部会講習会申し込み」と必ず標記し、参加者氏名、年齢、所属、連絡先(住所、電話番号、FAX番号)を明記して、下記お申し込み下さい。

(注) キャンセルは整理の都合で受け付けません。参加できなくなった時はできるだけ代理の方を立ててください。

申し込み、問い合わせ：101 東京都千代田区神田駿河台1-5

日本化学会 情報化学部会 講習会係

電話03-3292-6163、FAX 03-3292-6318

支払方法：1.郵便振替 口座/東京7-6058 社団法人 日本化学会、

2.銀行振込 第一勧銀神田支店 普通1073505 (社)日本化学会(振込日を連絡下さい)、3.現金書留

(いずれも第3回情報化学部会講習会と明記して下さい。)

なお、情報化学部会の入会(日本化学会員、非会員共 年会費2,000円)は常時受け付けておりますので、非会員の方は是非加入されることをお勧めします。問い合わせは上記「情報化学部会」までお願いします。

以上

計算化学・情報化学サマーセミナー

主催 日本化学プログラム交換機構、日本化学情報化学部会

日本化学情報化学部会では毎年夏期セミナーを開催してきましたが、今年は日本化学プログラム交換機構（JCPE）との共催で、セミナー名も上記のようにして行うことになりました。内容は「分子軌道計算のあれこれ」と題し、分子軌道計算関係の先生を多数お招きし、分子軌道計算の色々な面からの講演会を企画しました。分子軌道計算について深く勉強したいという方には最適の機会と思います。ふるって御参加ください。

日時 7月13日（火）～7月15日（木） 2泊3日

場所 栃木県鬼怒川温泉・ホテルたかはら

主題 「分子軌道計算のあれこれ」

講師 （お茶大理）平野恒夫、（広大理）今村詮、（阪市大理）西本吉助、
（お茶大理）細谷治夫、（埼大工）時田澄男、（お茶大理）長嶋雲兵、
（兵庫教大教）足立裕彦、 その他（依頼中を含む）

定員 100名（先着順、満員になり次第締切）

参加費 情報化学部会々員（含、JCPE個人会員）20,000円、日本化学会々員（含、
情報化学部会法人会員、JCPE法人賛助会員）30,000円、その他50,000円
（宿泊費等一切を含む）

申込方法 下記申込書（コピーしてお使いください）に記入の上、郵送またはFAXにてお
申し込み下さい。

申込および問い合わせ先

〒113 東京都文京区弥生2-4-16 学会センタービル 化学情報協会内
日本化学プログラム交換機構 電話 03-3816-3581,3582 FAX 03-3816-7826

計算化学・情報化学サマーセミナー参加申込書

住所	〒
所属	
氏名	
電話番号	
会員種別	情報化学部会々員・日本化学会々員・その他（いずれかの分類に○）

注意：参加申込を受け付けた方には請求書をお送りしますので、それから参加費用を振り込んで下さい。

関 連 行 事

第 2 1 回 構造活性相関シンポジウム

主催 日本薬学会医薬化学部会

共催 日本薬学会中四国支部、日本化学会、日本農芸化学会、
日本農薬学会、構造活性相関懇話会

日時：11月24日(水)～26日(金)

会場：徳島大学大塚講堂、長井記念ホール、歯学部(徳島市庄町1丁目78-1)。

【交通】JR蔵本駅下車 徒歩5分、JR徳島駅下車 バス15分

なお、本討論会は第15回情報化学討論会と併催されます。

講演申込締切：7月9日(金)(郵送の場合)、7月23日(金)(電子メールの場合)

今回、従来の郵送による申し込みと電子メールによる申し込みを受け付けます。

学会等のデータベース化に伴い、なるべく電子メールでの申し込みをお願い致します。

【講演申込】〒101 東京都千代田区神田駿河台1-5 化学会館内 情報化学・構造活性
委員会 講演申込係 電話(03)3292-6190(担当：石鍋、清水)、FAX(03)3292-6318
電子メール：RGD03172@niftyserve.or.jp

【郵送の場合】1件ごとにA4版(縦)の用紙に横書きで次の事項を記入のうえ、上記
講演申込先宛てお送りください。(a)部会名：構造活性、(b)演題、(c)発表者と所属(講演者
に○印、全員にふりがな、所属が複数の場合は誰がどこの所属かわかるように記入してく
ださい)、(d)キーワード(全角40字、10個以内)、(e)発表形式：口頭、(f)連絡先。

【電子メールの場合】郵送の場合と同様の事項を電子メールにてお送りいただきますが、
詳細資料を上記講演申し込み先または下記問い合わせ先宛てご請求下さい。

講演要旨締切：9月18日(土)

講演1件についてA4版用紙4枚。他に半ページの英文要旨を添付して下記宛てお送りく
ださい。記入要領の詳細は申し込みを受理したのち送付します。

懇親会：11月25日(木)18時から。徳島大学蔵本学生会館にて第15回情報化学討論
会と合同で開催します。会費：予約5,000円、当日6,000円

参加登録費(情報化学討論会と共通)：

予約(一般)6,000円、(学生)4,000円。当日(一般)7,500円、
(学生)5,500円。講演要旨集(情報化学討論会要旨集と合本)1冊を含みます。

講演要旨集の前送を希望される方は送料として1,000円追加。

参加登録、懇親会予約、要旨集前送申し込み締め切り：11月5日(金)

参加申込方法：A4版(縦)用紙に横書きで、氏名、勤務先、住所(TEL,FAX)、学生は
指導教官名、懇親会出欠、要旨集前送の有無をご記入の上、下記宛てお送りください。

銀行口座：銀行名 第一勧業銀行徳島支店

口座番号 普通預金 645-1431070

名義 情報活性委員会

実行委員会：寺田弘(徳島大学薬学部)、月原富武(徳島大学工学部)

講演要旨、参加予約申し込み、問い合わせ先

770 徳島市南常三島町2-1 徳島大学工学部

月原富武、森本幸生(TEL&FAX 0886-55-3160 ダイヤルイン)

データベース作成における機械翻訳の利用

日本科学技術情報センター
佐藤雅之

1. はじめに

近年、日本においても英語で書かれた論文が徐々に増えてきているが、まだまだ日本語で書かれた文献の量のほうが遥かに多い。これらの和文文献を国際レベルで流通させ、日本の優秀な技術の普及を図り、技術大国として国際社会に貢献する必要がある。また、1980年代の初めには、欧米諸国と日本の間に通商摩擦が起こり、その数年後には科学技術情報の分野にまで、その摩擦が拡大した。特に米国からは、「日本は欧米の研究成果を利用するばかりではなく、外国から得ているのと同じだけの恩恵を外国にも与えるように、先ずは、日本の科学技術情報を英語で提供すべきである」という強い批判を受けた。このような要求に答えるため、JICSTでは日本の科学技術情報の海外普及活動の一つとして、JICST-Eという英文データベースを作成し、オンライン情報検索システムに搭載して、1986年10月から提供してきた。

JICST-Eファイルは、国内刊行物に発表された、医学を含む科学技術全般の文献及び、国内の公共試験研究機関が発行した研究成果資料の全てを対象とし、そのタイトル、抄録、書誌事項を英文化したものである。このファイルは、1985年1月からデータの蓄積を開始し、1993年3月現在、約160万件の大型ファイルとなっている。

JICSTでは1990年7月以降、機械翻訳システムを利用してこの英文データベースを作成している。ここでは、機械翻訳が英文データベース作成にどのように使われているか、また、どのように使わなければいけないかという事を中心に述べてみたいと思う。

2. 人手による翻訳から機械援助翻訳へ

英文データベースを作成する手段として、JICSTでは従来次の2つの方法を使っていた。

- ① 著者が記述した英文標題や英文抄録があれば、英文標題については、そのまま、英文抄録については著作権者の利用許諾を得てから利用する。
- ② 英文標題や英文抄録のないものについては、日本語の標題や日本語抄録を人手で英語に翻訳し、それを利用する。

①の方法で英文抄録を利用する場合には、利用は原則として有料であり、入力に際して、文末に (author abst.) の文字を付加し、2000字を越える抄録については、文単位で文字を削除して (abridged author abst.) の文字を付加している。1991年3月までは②の方法で和文を英文に翻訳していたが、翻訳のための費用がかなりかか

るし、時間もかかり生産性が上がらない。生産性を上げるためには、機械翻訳システムを活用してより安く、より速く英文抄録を作成できるようにしなければならないということになり、人手による翻訳の代わりに、JICSTで開発した日英機械翻訳システムを使った翻訳を1990年7月から開始した。1991年4月以降は、人手による翻訳はなくなり、②の方法は全て機械翻訳を利用することになった。

現在の機械翻訳システムは完全なものではなく、後述するような前編集・後編集が必要となり、むしろ機械援助翻訳と言うべきものかもしれない。それにもかかわらず、あえて機械翻訳システムを利用した理由は、安く、速く、大量に翻訳することができるからである。

3. 機械翻訳システムの構造

JICSTでは、1986年4月より4年の歳月をかけて MAJESTIC (Machine Aided JICST's Exclusive System for Translating Information Comprehensively) と呼ばれる実運用日英機械翻訳システムを開発した。このシステムは、1982年から1985年までの4年間にわたる科学技術振興調整費による研究プロジェクト (通称Muプロジェクト) の成果を利用して開発した。

MAJESTICはC言語で記述されていて、約55万語の辞書を持ち、日立のメインフレームM-660K上にインプリメントされている。前編集と後編集は日立のワークステーション2050/32を使用している。システムは前編集部、翻訳部、後編集部の3部から構成されている。

ここで言う前編集と後編集とは機械翻訳システムによる翻訳作業の前後に人手により行う作業を指す。人間が翻訳をする場合にも、翻訳者が翻訳すべき文章を読んで、どういう意味で書いたかわからない部分があると、著者に問い合わせて文意を確認することがよくある。これが機械翻訳でいう前編集に相当する。また、高品質の翻訳を行う必要があるときには、人間の場合、まず粗訳を行い、その後にベテランの翻訳者やネイティブスピーカーが校閲や文体の統一などを行い、読み易い翻訳にする。それと同じように、機械翻訳でも、間違っている部分を直していくという後編集が必要になる。

翻訳部では、入力文章を解析、変換、生成していく。日本語文の場合には、日本語文法によって解析し、主語、述語、目的語などを決定し、日本語から英語に単語を入れかえ、文の組み立て順序を入れかえるというプロセスを経て英語文が出てくることになる。

文章の解析には文法規則と辞書が、日本語から英語に変換する場合には対訳辞書が、英文を作り出すためには英語の辞書と文法が必要になる。従って、辞書と文法がどの程度完備しているかが、機械翻訳システムの性能を決めることになる。

4. 前編集

日本語の解析が難しい理由の一つは、日本語には、主語や目的語の省略、係り受けの曖昧さ、特殊な言いまわし等が存在する点である。我々が文章を読む場合には、

省略や曖昧さは、前後の文脈から推定したり、補完したりできるが、現在の機械翻訳システムの多くは文脈の解析はできないので、翻訳の質を上げるためには、入力日本語を機械翻訳システムが理解できるような文に修正するのが今のところ一番良い方法である。このような作業が前編集作業である。もちろん、日本語の漢字の誤りを訂正したり、送り仮名の不統一を無くすといった事もこの作業に含まれる。

MAJESTICが翻訳対象としている抄録には、連用中止法や名詞句の並列を用いて、長く続く文が多い。また、主語を省略した表現や、1文中に動詞がたくさん使われている場合も多い。そこで前編集では、①原文の文意を損なわない範囲で、できるだけ簡潔な表現で、1文80字程度の「単文化」を図ること。②曖昧な係り受けなどがない明快な文とすることを作業のポイントとしている。

5. 後編集

入力日本語を理想的に前編集し、システムの翻訳メカニズムや文法規則が完璧で、十分な規模の辞書があれば、高品質の英文が得られても良いはずであるが、現実には必ずしも満足のいく翻訳結果が得られない。このような品質の英文は、文意を理解できる程度までに人手で修正した後に英文データベースとして提供しなければならない。この修正作業が後編集作業である。

後編集をどのレベルまで行うかは、出来上がった英文をどう使うかによって違ってくるが、Ian M. Pigott[1]によると、後編集のレベルは次の4通りに分けることができる。

- ①後編集なし : 情報のスキランニングが目的の場合。機械翻訳結果に少しくらいの重大なエラーがあったとしても、その分野のエキスパート達ならば、直ちに理解できるので後編集はしない。
- ②部分的後編集 : 誤った情報を伝える恐れのある重大なエラーのみをなくす目的で行う後編集。
- ③迅速後編集 : 構文と用語に関しては十分に正しい翻訳を目的とする後編集。文体に関しては、殆ど注意を払わない。
- ④完全後編集 : トップレベルの品質の翻訳を作るための手助けとして機械翻訳を使う場合。機械翻訳結果は粗訳として利用される。

これらのどのレベルの後編集を選ぶかは、機械翻訳結果の品質（入力文の質によって大いに変わることがある）や翻訳すべきテキストが属する技術分野の辞書のカバレッジなどによって違ってくる。

JICSTが提供する英文抄録は、その利用目的や、作成コスト、期間から考えると、④のレベルの後編集を行ってまで高品質にする必要はなく、③のレベルの後編集で十分であると考えられる。機械翻訳をうまく活用するためには、経済的な後編集を行わなければならない。つまり、全文をリライトしようとするのではなく、むしろ、最小限の変更で許容可能な原稿を作るような後編集を目指すべきである。そうすれば、従来の人手による翻訳に比べて、約2/3の費用で翻訳結果を入手することができる。

6. 運用実績

1992年度に、MAJESTIC を利用して翻訳することになっている文献数は、標題のみの翻訳が約76,000件、標題+抄録の翻訳が約15,000件である。したがって、年間約91,000件の文献を機械翻訳していることになるが、これはJICST-Eファイルの年間蓄積量の4割弱に相当する量である。MAJESTIC の処理能力を考えると、現在の処理量は決して多いとは言えないが、英文データベース作成のための予算的制約や、前編集・後編集のマンパワー的な制約などにより、処理量が急激に増加することは期待できない。

7. おわりに

現在実用化されている機械翻訳システムの翻訳メカニズムは、人間が実際に脳の中で行っている翻訳メカニズムとは本質的に違っている。人間と同じメカニズムで翻訳するシステムは、当分の間は出現しそうにもないので、現在の機械翻訳システムを少しずつ改良しながら、使い込んでいくしかないだろう。

JICSTにおいても、MAJESTIC の改良作業は、実用化された現在でも継続されており、今後も続けていかなくてはいけない。何故ならば、たとえ理想的な前編集ができたとしても、全ての機械翻訳結果を後編集なしでそのまま使えるわけではないこと、大量処理を前提とする実運用の世界では、理想的な前編集はとうてい望めないこと、翻訳対象分野の広がりや、種々の文体、表現に対応するために、辞書情報や文法規則を拡充していかなくてはいけないこと等のためである。

今のところ、どんな文章でも翻訳できるような機械翻訳システムは存在しないのであるから、システムを如何に上手に使っていくかが大切な事である。機械翻訳システムの癖、長所、短所を熟知して使いこなせば、十分に効果を発揮し、人手だけにたよる翻訳よりも速く、また安くできるのである。

参考文献

- (1) Pigott, Ian M. Document Preparation and Post-editing for Machine Translation

日本特許情報機構の特許情報サービス について

(勸)日本特許情報機構 パトリス部 大山 勝弘

1. はじめに

日本特許庁（以下特許庁という）は、本年1月からCD-ROMを媒体に使った公開公報の発行を開始した。毎週2回、各1枚発行されるこのCD-ROM公開公報には、出願後18ヶ月経過した特許、実用新案合わせて5,000件が収録されている。その内容は、全文の文字データと全図面のイメージデータがミックストモードで記録されている。

このように、特許情報の分野においても新しい媒体を使用した情報の流通が始まり、今後のデータベース構築とその利用法の研究開発等が活発化していくものと期待される。

本記事では、そうした情勢も入れながら、最近の特許情報の提供の流れを紹介したい。

2. ペーパーレス計画

技術開発、技術革新により世の中は急速に変化し、高度情報化社会の実現をめざして、情報処理に関係あるいろいろな分野でいくつかの大きな計画が進行している。その中で、特許行政における最大の関心事は特許庁が昭和59年度に開始したペーパーレス計画である。

特許庁は、工業所有権制度の改革、工業所有権制度の国際的ハーモナイゼーション（調和）などの施策とともに、迅速的確な権利付与を図るべく、ペーパーレス計画を始めとした総合的な施策を推進している。このペーパーレス計画とは、従来ペーパー（書面）で行ってきた工業所有権の出願・受付・審査・審判・登録等の業務および工業所有権情報提供サービスを、コンピュータを利用して行うことによりこれらの業務の効率を高めようとするものである。現在までに、特許文献の光ディスク（以下ODという）入力、特許・実用新案の電子出願、電子公報発行、Fターム検索による審査等が実施に移されている。

これらのペーパーレス計画の成果物は、(勸)日本特許情報機構（本部：東京都江東区東陽4-1-7。以下Japioという）を通して一般にも公開、提供されている。

3. Japioの特許情報サービス

Japioは、企業活動の多様化や技術開発の活発化を反映して年々増加し高度化・複雑化する内外国の特許情報を、総合的に収集・処理・加工し、多くのユーザーに利用しやすい形態で提供する非営利の特許情報専門機関である。

Japioは、公報の販売・複写から電算機処理による加工製品まで、特許情報に関するあらゆるサービスを提供している。その特長は広範な情報源とサービスの多様性である。

すなわち、取り扱う情報の広さ（国内の特許情報から、外国の特許情報まで。そのうち国内については、特許、実用新案、意匠、商標をすべてカバー）、加工段階の多様さ（一次情報＜公報＞から、二次情報＜抄録・索引＞、さらにデータベースまで）、媒体の多様

性（紙、マイクロフィルム、磁気テープ、フロッピーディスク、CD-ROM、OD、さらにオンラインまで）等である。

具体的なサービスの種類を列挙すると以下の通りである。

- ・特許公報類販売（紙公報、CD-ROM公報、外国の特許公報類の頒布・取り寄せ等）
- ・公報類の複写サービス
- ・抄録誌（公開特許全件）、索引誌（特許、実用新案）の発行
- ・データベースを利用した検索サービス（オンライン検索、バッチ検索等）
- ・電子媒体を利用したデータ提供サービス（磁気テープ、CD-ROM、OD等）
- ・電算機システム販売（ハードウェア、ソフトウェア）
- ・調査・翻訳サービス
- ・閲覧サービス、その他

これらのうちデータベースを利用したオンライン検索サービスについては別項で紹介するので、ここでは最近の新しい電子媒体を利用したサービスに関して簡単に説明する。

・ CD-ROM

特許庁から交付された公報イメージデータをもとに、昭和62年以後に発行された公開特許・実用新案公報全頁をCD-ROMに記録して販売している。記録方式は、文字情報・図面情報とも全てファクシミリモードである。図面情報には化学構造式も含まれる。

全技術分野全件を収録し、月に数回、公開日から約6週間遅れで公開番号順CD-ROMを作成してきたが、平成5年1月発行の公開公報からは、特許庁によってCD-ROM公報が発行されるようになったのを機に、Japio独自の公開番号順CD-ROMは平成4年12月分迄で作成を終了した。

・ OD

特許庁から交付された公報イメージデータ中の公開特許・実用新案公報、公告特許・実用新案公報をもとに、顧客から指定された分野、期間等に限定してデータを抽出し、指定されたOD装置の機種に合った仕様に再加工して販売している。新規発行分の追記サービスも行っている。なお、これを利用するためのOD装置は特定の機種に限定されている。

記録方式は、文字情報・図面情報とも全てファクシミリモードであり、図面情報には化学構造式も含まれている。

ODには、番号、日付、国際特許分類（以下IPCという）、出願人コードの他に、利用者の社内分類もインデックス化できるようになっているため、これらを使って検索ができるようになっている。検索結果から、OD中の図面を含む全文をディスプレイ表示あるいはプリントアウトする。

・ Japioの閲覧サービス

Japioの本部には一般利用者向けに無料閲覧室が設けてあり、そこでは特許庁発行のCD-ROM公報およびJapio作成のCD-ROMを利用して公開特許公報、公開実用新案公報の閲覧ができるようになっている。CD-ROMは、閲覧室内に開架式で配列されており、利用者は自由に備えつけのCD-ROMビューア装置で

内容の閲覧ができる。

4. 特許情報のデータベース

Japioが作成するデータベースと蓄積状況は表1に示す通りである。

基本的な事項は特許庁から交付されたデータを使用し、それにJapioが独自に作成した技術内容データ（抄録、図面、キーワード、分類等）を追加蓄積している。図1、図2に、フリーキーワードの一部と固定キーワードの一部をそれぞれ示す。

また、対応特許を検索するのに有効な外国の特許情報のデータベースは、ヨーロッパ特許庁（EPO）との間に相互の協定を結んで、それによって入手した磁気テープデータをもとに作成している。データ項目には、ケミカルアブストラクト抄録番号やダーウェント抄録番号も含まれる。

表1. データベース一覧

(平成5年2月末日現在)

種類	情報の種別	データ蓄積期間	累積蓄積件数
日本特許	公開	発行日 1971. 7. 1 ~ 1993. 2. 12	4,614,050 件
	公告	発行日 1955. 1. 1 ~ 1993. 2. 12	1,615,300 件
	審判	請求日 1982. 9. 1 ~ 1992. 12. 18	186,128 件
	登録	登録日 1967. 6. 19 ~ 1992. 11. 27	1,087,630 件
日本実用新案	公開	発行日 1971. 7. 1 ~ 1993. 2. 12	3,532,850 件
	公告	発行日 1960. 1. 1 ~ 1993. 2. 12	1,459,780 件
	審判	請求日 1982. 9. 1 ~ 1992. 12. 18	113,115 件
	登録	登録日 1974. 11. 20 ~ 1992. 11. 25	883,904 件
日本意匠	登録公報	発行日 1965. 1. 1 ~ 1992. 11. 12	802,575 件
	審判	請求日 1982. 9. 1 ~ 1992. 12. 18	11,122 件
	登録	登録日 1968. 5. 11 ~ 1992. 9. 29	520,081 件
日本商標	出願	発行日 1888. 10. 12 ~ 1992. 12. 11	2,186,768 件
	公告	発行日 1921. 6. 20 ~ 1993. 2. 12	2,388,428 件
	審判	請求日 1982. 9. 1 ~ 1992. 12. 18	78,782 件
	登録	登録日 1902. 4. 21 ~ 1992. 11. 30	1,955,849 件
外国特許 (INPADOC)	特許	発行日 1968. 1. 1 ~ 1992. 12. 31 (国により蓄積期間等に相違あり)	20,642,327 件

5. 特許情報オンラインサービス

前項で紹介したデータベースを使って、Japioが提供している特許情報オンラインサービスシステムをPATOLIS®と呼んでいる。

コマンド方式を主体とするオンライン検索システムで、基本的な検索方式はインデックスサーチ（IPC、出願人、フリーキーワード、抄録キーワードなど約80項目が検索可能）であるが、抄録文テキストサーチ（文字列のストリングサーチ）の機能も備えている。

化合物等の検索には、IPC、フリーキーワード、FI（特許庁ファイルインデックス）あるいはFターム（File Forming Term）等を利用する方法があるが、フリーキーワードの場合は、（ポリ AND ビニル AND アルコール）のように短い単語の組合せで質問式を作成することが必要である。また、インデックスサーチによる抄録キーワード検索では、各キーワードは抄録中での出現位置を示すアドレスを持っているので、隣接演算、近接演算の機能も使用できる。隣接演算は複合語を検索するのに有効である。入力日本語（漢字、ひらかな、カタカナ混じりの表記）の用語で実行できる。異なる項目同士の組合せによる検索も可能である。

回答出力は、オンライン上では、審査経過情報付きや抄録付きなど各種の回答様式で随時見ることができ、オフライン出力では抄録・図面データ付きの様式などを郵送で入手できる。また、ファクシミリ出力を利用すると、オフライン出力と同じ抄録・図面データ付きの様式などをその日のうちに入手できる。図3に出力例を示す。

PATOLIS®はまた、特許庁が推進しているペーパーレス計画の数々の成果物を有効に提供・活用する役割を遂行することを大きな目的のひとつとしている。

その中の一つであるFターム検索は、技術分野（テーマ）毎の詳細な技術観点別分類表を作成して、それによりデータ解析と蓄積、検索を行うものである。表2に示したのは、PATOLIS®により公開されたFターム検索システムの中の化学分野のテーマだけを抜粋したものである。今後も、半年毎に新しいテーマが追加公開されていく予定である。

6. おわりに

Japioが行っている特許情報サービスは、特定の出願に関する情報をスポット的に必要としている顧客に対応するものから、社内に特許情報のデータベースを構築して独自の情報システムにより継続的に運用している顧客への対応まで幅広い。

また、特許庁が実施する特許制度の変更等にも迅速、柔軟に対応して特許情報提供活動を停滞させることなく継続していくことも要求される。

今後も、各方面からのご指導を戴きながら、特許情報サービスの普及に努めていきたい。

7. Japioサービスに関する問い合わせ先、資料請求先

〒100 東京都千代田区霞が関3-4-2 商工会館・弁理士会館ビル
（財）日本特許情報機構 サービス部

TEL. 03-3503-6181 FAX. 03-3580-7164

表 2. F タームテーマ一覧 (IPC Cセクション関係)

平成 4 年 1 0 月 3 1 日時点

テーマコード	テーマ名称	関係IPC	テーマコード	テーマ名称	関係IPC
4G040	水素、水、水素化合物	C01B	4C057	糖類化合物	C07H
4G042	酸素；オゾン；酸化物一般	C01B	4J006	ポリハロゲン化オレフィン	C08F
4G002	鉄化合物	C01G	4J015	重合触媒	C08F
4D003	生物膜廃水処理	C02F	4J016	重合後の処理	C08F
4D024	吸着による水処理	C02F	4J019	オレフィン系重合体	C08F
4D025	イオン交換による水処理	C02F	4J020	スチレン系重合体	C08F
4D027	生物学的処理一般	C02F	4J022	酸素含有重合体	C08F
4D028	活性汚泥処理	C02F	4J023	アクリル系重合体	C08F
4D029	活性汚泥処理における曝気装置	C02F	4J025	江、アセリ、炭化水素他系重合体	C08F
4D034	水、廃水または下水の加熱処理	C02F	4J027	マクロモノマー系付加重合体	C08F
4D037	物理的水処理	C02F	4J001	ポリアミド	C08G
4D038	特定物質の除去	C02F	4J005	ポリエーテル	C08G
4D039	中和・改質	C02F	4F006	高分子成形体の被覆	C08J
4D040	嫌気、嫌気好気又は生物に特徴ある処理	C02F	4J002	高分子組成物	C08K
4D045	スケール防止	C02F	4J004	接着テープ	C09J
4D046	水・廃水の多段階処理	C02F	4H017	シーリング材組成物	C09K
4D049	殺菌剤による水の殺菌処理	C02F	4H018	抗スリップ物質	C09K
4G020	ガラスの成形	C03B	4H020	対水表面処理用物質	C09K
4G021	ガラス繊維の製造、処理	C03B	4H026	土壌改良剤および土壌安定剤	C09K
4G001	セラミック製品	C04B	4H027	液晶物質	C09K
4G032	セラミック製品(2)	C04B	4H028	防火用物質	C09K
4G033	セラミック製品(3)	C04B	4H012	コークス工業	C10B
4G016	熔融スラッグの処理	C04B	4H029	炭化水素油の製造、分解及び精製	C10G
4G017	水硬性セメントの製造一般	C04B	4H015	固形燃料および燃料付随物	C10L
4G018	磁性セラミックス	C04B	4H004	潤滑性組成物	C10M
4G019	多孔質人造石または多孔質セラミック製品	C04B	4H003	洗浄性組成物	C11D
4G025	粘土製品	C04B	4B028	発酵液の蒸留、酒類加工、食酢及びビール	C12C
4G026	セラミックスの接合	C04B	4B015	酒類	C12G
4G028	人造石、天然石の後処理	C04B	4B029	微生物・酵素関連装置	C12M
4G030	酸化物セラミックスの組成(1)	C04B	4B024	突然変異または遺伝子工学	C12N
4G031	酸化物セラミックスの組成(2)	C04B	4B033	酵素、微生物の固定化、処理	C12N
4C005	ペニシリン系化合物	C07D	4K012	鉄の製造	C21B
4C012	トリアゾール系化合物	C07D	4K015	溶鉱炉	C21B
4C013	トリアジン系化合物	C07D	4K002	炭素鋼または鉄鋼の製造	C21C
4C014	ジアゼピン系化合物	C07D	4K013	熔融状態での鋼の処理	C21C
4C031	キノリン系化合物	C07D	4K014	鉄鉄精製；鉄鉄製造；転炉法以外の製鋼	C21C
4C032	1, 2-ジアゾール系化合物	C07D	4K032	鋼の加工熱処理	C21D
4C033	チアゾール系化合物	C07D	4K033	電磁鋼板の製造	C21D
4C037	フラン系化合物	C07D	4K001	金属の製造または精製	C22B
4C049	プリン、プテリジン系化合物	C07D	4K016	合金又は金属化合物を含有する合金	C22C
4C055	ピリジン系化合物	C07D	4K019	非鉄合金の製造	C22C
4C019	1, 2-ジアジン系化合物	C07D	4K020	複合金属又は合金の製造	C22C
4C020	1, 3-ジアジン系化合物	C07D	4K022	化学的被覆	C23C
4C044	1, 4-ジアジン系化合物 (化合物)	C07D	4K026	金属の化成処理	C23C
4C015	5員環以上窒素含有飽和複素環式化合物	C07D	4K027	熔融金属による被覆	C23C
4C036	6員以上のNS含有複素環式化合物	C07D	4K028	金属質材料の表面への固相拡散	C23C
4C022	二個以上酸素原子を含む複素環式化合物	C07D	4K029	物理蒸着	C23C
4C023	硫黄原子を含む複素環式化合物	C07D	4K030	CVD	C23C
4C034	その他の1N系複素環式化合物	C07D	4K011	化合物または非金属の製造のための電極	C25B
4C050	その他の窒素含有縮合複素環式化合物(3)	C07D	4K021	非金属・化合物の電解製造；その他装置	C25B
4C064	その他のN系縮合複素環式化合物(1)	C07D	4K023	電気鍍金、そのための鍍金浴	C25D
4C035	他の環と縮合した1, 3-ジアジン環	C07D	4K024	めっき方法・物品	C25D
4C056	N, O含有複素環式化合物	C07D			

各テーマ毎の詳細は、「Fタームリスト」または「Fターム解説書」(有料販売中)に記載されています。

カルボキシ塩 (F042459)		カルボジミド (F042500)	
カルボキシル塩基 (F042460)		カルボスルファン (F042503)	
カルボキシル化 (F042462)	←カルボキシ化	カルボチオ酸 (F042508)	
カルボキシル化剤 (F042463)		カルボツクス (F042509)	
カルボキシル化物 (F042464)		(カルボナート)	→カーボネート (F043152)
(カルボキシル化法)	→カルボキシル化	(カルボナート類)	→カーボネート (F043152)
	(F042462)	カルボナミド (F042510)	
カルボキシル価 (F042461)		カルボニ	
カルボキシル基 (F042465)	←CO↓2H, COOH,	(カルボニール)	→カルボニル (F042521)
	カルボキシ基, カルボキ	カルボニウム (F042513)	
(カルボキシル基型)	シル基	カルボニウム塩 (F042514)	←カーボニウム
	→カルボキシル基	カルボニツク (F042516)	
(カルボキシル基系)	(F042465)	カルボニフイツク (F042518)	
	→カルボキシル基	カルボニル (F042521)	←カーボニール, カーボニ
	(F042465)		ル, カルボニール
カルボキシル基側 (F042467)		カルボニル塩 (F042522)	←CO塩
カルボキシル基量 (F042468)	←COOH量	カルボニル化 (F042524)	←カーボニル化
(カルボキシル型)	→カルボキシル (F042458)	カルボニル化剤 (F042525)	
(カルボキシル系)	→カルボキシル (F042458)	(カルボニル化法)	→カルボニル化 (F042524)
(カルボキシル酸)	→カルボン酸 (F042567)	カルボニル価 (F042523)	
カルボキシル酸塩 (F042470)		カルボニル基 (F042531)	←C=O基, CO基, カル
カルボキシル性 (F042472)		(カルボニル型)	ボニ基
カルボキシル側 (F042473)		(カルボニル系)	→カルボニル (F042521)
カルボキシル体 (F042474)		(カルボニル酸)	→カルボニル (F042521)
カルボキシル端 (F042475)		(カルボニル側)	→カルボン酸 (F042567)
カルボキシレー		カルボニル粉 (F042533)	→カルボニル (F042521)
カルボキシレート (F042476)	←カーボキシレート, カル	(カルボニル類)	
	ボキシレート, カルボキ	カルボニロキシ (F042534)	
	シラト, カルボキシレイ	カルボニン (F042536)	
	ト, カルボキレート	カルボニン化 (F042537)	
		(カルボニ基)	→カルボニル基 (F042531)
カルボキシレート塩 (F042477)		カルボネ	
カルボキシレート化 (F042478)	→カルボキシレート	(カルボネート)	→カーボネート (F043152)
(カルボキシレート基)	(F042476)	(カルボネート類)	→カーボネート (F043152)
(カルボキシレート型)	→カルボキシレート	カルボノ (F042539)	
(カルボキシレート系)	→カルボキシレート	カルボノキシ (F042540)	
(カルボキシレート類)	(F042476)	カルボヒドラーゼ (F042543)	
(カルボキシレイト)	→カルボキシレート	カルボフェラス (F042544)	
	(F042476)	カルボポール (F042546)	
(カルボキシ化)	→カルボキシレート	カルボマイシン (F042547)	
	(F042476)	カルボミン酸 (F042549)	
(カルボキシ基)	→カルボキシル化	カルボメート (F042551)	
	(F042462)	(カルボモイル)	→カルボモイル (F042325)
(カルボキシ型)	→カルボキシル基	カルボラクトン	
(カルボキシ酸)	(F042465)	カルボラン (F042556)	
カルボキシ側 (F042485)	→カルボキシ (F042447)	カルボリ (F042557)	
カルボキシ量 (F042486)	→カルボン酸 (F042567)	カルポリン (F042558)	
カルボキシニウム (F042487)		カルボン (F042559)	
(カルボキレート)	→カルボキシレート	カルボン塩 (F042563)	
	(F042476)	カルボン化 (F042565)	
(カルボクサミド)	→カルボキサミド	カルボン基 (F042566)	
	(F042442)	(カルボン型)	→カルボン (F042559)
カルボククロメン (F042489)		カルボン酸 (F042567)	←カルボキシル酸, カルボ
カルボコン (F042490)		カルボン酸 (塩)	キシ酸, カルボニル酸
カルボキサミド (F042491)		カルボン酸塩 (F042570)	
カルボシ (F042492)		(カルボン酸塩型)	→カルボン酸塩 (F042570)
カルボシキル基 (F042496)		(カルボン酸塩系)	→カルボン酸塩 (F042570)
カルボシステイン (F042497)		(カルボン酸塩類)	→カルボン酸塩 (F042570)
カルボジ		カルボン酸化 (F042572)	
		カルボン酸化剤 (F042573)	

図1 フリーキーワード一覧表 (部分)

固定キーワード表(分野別)

共 通		金 属		
(R001)	レーザー	R029	原子力製鉄(←直接製鉄・直接還元)	
R002	レーザー	R030	エレクトロスラグ技術(ESR)	
R003	電子ビーム	R031	粉末冶金(←サーメット、超硬合金)	
R115	X線応用(←検査等)	①	R032	繊維強化複合金属(FRM)
R004	プラズマ	R033	マルエージング鋼	
R005	圧電・強誘電体	R034	低温用材料	
R006	超電導	R134	形状記憶合金	
R007	超音波	R135	アモルファス金属	
(R008)	磁気浮上	R035	精密鑄造(←シェルモールド法、ロストワックス法)	
R009	ホログラフィ	R036	自動溶接	
R010	ホトクロミズム化合物	化 学		
R011	液晶	R037	排煙脱硝(←NO _x 用触媒)	
R012	光学繊維(オプティカルファイバ)	R038	排煙脱硫(←SO ₂ 用触媒)	
R013	マイクロカプセル	R117	液化ガス(LNG、LPG)(←貯蔵・輸送)	
R014	マイクロフィルタ	R039	海水淡水化	
(R015)	水中燃焼	R122	繊維強化セメント	
R016	帯域溶融	R123	繊維強化セラミックス・繊維強化ガラス	
R017	粉体化技術	R113	フルラン	
R018	流動層(床)	R127	固定化酵素	
R019	エアゾール	R040	強化プラスチック(FRP)	
R020	真空技術	R041	分解型プラスチック	
R021	高压技術	R042	親水性(水溶性)プラスチック	
R022	超低温	R043	光導電性プラスチック	
R023	流体素子	R044	感光性樹脂(←フォトレジスト、エレクトロレジスト)	
(R111)	逆浸透	①	R045	嫌気性樹脂
R120	限外濾過(UF)②(←逆浸透①)	⑥	R046	ガスバリア樹脂
R121	磁性流体	③	R118	難燃性樹脂(防炎、自己消火性)
(R900)	超材料、超技術	①	R119	耐熱性樹脂
食 品		R047	リキッドラバー(液状樹脂)	
R024	人造肉	R124	エポキシ樹脂	
R025	ダイエットフード	R125	ポリカーボネート樹脂	
R026	インスタント食品	R048	合成木材	
R027	新蛋白源(←クロレラ)	R049	無公害洗剤	
(R028)	宇宙食	R126	粉末塗料	
		医 薬		
		R050	ピル(径口避妊薬)	
		R051	制ガン剤(←抗腫瘍剤・抗白血病剤)	
		R133	インターフェロン	
		医 療		
		R085	自動検診	
		R086	人工臓器(←人工血管)	

(注) ①は昭和51年2月より蓄積。
 ②は昭和51年4月より蓄積。
 ③は昭和51年8月より蓄積。
 ④は昭和52年3月より蓄積。
 ⑤は昭和58年8月より蓄積。
 ⑥は昭和59年1月より蓄積。
 ⑦は昭和55年1月より蓄積。
 (R...)は現在廃止のコード。

図2 固定キーワード表(部分)

P 開 2- 785 出 63-325701

H02.01.05 S63.12.23

2 5-ジメルカプト-1,3,4-チアアジアゾールを含有する潤滑剤組成物

チバ ガイギー AG (ヒューゴ カメンジント) <優CH 87.12.23 87 5033>

〔目的〕潤滑油等及び新規なジメルカプト-チアアジアゾールとを含有してなる、極圧及び耐摩耗特性、酸化防止特性、腐食特性の優れた標記組成物。

〔構成〕(A) 潤滑剤もしくは圧媒油、及び(B) 式〔R¹~R⁴はC₁₋₁₂アルキル、C₃₋₁₂アルケニル、(置換)フェニルC₁₋₁₂アルキル、R¹~R⁴はそれらが結合しているNと一緒になって(置換)モルホリン、ピペリジン、ピロリジン、ピペラジン、メチルピペラジン、ペルヒドロアゼピン; X, X¹はH, C₁₋₂₀アルキル、フェニル、フェニルC₁₋₁₂アルキル、C₁₋₁₂アルコキシ、OH, ジ-C₁₋₁₂アルキルアミノ、ハロゲン置換フェニル等〕で示される2,5-ジメルカプト-1,3,4-チアアジアゾール、とを含有してなる目的の組成物。尚、式の化合物は新規物質で、1モルのヒドラジン等と2モルの二硫化炭素をアルカリ媒体中で反応させ、続いて酸性化して得られる。

<ジメルカプト, チアアジアゾール, 含有, 潤滑剤 組成物, 潤滑油, 新規, ジ, メルカプト, 極圧, 耐摩耗 特性, 酸化 防止 特性, 腐食 特性, 組成物, 潤滑剤, 圧媒油, B, 式, C↓1, アルキル, C↓3, アルケニル, 置換, フェニル, 結合, N, モルホリン, ピペリジン, ピロリジン, ピペラジン, メチル ピペラジン, パー ヒドロ, アゼピン, H, アルコキシ, OH, アルキル アミノ, ハロゲン 置換 フェニル, 目的, 化合物, 新規 物質, 1モル, ヒドラジン, モル, 2 硫化 炭素, アルカリ 媒体, 反応, 酸性 化> [141,146,222]

C07D285/125

C10M135/36

#C10N 30:06

#C10N 30:10

#C10N 30:12

#C10N 40:08

C07D285/12

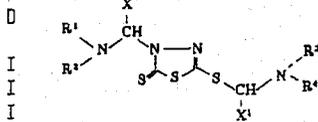
C10M135/36

C10N 30:06

C10N 30:10

C10N 30:12

C10N 40:08



P 開 2- 45403 出 1-160996

H02.02.15 H01.06.26

ジフルオロメチル-チアアジアゾール-オキシアセトアミドの選択除草剤としての使用

バイエル AG (ハインツ フェルスター, ハンス-ヨアヒム ザンテル, ロベルト アール シュミット, ハリイ ストラング) <優DE88.06.27 88 3821597>

〔目的〕新規化合物を含むジフルオロメチル-チアアジアゾール-オキシアセトアミドを含有してなる、発芽前法を用いて単子葉及び双子葉作物における単子葉雑草の防除に特に適した選択除草剤。

〔構成〕式 I (R¹はH, アルキル, アルケニル, アルキニル, アラルキル; R²はアルキル, アルケニル, アルキニル, シクロアルキル, アリール, アルコキシ等; R³とR⁴は結合するNと一緒になって複素環を形成) の化合物を含有させる。式 I の

化合物において、R¹がC₁₋₁₂のアルキル、R²が(置換)アリールである化合物、例えばN-メチル-2-(5-ジフルオロメチル-1,3,4-チアアジアゾール-2-イル-オキシ)-アセトアニリド、は新規であり、式IIの5-ジフルオロメチル-2-メチルスルホニル-1,3,4-チアアジアゾール(新規物質)を式III(A¹はC₁₋₁₂アルキル; A²はアリール)のヒドロキシアセトアミドと反応させることによって得られる。

<ジフルオロメチル, チアアジアゾール, オキシ, アセトアミド, 選択 除草剤, 使用, 新規 化合物, 含有, 発芽前法, 単子葉, 双子葉 作物, 単子葉 雑草, 防除, 式, †1, H, アルキル, アルケニル, アルキニル, アラルキル, シクロ アルキル, アリール, アルコキシ, 結合, N, 複素環, 形成, 化合物, C↓2, 置換, メチル, ジアジ, アゾル, イル, アセト アニリド, 新規, メチル スルホニル, チア アジアゾール, 新規 物質, ヒドロキシ, 反応><R013 > [144,141]

A01N 43/82

101

C07D285/125

C07D285/13

C07D417/12

C07D285/12

C07D285/12

C07D417/12

A01N 43/82

C

D

101B

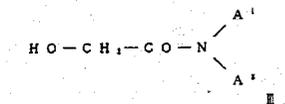
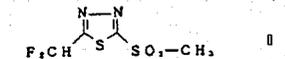
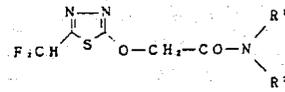


図3 回答出力例 (FAX出力)

日本の研究者はCCSで困っている？

(株)化学工業日報社 黒坂 厚

コンピューターケミストリーシステム(CCS)は、1990年代に入って着実に国内に浸透しつつあるが、その歩みは関係各界の期待とは裏腹に非常にゆっくりとしている。CCSで日本は米国に5-10年遅れていると、5年前から(あるいはもっと前から)いわれているが、最近の米国のトップユーザーの利用の高度さをつぶさにみるにつけ、少しベシミスティックかもしれないが、果たしてあと5年してもいまの米国の利用レベルに達することができるだろうかと、不安が募ってくる。日本と米国では一体どこがどのように違っているのだろうか。

◇「日本のユーザーは困っていなかったのかもしれない」◇

コンピューターの世界にX/Openという組織がある。富士通からの出向でこの日本事務所長を務めていた田中幹雄氏(今年の1月に富士通に戻られました)が、任期中の4年間を振り返ってこのように述べている。「日本のユーザーは困っていなかったのかもしれませんが」と。このことについて、少し考えてみたい。日本のCCSの現状を理解するのに役立つような気がする。

◇オープンシステムの経緯に学ぼう◇

X/Openは1984年に英国で設立された。株式会社組織になっているが、実際には非営利の業界団体だ。その使命は、UNIXを中心としたオープンシステムの推進。OS(基本ソフト)とアプリケーションの間に介在するさまざまなインターフェースを規定し、異なるプラットフォーム間でのアプリケーションの移植性、オープンなネットワーク環境での相互運用性などを実現することである。

組織的には、米国のIBMやAT&T、サン・マイクロシステムズ、ヒューレット・パカード、欧州のICL、ブル、シーメンス、フィリップス、オリベッティ、日本からは富士通、NEC、日立製作所など、世界の主要なコンピューターメーカーが株主会員となっているほか、一種の諮問機関としてソフトウェア会社を組織した「ISVカウンシル」、エンドユーザーを組織した「ユーザーカウンシル」などから成り立っている。ここで取り上げたいのはこの「ユーザーカウンシル」の存在である。

X/Openの「ユーザーカウンシル」には約30社が加盟している。化学に関係がありそうなところだけでも、アモコ、アーク・オイル&ガス、デュボン、イーストマン・コダック、エクソン、GE、シェル、テキサコなど意外に多い。ところが日本企業となると、化学会社に限らずまったくのゼロだった。一昨年にNTTが入っただけである。冒頭に紹介した元X/Open日本事務所長の田中

氏の発言はこのことに関してなのである。田中氏もなんとか日本企業に入ってもらおうと奔走したという。しかし、残念ながら成果はほとんどあがらなかったわけだ。X/Openと同様のオープンシステムの業界団体で、OSF（オープンソフトウェアファウンデーション）というものがある。これにも「エンドユーザーカウンシル」があるが、日本に関してはX/Openの「ユーザーカウンシル」と同じ状況だ。

オープンシステムを推進するコンピューター業界の団体に、なぜユーザーが入らなければならないのか、もしかしたら奇異に思うかもしれない。これは、オープンシステムのムーブメントが発生した歴史的経緯をみればはっきりする。オープンシステムの核であるUNIXは1969年に研究用のOSとしてAT&Tベル研究所で誕生した。1979年のバージョン7を境に、かつてUNIXの二大潮流と呼ばれたAT&TのUNIXシステムV（SVR）とカリフォルニア大学バークレー校のBSDに分かれていく。米国の大学や研究機関などのOS研究者の間に広く流布し、これが後に「/usr/group」と呼ばれるユーザー団体（現在ではUNIFORMと称している）の母体になり、さらにIEEE（米国電気電子技術者協会）によるPOSIX標準の確立へとつながった。POSIXは、/usr/groupの後押しでIEEEが制定した移植性の高いOSインターフェースの国際規格。これが現在のオープンシステムの原点になったものであり、その原動力は紛れもなくコンピューターユーザーの力だった。コンピューターのOSが数ある中で、ほとんど唯一、UNIXだけがユーザーによって作り育てられてきたOSなのである。

オープンシステムは、コンピューターメーカーのビジネス戦略として登場したのではなく、その主役はユーザーだった。これ以上詳しい経緯は省くが、その後のUNIXがコンピューター業界の勢力争いの材料となり、ユーザー無視のせめぎ合いを続けたことから、ついにユーザーの怒りが爆発、オープンシステムの各団体は反省を余儀なくされ、その結果として「ユーザーの声を最大限尊重します」という新しい公約を掲げさせられたのである（日本ではあまり知られていないが、実際にいくつかのユーザーグループが結成され、大規模に抗議運動を展開したらしい）。それで、「ユーザーカウンシル」なるものがスタートしたわけである。では、ユーザーはなぜオープンシステムを求めたのであろうか。

◇アメリカのユーザーは困っていた◇

化学業界の仲間であるデュボン社のヘンリー・モルノー氏に登場してもらおう。モルノー氏はデラウェア州ウィルミントンにあるデュボン中央研究所内の「サイエンティフィック・コンピューティング・ディビジョン」（SCD）のディレクターである。すでに述べたように、デュボンはX/Openの「ユーザーカウンシル」にもOSFの「エンドユーザーカウンシル」にも入っている。

「化学会社がなぜオープンシステムを推進するのですか」

「研究所のシステム化として、1950年代にIBMのマシンを使い始め、60年代になってDECを導入した時にたいへんな苦勞を味わった。二種類のマシ

ンをつなげるのは至難の業だった。メーカー固有のプロプライエタリー（オープン
の反対語）なシステムをつなぐ手間、このようなものこそユーザーにとって削
減すべきものだ」

「コスト的な観点からはどうでしょう」

「情報化投資はこれまでは増える傾向にあり、危惧されていた。現在は、ダウ
ンサイジングとオープンシステムによって横ばいに抑えることができている。い
まの全社的な情報化投資額は10億ドルだが、この方向をさらに進めることで予
算を8億ドルまで削減するのが目標である」

「実際にオープン化団体に自ら参加する理由は何ですか」

「これらのオープン化団体はコンピューターベンダーが主導であり、もっとわれ
われユーザーの声を反映させるために、自らが参加して積極的に活動していく
必要がある」

つまり、デュポンは「困っていた」といえないだろうか。そうなのである。米
国のユーザーはそれまでのプロプライエタリーな世界に困っていたので、オープ
ンに走ったのである。解決策を模索したら、それがオープンシステムになったと
いうわけだ。

元X/Open日本事務所長の田中氏は「日本のユーザーさんはまったく逆な
んですよ」と述べる。「まずオープンシステムという目標があって、それに向か
っていかねばならないと考えている」。米国ユーザーとは発想が逆であり、
それで「日本のユーザーは困っていないのかもしれない」という感想をもったの
だという。

◇みんなで困ろう◇

なぜコンピューターを使うのかを考えてみよう。当然、仕事を行う上で発生す
る問題を解決するために使うのである。困っていることがあるから、コンピュー
ターを使おうと思うのである。これが「ソリューション」という考え方の基本だ
ろう。CCSについても同じだ。いま、CCSに取り組もうとしている国内の企
業や研究者は、もしかしてCCSを「向かうべき目標」と考えていないだろう
か。CCSは「目標」ではなく、問題を解決するための「道具」である。

米国では、一般の研究者の間にまでCCSが浸透し始めている。いわゆる実験
やさんもCCSを使い始めているのである。彼らにも何か困った問題があり、C
CSがその解決に役立っているはずだ。そうでなければ、誰がCCSを使うだろ
う。何も困っていない人が、どうしてCCSを使おうと思うだろう。

日米の大きな違いの一つはここにあるような気がする。もしかしたら、日本の
研究者、化学者は困っていないのではないだろうか。この機会に、研究者一人ひ
とりが「自分は研究で困っているだろうか」「その解決策がCCSの中にないだ
ろうか」と問うてみるのもいいかもしれない。そして、研究を進めていくために
「CCSがなければ困る」という認識でみなが一致した時、日本のCCSは本当
に隆盛の時を迎えるといえそうだ。

でも、自分が困っていることに気付かない人ばかりだったら、――。悲劇
だ…。

ケモメトリックス関連ソフトウェアの利用状況の
アンケートと執筆協力をお願い

化学ソフトウェア学会
刊行物編集委員会

編集幹事：吉村忠与志、藤枝 修子

本学会では、化学ソフトウェアシリーズ1として、「実践・パソコン 分子科学計算」を企画・出版しました。現在、会員には2割引き（送料は出版社負担）で安く販売していますので、ぜひお求めください。分子科学計算におけるソフトウェアを利用するうえでの実践書であります。

ところで、シリーズ2として「実践・パソコン ケモメトリックス」の企画を進める中で、ケモメトリックス関連ソフトウェアの利用状況を調査することとなりました。それは、より実践的に活用されているソフトウェアを発掘することであり、日本での利用度の高いもののソフトウェア実践書を出版するためです。将来の発展が期待されるケモメトリックス関連ソフトウェアについて、その存在と、利用の実態をソフトウェアの国産・海外産を問わず、幅広く紹介したく存じます。同書籍にソフトウェア一覧や入手に関する情報をまとめ、皆様の研究や事業の一助とさせていただけるようにすると同時に、記事の正確を期すべく、ここにアンケート調査をしています。アンケートに回答してくださる方は、アンケート用紙を送付しますので、事務局までご一報下さい。また、執筆にご協力をお願いできる方は、その旨を明記してください。

アンケートそのものについては、厳重に管理して機密保持に努めますので、ご協力のほどよろしくお願い申し上げます。ご多忙とは思いますが、至急ご一報ください。

連絡先

〒916 鯖江市下司町 福井高専内

化学ソフトウェア学会刊行物編集委員会 宛て

tel: 0778-62-1111, fax: 0778-62-1108

CHEMICS研究会

「第2回特別講演会」のお知らせ

この度、CHEMICS研究会では、コンピュータケミストリー分野で世界的にも著名な先生方をお招きし、下記の要領にて講演会を開催致します。

今回は、欧米、国内においてもリーダー的な存在である各企業の、コンピュータケミストリーへの取り組みを中心とした講演会を企画致しました。

CHEMICS研究会会員以外の方々のご参加も歓迎致しますので、ふるってご参加下さい。

日 時：平成5年5月31日(月) 9:30～17:00

会 場：日本化学会 化学会館7Fホール

東京都千代田区神田駿河台1-5

参加費：¥3,000- (CHEMICS研究会会員は無料)

<プログラム>

1) 挨拶

CHEMICS研究会会長 豊橋技術科学大学 佐々木愼一学長

2) 「BASF社およびヨーロッパにおけるコンピュータケミストリー」

BASF社 中央研究所 Hadicke博士

3) 「Dupont社およびアメリカにおけるコンピュータケミストリー」

Dupont社 中央研究所 Dixon博士

4) 「三菱化成(株)におけるコンピュータケミストリー」

三菱化成株式会社 計算科学研究所 岡崎慶二氏

5) 「住友化学工業(株)におけるコンピュータケミストリー」

住友化学工業株式会社 筑波研究所 吉田元二氏

<お申し込み方法>

1)氏名、2)会社名または学校名、3)部署、4)住所、5)TELおよびFAX番号を明記の上、下記のCHEMICS研究会事務局までFAXにてお申し込みください。

<お問い合わせ/お申し込み先>

CHEMICS研究会事務局

長瀬産業(株) 開発室 山口裕樹宛

東京都中央区日本橋小舟町5-1

TEL 03-3665-3861

FAX 03-3665-3890

<定員>

50名 (定員になり次第締め切らせていただきます。)

*懇親会も予定しております。

文献紹介

[コンテンツリスト]

JOURNAL OF COMPUTER-AIDED MOLECULAR DESIGN

□ VOL.6, NO.6, 1992

- Structure-activity relationship between the 3D distribution of the electrophilicity of sugar derivatives and their cytotoxic and antiviral properties
Alessandra Ricca, Jean M.J. Tronchet and Jacques Weber (Geneva, Switzerland) 541
- Comparative molecular field analysis of CCK-A antagonists using field-fit as an alignment technique. A convenient guide to design new CCK-A ligands
Sylvain Rault, Ronan Bureau, Juan Carlos Pilo and Max Robba (Caen, France) 553
- A comparison of progestin and androgen receptor binding using the CoMFA technique
Deborah A. Loughney and Charles F. Schwender (Raritan, NJ, U.S.A.) 569
- Surface comparisons of some odour molecules: Conformational calculations on sandalwood odour V
Gerhard Buchbauer, Susanne Winiwarter and Peter Wolschann (Wien, Austria) 583
- LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads
Hans-Joachim Böhm (Ludwigshafen, Germany) 593
- A fast and efficient method for 2D and 3D molecular shape description
Guy W. Bemis and Irwin D. Kuntz (San Francisco, CA, U.S.A.) 607
- AM1-SM2 and PM3-SM3 parameterized SCF solvation models for free energies in aqueous solution
Christopher J. Cramer and Donald G. Truhlar (Aberdeen Proving Ground, MD and Minneapolis, MN, U.S.A.) 629

JOURNAL OF COMPUTATIONAL CHEMISTRY

□ Feb. 1993, Vol. 14, No. 2

Parallelization Strategies for Molecular Simulation Using the Monte Carlo Algorithm <i>Douglas M. Jones and Julia M. Goodfellow</i>	127
Rapid Empirical Calculation of the First (n or π) Ionization Potential of Organic Molecules <i>W. Hanebeck and J. Gasteiger</i>	138
Search for Nondegenerate Real Vertex Invariants and Derived Topological Indexes <i>Alexandru T. Balaban and Cornel Catana</i>	155
Microscopic and Semimicroscopic Calculations of Electrostatic Energies in Proteins by the POLARIS and ENZYMIK Programs <i>Frederick S. Lee, Zhen Tao Chu, and Arieh Warshel</i>	161
<i>Ab Initio</i> Study of the Ground and Excited States of HCP and Its Isomer HPC <i>Elisheva Goldstein, Suqian Jin, M. Robyn Carrillo, and Robert J. Cave</i> ...	186
Monte Carlo Simulations of the Hydration of Substituted Benzenes with OPLS Potential Functions <i>William L. Jorgensen and Toan B. Nguyen</i>	195
Monte Carlo Simulations of Pure Liquid Substituted Benzenes with OPLS Potential Functions <i>William L. Jorgensen, Ellen R. Laird, Toan B. Nguyen, and Julian Tirado-Rives</i>	206
Use of Effective Core Potentials for <i>Ab Initio</i> Calculations on Molecular Siloxanes and Silicates <i>Clarke W. Earley</i>	216
MORMIN: A Quasi-Newtonian Energy Minimizer Fitting the Nuclear Overhauser Data <i>Joël Pothier, Jacques Gabarro-Arpa, and Marc Le Bret</i>	226
3-D QSAR for Intrinsic Activity of 5-HT _{1A} Receptor Ligands by the Method of Comparative Molecular Field Analysis <i>Atul Agarwal and Ethan Will Taylor</i>	237
Fast Generation of Molecular Surfaces from 3D Data Fields with an Enhanced "Marching Cube" Algorithm <i>W. Heiden, T. Goetze, and J. Brickmann</i>	246
Further Comments on the Lack of Homoaromaticity in Triquinacene <i>Andrew Holder</i>	251

Electrostatic Solvent Effects on the Electronic Structure of Ground and Excited States of Molecules: Applications of a Cavity Model Based upon a Finite Element Method <i>Thomas Fox, Notker Rösch, and Randy J. Zauhar</i>	253
Vibrational Molecular Force Field of Model Compounds with Biologic Interest. II. Harmonic Dynamics of Both Anomers of Glucose in the Crystalline State <i>Manuel Dauchez, Philippe Derreumaux, and Gérard Vergoten</i>	263
Observations Concerning the Treatment of Long-Range Interactions in Molecular Dynamics Simulations <i>K. Tasaki, S. McDonald, and J.W. Brady</i>	278
Is Tetrahedral H_4^{2+} a Minimum? Anomalous Behavior of Popular Basis Sets with the Standard p Exponents on Hydrogen <i>Mikhail N. Glukhovtsev, Paul von Ragué Schleyer, Nicolaas J.R. van Eikema Hommes, Jose Walkimar de M. Carneiro, and Wolfram Koch</i>	285
Conformational and Energetic Effects of Truncating Nonbonded Interactions in an Aqueous Protein Dynamics Simulation <i>Jeanmarie Guenot and Peter A. Kollman</i>	295
AMBERCUBE MD, Parallelization of AMBER's Molecular Dynamics Module for Distributed-Memory Hypercube Computers <i>Stephen E. DeBolt and Peter A. Kollman</i>	312
Stochastic Searches for Lactone and Cycloalkene Conformers <i>Martin Saunders and Hugo Alejandro Jiménez-Vázquez</i>	330
Rapid Approximation to Molecular Surface Area via the Use of Boolean Logic and Look-Up Tables <i>Scott M. Le Grand and Kenneth M. Merz, Jr.</i>	349
Conformational Dependence of the Electrostatic Potential-Derived Charges of Dopamine: Ramifications in Molecular Mechanics Force Field Calculations in the Gas Phase and in Aqueous Solution <i>Joseph J. Urban and George R. Famini</i>	353
Fitting of Nonlinear Regressions by Orthogonalized Power Series <i>Milan Randić</i>	363
Reconsideration of Solvent Effects Calculated by Semiempirical Quantum Chemical Methods <i>Miroslaw Szafran, Mati M. Karelson, Alan R. Katritzky, Jacek Koput, and Michael C. Zerner</i>	371

JOURNAL OF CHEMICAL INFORMATION

AND COMPUTER SCIENCES

VOL.33, Jan. / Feb., 1993

- 1 ACS/CODATA Symposium on Numeric/Factual Materials and Chemical Databases
J. G. Kaufman
- 2 Physiognomy of Numeric/Factual Chemical and Material Property Data
J. Gilbert Kaufman
- 6 Problems in the Computerization of Chemical Information: Capture of Tabular and Graphical Data
J. H. Westbrook
- 18 Recent ASTM Standardization Developments for Chemical Information
Charles E. Gragg
- 22 Socioeconomic Aspects of Materials Data: Serving the User
Anthony J. Barrett
- 27 Composition: A Critical Property for Chemical and Material Databases
Charles W. Moulton
- 31 Operation of an International Data Center: Canadian Scientific Numeric Database Service.
Gordon H. Wood, John R. Rodgers, and S. Roger Gough*
- 36 Numerical Databases on Materials Property Data at CINDAS/Purdue University
C. Y. Ho and H. H. Li*
- 46 Technical Features of the Chemical and Materials Property Data Network Services on STN International
*V. J. Dragò and J. G. Kaufman**
- 52 SpecInfo: An Integrated Spectroscopic Information System
Andreas Barth
- 59 Personal Computer Programs To Generate Reference Diffraction Patterns
J. T. Staley
-
- 70 Crystallographic Studies of Inter- and Intramolecular Interactions Reflected in Aromatic Character of π -Electron Systems
Tadeusz Marek Krygowski
- 79 New Method for Rapid Characterization of Molecular Shapes: Applications in Drug Design
R. Nilakantan, N. Bauman, and R. Venkataraghavan*
- 86 Computers in the New Drug Application Process
*Joel F. Studebaker**
- 95 A Nonunique Path Connectivity Matrix
*Donald J. Polton**
- 102 Rapid Conversion of Molecular Graphs to Three-Dimensional Representation Using the MOLGEO Program
Ekaterina V. Gordeeva, Alan R. Katritzky, Vladimir V. Shcherbukhin, and Nikolai S. Zefirov*

- 112 Rapid Evaluation of Shape Similarity Using Gaussian Functions
*A. C. Good and W. G. Richards**
- 117 Notes on Constant-Isomer Series
Jerry Ray Dias
- 128 Automatic Assignment of Molecule Keywords
Martin J. Schweiger
- 131 FATE, the Environmental Fate Constants Information Database
Heinz P. Kollig, Brenda E. Kitchens, and Karen J. Hamrick*
- 135 Displaying the Structure of Molecules by Multidimensional Plots of Their Torsion Angles
Brian T. Luke
- 143 Design of Molecules from Quantitative Structure-Activity Relationship Models. 1. Information Transfer between Path and Vertex Degree Counts
Lemont B. Kier, Lowell H. Hall, and Jack W. Frazer*
- 148 Design of Molecules from Quantitative Structure-Activity Relationship Models. 2. Derivation and Proof of Information Transfer Relating Equations
Lowell H. Hall, Lemont B. Kier, and Jack W. Frazer*
- 153 A Novel Definition of the Wiener Index for Trees
*Bojan Mohar, Darko Babić, and Nenad Trinajstić**
- 155 CONCORD and CAMBRIDGE: Comparison of Computer-Generated Chemical Structures with X-ray Crystallographic Data
Marissa A. Hendrickson, Marc C. Nicklaus, George W. A. Milne, and D. Zaharevitz*
- 164 Image of Analytical Chemistry As Reflected in the Analytical Abstracts Database: Journal Coverage, Concentration, and Dispersion of the Analytical Literature
T. Braun, W. Glänzel, Hajnalka Maczelka, and S. Zsindely*
- 174 Rapid Construction of Data Tables for Quantitative Structure-Activity Relationship Studies
John S. Delaney, Anne Mullaley, Graham W. Mullier, Graham J. Sexton, Robin Taylor, and Russell C. Viner*

CICSJ Bulletin

Published Bimonthly by Division of
Chemical Information and Computer Sciences
The Chemical Society of Japan

ケモトリックス 特集号

日本化学会
情報化学部会

Volume 11, Number 3
May 1993

目 次

特集：ケモトリックスワークショップから

ワークショップ報告	世 話 人	2
測定値の不確かさの Fuzzy 回帰法による取扱いと定量分析への応用	北森 武彦	3
たばことケモトリックス	榊 武彦	5
ニューラルネットワーク事始め	清水 昇市	9
FT-NIR によるパターン認識手法と製造現場におけるその利用	桜井 香織	14
FT-IR へのケモトリックスの応用	新見 栄	18
C-13 NMR 化学シフトデータの PLS 解析	石山 純一	23
Novel Kalman Filtering Techniques in Chemometrics	Zhiliang LI	27

部 会 行 事

第16回情報化学討論会 案内	
計算化学・情報化学サマーセミナー 案内	
情報化学部会ワークショップNo8:	
第一線若手研究者によるコンピュータケミストリー活用事例 案内	

部 会 記 事

平成5年度情報化学部会の課題	
情報化学部会収支報告書	

年 会 記 事

日本化学会第65春季年会(東京)記事	
化学物質と特許の現状にまつわる諸問題/植田 賢三	

関 連 記 事

第21回構造活性相関シンポジウム	
------------------	--

記 事

機械翻訳入門	石塚 英弘	45
研究開発におけるコンピュータ化学の活用	石田 嘉明	54

文 献 紹 介

コンテンツ・リスト:	
Journal of Chemical Education (Vol.70, No.1~2)	
Journal of Chemical Information and Computer Sciences (Vol.33, March/April)	

編 集 後 記

特集：ケモメトリックスワークショップから

日本化学会情報化学部会 第4回ケモメトリックスワークショップ 報告

過去においては、第1回をお茶の水女子大学で開催し、その後1年に1回ずつではあるが、豊橋技科大学（世話人：宮下芳勝氏）、大阪大学（世話人：高木達也氏・山川眞透氏）で行われ、今回のワークショップは第4回目である。

平成5年2月8日（月）に（社）日本化学会の化学会館で10時半から18時まで開催された。CICSJ Bull. の読者に対して今さらケモメトリックスの説明は不要であるが、このワークショップでは毎回参加者が増加しているものの、ケモメトリックスは専門家の手を離れて、実用段階に入ったという程には、残念ながらまだ普及していないのが現状である。しかし、理論的な基礎研究とともに、いろいろな方面で多彩な応用を展開するには、使いやすく、有能なソフトの普及が不可欠であろう。

今回のワークショップでは、ユーザーサイドから、こんなソフトが欲しいとか、こんなことがしたいという希望を出し、ソフトメーカーサイドには、その希望、需要に耳を傾けてもらって、今後のソフト開発の参考にしてもらい、「現状ではこんなソフトで、この程度までは実現できます」という情報提供をしてもらった。さらに、測定機器メーカーもユーザーの希望をかなえるには、ソフトメーカーにどんな測定データをどのように提供すればよいかといった3サイドの各立場での意見を出し合うための討論の場を作った。

特別講演を含めて8件の問題提起があった。プログラムは CICSJ Bull., 10, (No. 6, Nov.), 37 (1992). を参照されたい。今回のワークショップでは、ユーザーサイド以外にソフトウェアメーカーと測定機器メーカーのご協力により、まず30分の講演形式による趣旨・内容の説明と、機材持込みによるデモンストレーションを行った。エスピーエスエス(株)、ニチメンデータシステムス(株)、ブランルーベ(株)、(株)パーキンエルマージャパンの4社のご協力をえた。もう一つの今回の配慮は、ややもすると講演者からフロアの参加者へという一方の情報伝達に終わりがちなワークショップを改善して、参加者全員による相互の情報交換の場を作りたいと考えた。さらに時間の節約を考慮して、デモンストレーションとミニ懇親会をお茶の時間に一箇所ですべて同時進行して約1時間半にわたって行った。スペースのわりには参加者が多かったが、参加者が分散したので、デモンストレーションも比較的ゆったりと見ることができたと思う。

参加者は87名であったが、1社から8名参加された企業もあった。参加業種は、食品、香料・化粧品、製粉、石油、医薬品、コンピュータメーカー、電気機器メーカー、ソフトハウス、化学系、販売系、通産省や農水省の研究所や試験場、大学関係などの極めて多彩な業種関係者が集った。参加者の業種分野が着実に1回ごとに広がっていることを、今回も実感しながら、散会した。

今回のワークショップの話題を特集としてまとめていただけることになった。今後、ケモメトリックスがもっと広く応用されることを願って、問題提起をされた講師の方々に重ねてお手を煩わせた。その他でもいろいろご協力下さった各位に感謝の意を表したい。

(キッコーマン(株) 相島鐵郎、お茶の水女大理 藤枝修子)

測定値の不確かさのFuzzy回帰法による 取り扱いと定量分析への応用

東京大学工学部

下坂琢哉、北森武彦、澤田嗣郎

【序論】 測定法の超高感度化、必要とする情報の多様化に伴い、測定値中の誤差が推定値の結果に多大な影響を与えることが多くなってきている。例えば、生体物質を超高感度で計測できる免疫分析法は、免疫反応の不安定さにより誤差を生じその定量的な取扱いが不可欠であるが、検量線が曲線になり従来法では困難である。また、レーザー蛍光法は、観測角度を変えて測定し逆積分することで、非破壊深さ方向濃度分布分析の可能性が示されているが[1]、誤差が逆積分変換法の結果に大きな影響を与える。しかし、個々の誤差を考慮し逆積分変換する方法はできていない。一般的に、誤差は統計的手法により処理されるが、複雑な系では煩雑になりやすく、個々の誤差を考慮にいたる解析法は少なく、誤差をより柔軟に取扱う方法が必要である。そこで、測定値に誤差があることを「あいまい」であると考え、ファジィ理論を用いて、柔軟かつ定量的に取扱うことを試みた[2]。まず、簡単な場合として、ファジィ線形回帰法[3]を直線の検量線に応用し、実測値を用いて妥当性を検討した。次に、多様なデータに対応できるようにニューラルネットワークの関数近似能力を利用して曲線に拡張し[4]、免疫分析法・レーザー蛍光法に応用し、滑らかでないデータには区分的に近似するアルゴリズムを導入した。

【理論】 図1に免疫分析における測定値を示す。信頼性は、各濃度における信号強度の平均値 a で最も確からしく、 a から離れるほど信頼性は低くなる。そこでこの信頼性を数値で表し、 a での信頼性に1を割当て、 a から離れるほど小さな数値を割当てると、例えば、図1に示した二等辺三角型になる。この時、信頼性は測定値の関数になり、ファジィ理論では、これをメンバシップ関数と呼び、この様に表した測定値をファジィ数 (FN) と呼ぶ。FNで表された測定値に直線の検量線を当てはめるとき、ファジィ線形回帰法 (FLR) を用いる。FLRとは、図1の様に推定区間がすべてのFNで表した測定値の底辺を含み、かつ区間の長さの和を最小とする方法である。この問題は線形計画問題に帰着され、これを解くと従来の線形回帰法と異なり、図1に示すような領域が導出された。この領域を検量域と呼ぶことにする。

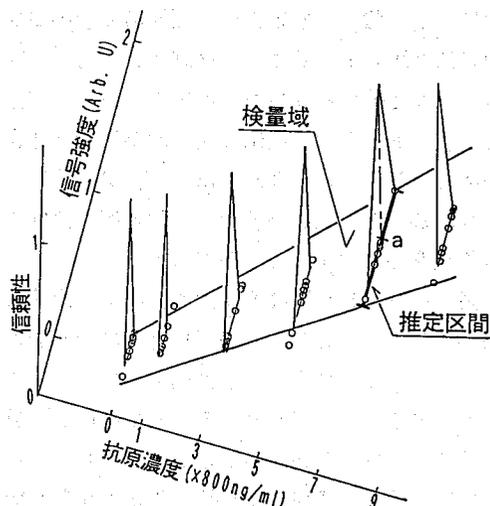


図1 測定値のファジィ数による表現と検量域

さらに、ニューラルネットワーク (NN) を利用し、ファジィ回帰法を曲線に拡張した。求める検量域は、全てのデータが含まれる曲線で挟まれた領域であり、この曲線の近似に NN を利用した。NN は信号強度を表す濃度についての関数で、構造はシグモイド関数の中にシグモイド関数の線形結合が入った形である。また、係数は、ファジィ回帰法に適合するように NN の持つ学習機能により自動的に決定させた。

【結果と考察】 図1に、免疫分析法の直線部分の測定データから求めた検量域を示す。測定データ

は濃度が高いほどばらつきが大きい、検量域も同様に高濃度側で幅が広い。このことから、検量域は濃度・信号強度・推定値の信頼性の3者の関係を定量的に表していると言える[5]。従来の線形回帰法では、このような誤差の濃度依存性を表すことはできない。また、生物学的親和性を利用する免疫分析法では、化学量論的な誤差要因は反応速度の不安定さやpH、温度等により決まり、濃度と比例関係があると言われている。図1の検量域は、この様な誤差要因を定量的に反映したものと考えられる[6]。

図2に、NNを用いて曲線に拡張したファジィ回帰法を、より広範囲の濃度で測定した免疫分析の実測値に応用した結果を示す。濃度と信号強度の関係はシグモイド型曲線の関係にあり、測定値の誤差は各濃度で異なる。検量域は免疫分析に特有なシグモイド型曲線を良く近似できている。各濃度での測定値のばらつきは変化に対応して検量域の幅が変化していることから、直線の場合と同様に濃度・信号強度・推定値の信頼性の3者の関係を定量的に表していることが検証された[7]。この様に、測定値に曲線をあてはめその誤差を定量的に表現することは、従来の統計的手法では非常に困難である。また、今回導出した検量域はNNにより関数の形で表されるので、例えば、曲線でも信号強度から容易にある信頼性における濃度および誤差範囲を推定することができる。

図3に、蛍光物質濃度が深さ方向に一次的に減衰するLB膜をレーザー蛍光法で測定したときの理論値をプロットした。このデータには折れ曲がり点があり、図3に示すように従来のNNによるファジィ回帰法で近似することは困難であった。そこで、折れ曲がり点を接点として、区分的に近似するように従来法を拡張した。その結果を図3の実線で示す。折れ曲がり点付近も精度良く近似できた。検量域は解析的な関数の形として表され、高度な他のデータ処理法と結び付けやすく、誤差を定量的に取扱うデータ処理法に発展することが期待できる。特に、本法を誤差に影響されやすい逆積分変換法と組み合わせることで非破壊3次元分析が可能に成ることが期待できる。

【参考文献】

- [1] X. Z. Wu, T. Kitamori, N. Teramae, and T. Sawada, Bull. Chem. Soc. Jpn. 64, 1757(1991)
- [2] 下坂琢哉, 修士論文(東京大学工学部工業化学科), (1993)
- [3] H. Tanaka, S. Uejima, and K. Asai, IEEE Trans. Systems Man Cybernet. SMC-12(6), 903(1982)
- [4] 石淵久生, 田中英夫, 計測自動制御学会論文集, 28(1), 116(1992)
- [5] T. Shimosaka, T. Kitamori, A. Harata, and T. Sawada, Anal. Sci., 7, Supple., 1389(1991)
- [6] 下坂琢哉ら, 第53回分析化学討論会講演要旨集, 105(1992)
- [7] 下坂琢哉ら, 日本化学会第64秋季年会, 868(1992)

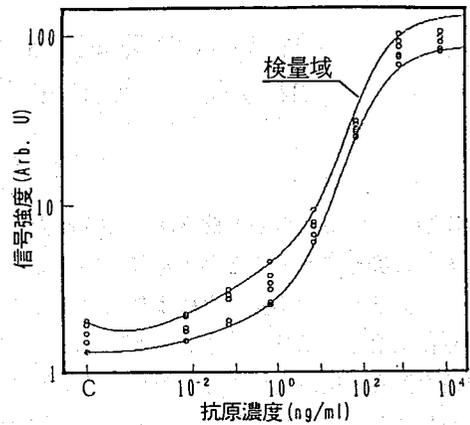


図2 免疫分析法による微量抗原の検量域

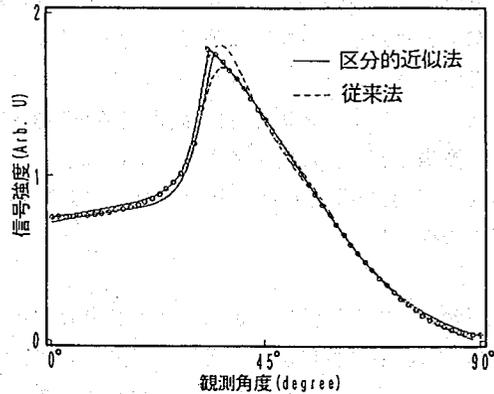


図3 レーザー蛍光法による理論データの検量域

たばことケモメトリックス

日本たばこ産業(株) たばこ中央研究所 榎 武志

1. はじめに

たばこは香りや味を楽しみニコチンを摂取する嗜好品で、3000万人の愛煙家に支持された商品である。たばこにとって、その香りや味(香嗅味)は重要な評価基準である。たばこの香嗅味は主に官能検査により判断されている。そこで、この官能検査によるたばこの評価に補助的な情報を提供することを目的に、たばこの内容成分や煙成分を分析し、その分析値から、たばこの香嗅味特性を推定したり、葉たばこ品種や製品たばこの銘柄を分類するケモメトリックス研究をおこなっている。図1に、たばこ研究でのケモメトリックス利用の考え方を図示した。

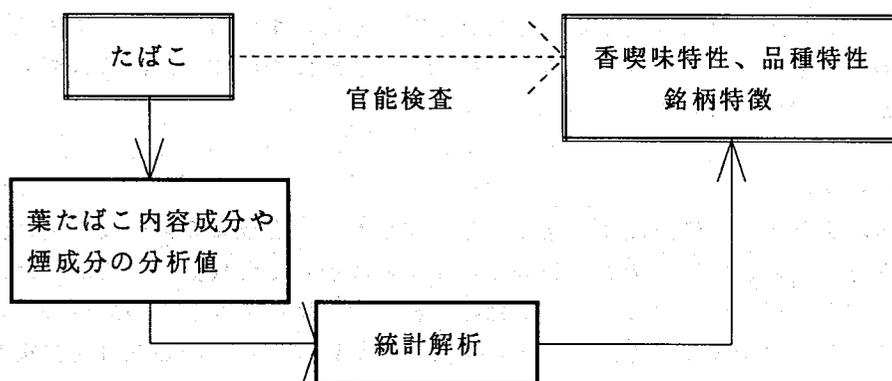


図1 たばこ研究でのケモメトリックス

葉たばこの内容成分や煙成分は多種多様な成分で構成されている。たとえば、たばこの煙には同定されているだけでも約4000種以上の化合物が含まれている。これらの化合物が複雑に組合わさって、たばこの香嗅味が形成される。また、これらの化合物には1個でたばこの香嗅味を代表するキー化合物は存在しない。したがって、煙成分の分析値から香嗅味を評価・分類するには、各ピークの相互関係を中心に解析する必要があった。このため、統計解析手段として多変量解析の考え方を利用して仕事を始めた。

本文では、これまでのケモメトリックス研究における経験のうち、煙成分の分析値を使用して葉たばこ品種や製品たばこの香嗅味特徴を分類する「パターン認識」研究を紹介し、ケモメトリックスに対する私見を述べてみたい。

2. ケモメトリックスシステム

たばこの煙成分分析値を解析するケモメトリックス環境について紹介する。現在、主にガスクロマトグラフを利用して分析した煙成分の分析値を解析対象としている。例えば、煙成分のうち中・高沸点化合物であるセミボラタイル成分を分析すると、ガスクロマトグラム上には300本以上のピークが認められる。それらのピーク面積値をインテグレータで測定

し、そのデータをフロッピーディスクに記録する。さらに、フロッピーディスクからデータを所内のネットワーク経由で大型コンピュータ（DEC社製VAX）に転送する。

転送されたデータの蓄積管理には日本BBN社製のRS/1システムを利用している。また、多変量解析にはBMDPシステムを使用している。RS/1システムにはデータの蓄積管理の機能のほかに、統計解析機能とグラフ作成機能がある。この機能は、説明変数のデータや解析結果をグラフとして表示する際に使用する。数字そのものの並び方からそのデータ構造を理解することはほとんど不可能に近い。したがって、機会あるごとに、グラフを作成することは解析を成功させる鍵ともいえる。グラフ作成機能はケモメトリックス研究にとって重要な機能である。さらに、RS/1システムにはデータをBMDPシステムに転送する機能もあるため、ケモメトリックス研究を実施するにRS/1システムは有効な道具である。

3. ケモメトリックス研究の流れ

ケモメトリックス研究には4個の重要なステップがある。すなわち、データ測定、データ評価、モデル検索、モデル評価である。これらの4段階について、たばこ研究で実施している内容と共に紹介したい。

3. 1. データ測定

たばこの煙は葉たばこの燃焼により発生する。すなわち、燃焼時の熱により葉たばこ成分は分解、揮発、蒸留、反応、吸着、脱着の複雑な過程を経て煙成分になる。したがって、発熱量に関係する燃焼条件により同一の葉たばこ組成のたばこでも、その煙成分に差がある。このため、その香嗅味にも差が生ずる。よって、複数なたばこの煙成分を分析して比較する場合には、それぞれのたばこの燃焼条件を一定にするように喫煙パターンを共通化している。たとえば、各社の製品間で比較するニコチンやタール値の場合には、共通の喫煙パターンで測定することが義務づけられている。これが標準喫煙条件で、1分間に1回、2秒間で35mlを吸煙する条件である。われわれの煙成分の分析においても、この標準喫煙条件に従っている。しかし、喫煙パターンを揃えても、たばこの燃焼条件はたばこの物性値に大きく影響され、分析値の変動を生ずる。

このように、変動が大きいデータを解析対象とするときには、できるだけ制御可能な変動を抑える必要がある。たばこの煙成分分析値の変動を抑えるために、第一に分析装置の管理をおこなっている。分析装置の再現性をできるだけ高めるために、装置の整備は必要不可欠である。最近では分析業務が自動化され、装置の整備をメーカー任せになっているケースを見る。これでは、解析結果に問題が生じた時に、その原因がデータそのものによるのか、解析方法なのかかわからない。少なくとも、データ測定の問題点は自分で解決してから、統計解析を再度実施すべきである。コンピュータの発達により、容易に多変量解析が実施できる時代だからこそ、解析対象のデータにはこだわりたい。

3. 2. データ評価

分析が終了したデータに対して、直ちに多変量解析プログラムで解析することは危険である。つまり、単変量としてのデータを十分に評価してから多変量解析を実施すべきである。解析を始める前のデータ評価は、目的の解析結果をえるための近道でもある。具体的

には、探索的データ解析の手法を使用して評価をおこなう。

また、データの形態としての、説明変数の数と実験繰り返しの数の関係にも注意する。通常、繰り返し数は使用する説明変数の数より多くなくては多変量解析は実施できない。理想的には繰り返し数は説明変数の5倍以上あることが望ましい。しかし、現実には実験の性質や試料数の関係で、データ群の構造がそうにならないことが多い。そこで、説明変数の数と繰り返し数の関係により、そのデータ群の解析結果の意味が異なってくる。

すなわち、繰り返し数が説明変数の3倍以下であるデータ群についての解析結果は、データ構造を表記するだけの内容と理解している。一方、解析結果を未知試料の分類推定等に利用するため、より信頼性の高い解析結果を必要とするときには、繰り返し回数は説明変数の3倍以上が必要である。そこで、少ない繰り返し実験数の多量な説明変数から、実際の解析に使用する説明変数を選択する作業をおこなう。この変数選択の解析には、その後の解析作業と同じだけの時間をかけている。たばこ研究ではこの変数選択の条件として以下の3点を設定している。すなわち、モデルに組み込む際に重要となる説明変数の安定性を示す指標としての再現性（同一試料内）と説明変数間相互の相関係数。そして、繰り返し分析試料内で観測されるという普遍性である。ある特定の試料でのみ認められる説明変数は排除するということが普遍性を保証する鍵である。

3. 3 モデル検索

たばこ研究でのケモメトリックスで使用している多変量解析手段は、目的により2つに分類される。すなわち、葉たばこの品種やたばこの銘柄を分類する「パターン認識」^{1, 2)}と、香喫味評価の官能検査値を推定するモデルを検索する「重回帰分析」³⁾である。本文では、たばこ研究における「パターン認識」の解析手順を述べる。

説明変数を評価・選択したデータ集合体について、最初にクラスター分析を実施する。この解析では全説明変数に含まれている情報をもとに、分類などの外的基準を加えずに試料を分類する。解析は各試料の類似度の計算をおこない、説明変数のパターンで解析対象の試料群の集合体が形成される。その集合体が目的とする分類に一致した場合、解析対象のデータ群にその情報が含まれていることが明らかになる。しかし、たまにはその集合体が全く予期しない分類パターンを示すときもある。したがって、このクラスター分析の結果を見るときがケモメトリックス研究をおこなっているときの、最も愉しくもあり、緊張する瞬間でもある。

次に、説明変数に含まれる情報を抽出するために、主成分分析をおこなう。主成分分析もクラスター分析と同様に、外的基準を与えずに解析する。解析結果は、固有値1以上の主成分の因子負荷量と各試料の主成分得点である。各試料の主成分得点をクラスター分析で認められた分類にしたがって作図する。その結果、クラスター分析の場合と同様な試料群の集合体が認められた場合、先の全説明変数により示された分類を、より少ない説明変数で再現できたことになる。そして、その時に関係した主成分の因子負荷量から、目的の分類に関係する説明変数が明らかになる。

さらに、積極的に分類をおこなうために必要なモデルとしての判別式を、段階的判別分析により検索する。判別式に最低必要な説明変数を選出するには、交差検証法の判別結果を参考にする。そして、なるべく最低の説明変数を組み合わせて最大の効果を示す判別式

を求める。また、実際の利用に際して分析の容易さも考慮にいれる場合もある。このようにして選出した説明変数の関連性を、さらに正準判別分析を使用して解析する。その結果、各試料の正準量を図示して、目的の試料分類がえられ、前出の主成分得点の分布と同一になった時点で、目的とする分類を実現するためのモデルについての検索が完了する。

3. 4 モデルの評価

検索により得られたモデルの安定性、頑健性を次に評価する。評価方法としては内部法と外部法がある。内部法は、解析に使用した全試料のうち90%を使用して、目的の分類を最高の正解率で達成する判別式を計算する。そして、解析に使用しなかった残りの10%試料を得られた判別式に当てはめ、各試料の正解率を計算する。これをランダムに10回繰り返す。この時の総合判別正解率と10個の判別式での各説明変数の係数の再現性を比較する。もし、データの中に全体に大きく影響するはずれ値が含まれている場合には、判別正解率や係数に大きな変化が生ずる。

一方、外部法は新たに実験をおこなって得たデータを使用して、判別正解率を調べる方法である。理想的には、外部法で評価するのが一番である。しかし、試料数が限られていたり、再度同じ条件での試験に時間がかかる場合が多い。したがって、内部法を使用してモデルの評価をすることが多い。

これらの評価で全て問題が無いと確認されて初めてパターン認識の解析が終了となる。実際の解析作業では、1回の計算で終ることはほとんどない。異常な結果が生じた場合には、全段の検討へ遡って検討する。

4. おわりに

ケモメトリックスの手法を利用してたばこ煙成分の研究をおこなってきて一番感じていることは、その解析手段としての有効性である。煙成分は複雑で不安定な情報である。しかし、ケモメトリックスにより煙成分に含まれる情報に一步近づくことができたと思う。

しかし、注意しなくてはいけないことがある。つまり、ケモメトリックスで抽出された説明変数と目的とする分類関係の直接的な因果関係は、実際にその説明変数として使用した煙成分が原因で、その結果分類が生じるという証明をするまでは、即断できない。つまり、因果関係がはっきりしないうちは、ケモメトリックスの結果は「鏡像」にすぎない。コンピュータの発達でケモメトリックスはだれでも使用できる手段となった。しかし、ケモメトリックス手法は魔法ではない。便利な道具ほど、使い方を誤ると大きな被害が生じる。「ごみのデータから、意味ある結果は得られない」という鉄則を肝に銘じて今後も、ケモメトリックスを手段として利用していきたい。

参考文献

- 1) Y. Sone, T. Sakaki, H. Takahashi, M. Fujiwara, Y. Ishizu, CORESTA p88 (1992).
- 2) T. Sakaki and P. C. Jurs, Pittsburg Conference (1992).
- 3) 完成計測先端技術集成、栗山洋四、江原勝夫、山根木正人編、p262、サイエンスフォーラム(1991).

<<ニューラルネットワーク事始め>>

ニチメンデータシステム株式会社
AI営業第2課 清水 昇市

0. はじめに

ニューラルネットワークブームが沸き起こり、少なくとも「ニューラルネット」といえば「ええ、聞いた事がありますよ。」とか「洗濯機やエアコンで使われているヤツでしょう」という人が大多数になってきた今日このごろですが、「一般化デルタルールについて説明して下さい」とか「バックプロパゲーションネットワークにどのようなデータを与えれば良いでしょうか」などの質問をすると「. . .」となってしまう人がまだ多いのではないのでしょうか。要は言葉だけの人気が行先して、実体が伴っていないというのが実状だと思います（某野球選手のようなのですが. . .）。そこで今回はそんな人のために「これなら自分の仕事に役に立つかもしれない」と思っただけのような糸口を提供したいと思えます。ただ話をわかりやすくするためにやや厳密性が欠けている点にご容赦願います。（なお、上の2つの質問に答えられる人は以下の文章を読むのは時間の無駄です。予めご注意ください。）

1. なぜニューラルネットワークなのか？

結論からいうと、現状の自分の仕事に100%満足している人なら新しい技術を評価しようとする事もないでしょう。しかし実際にはほとんどの人が、今やっている仕事を「もっと早く」とか「もっと正確に」などもっと効率的にやりたいと思っています。だからこそ「ファジー」「GA（遺伝アルゴリズム）」「統計解析」. . .等今まで利用した事がないものでも「これは役に立ちそうぞ」と思えば本屋さんに行って本を買ってきて勉強するのでしょう。そういう意欲的な人は是非「ニューラルネットワーク」を勉強してみてください。きっとこれは便利だなと感じていただけると思います。それではニューラルネットワークを使うとどんなメリットがあるかを列挙してみましょう。

- 1) アルゴリズムを考える必要がない（自己学習機能がある）。
- 2) 汎化能力がある。
- 3) 入出力が非線形の複雑なモデルも作成できる

どうです？何となく使ってみたいという気持ちになりませんか？

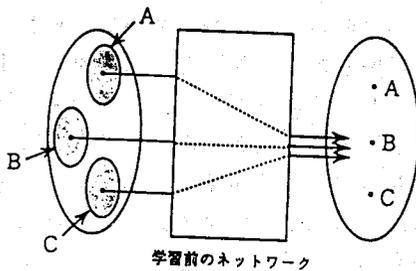
2. ニューラルネットワークには何が必要か？

早速その気になってシミュレーションソフトを購入し自分のパソコンにインストールしたもののニューラルネットワークはいつまで経っても問題解決をしてくれそうにありません。何故でしょう？何か重要なものが足りないようです。何が足りないのかお分かりですか？そうです。「データ」が必要なのです。「なーんだ。当たり前じゃないか。そんな事言わ

れなくても分かっているよ！」と思っている貴方。今お持ちのデータをどの様にニューラルネットワークに与えればいかご存じですか？ここがポイントです。適切なデータセットを与えるとニューラルネットワークはあっという間に問題を解決してくれますが、そうでない場合には、いつまで経っても望ましい答を返してはくれません。前項でニューラルネットワークのメリットを述べましたがこれらも適切なデータセットが与えられるのが前提条件になっています。

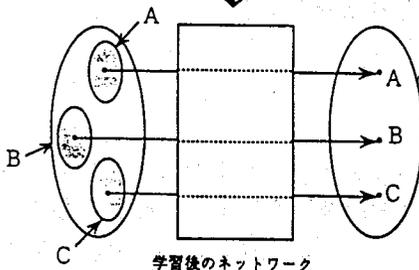
3. ニューラルネットワークは「学習する」？

何が必要なのかが分かった所で、まずニューラルネットワークがどのような仕組みで動作するのかおさらいしておきましょう。前項で述べたように適切なデータが与えられればニューラルネットワークはそのデータの中から入出力関係を自動的に見つけてくれます。その動作概要を表したのが図1と2です。初期状態のニューラルネットワークは入出力関係についてまったく知らない状態なのでどんな入力を与えられても分からないと答えます（図1のa、図2のa）。そのニューラルネットワークに入力に対し望ましい出力（教師信号）を与えてやるとニューラルネットワークは内部にある「重み」と呼ばれるパラメータを徐々に調節して出力値を教師信号の値に近づけていきます（図1のb、図2のb）。このパラメータを調節する操作の事を「学習」と呼んでいます。調節の仕方は「学習則」（代表例：一般化デルタルール）と呼ばれるルールによって規定されており、学習則がネットワークに与えられる「データ」に応じて「重み」パラメータを自動的に調節（自己学習機能）していきます。この様にニューラルネットワークを使えば「入出力関係がどのようなアルゴリズムで決まっているのか」を考える必要がありませんので、複雑な入出力関係も難しい計算や解析をする事無く事例データから導く事ができます。



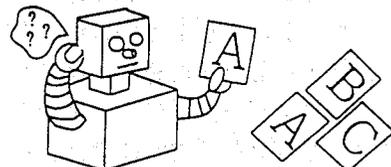
学習前のネットワーク

図1-a



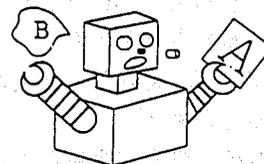
学習後のネットワーク

図1-b



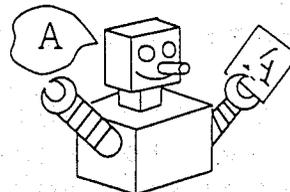
学習していないニューロン君は何もわからない

図2-a



学習途中のニューロン君は間違った答えをだすが徐々に賢くなっていく

図2-b



学習を終了したニューロン君は多少汚れていても正しく答えをだす

図2-c

さらに良い事に学習がうまく終了したニューラルネットワークは学習したデータだけでなく未学習のデータに対しても望ましい値を出力してくれます(図2のc)。この様に学習していないデータに対してもそれなりの答を出してくれる能力のことを「汎化能力」と呼んでいます。どうです。ニューラルネットワークはなかなかやるとは思いませんか？

4. データをどの様に準備するか？

今までの話だけですと、ニューラルネットワークはまるで魔法のごとくデータを与えておけば勝手に入出力関係を見つけてくれるように錯覚してしまいそうですが、そこまで現実には甘くありません。データをせっせとニューラルネットワークに与えてもいっこうに期待する入出力関係を見つけだしてくれない事があります。この原因の1つにデータの質の問題があります。例えば図2で”A”だと教えるべきところを”B”だと教えているようなデータが混ざっていたらどうしてもそれに従って間違った関係を学習してしまいます。また、大きさがまったく異なる文字データを何も正規化しないで学習させてもやはり期待するような結果は出てこないでしょう。また学習するためのデータセットが2、3ヶしかなくて未学習のデータセット5、6ヶに正しい結果を期待するのも無理な相談です。つまりニューラルネットワークにしっかりと学習させようとするなら、それなりの学習データを用意してやらなければならないのです。これは人が勉強するのに、いい教材を使わなければいつまで経っても学習効果が出て来ないのと同じです。それではいい教材かどうかを見分けるポイントは何でしょうか？そのためのチェックポイントをお教え致します。

- ・データの件数が豊富にある
データは多ければ多いほどより正確にモデルを作成する事ができます。
- ・データに一貫性がある
所々のデータが欠落している、測定条件が異なるデータが混じっている、同じ入力データに対して異なる教師信号があるといった一貫性のないデータでは学習は上手くいきません。
- ・データに偏りが無い
”A”、”B”、”C”の3文字を学習させるのに、それぞれの学習データ件数が100、20、3では”A”は正しく学習できてもその他は正しく学習できません。
- ・適切な前処理が行われている
サイズや位置の正規化を始めとし、シンボリックデータの数値コード化、偏った分布を修正する各種変換などニューラルネットワークが学習しやすい形にデータを前処理しておく必要があります。
- ・適切な特徴量が抽出されている
問題によって何を特徴量にすればいいかが異なります。専門家の人の意見を聞いたり、統計解析等の手法を使って重要な因子を取り出す必要があります。

5. データは準備したけど...

さて注意事項を守ってデータが準備できたので早速学習をさせてみようと思ったら、また

新たな疑問点がわいてきました。「どの様なネットワークパラダイムを使えばよいのだろうか？」という疑問です。これに対しては少々乱暴ですが「まずバックプロパゲーションを使って下さい」が答です。何故「バックプロパゲーション」なのか？答は「今までに一番多く使われているから」です。バックプロパゲーションネットワークは分類、予測、モデル作成、フィルタリング、制御など幅広い分野に適用する事ができ、かつ、かなり複雑な非線形な入出力関係も表現する事が可能です。今までに発表されている事例をみても60%から70%ぐらいがこのタイプを利用しています。つまり「適用範囲が広く、かつ、強力だから」これが「一番多く使われている」理由でしょう。しかしバックプロパゲーションも万能ではありません。「学習回数が多くかかる」「ローカルミニマムに陥る」等の問題のほか対象とする課題によっては「誤差が収束しにくい」といった問題も発生します。これらの問題に対してはネットワークパラメータをチューニングしたりネットワークパラダイムを変更することにより対応することになります。するとさらに「パラメータはどの様にチューニングするか？」「ネットワークパラダイムはどの様にして選ぶか？」といった疑問点が出てきます。これに対しては大まかなガイドラインはありますがこれといった決定打はなく、ある程度試行錯誤しながら最適値を見つけていくことになります。ここ2、3年の間にバックプロパゲーションの持つ弱点を補う能力をもった新しいネットワークパラダイムも開発されてきており、開発者の負担は徐々に軽くなってきています。

6. 弘法筆を選ぶ？

以上ニューラルネットワークを学習する際の留意点について述べてきましたがもう1つ重要なポイントがあります。それはネットワークを学習させる環境、つまり「シミュレーションツール」です。通常ニューラルネットワークシステムの開発は試行錯誤しながら行うため、ネットワークの作成、編集、学習、評価といった一連の作業がスムーズに行える作業環境が必要です。その点弊社で販売しており既に数多くの方にご利用頂いております米国ニューラルウェア社製の「NeuralWorks」シリーズのニューラルネットワークシミュレーションソフト「Professional II/PLUS」はユーザインターフェイスに優れ、初めての方でもすぐに使えるように簡易メニューを取り揃えているほか、数多くのネットワークパラダイムを用意してあるので問題に応じて最適な選択をすることができます。ネットワークの評価に欠かせないモニタ機能もワンタッチで作成できるなどシミュレータとして必要な機能をほとんどカバーしており、試行錯誤の作業をあまり苦勞しないで行う事ができます。更に「Designer Pack」と呼ばれるソフトを使えばシミュレーションした結果からANSI標準のC言語で関数プログラムを作成してくれるので、作成したニューラルネットワークを様々な動作環境で自由に利用することも可能です。また「DataSculptor」と呼ばれるソフトは学習およびテストデータを作成するのに威力を発揮します。「DataSculptor」を使えば複数のデータソースからのデータの収集、統計量の計算、グラフ表示、各種データ変換、レコードの並び替えおよび抽出などの作業を1つのパッケージソフトで実行することができます。しかも各種機能はオブジェクトと呼ばれるアイコンになっており、アイコンどうしを結合するだけで一連の処理手順を定義することができるので、プログラミングが苦手な方でも上級プログラマが作成したプログラムと同等な処理が実現可能です。

7. おわりに

よいニューラルネットワークを作成するには「良いデータ」と「良いツール」を揃える事が極めて重要であることが以上の説明で判っていただけたと思います。逆に言えば、「良いデータ」と「良いツール」さえ用意すればニューラルネットワークを作成するのは非常に簡単です。今、あなたがお持ちのデータの中に宝の山が埋もれているかも知れません。その宝の山を掘当てるための道具を弊社では提供していますので是非いろいろ試してみてください。これからニューラルネットワークに取り組んでみたいという方には、弊社主催の入門セミナー、体験セミナーがお薦めです。またご希望に応じてツールのデモンストレーションや説明等も行いますので興味をお持ちの方は以下の問い合わせ先までご連絡下さい。ツールのカタログ、説明資料、セミナーの案内書等を送付いたします。

問い合わせ及び資料請求先：ニチメンデータシステム株式会社 AI営業第2課
「NeuralWorks」サポート係 TEL：03-3864-7692

注：今回の説明は教師信号を必要とするバックプロパゲーションタイプのネットワークを想定して行いましたが、ニューラルネットワークには動作の仕方、特性などが異なる相互結合型（ホップフィールドタイプ）や教師信号を必要としないタイプのネットワークもあります。それらに対しては今回説明した内容が当てはまらない所もありますのでご了承下さい。

【参考図書】

もう少しニューラルネットワークについて知りたい方に対し参考図書をご紹介します。

○ニューラルネットワークを初めて勉強する方

「ニューラルコンピュータ 脳と神経に学ぶ」、合原 一幸、東京電機大学出版局

○ニューラルネットワークの理論を体系的に勉強してみたい方

「ニューラルネットワーク情報処理」、麻生 英樹、産業図書(株)

○最近のニューラルネットワークの動向を知りたい方

日経インテリジェントシステム別冊 1992年夏号、日経BP社

○ニューラルネットワークの応用事例を知りたい方

「ニューラルネットワークス応用事例集」WHITE SERIES NO. 98、阿部 洋一、合原一幸 監修、(株)トリケップス

以上の他にも最近では多数のニューラル関連図書が出版されていますので書店等でご自分の興味に合った図書を探してご覧下さい。

FT-NIRによるパターン認識手法と製造現場におけるその利用

ブラン・ルーベ株式会社
桜井香織

1. NIRスペクトルはどんな情報を含んでいるか

周知の通り、赤外線エネルギーはちょうど分子振動のエネルギーに相当するので、物質に赤外線を照射すると吸収が起こる。近赤外吸収スペクトル（NIRスペクトルという）は、最も可視光に近い赤外領域で測定される分光データであり、分子振動でも倍音と結合音による吸収に由来している。

NIR吸収スペクトルは、試料を希釈しなくても測定できる。従って、情報として、試料の化学的な組成や分子構造に関わるものと表面状態や結晶構造などの物理的な性質に関わるものを同時に含んでいる、と期待される。図1は粒度分布の違う結晶セルロースの拡散反射スペクトルであるが、粒度分布の差は、縦軸方向のオフセットとして非常に顕著に現れている。ところが一方、横軸方向では、倍音、結合音の吸収が複雑にオーバーラップするため、吸収バンドがブロード化して特性吸収波長を決定しにくい一面を示している。

工業計測においては、試料を前処理せずに測定できる簡便さは大きな魅力である。また、多くの化学物質を何度も使用したり製造するとき、その品質を化学的性質と物理的性質の両方について（あるいはそのどちらかについて）同定し確認するのは非常に重要なプロセスである。しかし、従来のスペクトルマッチングでは、NIRスペクトルの微少な差異を検出するには不十分であった。

そこで、当社では、ケモトリックスを応用し、NIRスペクトルから有用な情報を抽出してより迅速に品質の同定確認が行えるスペクトルパターン認識手法を、FT-NIR分光装置というハードウェアとともに開発した。

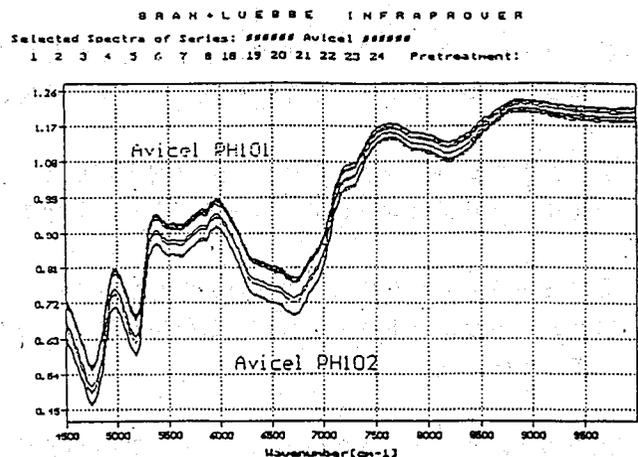


図1. 粒度分布によるスペクトル変化
(結晶セルロース、平均粒度径100 μ 、200 μ)

2. スペクトルパターン認識手法

あらかじめ、品質確認をするべき物質のリファレンススペクトルをライブラリーモデルとして記憶させ、新たな試料のスペクトルと比較して、判別及び同定確認するが、1つのライブラリーに複数の物質を格納し、それらの中で判別を行う。

NIR吸収スペクトルデータには、前述した品質確認試験に役立つ情報と不要な情報が混在しているので、まず、これらを分割したい。不要な情報とは、サンプリングの際発生する誤差や機器のノイズなどである。当社では、主成分分析 (Principal Component Analysis、以下PCA) を応用してこれを行っている。

スペクトルは、約500の波長で採られた吸光度によって構成されている。そこで、おのおのの吸光度を座標軸とした新しい変数空間を作って、データの記述の仕方を変えてみる。この多次元空間では、ひとつのスペクトルデータが一点で表現される。そして、複数物質 (リファレンス) のスペクトルデータはそれぞれ特異な位置に分布する。この位置のパターンをリファレンス物質の「判別モデル」として記憶させ、未知の試料がこのモデル中の物質に帰属されるか否かを同定する。しかし、ここでは吸光度空間 (約500次元の超空間) をそのまま使わず、物質間の違いを良く表すことができるいくつかの新しい座標軸を選び出す。これが主成分軸であり、この新しい座標軸を見つけたことは、スペクトルデータが含んでいる変動を寄与率の高い順に分解することに等しい。これによって、物質の化学物理的特性に由来しているデータ成分と不要な残差成分を分離することができる。また、約500次元もあったデータを、ただかか十数個の次元に圧縮することになる。

物質の品質の同定確認の基礎となるパターンモデルは、これらの新しい主成分軸のなかから個々の物質をもっともよく分離した位置にするものを選択して決定する。モデル化のために必要な主成分数は、1つでも良いこともあり、もっと多くを必要とすることもある。図2に示したように、この主成分空間において、各々のリファレンス物質の占有する空間容積を求めておき、同定確認の為の数值基準として利用する。

さらに、異物質の混入や測定の際の何らかの予想していない要因による誤差などを見逃すことがないように、残差となったスペクトル部分はリファレンスと未知試料との吸光度差を利用した従来型のスペクトルマッチング手法でダブルチェックをおこなっている。

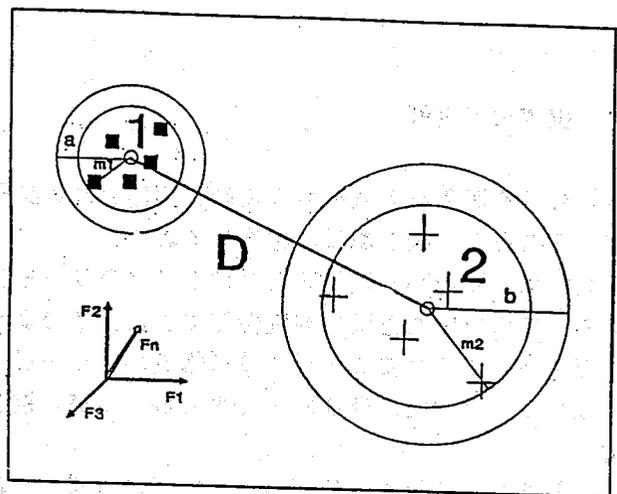


図2. 近赤外吸収スペクトルの主成分分析概念図

実際には、品質の確認をしたい物質の化学物理的性質が大きく異なっている場合と、非常に類似している場合とを想定し、それぞれ少し異なったモデル計算を行っている。後者の場合、わずかなスペクトルデータの差異を検出しなければならないので、モデルを確立するためにできるだけ多くのリファレンスデータを使って主成分空間が決定され、そこで各物質によって占有される空間容積を求める(図3)。この考え方は、Soft Independent modelling of Class Analogy、略してSIMCAと呼ばれる方法を応用したものである。

他方、前者の場合には、各物質の重心となるデータによって主成分空間が決定され、各々の重心の回りでの占有空間面積を求める。このような物質同士は、一般に、スペクトル自体明らかに異なっていることが多く、データ変動量が大きいのでそれらを個々の寄与に分離しやすい。そして、同一物質データ内で起こり得るデータ変動は異種物質間の変動量に比べれば、小さいものである。ゆえに、重心データによる主成分空間で十分なのである。また、この方法は計算処理時間の短縮にもなっている。

長期的に安定で信頼性の高い実用的な判別モデルを確立するための必須条件は、リファレンスデータが予想される試料母集団をきちんと代表していることである。また、いったんモデルができあがっても、別に用意したデータで将来的な性能を評価しなければならない。

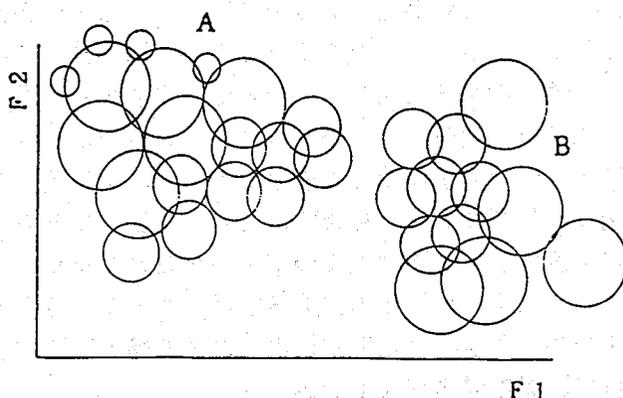


図3. 類似物質の為の
判別確認モデル概念図

3. 実際の応用例

化粧品や医薬品の原材料入荷時に実施される品質チェックに日常的に使用されているので、それらの一例を簡単に紹介しておく。

図5は、賦形剤や結合剤などのライブラリーモデルである。この場合、2主成分を使ってすべての物質を正確に判別確認できた。このようなライブラリーは、幾つでも作ることができるので、必要に応じて試料物質のグルーピングを行って記憶させることもできる。

図6は、化学的にほぼ同一と考えられるが、粒度分布が異なる結晶セルロースのモデルである。

原材料だけでなく、製品規格の検査にも利用されている。ラミネートフィルム上のポリマーコーティング量の異なる4系統の製品の品質確認検査のためのモデルを図7に示した。これらのモデルは、評価用の試料群を100%まちがいがなく判別することができた。

未知試料がモデルのどの物質にも帰属しない場合は、もっとも近似している物質に対し、どのくらいの類似度であるかと、スペクトル残差を参考数値として与えてくれる。

当社の装置では、液体透過測定から拡散反射測定までがオプションなしで可能なので、原料の受け入れ、中間工程での予備的チェック、最終製品の品質保証まで、工業計測のあらゆる段階で使われる。現在では、バーコード入力によって品目ID、製造バッチ番号、コンテナ番号などを認識させ、分光学の知識を持たない人でもこのシステムによって数値で次々と品質チェックをこなして、トラブル防止、工程ダウンタイムの軽減が実現されている。

図4. ライブラリーモデル例
(医薬品原料)

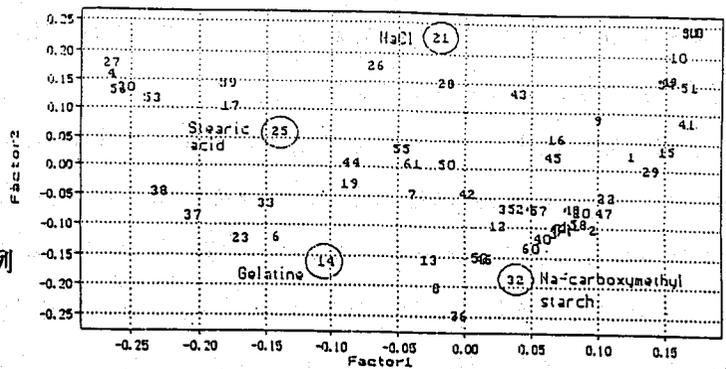


図5. 類似物質判別モデル例
(結晶セルロース)

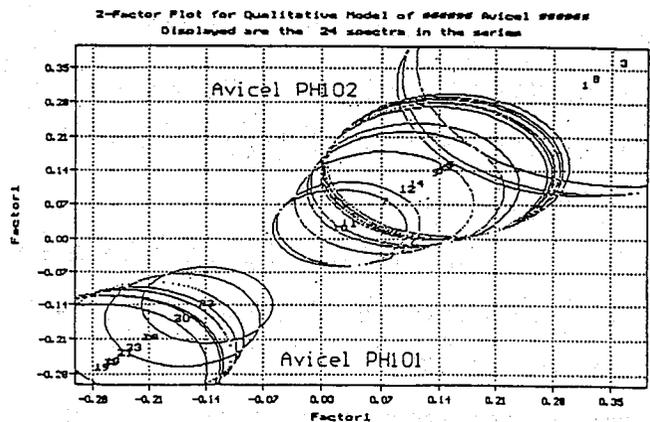
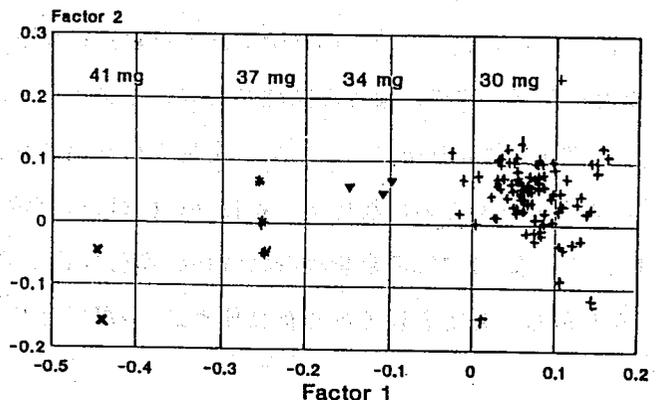


図6. 製品規格判別モデル例
(ラミネートフィルム)



<参考文献>

- 1) N.K.Shah, P.J.Gemperline:Anal. Chem.:61, (1990), 138-144
- 2) 相島鐵郎:ケモメトリックス、丸善(1992)

FT-IRへの計量マトリックスの応用

(株)ハルキョウ エルマージャパン 応用技術部 新見 栄

1. はじめに

赤外分光法を用いての物質の定性・定量分析は古くから行なわれています。近年の赤外分光光度計(FT-IR)、データ処理システム(コンピュータ)の進歩に伴い、さらに赤外分光法の定量分析への応用が広まってきています。

従来、赤外スペクトルを利用した定量分析は、検量線法、カーブフィッティング法が用いられてきました。カーブフィッティング法は、赤外スペクトルの全測定ポイントを用い最小自乗法により検量マトリックスを作成する手法で、スペクトル中あるピークの高さ、あるいは面積を用いて定量を行なう検量線法に比べ、より科学的精度の高い定量手法であると言えます。しかし、スペクトルはノイズも含んでおり、また吸収の存在しない波数領域もあります。従って、この吸収波数に関する情報、ノイズに関する情報、両方を含む赤外スペクトルを用いて定量を行なうには、これらの中から必要な情報のみを選択する必要があります。そこで、必要な情報のみを選択し、互いに直交するスペクトルを作成、定量分析を行なう、因子解析法を用いた定量分析が行なわれるようになりました。

ここでは、因子解析法の1つである主成分分析法を応用した定量分析ソフトウェア-QUANT+-の概要と、このソフトウェアを用いた分析例を紹介します。

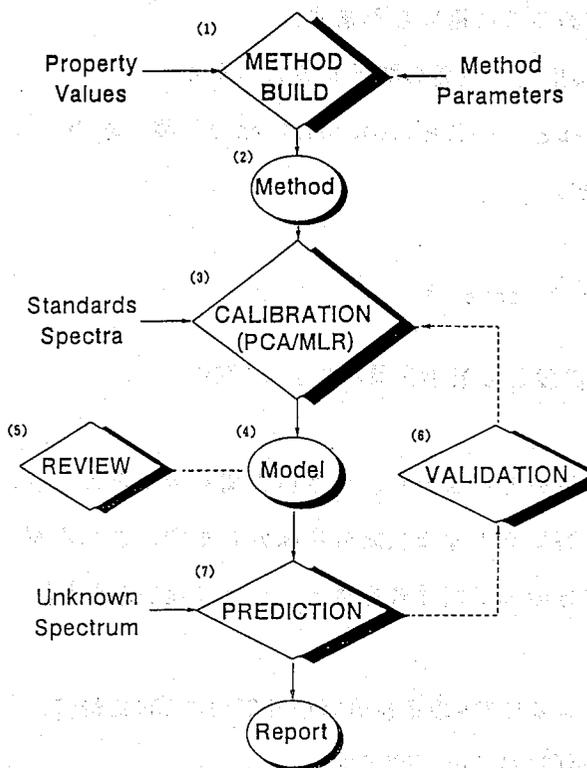
2. QUANT+ソフトウェアの概要

QUANT+は多変量解析に適用できる計量マトリックス定量分析ソフトウェアです。

従来の分析法は、サンプルの(全てあるいは注目する)成分ピークが分離されていることが必要でしたが、QUANT+はこの条件を必要としません。

QUANT+は主成分回帰によって、定量を行ないます。この計量マトリックス手法は、calibrationスペクトルと各成分あるいは注目する特性-property-に相関を求めます。そしてこの相関関係を用いて、未知サンプルの定量-prediction-を行ないます。この手法はできるだけ多数のスペクトルを利用し、分析者はその分析に伴うエラーや異常サンプルを確認することができます。

図1にQUANT+の概略図を示します。



(1)METHOD BUILDで分析に必要なパラメータ(波数範囲、ステップ、解析手法等)を設定し、(2)Methodで記憶させます。
 (3)CALIBRATIONで実際に解析を実行させます。ここで求められた(Model)は(5)REVIEWで確認することができ、必要に応じてMethodの設定を変更(改良)、最適なModelを得られるまでこれを繰り返します。そして、満足な結果を取得後、(7)PREDICTIONで未知サンプルの定量分析を行ないます。

図1 QUANT+概略

主成分回帰分析-Principle Component Regression Analysis-を用いたQUANT+の主ステップである(3)CALIBRATIONは、2つのステップからなります。まず、主成分分析-Principle Component Analysis-を行ない、各スペクトルを最小数のfactor(主成分)の重量平均で表わします。このPCAは、次のマトリックスで表わされます。

$$\begin{matrix} & & n_s & & & n_f & & n_s \\ & & \left[\begin{array}{c} \text{DATA} \\ \text{MATRIX} \end{array} \right] & = & \left[\begin{array}{c} \text{VARIABLE} \\ \text{LOADINGS} \\ \text{MATRIX} \end{array} \right] & \left[\begin{array}{c} \text{SAMPLE} \\ \text{SCORES} \\ \text{MATRIX} \end{array} \right] \\ n_v & & & n_v & & n_f & & \\ & & & & & & & \end{matrix}$$

n_s : サンプル数 n_f : factors数(PCs)

n_v : 変数の数(通常各スペクトルのデータポイント数)

SAMPLE SCORES MATRIXはスペクトル間の関係を表現し、次のステップである重回帰分析-Multiple Linear Regression-で各propertyの相関を求めるのに用いられます。

他の2つのマトリックスは、異なる振動数における強度(ピーク)間の関係を表わします。

MLRは、calibrationスタンダードの主成分Scoresと、それらのcompositionの相関を求め、次式より未知サンプルの濃度の計算を行ないます。

$$\text{濃度} = k_0 + k_1 * \text{score}(1) + \dots + k_n * \text{score}(n)$$

k_0, k_1 : calibrationサンプルのabs.と濃度との最小自乗法による定数

この段階(計算)は、全てのfactor(主成分)を用いて行なわれ、小さい相関を示すだけの主成分がもつpropertyを見つけます。そして同時にその様な主成分を除外します。これを繰り返すことによって、重要な相関を示す主成分のみで回帰式を構成し、より良いModelを求めます。

こうして求めた回帰式が、最終目的である未知サンプルの定量分析(7)PREDICTIONに最適であるか、また十分であるかを確認するステップが(6)VALIDATIONです。

すなわち、(3)CALIBRATIONによって得た回帰式の評価を行ない、未知サンプルの定量分析に最適の回帰式を求めることを行ないます。

QUANT+は2種類のVALIDATION機能を持ちます。1つはindependent validationです。これはcalibration Modelをvalidationサンプルセットで評価する手法で、例えばcalibration結果にvalidation結果が匹敵しない場合、その回帰モデルに採用されたfactorになんらかの問題があると判断することができます。この手法を行なうのに適したサンプルがない場合、full cross validation手法を用います。この手法はcalibrationからあるサンプルを除き、残りのサンプルで求めたcalibrationを評価します。サンプルを取り除くことでcalibrationセットのバランスを崩し、歪んだModelを作成、評価し、サンプルを取り除いていないcalibrationセットでの結果と比較することでModelの評価を行ないます。得られたvalidation結果より、Modelに適さないサンプルや異常サンプルを容易に確認することができます。

QUANT+は、IR(赤外分光)、UV/VIS(紫外/可視分光)、FL(蛍光分光)、MS(質量)、その他の分析手法で得たスペクトルを用いて分析を行なうことができ、また用いるデータのフォーマットもASCII等に対応したソフトウェアです。

3. 分析例

3-1. はじめに

オクタン価は、ガソリンのアンチノック性を数量的に表わす指数で、ガソリン機関燃料の中で最も重要な性質です。すなわち、オクタン価の高いガソリンほどアンチノック性が高く、機関内で出力低下の原因となるノッキング現象(異常燃焼)を起こしにくい性質を示します。このアンチノック性は、その化学的成分によって大きな差があり、分枝状脂肪族と芳香族炭化水素が最も高い、そして直鎖状脂肪族炭化水素が最も低いアンチノック性を示します。

試料ガソリンの測定は、現在、国際的に規格化されたCFRインジコンで行なっています。⁽¹⁾ アンチノック性の高い代表的成分としてイソオクタンを、アンチノック性の低い代表的成分としてn-ヘプタンを採用し、それぞれのオクタン価を100,0と定めます。そして、両成分を任意の割合に混合し、測定、イソオクタンのvol%の数値でそのオクタン価を表わします。

この手法による測定には1検体の分析に30分を要し、また精度を必要とする場合には、熟練者による測定が要求されます。今回、容易かつ迅速に測定を行なえるFT-IRを用いてガソリンのオクタン価の決定を試みました。

3-2. 実験

測定には、キヤシバ製System2000FT-IR分光装置を用い、データ処理に、QUANTソフトウェアを用いました。アクセサリとして、スペクトラテック社製サークルセル(IRE:ZnSe)を用い、内部多重反射法で測定を行ないました。測定条件は、分解能 8cm^{-1} 、測定波数範囲 $4000\sim 700\text{cm}^{-1}$ 、検出器DTGS、測定回数16回です。スタンダードサンプルとして、オクタン価82.4~98.5の試料ガソリン69点を測定しました。

3-3. 結果及び考察

図2に、オクタン価の異なる3種類のスタンダードサンプルのIRスペクトルと、オクタン価未知のサンプルのIRスペクトルを示します。芳香族化合物の存在を示す吸収($3028, 1608, 1496, 796, 720\text{cm}^{-1}$)と、アルフィン化合物の存在を示す吸収($2928, 2872, 1460\text{cm}^{-1}$)を、それぞれ確認します。測定手法が内部多重反射法であるため、特に芳香族の面外変角振動領域($900\sim 700\text{cm}^{-1}$)の吸収がより強いものとなっています。

図3は、次式で計算されるガソリンのオクタン価の相関スペクトルです。

$$P = 1/N \sum_{i=1}^{ns} (P_i - P')(S_i - S')$$

P : 相関スぺクトル

P_i : i番目のガソリンのオクタン価

P' : ガソリンのオクタン価の平均値

S_i : i番目のサンプルのIRスぺクトル

S' : 平均IRスぺクトル

ns : サンプル数

$$N : S' / \sum_{i=1}^{ns} (P_i - P')^2$$

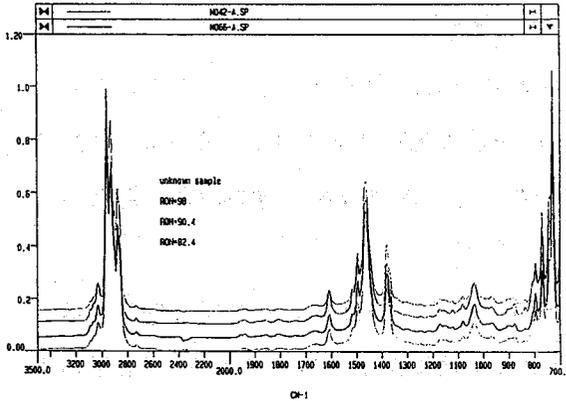


図2 ガソリンのIRスぺクトル

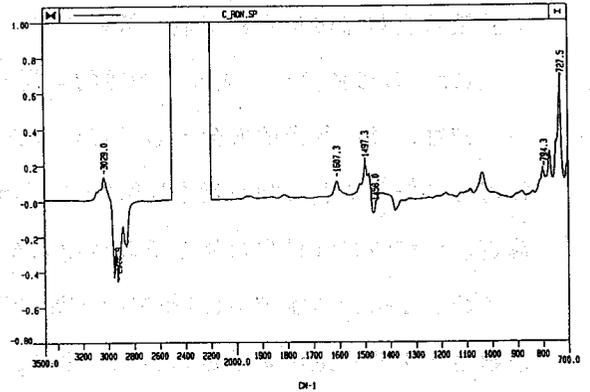


図3 ガソリンのオクタン価の相関スぺクトル

芳香族の特徴を表わすピークは正を、アルフィン(脂肪族炭化水素)の特徴を表わすピークは負を示しています。これは、芳香族炭化水素は高いオクタン価を示し、脂肪族炭化水素は低いオクタン価を示すことに対応します。(2)

RHT+ (Revision 2.00) METHOD RONM4 Ver.: 1, I.D.: 3327

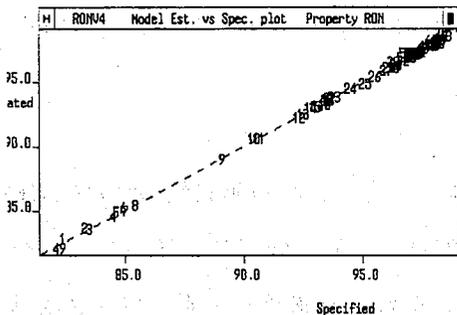


図4に、QUANT+を用いて計算されたオクタン価の値 (Estimated) と入力値 (Specified) の比較を示します。このModelは相関係数 0.9988、標準誤差 0.1964をもち、非常に高い定量精度を得ることができました。

参考文献

(1) Annual Book of ASTM Standards, Petroleum Products, Lubricants and Fossil Fuels (ASTM, Philadelphia, 1990), vol. 05.04, pp. 13, 21, 26.

(2) P.M. Frederics, J.B. Lee, P.R. Osborn and D.A.J. Swinkels, Appl. Spectrosc. 39, 311 (1985).

1. はじめに

C-13 NMRスペクトルは、有機化合物の炭素骨格そのものの情報を与えるので、立体構造の解析には非常に有用であることが知られている。しかしながら、化学シフトから立体異性体を判別する際には、化学者の専門の知識を必要とし簡単には区別できない場合が多い。例えば、2, 3, 4位にメチルまたは t -ブチル基を持つ1-メチルシクロヘキサノールの cis -体と $trans$ -体の環炭素の化学シフトを表1に示したが、これら12種類の立体異性体をすべて並べると2-および4-アルキル置換体の場合には

cis -体が、3-アルキル置換体の場合には $trans$ -体が概して高磁場にシフトしていることが分かるが¹⁾、この中の一つのデータのみがある場合に、この化合物がどのような立体配置にあるかを予測することは、非常に困難を伴うだろう。このようなときに、化学シフト値を多変量とし、統計的な解析をすることにより立体配置の判別に有用な情報が得られるのではないかと考え、初めにこの系の主成分分析を行なってみた。

Table 1. C-13 Chemical Shifts of 1-Methylcyclohexanols

	C-1	C-2	C-3	C-4	C-5	C-6
c	71.0	40.5	30.7	26.1	22.1	40.1
t	73.0	42.3	32.2	25.5	24.3	41.5
c	73.8	54.8	24.5	27.1	22.1	44.6
t	75.5	56.6	27.2	27.2	24.8	46.9
c	71.2	49.5	30.6	34.8	23.8	40.2
t	69.8	47.7	27.8	34.6	21.7	38.4
c	71.6	42.0	45.8	26.9	24.0	40.0
t	70.1	40.0	42.8	26.6	22.1	38.6
c	68.8	38.6	30.5	31.4	30.5	38.6
t	70.6	39.8	32.4	31.9	32.4	39.8
c	68.9	39.4	22.7	47.7	22.7	39.3
t	71.2	41.0	25.1	47.9	25.1	41.0

2. 1-メチルシクロヘキサノール類の主成分分析

12種類のそれぞれの立体異性体(図1)の環炭素のC-13化学シフトから、母核である t -ブチルシクロヘキサノールまたはメチルシクロヘキサノールの環炭素のそれらを引いた $\Delta\delta$ 値を用いて主成分分析を行なった。その結果、第一(t_1)と第二(t_2)主成分および t_1 と第三(t_3)主成分の二次元スコアプロットを取っても十分に区別

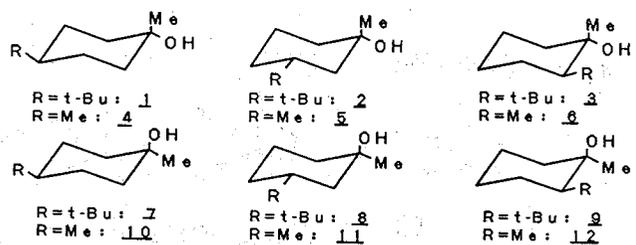
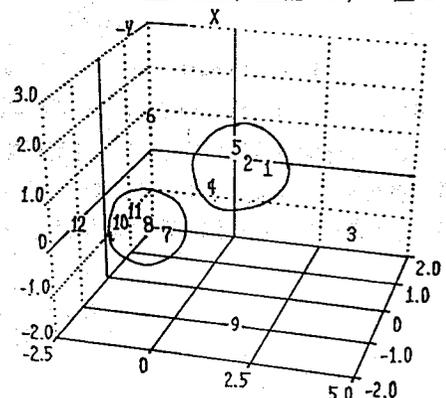


Fig. 1. Stereoisomers of 1-Methylcyclohexanols

することができなかったが、 t_1 , t_2 および t_3 の三次元プロットを取った場合に、メチル基がアキシャル位にある1から6の立体異性体と水酸基がアキシャル位にある7から12の立体異性体をきれいに二つの領域に識別することができた(図2)。と同時に、1, 2, 4, 5および7, 8, 10, 11の3, 4位にアルキル基を持つ場合にはそれぞれのクラスターを形成し、2位に置換基を有し1位の置換基と大きな立体的相互作用を持つ3, 6, 9および12の化合物は、これらのクラスターとは離れて存在しその特殊な状況をも併せて示していることが分かる。このように、C-13化学シフトデータを利用したケモメトリックス手法は有機化合物の立体異性体の識別に有用であることが分かったので、次に多変量解析の中では比較的新しいPLS (Partial Least Squares) 法²⁾を用いて、ノルボルナン類およびシクロヘキサノール誘導体の立体異性体の識別を行なった例を紹介する。

Fig. 2. PCR Score Plot (t_1, t_2, t_3) of 1-Methylcyclohexanols

3. PLS法

ノルボルナン類は、各々の基本化合物と炭素骨格の化学シフト差 ($\Delta\delta_n$) により、7次元ベクトル \mathbf{x} として表すことができる (式1)。

$$\mathbf{x} = (\Delta\delta_1, \dots, \Delta\delta_7) \quad (1)$$

$\Delta\delta_n$ で表されたノルボルナン類を行列 X とし、その立体配置を *exo* を1, *endo* を0とする目的変数ベクトル \mathbf{y} で表す。PLS解析の最適化の基準は X 空間の潜在変数 t が \mathbf{y} と最大分散を持つと同時に、 \mathbf{y} と t との相関を大きくすることである。行列 X は、

$$\mathbf{x} = \sum^A t_h \mathbf{p}'_h + \mathbf{E} \quad (2)$$

と書ける。

ここで \mathbf{p}_h はローディングベクトル、 \mathbf{E} は残差行列である。目的変数 \mathbf{y} はこの t_h を用いて

$$\mathbf{y} = \sum^A \rho_h t_h + \mathbf{f} \quad (3)$$

と書ける。ここで ρ_h は第 h 成分のモデルであり \mathbf{f} は \mathbf{y} の残差である。またモデル式の成分数 A はクロスバリデーションによって求める。

4. ノルボルナン類のPLS解析³⁾

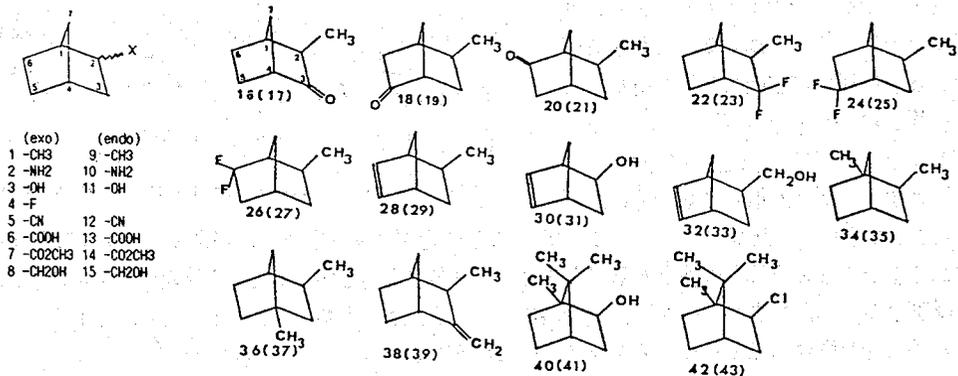


Fig. 3. Calibration (1~15) and Validation (16~43) Sets of Norbornanes

初めに、Calibration Setとして立体配置 (*exo*, *endo*) 既知のノルボルナン類1~15 (図3) について立体配置分類のためのPLS解析を行なった。化学シフト差 $\Delta\delta$ の7次元データの各値を平均0、標準偏差が1となるようにオートスケールしPLS解析を行なった結果、クロスバリデーションで2成が有意であった。PLS法では説明変数 X_k の有意の程度をモデリング力 ϕ_k によって表す (式4)。

$$\phi_k = 1 - \sqrt{\frac{\sum e_{ik}^2 / (n-A-1)}{\sum X_{ik}^2 / (n-1)}} \quad (4)$$

ここで、 e_{ik} はデータ行列 X の残差の ik 成分、 n はサンプル数である。即ち、 ϕ_k が1なら X_k の全ての

情報がモデリングに役立ち、0なら全く役に立たないことを示す。式4によるモデリング力の計算結果、PLSモデルは2成分で全ての変数が有意になり、特に置換基の位置C-2と γ 位のC-4, C-6, C-7の情報が重要なことが明らかになった。

PLSモデル式は、式5の様に2成分で表現される。

$$y=0.537t_1+0.208t_2 \quad (5)$$

このモデル式により全分散の92.5%を説明でき、またクロスバリデーションによる説明分散は91.2%となった。また式5はスケーリングした化学シフト差によって、式6の様に表すことができる。この式より求めた識別関数の値を表2に示した。これより明らかなように、exo/endoの立体配置の分類は良好に行なわれる。図4に化合物1~15のPLSスコアプロットを示した。

$$y=0.052\Delta\delta_1-0.017\Delta\delta_2+0.022\Delta\delta_3-0.283\Delta\delta_4-0.126\Delta\delta_5+0.421\Delta\delta_6-0.377\Delta\delta_7 \quad (6)$$

次にローディングP1とP2のプロットを図5に示す。この図より、 α , β 位のC-1, C-2, C-3および γ 位のC-4, C-7はそれぞれのクラスターを作り、それぞれの化学シフトは類似の挙動を示していることが分かるが、同じ γ 位でもC-6は独自の挙動を示し、また δ 位のC-5は4, 7位と似た方向性を持つことが分かる。

Validation Setとして化合物16~43までの28の化合物の立体配置を予測するために、これらの母核との化学シフト差をモデル式5に代入して y の予測値を求めた。その結果、exo体は1付近の、endo体は0付近の値を示し、0.5をシキイ値とすると全ての立体配置の予測を正しく行ない、式5は予測式として有意であることが分かった。図4には16~43におけるPLS潜在変数 t_1 と t_2 のスコアプロットを1~15と併せて示してある。これより、Validation Setのノルボルナン類はCalibration Setの付近に位置し、アシンメトリックなデータ構造を示している。このようにPLS解析を行なうと、クラス分類の様子を良好に反映したデータ構造を容易に得ることができ、先に開発されたSIMCA (Soft Independent modeling of Class Analogy) 法⁴⁾の数学的な複雑さを回避することもできる。

5. シクロヘキサン誘導体のPLS解析

シクロヘキサンのアキシャル位およびエクアトリアル位に種々の官能基を持つ化合物の環炭素の化学シフトデータを用いてノルボルナン類と同様にPLS解析を行なった。その結果、置換基から β , γ , δ 位にある $\Delta\delta$ 値を使って、それらのスコアプロットをした結果、アキシャルおよびエクアトリアル位に置換基を持つ立体異性体は良好に識別されると同時に(図6)、ハロゲン置換基の場合にはI, Br, Cl, Fと周期率表の順に上から並び、単に立体化学的な情報だけでなく電子的な情報をもPLS解析から得られることが明確に確認された。

6. おわりに

PLS法を使ったC-13 NMR化学シフトの解析により、立体化学に関する知見を得るのみならず電子的な情報も得られることが明らかになった。さらに、PLS法は他のケモメトリックス手法と同様に化学シフト帰属の有力な補助手段となることが判明した。

最後に、本研究に於けるPLS解析をご指導下さった豊橋技術科学大学の宮下芳勝助教授に深く感謝します。

Novel Kalman Filtering Techniques in Chemometrics

Zhiliang LI^{*} Department of Knowledge-based Information Engineering, Toyohashi University of Technology, Tempaku-cho, Toyohashi 441, JAPAN / ^{*}On leave from Department of Chemistry and Chemical Engineering, Hunan University, Changsha 410082, P.R.CHINA,

This lecture will describe the Kalman filter(KF) in chemometrics and cover the following contents: introduction, algorithms and applications. It will mainly focus on some novel KF: fading(FE), parallel/networked(PF,NF) and parallel fading KF(PFF).

Introduction

At first, I would like to give a brief history of KF, more than 30 years. In 1960, a recursive optimal filtering algorithm was developed based on the least-squares by Kalman[1] and then called the Kalman filter. In 1976, Seelig[2] used extended KF(EKF) to voltammetry, which can be regarded as the beginning of applications of KF into chemistry. After 3 years, Poulisse[3] applied the conventional KF (CF) to spectroscopy, In 1984, Rutan and Brown[4] employed the adaptive KF(AKF) to multicomponent analysis, They and other chemists wrote several reviews in 1986-1988[5-7]. In China, we investigated on the performance and applications of KF in analytical chemistry[8-12], for instance, to resolve overlapped spectroscopic[8], fluorescent[9] and chromatographic responses, to interpret chemical equilibria[10], to correct titration errors[11], to perform valence speciation[12]. In Japan, Hayashi's group used KF to chromatography[13]. In the early of 1990s, Wentzell's team proposed the parallel Kalman filters network(PKFN) to peak purity analysis[14]. Recently, we have developed several novel filters: FE[15], PF or NF[16,17], PFF[18] and used them to multicomponent analysis, impurity detection, elimination of nonlinear effects, multivariate QSAR studies, kinetic and pharmacokinetic data analysis.

In order to solve multivariate chemical problems, we use various data processing techniques. Those can be divided into two ways: one is the batch way, such as multiple linear regression (MLR), principal component regression (PCR), and partial least squares(PLS); another is recursive way: this means to process data step by step. The most useful technique is KF. KF as a new alternative method is widely used and paid more and more attentions. There are various algorithms covering CF, AKF, EKF, and plus three novel filters FE, PF or NF and PFF.

KF needs no matrix inversion, uses very small memory and estimates while measurement. It can estimate both state and parameter, give rapid calculations and accurate results and perform the real-time or on-line data processing. Of course, there are some limits: convergence may not be obtained if the system is collinear or not observable; proper initial guesses are required.

There are three types of estimation in the time domain: (a) prediction, (b) filtering and (c) smoothing. The symbol $x(k/j)$ means the state estimation at time k based on the measurement at time j . They are called prediction, filtering and smoothing When $j < k$, $j = k$ and $j > k$, respectively. KF includes those three estimation stages. But usually, it consists of the first two steps: prediction and filtering.

Model and algorithm

What is KF? The system under consideration and the algorithm equations are given in Table 1. Here $x(k)$ is the state/parameter vector, $z(k)$ is the measurement, $F(k,k-1)$ is transformation matrix, $h'(k)$ is the functional vector; $y(k)$ is the innovation which is very important to reflect whether the convergence is obtained or not and the model is correct or not; $w(k)$ and $v(k)$ are the dynamic and measurement noises. The noises are assumed to be random with zero-mean, white gaussian distributions. KF gives the recursive optimal estimation.

Table 1. The Conventional and Fading Kalman Filtering Algorithms

Dynamic (state-space) model:	$x(k)=F(k,k-1)x(k-1)+w(k), x(k) \in \mathbb{R}^{(n)}$
Measurement (process) model:	$z(k)=h'(k)x(k)+v(k), z(k) \in \mathbb{R}^{(1)}$
Initial guesses:	$E[x(0)]=x(0/0)=x_0$ or $0, E[(x(0)-x_0)(x(0)-x_0)'] = P(0/0)=P_0=p_0 I$
State estimate extrapolation(prediction):	$x(k/k-1)=F(k,k-1)x(k-1/k-1)$
Error covariance estimation extrapolation:	$P(k/k-1)=F(k,k-1)P(k-1/k-1)F'(k,k-1)+Q(k-1)$ with fading factor: $P(k/k-1)=f(k)F(k,k-1)P(k-1/k-1)F'(k,k-1)+Q(k-1), f(k)>1$
Innovation (scalar):	$y(k)=z(k) - h'(k) x(k/k-1)$
Kalman gain (vector):	$g(k)=P(k/k-1)h(k) [h'(k)P(k/k-1)h(k)+r(k-1)]^{-1}$
State estimate update(filtering):	$x(k/k)=x(k/k-1) + g(k)y(k)$
Error covariance estimation update:	$P(k/k)=[I-g(k)h'(k)]P(k/k-1)[I-g(k)h'(k)]'+g(k)r(k)g'(k)$

Now let us take a very simple example of KF for linear calibration, $a(k)=b*c(k)+v(k)$, in order to explain how to use KF. Here $a(k)$ is absorbance, $c(k)$ the standard concentration, $v(k)$ the noise with variance $r=0.01$, $x(k)=b$ the calibration slope, the parameter to be estimated. The model information and initial guess are given as $r(k) = 0.01$, $Q(k) = 0$, $x(0/0)=0$, $P(0/0)= 1.00$. At first ($k=1$), we obtain the extrapolation $x(1/0)$, and $p(1/0)$, the innovation $y(1)$, the gain $g(1)$, and the filtering estimates $x(1/1)$ and $p(1/1)$. This process is repeated until the end. Finally we estimates the slope $b=x(4/4)=0.19994$ and the error covariance $p(4/4)=0.0003332$ (see Table 2 for details).

Table 2. An example calculated by CKF algorithm

k	z(k)	h(k)	extrapolation		innov.	gain	update estimation	
	a(k)	c(k)	x(k/k-1)	P(k/k-1)	y(k)	g(k)	x(k/k)	P(k/k)
1	0.212	1.0	0.0000	1.000E00	0.2120	0.9901	0.2099	0.0099099
2	0.396	2.0	0.2099	9.901E-3	-0.0238	0.3992	0.2004	0.0019960
3	0.620	3.0	0.2004	1.996E-3	0.0176	0.2141	0.2044	0.0007138
4	0.784	4.0	0.2044	7.138E-4	-0.0336	0.1333	0.1999	0.0003332

KF has been widely used in chemistry especially in analytical chemistry[5-18]. Now let us discuss the novel KF.

The Novel Kalman Filters

The Networked or Parallel Kalman Filter. In contrast to CF which uses only one unitary model, the novel PF or NF technique employs a set of various models, i.e. multiple models, see Figure 1. PF was used for impurity detection at the level of less than 1% even low to 0.1% of the primary compounds[16].

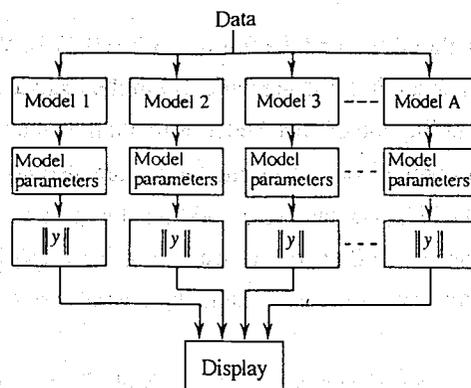


Fig.1. Parallel Kalman filter

The fading KF. FF is a modified algorithm by introducing a fading factor(f) to CF, see Table 1 for details. Usually f is greater than or at least equal to 1. If f is always taken as 1, FF is reduced to CF. How to select the fading factors? which is a very important problem. There are two ways: one is taken as constants, for example, $f=1.02, 1.05, 1.2, \dots$, etc.; another is taken as changables, i.g., $f(k)$ is k -dependent. On the CAMSE'92 in Yokohama [15], we have made some primary conclusions: FF can convergence faster than CF; the accuracy is slightly less than that of CF; the memory for FF is also very small; the fading algorithm needs no matrix inversion, too; the estimation errors increase with system noises and fading factors; the changable fading factors are better than constants and the optimal adaptive fading factor is the best. The later is quite complicated and omitted due to the space.

We have used FF to solve the following problems: state/parameter estimation; multicomponent analysis; equilibrium interpretation, kinetic analysis including the rate-differential kinetic analysis and pharmacokinetic data analysis; multivariate QSAR studies.

Table 3. Estimation of parameters for simulated data by KF

f	converg fitting		parameters estimated			relative deviation(%)		
	at point	RSD%	A_1	A_2	C	e_1	e_2	e_3
1.00	11/12	0.27	0.3982	0.4002	0.2003	-0.45	0.05	0.15
1.10	89	0.28	0.3917	0.4045	0.1994	-2.08	1.10	-0.30
1.20	89	0.30	0.3852	0.4070	0.1936	-3.70	1.75	-3.20
0.90	11/12	0.29	0.4034	0.3944	0.2024	0.85	-1.35	1.20
0.70	16/17	1.07	0.3842	0.4314	0.1818	3.95	7.85	-9.05
$1+(m-k)/m$			0.3968	0.4049	0.2006	-0.80	1.02	0.30
the optimal f			0.3946	0.4026	0.2004	-1.35	0.65	0.02

Here, we will focus on application to the kinetic analysis. The model for first-order kinetic reaction can be expressed as this equation: $s(t)=A \exp(-Bt) + C$. For the differential-rate kinetic analysis, we have $s(t)=\sum[A_i \exp(-B_i t)] + C$. In order to demonstrate the performance of FF, we have done the simulation studies. For example, the real parameters are $A_1=0.4, B_1=0.20, A_2=0.4, B_2=0.05, C=0.2, \text{noise}=0.5\%$. The model parameters are $B_1=0.20, B_2=0.05$, and the initial guess was taken as $x(0/0)=0, r=10^{-4}, P(0/0)=10^4 I$. A_1, A_2 and C are the state parameters to be estimated. Both FF and CF give the parameter estimation. When $f=1$, the estimated parameters is given by 0.3982, 0.4002,

0.2003, the relative error are within 0.5% and the fitting RSD is 0.27%. When $f=1.2$, $A_1=0.3917$, $A_2=0.4045$ and $C=0.1994$, which are in good agreement with the real values and with the results by the CF. The optimal adaptive fading factor gave better estimation than all constants and other adjustable fading factors. For comparison, the fading factors are also taken as less than 1. It is found that this type of KF, called XF, can also give good results. But, when fading factors are taken too large or too small, for instance, $f=2.0$ or 0.6 , the estimation error will become very large. Effects of parameter number and model errors on the FF estimation are also examined. Part of results are shown in Table 3.

FF is also used in pharmacokinetic data analysis. The models for pharmacokinetics are expressed by $s(t)=\sum[A_i \exp(-\alpha_i * t)]$ and $s(t)=A \exp(-\alpha * t)+B \exp(-\beta * t)$ for the multi- and two-compartment systems, respectively. Here α , β are the rate constants of pharmacokinetics; A and B as well as Cp^0 are the parameters to be estimated. The algorithm for FF is given by $x(0)=0$, $r=10^{-4}$, $P(0)=10^4$, $f \sim 1$.

The pharmacokinetic process of sulfisoxazole[19] and fitting results based on two-compartment model by reference[19], FF and CF can be seen in Figure 2. All fitting errors are very small.

The results of subjects 1 to 7 given by reference[18] are listed here. It is needed to point out that subjects 6 and 7 are involved some problems: number 6 did not lend itself to calculate parameters A and α ; number 7 was given too large values, $A=222.5$, which is much greater than that of any other subject. We consider it as the outlier. In this case FF and CF give more reasonable results (Table 3).

Table 3. The estimated Pharmacokinetic parameters

Subject	1	2	3	4	5	6	7	mean
conventional procedure[19]								
A(mg/ml)	95.92	94.42	69.63	84.60	83.50	-	222.5	108.4
B(mg/ml)	175.0	124.2	129.1	181.2	156.2	161.3	137.9	152.1
Cp^0	270.9	218.6	198.7	265.8	239.7	-	360.4	259.0
CF (f=1.0)								
A	95.71	90.36	76.42	85.14	87.52	56.97	92.92	83.6
B	171.9	126.7	122.0	181.9	152.1	177.9	164.2	156.7
Cp^0	277.6	217.0	198.4	267.0	239.6	234.9	257.1	239.4
FF (f=1.2)								
A	94.91	95.45	79.41	86.43	86.96	73.15	109.35	88.4
B	172.1	125.1	120.6	181.3	152.3	172.2	158.6	148.3
Cp^0	267.0	220.6	200.4	267.7	239.3	245.4	268.0	239.0

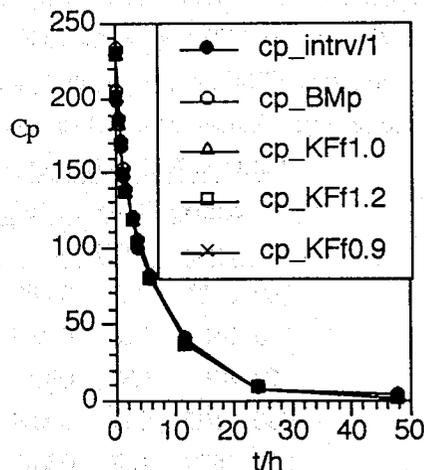


Figure 2 Experimental (1) and fitting curves estimated by [18](2), CF(3), FF(4) and XF(5).

The Parallel Fading Kalman Filter (PFF). PFF consists of multiple models and several fading factors. The algorithm and effects of models and factors were examined. PFF with 10 models and 7 fading factors was applied to QSAR studies. An evaluation function for 2,4-diamino-5-(substituted-benzyl) pyrimidine DHFR inhibitors is $RSD(\%) = \text{SQR}(\sum(\text{Act}_{\text{obsd}} - \text{Act}_{\text{pred}})^2 / (N_{\text{obj}} - 1))$. The best two modelling QSAR equations and $RSD(n=68)$ are obtained as follows:

$$(a) \text{Act} = 0.938MR'_5 + 0.895MR'_3 + 0.822MR'_4 - 0.221MR'_4^2 + 1.580\pi - 1.780\log(\beta 10^\pi + 1) + 6.647$$

$$(b) \text{ Act} = 6.904MR'_5 - 17.735MR'_5{}^2 + 13.808MR'_5{}^3 + 0.912MR'_3 + 0.806MR_4 - 0.217MR_4{}^2 + 1.420\pi - 1.610 \log(\beta 10^x + 1) + 6.150$$

The comparison of PKF, neural networks(NN), MLR, PCR and PLS is done in order to compare with each other. The observed and calculated activities of the pyrimidines are in good agreement with each other. The number of outliers (NO), whose deviation between the observed activity and the calculated activity is greater than 0.25 according to reference[19], and the number of erroneous estimates(NE), whose RSD is larger than 5%, are shown in Table 4. It is clear that PKF can also give a proper model for QSAR problems.

Table 4. RSD, NO and NE for various methods

Method	RSD%	NO	NE
PKF1	3.86	25	14
PKF2	3.78	25	14
NN	3.81	21(23)	13
MLR	3.86	26	14
PCR		32(6v)	
PLS		31(6v)	

Acknowledgement: The financial support from the Japanese Ministry of Education, Science and Culture (Monbusho), the helpful comments of Prof.Y.Miyashita and Prof.S.Sasaki and the kind assistants of Mr.H.Yoshida, Mr.M.Sakai, Mr. Y.Shiraiishi, Mr. O.Yaegashi, Mr. Y.Hasegawa, Mr. M. Kimura, and Dr. L.Shi are gratefully acknowledged.

References

1. Kalman, R.E. (1960) *Trans.ASME Ser.D:J.Bas.Eng.* **82**: 35
2. Seelig, P.F. (1976) *Anal. Chem.* **48**: 252
3. Poullisse, H.N.J. (1979) *Anal. Chim. Acta* **112**: 361
4. Rutan, S.C., Brown, S.D. (1984) *Anal.Chim.Acta* **160**: 87; (1985) *Anal.Chim.Acta* **167**: 39
5. Brown, S.D. (1986) *Anal. Chim. Acta* **181**: 1
6. Rutan, S.C. (1987) *J. Chemom.* **1** (1): 7; (1989) *Chemom.Intell.Lab.Syst.* **6**: 191
7. Li,Z.,Shi,L.(1987)*Proc.8thICCCRE*D-16; Li,Z.,Shi,L.,Li,M.,Yu,R.(1988)*Proc.2nd Conf.Comput.Anal.PL5*,pp1-18
8. Li,Z.,Shi,L.(1988)*Environ.Sci.* **9**: 56; Li,Z., Li,M, Shi,L, Lin,H, Yu,R. (1990) *Acta Chim.Sinica* **48**: 1101
9. Li,Z., Yu R.(1990) *Acta Chim.Sinica* **48**:1018; Li,Z., Li,M., Shi,L., Yu,R.(1990), *Chem.J.Chin.Univ.* **11**: 245
10. Li,Z., M.Li, L.Shi, R.Yu (1990), *Chem.J.Chin.Univ.*, **11**: 825
11. Li,Z., Qian,F., Shi,L., Li,M., Yu, R. (1990) *Acta Pharmac. Sinica* **25**: 35
12. Li,Z.; Yu,R.; Shi,L.; Xu,J.; Zhang,M.; Wang,Q. (1991) *Anal.Chim.Acta*, **248**: 257
13. Hayashi,Y., Shibusaki,T.,Matsuda,R.,Uchiyama,M.(1987) *Anal.Chim.Acta* **201**: 185, **202**: 187
14. Vanslyke,S.J., Wentzell,P.D. (1991) *Anal.Chem.* **63**: 2512
15. Li, Z., Miyashita,Y., Sasaki, S. (1992) *CAMSE'92*, **2nd**: H-4
16. Li, Z., Miyashita,Y., Yoshida,H., Sasaki,S. (1992) *15CICS&20SAR*, **26P015**: 122; *Anal.Chem.*, submitted.
17. Li, Z., Miyashita,Y., Yoshida,H., Sasaki,S. (1992) on FF, PF, NF and PFF, *unpublished*,
18. Kaplan,S.A., Weinfeld,R.E., Abruzzo C.W., Lewis,M. (1972) *J.Pharmac.Sci.* **61**: 773
19. So, S.-S., Richards, W.G. (1991), *J.Med.Chem.* **35**: 3201

部 会 行 事

第 16 回 情報化学討論会

主催 日本化学会情報化学部会

共催 日本化学会、日本薬学会、日本農芸化学会、日本分析化学会、
日本プログラム交換機構

日時：11月24日(水)～26日(金)

会場：徳島大学大塚講堂、長井記念ホール、歯学部(徳島市庄町1丁目78-1)。

【交通】JR蔵本駅下車 徒歩5分、JR徳島駅下車 バス15分

なお、本討論会は第21回構造活性相関シンポジウムと併催されます。

討論主題：A. 化学情報, B. 計算化学, C. ケモメトリックス, D. 化学用ソフトウェア
E. ラボラトリーオートメーション

講演申込締切：7月9日(金)(郵送の場合)、7月23日(金)(電子メールの場合)

今回、従来の郵送による申し込みと電子メールによる申し込みを受け付けます。

学会等のデータベース化に伴い、なるべく電子メールでの申し込みをお願い致します。

【講演申込】〒101 東京都千代田区神田駿河台1-5 化学会館内 情報化学・構造活性
委員会 講演申込係 電話(03)3292-6190(担当：石鍋、清水)、FAX(03)3292-6318
電子メール：RGD03172@niftyserve.or.jp

【郵送の場合】1件ごとにA4版(縦)の用紙に横書きで次の事項を記入のうえ、上記
講演申込先宛てお送りください。(a)部会名：情報化学、(b)演題、(c)発表者と所属(講演者
に○印、全員にふりがな、所属が複数の場合は誰がどこの所属かわかるように記入してく
ださい)、(d)キーワード(討論主題を含め全角40字、10個以内)、(e)発表形式：口頭、
ポスター(申し込み件数が多い場合には、ポスター発表に切り換えて下さるようお願いす
ることがあります)、(f)連絡先。

【電子メールの場合】郵送の場合と同様の事項を電子メールにてお送りいただきますが、
詳細資料を上記講演申し込み先または下記問い合わせ先宛てご請求下さい。

講演要旨締切：9月18日(土)

講演1件についてA4版用紙4枚。他に半ページの英文要旨を添付して下記宛てお送りく
ださい。記入要領の詳細は申し込みを受理したのち送付します。

懇親会：11月25日(木)18時から。徳島大学蔵本学生会館にて第21回構造活性相関
シンポジウムと合同で開催します。会費：予約5,000円、当日6,000円

参加登録費(構造活性相関シンポジウムと共通)：

予約(一般)6,000円、(学生)4,000円。当日(一般)7,500円、
(学生)5,500円。講演要旨集(構造活性相関シンポジウム要旨集と合本)1冊を
含みます。講演要旨集の前送を希望される方は送料として1,000円追加。

参加登録、懇親会予約、要旨集前送申し込み締め切り：11月5日(金)

参加申込方法：A4版(縦)用紙に横書きで、氏名、勤務先、住所(TEL、FAX)、学生は
指導教官名、懇親会出欠、要旨集前送の有無をご記入の上、下記宛てお送りください。

銀行口座：銀行名 第一勧業銀行徳島支店

口座番号 普通預金 645-1431070

名義 情報活性委員会

実行委員会：寺田弘(徳島大学薬学部)、月原富武(徳島大学工学部)

講演要旨、参加予約申し込み、問い合わせ先

770 徳島市南常三島町2-1 徳島大学工学部

月原富武、森本幸生(TEL&FAX 0886-55-3160 ダイヤルイン)

計算化学・情報化学サマーセミナー

主催 日本化学プログラム交換機構、日本化会情報化学部会

日本化会情報化学部会では毎年夏期セミナーを開催してきましたが、今年は日本化学プログラム交換機構（JCPE）との共催で、セミナー名も上記のようにして行うことになりました。内容は「分子軌道計算のあれこれ」と題し、分子軌道計算関係の先生を多数お招きし、分子軌道計算の色々な面からの講演会を企画しました。分子軌道計算について深く勉強したいという方には最適の機会と思います。ふるって御参加ください。

- 日時 7月13日（火）～7月15日（木） 2泊3日
場所 栃木県鬼怒川温泉・ホテルたかはら
主題 「分子軌道計算のあれこれ」
講師 （お茶大理）平野恒夫、（広大理）今村詮、（阪市大理）西本吉助、（九工大情報工）柏木浩、（埼大工）時田澄男、（横国大教）永瀬茂、（兵庫教大教）足立裕彦、その他（依頼中を含む）
定員 100名（先着順、満員になり次第締切）
参加費 情報化学部会々員（含、JCPE個人会員）20,000円、日本化学会々員（含、情報化学部会法人会員、JCPE法人賛助会員）30,000円、その他50,000円（宿泊費等一切を含む）
申込方法 下記申込書（コピーしてお使いください）に記入の上、郵送またはFAXにてお申し込み下さい。
申込および問い合わせ先
〒113 東京都文京区弥生2-4-16 学会センタービル 化学情報協会内
日本化学プログラム交換機構 電話 03-3816-3581,3582 FAX 03-3816-7826

計算化学・情報化学サマーセミナー参加申込書

住所	〒
所属	
氏名	
電話番号	
会員種別	情報化学部会々員・日本化学会々員・その他（いずれかの分類に○）

注意：参加申込を受け付けた方には請求書をお送りしますので、それから参加費用を振り込んで下さい。

情報化学部会ワークショップ

第一線若手研究者によるコンピュータケミストリー活用事例 ～ドラッグデザイン分野での有効活用例とその課題～

主催：日本化学会情報化学部会

共催：日本化学会・日本薬学会（予定を含む）

目的：計算化学・情報化学の研究活動への利用が盛んな製薬企業の若手研究者を中心に、実際の研究分野での創薬研究等への適用例を紹介して頂き、更に、手法・方法論・環境など、現状抱えている課題を提起するとともに、今後の夢を語って頂き、参加者との意見交流の場を持つワークショップを企画致しましたので、振るってご参加下さい。

世話人：三戸邦郎（三井東圧化学）、藤原巖（大日本製薬）、宮下芳勝（豊橋技科大）

日時：平成5年9月7日（火）10：00—18：00

会場：（社）日本化学会 化学会館

（〒101 東京都千代田区神田駿河台1-5 TEL 03-3292-6163）

定員：50名

プログラム：

「リードジェネレーションを目的とする新しい計算機手法

—DE NOVO DESIGNを中心に— 田辺製薬(株)・西端 芳彦氏

「PLS法とニューラルネットワークによるQSARモデリング」

興和(株)・長谷川 清氏

「COMFA法における統計学的基礎 —PLSとその周辺技術—」

大日本製薬(株)・富永 幸雄氏

「COMFAの利用を中心としたドラッグデザイン、実用性と問題点」

住友製薬(株)・西村 民樹氏

「COMFA法によるリセプターマッピング」中外製薬(株)・大田 雅照氏

参加費：無料（個人部会員）、3000円（非個人部会員）当日徴収

懇親会：講演終了後、会費2000円（予定：会員・非会員を問わず）当日徴収

参加申込み：名前、所属、連絡先、電話&ファクス番号、部会員か否かを明記の上、郵便又はファクスで、下記宛に、8月20日（金）まで。

〒247 横浜市栄区笠間町1190番地 三井東圧化学(株) 総合研究所

計算科学室 三戸邦郎（TEL 045-891-0317(ダイヤルイン), FAX 045-893-2643）

平成5年度情報化学部会の課題

大澤映二（豊橋技術科学大学）

この場所を借りて問題提起を行いたい。昨年のCICSJ第2号でも述べたが情報化学部会の会員数がここ数年600名台で推移していて横ばいである。これは日本化学会会員3.8万人中の2%弱に過ぎない。部会員数を増加させて会の存在意義を明らかにせよという日本化学会側からの強い要請がある。年会における情報化学のセッションに対する一般講演の申込応募がいつも少なく、研究発表だけではセッションが成立しないことは周知の事実である。また肝心の情報化学討論会への参加者も減っている。明らかに難問を抱えている。どうしたら良いかを考えてみたい。

まず、情報化学とは何であったかという問いに戻ってみよう。千原秀昭先生の定義[1]をお借りすると情報化学とは『情報を化学の研究に利用する技術を研究する化学』である。どちらかと言えば縁の下の力持ち的な性格を持ち地味ではあるが、この時代には必要で非常に重要な分野であることには間違いない。研究テーマは山積している筈である。情報化学の研究活動がもっと活発になる事が望ましいのは明らかである。何故停滞しているのか？

情報化学者が怠けているのだろうか？まず、かく言う私がこの問いに答えるべきであろう。しかし、私は情報化学の研究を行っていない。私の研究領域は計算化学である。それでは何故一見畑違いの私が部会長を仰せつかっているのだろうか？計算化学は情報化学の一部なのだろうか？千原先生は同じ文章の中で一つの術語が様々な意味を持つ例として情報化学を挙げ、情報化学が包含する分野としてつぎの5つを説明しておられる。

- (1) 文献検索、化合物構造検索
- (2) データベース
- (3) 物性推算システム、QSAR
- (4) 総合的分子設計システム
- (5) 計算化学

確かに最後に計算化学が入っている。実際、本部会には計算化学者がかなり加わっている。しかし、常識としては計算化学はむしろ理論化学や物理化学に近く、情報化学に含めるのは無理である。ア

アメリカ化学会でもDivision of Chemical Information (CINF)とDivision of Computers in Chemistry (COMP)は別々の部会になっていて、前者がいわゆる情報化学、後者が計算化学を受け持っているようである。日本の情報化学の中に計算化学が入っているのにはわけがあって、かつて計算化学者が極端に少なくて研究発表の場所が無かった頃、同じくコンピュータを道具として使う同業化学者として情報化学部会が我々を暖かく迎えてくれたので、今でも好意に甘えて居候をつづけているという独特の事情によっている。

私が情報化学の研究を行わない理由を長々と述べたが、情報化学部会全体としては計算化学の混在はどのように受け止められているだろうか？情報化学部会は他の二部会（コロイドと生態関連）に比べて化学産業界との関連が密接であり、事実相当数の部会員が企業におけるコンピュータ担当者である。いわゆるコンピュータエキスパートの業務はおよそコンピュータに関すること全般を包含する事が多く、実際には文献情報やデータベース以外に研究がらみの計算化学業務を抱え込んでいる事が多い。この意味では本部会に化学情報と計算化学が混在しても矛盾は無いと言えよう。

ところで日本化学会の目的が第一に化学研究の推進にあるのと同じく、情報化学部会の存在理由もまさに研究にあることを考えると、研究の低迷をこのまま放置しておくわけにはいかない。研究を盛んにする方策は昔から一つしかない。優れた、できれば若い研究者を引付けることである。計算化学に関して言えば、例えば分子シミュレーション討論会などで活発にMD、MCなどの研究発表を行っているような若手を本部会に引き入れたい。情報化学で言えば、情報科学会などに出ている自然科学系のキレル若手に本会に入ってもらいたいと思うであろう。

ただし、計算化学側から見ると情報化学部会という名前では、計算化学に興味はあるが情報化学には関心がないという人（とくに大学の研究者にこういう人が多い）を誘うのはためられる。むしろ名称を変更して欲しいと考える。一方、情報化学プロパーの研究者にとっては現在の名称は捨てがたい魅力があるだろう。しかし日化65春季年会の「二十一世紀に情報化学シンポジウム」においてパネル討論会でこの問題を論議したときには情報化学側から積極的な対案の提案はなかった。

「名は体を現わす」という言葉が示すように部会の名称は大切な問題である。情報関連研究発表の低迷、部会員の方の興味との乖

離という問題に対して、この一年間部会会員全員で解決策を考えようではないかというのが今回の提案である。折りを見てアンケートの実施等を常任委員会に計りたいと考えているが、良い考えがあれば是非聞かせて頂きたい。

[1] 千原秀昭、近畿化学工業界、1992年12月号、6～8頁。

情報化学部会収支報告書

(平成4年3月1日～平成5年2年28日)

[収入の部]

(単位：円)

科 目	予 算 額	実 績 額	差 異
部 会 費	4,685,000	3,821,610	△ 863,390
正部会員 (会 員)	(940,000)	(1,000,243)	(60,243)
正部会員 (非会員)	(120,000)	(117,533)	(△ 2,467)
学生部会員 (会 員)	(20,000)	(31,834)	(11,834)
学生部会員 (非会員)	(5,000)	(2,000)	(△ 3,000)
法 人 部 会 員	(3,600,000)	(2,670,000)	(△ 930,000)
事 業 費	250,000	595,000	345,000
広 告 掲 載 料	550,000	700,000	150,000
合 計	5,485,000	5,116,610	△ 368,390

[支出の部]

(単位：円)

科 目	予 算 額	実 績 額	差 異
B u l l e t i n 費	2,280,000	2,621,907	341,907
事 業 費	1,000,000	523,562	△ 476,438
会 議 費	752,000	689,006	△ 62,994
事 務 費	650,000	565,595	△ 84,405
事 務 負 担 金	803,000	803,000	0
合 計	5,485,000	5,203,070	△ 281,930

収 支 差 額	0	△ 86,460	△ 86,460
---------	---	----------	----------

第65回日本化学会春期年会（東京）記事

今回の春の年会は、日本大学の桜上水校舎で行われ、化学情報・計算機化学のセッションは3月28日のB3会場で、招待講演1件、普通講演10件に特別企画として「21世紀に向けての情報化学」というタイトルで化学会欧文誌の電子出版化についての専門委員の方々による経過報告と現状の紹介があり、その後パネル討論、総会という順の運びとなった。今年からの年会は種々新しい方針で始めることとなり、要旨の形式そのほか多くの新機軸を盛り込んだものの、まだプログラム編成側の意図を完全に実現するにはまだまだ難しい条件が山積している状態であり、今後時間を掛けて部会の方針に添うような企画を可能とするようつとめて行かなくてはなるまい。大体、プログラム編成に際して、折角の部会があるのにこれを飛び越して、直接担当者の選定が行われるというのはいささかならず奇妙である。当方は一セッションの内容（時間の割り振り、特別企画等）を全部部会に委任されるぐらいの度量を年会実行幹部に期待したのだが、諸般の事情で今回はそこまでは出来ないとのことであった。たとえば以前化学教育部門などで全部の講演をB講演扱いにして充分のディスカッションの時間をとった例もあったのだから、もう少し部会の独自の特色を織り込めるようなプログラムを作れるようになることが将来は望まれる。

今回の招待講演は、住友化学特許室の植田賢三氏にお願いして「化学物質と特許の現状にまつわる諸問題」というタイトルで30分間のお話をお願いした。年会の席上では現在までにあまり取り上げられた機会の少ない話題であったためか、「目新しい話題で興味深かった」という感想を述べられた参会者が何人か居られたが、やはりこの分野は大事な部門でもあり、今後も機会ある毎に取り上げるべきであろう。次の機会の特別企画に織り込んでよいかと考える。30分はいささか短すぎ、もっと時間をとって実のあるディスカッションができればよかったと考える。

10件の普通講演のタイトルは下記のようなものである。

- 1 B 3 0 7 : 多核NMRデータベース作製の研究 : Nb-93 (電通大・横浜市大) 山崎 昶、望月 桂
- 1 B 3 0 8 : PostScript対応化学エディタの開発 (豊橋技科大) 中野征司、高橋由雅
- 1 B 3 0 9 : 二重結合を奇数個有するポリエンの励起三重項状態 (筑波大化) 高橋 央宜、渡辺正行、菊池 修
- 1 B 3 1 0 : 分子静電ポテンシャルを用いた金属イオン捕捉位置の解析 (筑波大化) 山口 薫、菊池 修
- 1 B 3 1 1 : モリブデンクラスター錯体、 $\text{Mo}_6\text{X}_8(\text{PH}_3)_6$ (X=S, Se)のD.V-X α 法用いた分子軌道計算 (東大理・京大工) 井本英夫、足立裕彦
- 1 B 3 1 2 : 化学物質の分解性の予測 (3) ハロゲン化炭化水素の加水分解速度の予測 (化技研) 田辺和俊、都築誠二、杉江正昭
- 1 B 3 1 3 : 分子道力学法による流体の物性予測—MM2ポテンシャルの適用

化学物質と特許の現状にまつわる諸問題

住友化学工業(株)
特許室 植田賢三

1. はじめに

化学物質の発明は、特許法上は物の発明のカテゴリーに入り、その特許権の独占的効力は、その物質の生産、使用、譲渡（販売を含む）、貸与、譲渡又は貸与のために展示し又は輸入する行為に対して及ぶ。すなわち、これらの行為の中には産業上のあらゆる活動が含まれていると考えてもよく、非常に強い権利になる。これに対して、方法の発明ではその方法の使用が発明の実施とされるにすぎない。したがって、いわゆる物質特許制度の導入については、発展途上にある国々は自国の技術レベルとの関係上慎重にならざるを得なかった。我が国も1976年1月1日ようやくこの制度をとり入れたにすぎず、まだ17年の歴史しかもっていない。しかし近年、知的財産権（知的所有権とはほぼ同意）保護の世界的潮流の高まりと共に、世界の主要国ではインド以外殆んどどの国が物質特許制度を採用するに至っている。ブラジル、ブルガリアでも改正が予定されている。

ここで最近の特許を中心とする世界的な動向について簡単にふれる。

1980年代に入り、それまで独占禁止法の適用を強化してきた米国が、自国の産業を発展させる為に知的財産権を尊重する政策に転じ、これを契機に世界的にプロ・パテントの潮流が波及することになった。米国は、米国内はもとより、二国間又は多国間協議によって国際的にもその政策を推進している。また、WIPO（世界知的所有権機関）やGATTの場においても特許法等の運用を含めた国際的な調和が検討されてきており、懸案になっていたWIPOの特許調和条約もここ1～2年のうちには成立するメドがついてきた。その中には先願主義、出願の公開、特許期間の出願日からの起算等が明記されている。

一方、国内では特許法の改正や審査基準の改訂が進められており、基本的な発明に対しては広い権利が取得し易くなる方向、即ち、付与される特許の範囲が、明細書に書かれている実施例に基づいて決められる傾向の強かった、従来のいわゆる実施例中心主義的な審査に訣別して、実施例に拘泥することなく、特許請求の範囲および明細書の記載全体に即して、付与される特許の範囲が決められる欧米流の審査への移行が進められている。従って、今後は特許的にも益々創造性の高い研究が奨励されると共に、権利の強力な物質特許が一層重要となる時代になってくることが確実である。

以下、化学物質を巡り特許的にみて興味のある点について、主として侵害の観点から簡単に述べてみたい。

2. 物質の特定

物質特許の取得に際しては、特許明細書、中でも権利の対象となる特許請求の範囲（一般に「クレーム」という）のなかで、その新規化学物質を化学物質名や構造式により、あるいは物理・化学的性質により明確に特定しておくことは重要である。いわゆる実施可能要件、

即ち当業者（その分野の普通のレベルの研究者等）が容易に実施できるような記載が明細書には要求される。大切なことは、物理・化学的性質により特定されている物質の場合には、その分析方法をも明確にしておく必要があるということである。誰がどのような方法で行ってもいわゆる実験誤差程度しかないようなポピュラーな分析は別として、分析方法によって値がかなり異なってくる場合や、まだ分析方法が一般に知られていないような場合には、必要に応じて使用した機器の製造会社名を含め、明確に分析方法を記載すべきである。

高等裁判所まで争われた事件（'84判決）に、「10%以下のビニル含有量の1,4-ポリブタジエンの存在下でモノビニル芳香族単量体を重合する（要部のみ）」特許について争われた、いわゆる「耐衝撃性ポリスチレン事件」がある。ここではビニル含有量の分析方法が争点となったが、明細書には分析方法については全く記載がなかった。そして出願当時、赤外法は広く行われていたが、NMR法は知られていなかった。被告は赤外法での分析値11.4%を根拠に非侵害を主張し、原告はNMR法での分析値8.7%を根拠に侵害と主張した。結局第一審の東京地裁では、「出願当時当業者が容易に実施でき、普通に用いられていた方法である赤外法を採用すべき」とし、侵害なしとした。

控訴審の東京高裁では、「出願当時客観的にビニル含有量を測定する方法がなかったのだから、分析方法が示されていない本件では分析値を定めようがなく、その点で本発明を実施することが不可能であった」として、控訴を棄却した。

3. 不純物

化学製品には不純物として種々の物質を含む可能性がある。一般に、工業的規模での精製は、技術的にまた経済的な観点からも困難な場合が多く、不純物が他者の特許の権利に含まれる場合も起こりかねない。

東京地裁（'67判決）でこのような事件が争われた。「2,4-ジクロロフェニル-4'-ニトロフェニルエーテル(2,4体)を有効成分として含有する除草剤組成物」なる特許に対して、「2,4体を不純物として0.1~0.3%含む2,4,6-トリクロロフェニル-4'-ニトロフェニルエーテルを有効成分とした除草剤」が侵害かどうか争点であった。なお、本特許は既知物質の用途発明的なものであった。

判決では、「産業上実用に供しうるものが当然の前提となるから、有効成分量の下限をはるかに下回る成分を含むにすぎない除草剤についてはその技術的範囲に含まれない」として、侵害との訴えをしりぞけた。

妥当な判決と思われるが、物質特許の場合には微量であっても含まれていれば侵害になるとの意見もあり、特に欧米ではそのように考える特許弁護士も多く、ケース毎に慎重な検討が求められる。

4. 試験研究

特許法では「特許権の効力は、試験または研究のためにする特許発明の実施には、及ばな

い。」と規定されているが、ここで言う試験・研究とは純粹に技術の進歩を図ることを目的とするものであり、利益の追求がその目的となっているような場合には、上記試験・研究には該当しないものと考えられている。昨今、産業廃水の分析等の分析を業とする会社も多くなってきているが、そのような会社で、特許方法による分析を顧客の為に行なうことは、当然のことながら特許の侵害行為となる。

東京地裁（'87判決）で除草剤“グリホサート”を含む特許をその類似品の実施が侵害するかどうか争われ、結果は侵害となった。それに附随して、その類似品の農業登録申請の為の試験は前記の特許法で規定する試験・研究には当たらないとして、差止請求が認められた事件がある。一般的に、販売等の実施を前提とする段階の製品開発の為の試験・研究には注意が肝要である。

化学物質ではないが、この項との関連で注意を要するのは、最近発展のめざましいバイオの分野である。例えば遺伝子を微生物に導入する方法の発明の場合、一旦その特許方法を使用して目的とする組換え微生物を作成した後は、得られた微生物を増殖することが出来る。したがって実験室でその特許方法を一度実施すれば、その後の工業的实施、すなわち、その組換え微生物の増殖段階ではその方法特許を使用することはない。しかし、実験室での一回のみの使用であっても、その実施が商業的な製品の生産を目的とする等の利益を得ることを目的としたものである場合には「試験・研究のための特許発明の実施」には該当しないとする意見が多い。

5. 中間体と最終製品

中間体にのみ他者の特許がある場合には、その特許を実施しなければ最終製品の販売等は基本的には問題とならない。しかしその特許を実施した場合には、たとえ他者の最終製品に関する特許がなくても侵害となる。それが物質特許であれば製造方法の如何にかかわらず、その物質を中間体として経由していれば侵害となり、その中間体の特許が製造方法のみであれば、その製造方法を使用しなければ問題はない。

注意を要するのは間接侵害（特許侵害の前段階の行為）を構成する場合である。ある最終製品に他者の特許があり、その中間体には他者の特許がなかったとしても、その中間体の販売がその最終製品の特許の侵害とされる場合がある。それは特殊なケースではあるが、その中間体はその最終製品の製造にのみ用いられるものであって、他に用途がない場合である。

先に、方法の特許はその方法を使用すること自体が発明の実施とされるにすぎないと述べたが、一つ注意を要する事項がある。それは、物を生産する方法の場合には、その方法によって生産された物の使用、譲渡、貸与、譲渡又は貸与のための展示、又は輸入が実施行為になるという点である。先進国ではおよそ同様な規定を持っているので、特に輸入には気をつけなければならない。

例えば最終製品（販売される製品）の製造方法に関する特許が米国にはあるが日本にはな

いとす。従って日本で、その特許方法によりその製品を製造し販売等を行うことは全く法に触れない。しかしその製品を米国に輸出すると米国ではその特許を侵害することになる。

日本や西欧主要国では、その特許方法で直接最終製品が製造されていなければほぼ特許侵害とはされないものと判断される。しかし、米国ではもう少し拡張的に解されている。すなわち、特許方法によって製造された物と実質的に変わっていない物の販売や輸入等はその方法特許の侵害になると考えられる。米国がこのような法律改正を行なったのは1989年2月のことであり、その運用（判決例）が未だ必ずしも明確になっていないので、最終製品が特許方法によって製造されたものと非常に近い場合には充分の注意を払う必要がある。

6. コンピューター・ケミストリーの成果と特許性

コンピューター・ケミストリーはその進歩と相まって研究上の有効な武器となりつつあり、それを道具として使った研究成果も産み出されつつあるように思われる。ここでは、その成果と特許性について私見を述べてみたい。但し、その成果自体は特許性があるとの前提で考えることとする。

まずそのプログラム自体の自律性、即ち研究者が自らの頭脳をどの程度使ってプログラムと協同作業するかが、その成果の特許性（特に進歩性：当業者が容易に思いつかないこと）に関わってこよう。例えば、極端な例として、反応設計の場合を考えてみると、目的化合物を入力すればごく少数、例えば2～3個の答が出力されるとすれば、そのうち一つのものが発明性があったとしてもその成果はその研究者が行なった発明とは言えないのではなかろうか。この場合、発明者はコンピューターということになろう。他方、研究者が自分の知識を駆使して発明を完成していく中で、その一つの道具としてプログラムを使ってある成果に到達し得たような場合、即ち、そのプログラムを使う研究者によって答は全く異なってくるような場合には、その発明は研究者によってなされたと言えるのではなかろうか。

ここで問題になってくるのがそのプログラムの公開性である。プログラムの自律性の高い前者の場合、そのプログラムが一般に市販されていたらどのようなことになるか。おそらく最終的にはその出願は拒絶される可能性が非常に高くなるであろう。他方、後者の自律性の低い、即ちそれを使う研究者の発明力の影響の大きなプログラムの場合には、その公開性は問題にならないであろう。自律性の高い優れたプログラムが市販されるような時代がくると、それが特許性を否定する材料の一つとして使われる可能性も高くなってこよう。言い換えれば、そのプログラムに打ち克つような創作でなければ発明ではないとされる時代が来ると言えそうである。

最後に、我が国では先にも述べたように、いわゆる実施例中心主義はなくなる方向ではあるが、化学分野のように、実験してみなければ結果がわからないといわれているような分野では、少なくとも一つの実施例の記載は依然として必要であることを附言しておく。

以上

第 2 1 回 構造活性相関シンポジウム

主催 日本薬学会医薬化学部会

共催 日本薬学会中四国支部、日本化学会、日本農芸化学会、
日本農薬学会、構造活性相関懇話会

日時：11月24日（水）～26日（金）

会場：徳島大学大塚講堂、長井記念ホール、歯学部（徳島市庄町1丁目78-1）。

〔交通〕JR蔵本駅下車 徒歩5分、JR徳島駅下車 バス15分

なお、本討論会は第15回情報化学討論会と併催されます。

講演申込締切：7月9日（金）（郵送の場合）、7月23日（金）（電子メールの場合）

今回、従来の郵送による申し込みと電子メールによる申し込みを受け付けます。

学会等のデータベース化に伴い、なるべく電子メールでの申し込みをお願い致します。

【講演申込】〒101 東京都千代田区神田駿河台1-5 化学会館内 情報化学・構造活性
委員会 講演申込係 電話(03)3292-6190（担当：石鍋、清水）、FAX(03)3292-6318
電子メール：RGD03172@niftyserve.or.jp

【郵送の場合】1件ごとにA4版（縦）の用紙に横書きで次の事項を記入のうえ、上記
講演申込先宛てお送りください。(a)部会名：構造活性、(b)演題、(c)発表者と所属（講演者
に○印、全員にふりがな、所属が複数の場合は誰がどこの所属かわかるように記入して
ください）、(d)キーワード（全角40字、10個以内）、(e)発表形式：口頭、(f)連絡先。

【電子メールの場合】郵送の場合と同様の事項を電子メールにてお送りいただきますが、
詳細資料を上記講演申し込み先または下記問い合わせ先宛てご請求下さい。

講演要旨締切：9月18日（土）

講演1件についてA4版用紙4枚。他に半ページの英文要旨を添付して下記宛てお送りく
ださい。記入要領の詳細は申し込みを受理したのち送付します。

懇親会：11月25日（木）18時から。徳島大学蔵本学生会館にて第15回情報化学討論
会と合同で開催します。会費：予約5、000円、当日6、000円

参加登録費（情報化学討論会と共通）：

予約（一般）6、000円、（学生）4、000円。当日（一般）7、500円、
（学生）5、500円。講演要旨集（情報化学討論会要旨集と合本）1冊を含みます。

講演要旨集の前送を希望される方は送料として1、000円追加。

参加登録、懇親会予約、要旨集前送申し込み締め切り：11月5日（金）

参加申込方法：A4版（縦）用紙に横書きで、氏名、勤務先、住所(TEL,FAX)、学生は
指導教官名、懇親会出欠、要旨集前送の有無をご記入の上、下記宛てお送りください。

銀行口座：銀行名 第一勧業銀行徳島支店

口座番号 普通預金 645-1431070

名義 情報活性委員会

実行委員会：寺田弘（徳島大学薬学部）、月原富武（徳島大学工学部）

講演要旨、参加予約申し込み、問い合わせ先

770 徳島市南常三島町2-1 徳島大学工学部

月原富武、森本幸生(TEL&FAX 0886-55-3160 ダイヤルイン)

機械翻訳入門

図書館情報大学 石塚英弘

1. はじめに

機械翻訳とは英語のmachine translationの訳で、ある言語で書かれた文章を機械（実際にはコンピュータ）を使って別の言語の文章に翻訳することである。

こう言うと、機械翻訳の技術は既に完成し、その機械も実存すると思われるかもしれない。しかし、そうではない。機械翻訳は人間の素朴な願望から生まれた『見果てぬ夢』のようなものであり、その実現のために今も多くの研究開発が盛んになされている。そして、人間が行う「翻訳」では無意識のうちに如何に精妙な操作がなされているかが次第に明らかになってきた。その過程で得られた知見は、人間が使っている言語、これをコンピュータ科学の分野では特に「自然言語」というが、これをコンピュータで扱う研究、即ち「自然言語処理」の研究に進歩をもたらした。翻訳についてその機構が明らかになったことは学問的成果であるが、それとともに翻訳を行う人間の知的活動の素晴らしさを再認識することにもなったのである。

蛇足になるが、機械翻訳は何となく斜に見られることがある。翻訳結果が悪ければ「やはり、こんなものは役に立たないのだ。」と言われ、良ければ「機械が人間より巧くできる筈がない。巧くいかない例を考えてやろう。」となる。これは機械翻訳に限らず、人工知能研究の成果に対して、よく起こることである。どうも、人間の知的能力に係わることを機械にやらせようとすると、こう言われるようだ。自然界の機構を明らかにすることは自然科学の本来の目的であり、それは学問として歓迎される。しかし、人間の知的活動の機構を研究対象とすると、歓迎しない人が出てくるようである。閑話休題。

本稿では、機械翻訳の基本的な仕組みと原理、そしてその限界とそれを越える試みなどについて簡単に紹介する。機械翻訳の解説を本格的に書けば一冊の本になってしまう。実際、本もいくつか出版されているので、詳しくはそれらを読んでいただくとして、ここでは機械翻訳の概略を入門的に解説することにしたい。

2. 機械翻訳の仕組みと原理

機械翻訳という考え方は17世紀からあったという。各言語間を仲介する数値コードを考えることが流行し、大数学者かつ哲学者の Leibnizも、彼の哲学理論、单子論の中でその提案をしている。しかし、実際に翻訳するシステムの出現は、20世紀後半のコンピュータによる全く別の原理に基づいた機械翻訳の試みまで待たねばならなかった。

では、現時点における機械翻訳の基本的な仕組みと原理、そして、その仕組みを採った理由を述べることにしたい。

2.1. 中間表現を介した間接翻訳：中立言語方式と変換方式

機械翻訳では、元の文章を源文(source text)、訳文を目標文(target text)といい、元の言語を源言語(source language)、訳文の言語を目標言語(target language)という。英語から日

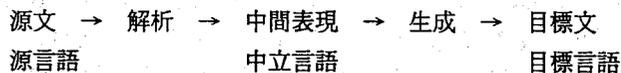
本語への翻訳ならば、源言語は英語、目標言語は日本語、源文は英語の文章、目標文は日本語の文章となる。

人間が翻訳する場合は源文を観て直接訳文を書く。しかし、コンピュータで翻訳する場合は、源文を目標文に直接変換するのではない。源言語と目標言語の間を取り持つ表現形式（中間表現）を設定し、源文を解析して中間表現を作り、それから目標文を生成するという間接的な方法を採用。その理由は二つの言語の違いによって生じる翻訳上の困難を避けるためである。

日本語、英語、独語、仏語、ロシア語、スペイン語、イタリア語、中国語等々、世界にはいろいろな言語がある。日本語と英語は文の構造が大きく異なっているが、ラテン語系の言語たとえばスペイン語とイタリア語はよく似ている。人間が翻訳する場合を考えても、よく似た言語間の翻訳は容易であるが、違いの大きい言語間の翻訳は難しい。人間は、それまで培った知識や経験を基にして、総合的に翻訳作業を行う。一方、コンピュータに翻訳をさせる場合は、システムティックかつ厳密なやり方でやらなければならない。そこで言語間の違いを解決するために、中間表現を使うわけである。

中間表現を使う方式には、中立言語(pivot language)方式と変換(transfer)方式の2つがある。図1を参照されたい。中立言語方式とは、源言語の特徴にも目標言語の特徴にも依存しない中間表現を用いる方式であり、この中間表現を記述する言語のことを中立言語という。なお、pivot languageは中間言語と訳されることもある。一方、変換方式とは、源言語に依存する中間表現と目標言語に依存する中間表現の二つを持ち、前者から後者への変換を行うことで言語に依存した問題点を解決する方式である。

・中立言語方式



・変換方式

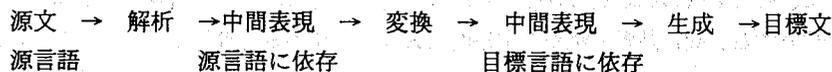


図1 中立言語方式と変換方式

あらゆる言語に対応できる中立言語が存在すれば、多言語間の機械翻訳が一つの中間表現を通してできることになる。しかし、厳密に中立言語と言えるものは未だ無く、各言語に依存する要素を無視できないのが現状である。一方、変換方式で多言語間の翻訳を行うと言語の数だけの中間表現とその間の変換が必要になる。実際の機械翻訳システムは基本的に中立言語方式か変換方式の一方を採用のもの、他方の要素も組み込んだ形にしているものが多い。

変換方式は、源文をどのレベルまで解析してから変換するかで、構文(syntax)レベル、意味(semantics)レベル、文脈レベル等に分かれる。図2を参照されたい。源文を構文解析して構文レベルの中間表現を得て変換するのが、構文レベルの変換方式である。構文解析だけでなく意味解析まで行って変換するのが意味レベルの変換方式であり、文脈解析までして変換するのが文脈レベルの変換方式である。

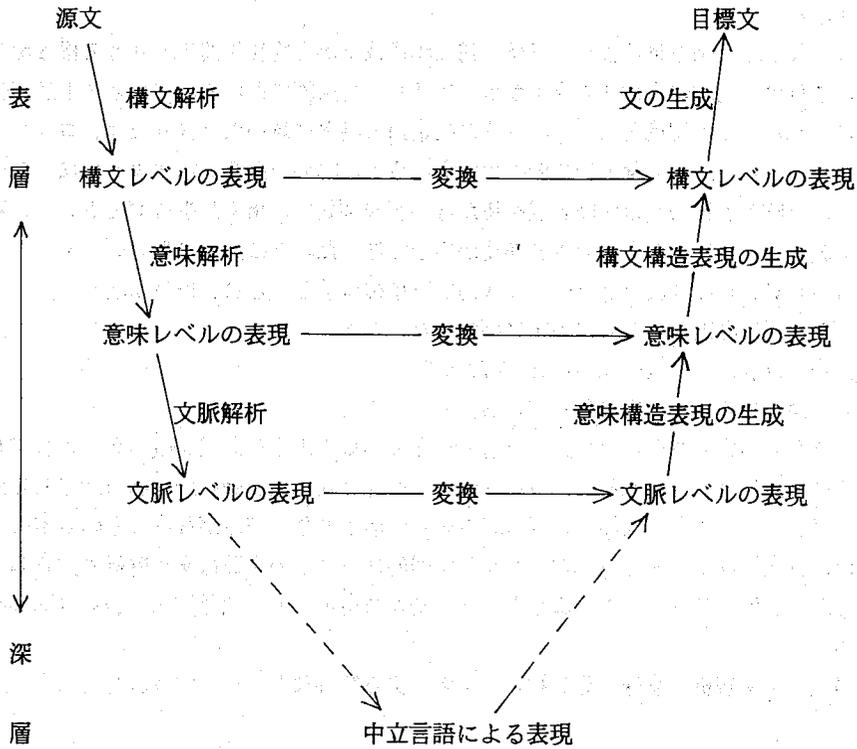


図2 機械翻訳の概念図

機械翻訳あるいは自然言語処理の分野では、「人間が言語で表現する場合は、まず頭脳の奥底で言語を越えた表現構造を構築し（これを「深層構造」(deep structure)という)、この表現構造を特定の言語の構造に従った表現構造（これを「表層構造」(surface structure)という）に変形して、書いたり言ったりする。また、言語を理解する場合は、文を解析して表層構造を作り、それを深層構造に変形して理解する。」という Chomsky の考え方がある。

この考え方に従えば、図2の上のレベルほど表層に近く、下のレベルに行くほど深層に近くなると言える。また、深層構造に近づくほど言語間の違いが少なくなるので、源言語に依存する中間表現と目標文依存の中間表現との違いも少なくなる。そのため、変換の寄与も減る。先に述べた中立言語方式では、中間表現は中立言語による表現1つだけであるが、この場合の中間表現は最も深いレベルであるから変換は必要ないと解釈することができる。

ところで変換方式の場合は、源言語と目標言語がよく似ているほど表層に近いレベルの変換で済むことになる。極端な場合、たとえば、東京弁と大阪弁のような方言の間であれば、構文レベルまで行かなくとも、単語レベルで済むと思われる。この場合、解析は源文の単語の認識（形態素解析）でよい。また変換は源文の単語を目標文の単語に置き換える操作でよく、その操作が生成処理も兼ねることになる。これならば簡単である。しかし、単語レベルの変換方式は源言語と目標言語の文法が全く同じで、単語だけが異なる場合しか有効でなく、そういう場合は殆どない。結局、変換方式では構文レベルは無論のこと、意味レベルもある程度は入っていないと使い物に

ならないのである。

使い物にならなかった例を挙げよう。実は、1950年代後半から約10年間アメリカで機械翻訳が試みられた。これは、1957年人類初の人工衛星がアメリカ合衆国ではなく、ソヴィエト連邦によって打ち上げられたことが契機となった。ソ連の技術力の高さに驚いたアメリカは、コンピュータによるロシア語から英語への翻訳システムの開発に莫大な額の補助金を出すようになったのである。しかし、当時の機械翻訳は中間表現を持たない直接翻訳で、構文解析はするものの、意味解析はせず、訳語対照表による単語の置き換えが中心であった。その方式で聖書の一節

"The spirit is willing, but the body is weak." (精神は強し、されど肉体は弱し。)

を英語からロシア語に翻訳し、さらに英語に翻訳したところ、

"The wodka is strong, but the meat is rotten."

(ウオトカ(酒)は強いが、肉は腐っている。)

がコンピュータから出力されたという。これ以外にも笑い話のような翻訳が幾つかなされてしまった。1964年に National Academy of Sciences は Automatic Language Processing Advisory Committee(ALPAC)を作ったが、ALPAC は「方式自体に問題があり、望みがない」とのレポートを出し、そのため資金援助は打ち切られた。しかし学問的には、この失敗は意味解析や文脈解析の必要性を強く示し、機械翻訳そのものよりも、その基本技術となる自然言語処理研究の進歩を促したのである。

つぎに、構文、構文解析、意味、意味解析、文脈、文脈解析などについて解説しておこう。

2.2. 構文・構文解析・文法

構文(syntax)とは文法的に見た文の構造であり、構文解析とは文を文法に従って解析し、構文を明らかにすることである。そのため、どのような文法を採用するかによって、構文(構文解析結果)の表現形式も異なることになる。

文法として何を採用するかを決定する要因としては、源言語への適合性と、コンピュータ処理への適合性が挙げられる。たとえば、文脈自由文法(context-free grammar)は英語には合っており、コンピュータ処理に適した文法であるが、日本語には合っていない。

英語は、主語となる名詞句の次に動詞句が来るし、名詞句の中では冠詞・形容詞・名詞の順となり、動詞句は動詞の次に目的語の名詞句が来るなど、句や単語の順序が決まっている。しかし、日本語では述語部が最後に来ることは決まっているが、それ以外の主語・目的語などの順序は自由である。文脈自由文法は句や単語の順序が決まっている場合に適しているので、英語には良いが、日本語には適していない。なお、英語には文脈依存性がある。たとえば、主語が三人称単数で時制が現在の場合動詞の末尾に"s"が付く。そのため、文脈自由文法では表しきれない部分がある。しかし、文脈依存文法ではコンピュータで処理できない。そこで、文脈自由文法の機能を一部拡張して文脈依存性を扱えるようにした文法が用いられる。また、英語には関係代名詞があるので、これを解析する機能も付加されている。これらの機能を持つ様々な文法が提案され、実際に用いられてもいるが、紙面の都合上これ以上の説明は割愛する。

日本語の文は、「係り受け構造」になっている。これは修飾・被修飾の構造のことであり、「係り」とは修飾語あるいは修飾句で、「受け」とは被修飾語あるいは被修飾句のことである。

構文解析した結果は木構造の形式で表す。これを構文木という。たとえば、「I read a book.」を文脈自由文法で解析した結果を図3に示す。文脈自由文法では、文は句で構成され、句は品詞または句で構成されている。この文の場合は現れていないが、動詞句のなかには前置詞句が現れることもある。前置詞句は前置詞と名詞句で構成される。文脈自由文法では句が構造上の重要な位置を占めているので、句構造文法の一つと見なされている。

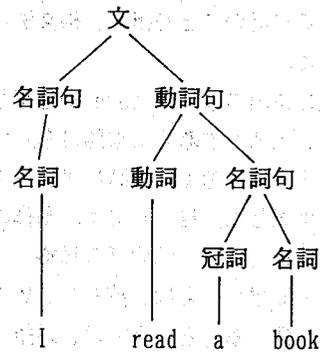


図3 文脈自由文法による構文木

日本語の文「昨日私は、母が買った本を読んだ。」の係り受け構造を図4に示す。

「昨日」と「私は」と「本を」は「読んだ」に係り、「母が」は「買った」に係り、「買った」は「本」に係る。

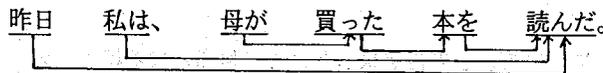


図4 係り受け構造の例

もしも、「私は昨日母が買った本を読んだ。」であれば、「昨日」が「買った」に係るのか「読んだ」に係るのかが曖昧となる。同様に「小麦色の帽子の女の子」では「小麦色の」は「帽子」だけでなく、「女の子」や、「女」に係るとする解釈もありうる。この種の例は、係り受け解析の難しさとともに、人間は意外に曖昧な文を書いていることも示している。

2.3. 意味・意味解析

セマンティクス(semantics)とは、一般には意味、または意味論のことであるが、ここでは意味を計算機処理可能な形式に構造化して表現したものをいう。また、意味解析とは構文解析結果を基に意味的に矛盾のない解釈結果のみを取り出して、構造化して表現することをいう。

なぜ、意味解析が必要なのか。構文解析のみでは、文法的には正しくとも意味的におかしい翻訳をする恐れがあるからである。たとえば、1950年代の機械翻訳の失敗例として有名なを示す例を挙げよう。「Time flies like an arrow.」『時は矢のように飛ぶ(光陰矢の如し)』について、likeを動詞、fliesを名詞の『ハエ』と解析し、Time fliesを複合名詞の『時バエ』と解釈した結果、『時バエは矢を好む』と訳したり、また、Timeを『時を計る』という動詞の命令形と判定し、『ハエの時を矢のように計れ』と訳したのである。2つの誤訳の内、前者はハエが矢を好むという常識に反する解釈をしている。また、後者は時を計る対象としてハエを許しており、文中の意味処理が不十分である。意味解析は上記の欠点を解決しようとするものである。

意味解析には格文法(case grammar)が使われ、解析結果はフレーム(frame)を使って表現されることが多い。格文法は次に述べるように、表層構造を表現する文法ではなく、意味が表現でき

よう考えられた文法である。また、frame は元来の意味は「枠」であるが、人工知能では知識を表現する枠組のことを指す。格文法はフレームと相性が良いので、この二つを組み合わせる用のである。

格文法は言語学者 Fillmore が考案した文法である。彼は日本語も含む色々な国の言語を研究した結果、主語を省略する言語はあっても述語の無い言語はないことに注目して、文の中の述語（動詞）を中心と考え、述語に対する文中の名詞（名詞句）の役割を格(case)と捉える文法「格文法」を考案した。格としては、動作の主体を示す「主格」(agent)、動作の対象を示す「対象格」(object)、手段を示す「道具格」(instrument)、場所を示す「場所の格」(location)などがある。ごく簡単にいえば、英語の5W1Hをより詳しくしたものである。たとえば、「太郎は鍵でドアを開ける」という文は、述語：開ける、主格：太郎、対象格：ドア、道具格：鍵、と解析される。同じ意味の英語の文 "Taro opens the door by the key." と "The door is opened by Taro with the key." も同じ解析結果となる。また、"The key opens the door." の場合も同様の解析結果となるが、主格は文中に明示されていないため、空白（おそらく人間）となる。このように、格文法は主語や目的語などといった表層の構造ではなく、意味を考えた構造を表現する。その点では、言語の違いを越えた文法といえる。

もっとも、格文法に基づいた解析を行うためには、

1) 述語（動詞）の辞書

各述語（実際には動詞）について、どの格が来るか、その目印となる格マーカー(case marker)は何かなどを書いたもの。

2) 名詞の辞書

各名詞について、その意味素性(semantic feature)を書いたもの。

の二つが必要になる。そして、いずれの辞書も計算機処理向きのものでなければならない。

例を挙げれば、「行く」という動詞の場合、主格、目的地を示す格、出発地を示す格などは来ても良いが、対象格が来るのはおかしい。どの格が来るかは動詞によって異なるのである。また、「行く」の目的地を示す格の格マーカーは助詞「へ」や「に」であるが、英語の"go"であれば前置詞"to"になる。そして、格マーカーとしては助詞や前置詞などといった文中に現れるマーカーだけでなく、意味的なマーカーも必要となる。たとえば、同じ主格でも「行く」の場合は人間でも動物でも良いが、「読む」の場合は動物ではおかしい。人間でなければならない。これら「人間」とか「動物」など意味的な素性を示すものを意味素性というが、格マーカーとしてこの意味素性も書いておかなければならない。

名詞の辞書の例を挙げれば、「太郎」については「人間」、「ドア」については「物」というように意味素性を与えることになる。

この二つの辞書を使って、格マーカーや意味素性の適合性を調べながら文中の名詞の格を判定していくわけである。

解析結果はフレームを用いて表現する。フレームは枠であり、その中にスロット(slot)という入れ物を複数持つことができる。そこで、図5に示すように、ある述語と格の関係の一つのフレーム

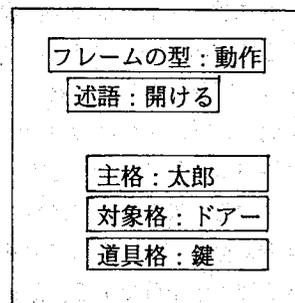


図5 格文法解析結果のフレームによる表現

で表現する。格の名前はスロットの名、その値はスロット値となる。

2.4. 文脈・文脈解析

ここでいう文脈とは文を越えた意味のことで、状況ともいう。これを解析するのが、文脈解析である。

簡単な例を挙げれば、2つの文があり、一方が結論で他方がその理由ということがある。また、指示代名詞がある時、それが指す対象が別の文のことがある。これらを解析するには文脈解析が必要である。今では、この解析はさほど難しくない。たとえば、先に述べたフレームを2つ用意し、一方を結論の文の、他方を理由の文のフレームにする。そして、結論の方のフレーム中に理由というスロットを用意し、そこに理由の方のフレームへのポイントを入れる。この手法によって2つのフレームは結論と理由の関係で結びつくので、この関係をコンピュータ上に明示することができる。

少し難しい例を挙げよう。一つの単語が複数の意味を持つことがある。たとえば、“cell”は生物学では細胞だが、電気化学では電池である。これを多義語というが、生物学の文脈か電気化学の文脈かを判定できるだろうか。さらに例を挙げれば、先程の“Time flies like an arrow.”の前後の文章の文脈解析をして、「おとぎ話」の文脈であれば『時バエは矢を好む』と訳し、「格言」の文脈であれば『時は矢のように飛ぶ（光陰矢の如し）』と訳せるだろうか。これは未だに困難な問題である。

2.5. 生成

生成は解析の逆の処理である。原理的にはこれでよいのだが、自然な文を生成するのは決して容易なことではない。我々が自然と感じる文とは何か、名文とは何か、などは厳密には明らかになっていない。そのため、そのような文を生成する仕方も解明されていない。もっとも、大人の間でも悪文を書く人はいるのだから、コンピュータにとって難しいのは当たり前かもしれない。

3. 機械翻訳システムの付加的機能：前編集、後編集、辞書追加機能

これまで、機械翻訳の仕組みと原理を説明してきたが、細かいところで色々困難な点があることが分かっていただけたと思う。実用システムとしては、あまり難しい処理を頑張って行うあまり、処理速度が遅くなるとは元も子もない。そこで付加されるのが、前編集(pre-edit)や後編集(post-edit)の機能である。

前編集とは、翻訳処理を行う前に源文を解析しやすいように書き換えることで、これはユーザがエディターを使って行うことになる。また、後編集とは、翻訳処理によって生成した文章を読みやすいように書き換えることである。自然な文章、まして名文を生成することは困難なので、ユーザが必要に応じて直すわけである。もっとも、前編集には源言語の、後編集には目標言語の豊かな知識が必要であることは言うまでもない。

なお、分野によって訳語が異なることが多い。機械翻訳システムに元から用意されている辞書にはないこともある。そこで、ユーザが追加できるようユーザ辞書追加機能が用意されている。色々追加することによって使いやすくなっていくようである。その点は日本語ワープロのユーザ辞書と似たようなものである。

機械翻訳システムは最近では実作業にも使われている。ユーザの話によると、完全な訳を出すの

ではなく、粗訳を出させて、それを適宜修正するのだそうである。人間が最初から全部翻訳するよりも、この方が速いという。この使い方が示す実態から、最近ではmachine-aided(またはcomputer-aided) translation と言う人も多い。

4. 最近の進歩と今後への期待

多義語の場合に適切な訳語を出す手法としては、語と語の共出現により判定する方法がある。これは、たとえば、“cell”の場合に生物学の分野の用語が共に出現していれば、細胞と訳し、電気化学の用語が共出現していれば電池と訳そうという考え方である。しかし、多数の用語との共出現を考えると話はそう単純ではない。最近、東芝からニューロ・コンピュータの手法を使って、多数の用語の相互間の影響を判定し、その結果によって最適の訳語を提案するシステムが市販されている。これは、源言語で書かれた文章を翻訳していく過程でニューロ・コンピュータの手法によって学習していくため、文脈を次第に変えられても追いついてきて、訳語を変えてくる。人間にとっては当たり前のことであるが、機械翻訳の仕組みに興味を持つ人には「さすが」と思わせるものがある。このシステムは昨年末、通産省から賞を貰ったそうである。

別の話題として、辞書の問題がある。格文法による解析のためには、先に述べた辞書が必要であるが、これを作るには多くの人手と資金が必要である。そこで、この辞書を作成する国家的プロジェクト(日本電子化辞書)が以前スタートしたが、つい最近終了した。その成果は公開されるとのことであるので、今後の進歩に大きく寄与するであろう。

また、関西学園都市にあるATRでは、国際電話の同時通訳システムを開発した。新聞発表にもなったので、ご存じの方も多いただろう。これは、単なる機械翻訳システムではなく、既存の技術ではあるが、会話を聞いて文として認識する技術や、文を生成するだけでなく音声にする技術も組み込まれており、全体として十分な速度で処理できる点が優れている。今のところ、処理できる内容は予約など簡単なことに限られているが、それにしても立派なものである。

その他、最近の進歩として挙げたいことは少なくないが、その説明に紙面を取られるので割愛する。参考文献を挙げておいたので、それを参照されたい。

より良い翻訳のためには、残された種々の問題を解決していく必要がある。たとえば、比喩の解析は難しい。同じ「走る」でも、「人が走る」「車が走る」は簡単だが、「虫酸が走る」は解析が困難である。また、常識や推論機能などをシステムに持たせることも必要である。機械翻訳にはいろいろな要求がありうるので、完成はしないだろうが、進歩していくことは確かである。(たとえ、完成したとしても人間はそれを認めたくないかもしれない。)最近、長尾京大教授らが中心となって機械翻訳協会が設立されたし、各社の取り組みも本格的になってきた。この分野は日本が世界をリードしている分野でもあり、今後が楽しみである。

5. おわりに

実は、編集部が筆者に用意して下さった頁数を遙に越えてしまった。おまけに指定された締切りも大幅に遅れてしまった。誠に申し訳ない。つつい書きすぎたと反省しているが、機械翻訳の概略とその難しさの一端を理解していただけたらうか。さらに興味を持たれた方のために参考文献を挙げておくので、詳しくはそれを参照されたい。

それにしても、人間が行う翻訳とは誠に素晴らしい知的活動である。そういませんか。

参考文献

・機械翻訳の解説書

長尾真、機械翻訳はどこまで可能か、岩波書店（1986）

著者の長尾京大教授はこの分野を代表する学者である。

野村浩郷編、言語処理と機械翻訳、講談社サイエンティフィク（1991）

野村、辻井、村木、新田、天野の5人による共著。各氏は大学や会社でこの分野を研究してきた中心人物である。

長尾真ほか編、機械翻訳サミット、オーム社（1989）

1987年に行われた国際会議、機械翻訳サミットの会議録を本にまとめたもの。付録として代表的各社の機械翻訳システムの概要が収録されている。

これ以外にもあるが、割愛する。

・自然言語処理の解説

長尾真、言語工学、昭晃堂（1983）

機械翻訳の基盤となる自然言語処理に関する解説書。

長尾確ほか、特集：自然言語処理技術の最近の動向、情報処理、33巻7号（1992）

情報処理学会の学会誌の特集。この種の話題は何回か採り上げられているが、最新のものを挙げておく。

・論文誌

情報処理学会自然言語処理研究会資料

研究会の発表論文集

情報処理学会論文誌

そのうちの機械翻訳や自然言語処理に関する論文が対象となる。

研究開発におけるコンピュータ化学の活用

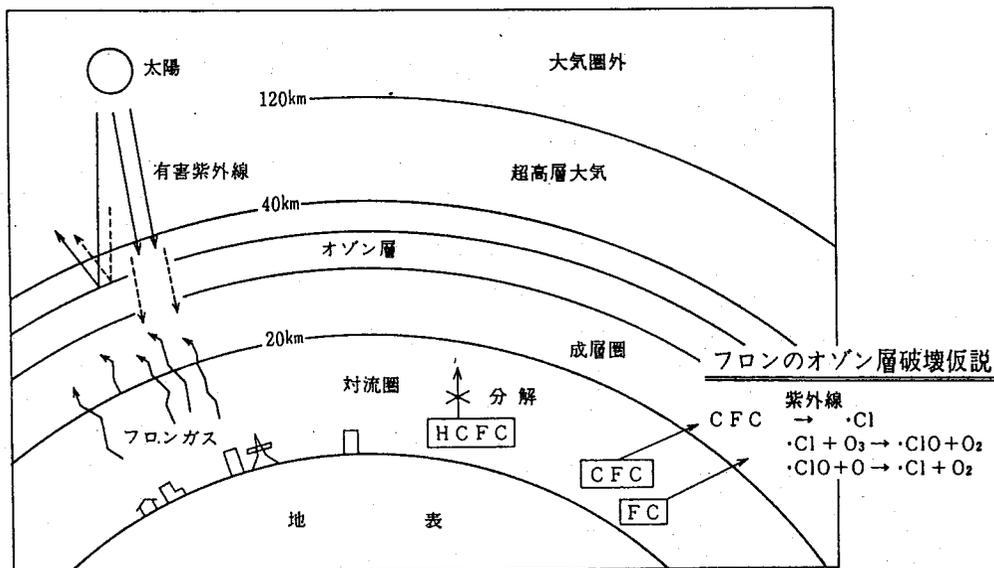
—フロン代替品の開発を例として—

旭硝子(株) 中央研究所
主幹研究員 石田 嘉明
(現在:旭硝子財団 担当部長)

1. はじめに — 代替フロン開発への動き

フロンは不燃性で、かつ低毒性の物質であり、優れた熱力学特性を有するため有用な物質として現代の社会に幅広く使用されて来た。その性質はクーラーなどへの冷媒用として、またウレタンなどのフォーム発泡剤用として、あるいは溶剤としてなど、DuPont社が1932年に本格的な生産を開始してから60年間に需要は大きく広がって来ていた。日本のフロンメーカー企業は旭硝子、ダイキン、昭和電工、三井フロロケミカル、セントラル硝子である。

しかし、フロン類の中でも、クロロフルオロカーボン(CFC)類は化学的に非常に安定であるため、地表で大気中に放出されたCFCは対流圏をゆっくりと上昇拡散し、約10年かかってその上の成層圏の中にあるオゾン層(地上30kmから40km付近)まで到達する。オゾン層は太陽からの強い紫外線を吸収し、地球を保護していると言われている。CFCはそこで紫外線的作用を受けて塩素ラジカルを生成し、塩素ラジカルがオゾンと連鎖的に反応してオゾン層を破壊する恐れがあると1974年カリフォルニア大学のローランド教授が発表した。オゾンホールが観測されたのは1984年であるがそれ以来、オゾン層保護の世界的活動が活発化したことは、広く知られているところである。



2. 代替フロン開発の方向性

モントリオール議定書などで規制及び全廃が決められているフロンを特定フロン（CFC-11, -12, -113, -114, -115 など）と呼ぶが、さらに特定のハロンと1,1,1-トリクロロエタンおよび四塩化炭素も規制対象である。地球全体の環境問題に係わるこれらの特定フロンを、国際的な理解と協力を得ること無しに代替品へ移行するのは、困難なことに思えた。おそらく、代替品の低公害性、低毒性、不燃性、などへの理解がまず第一に必要であるし、洗浄特性、熱力学特性などの機能性に対する理解も同時に得られなければならない。このような事情のもとで、代替物質探索の範囲は規制対象以外のフロンの中に安全な物質が残っていないかどうかを捜し出すことに、ひとつの方向性を求めたものと言える。

代替フロンを考える上で重要な指針となったことのひとつは、寿命を短くして大気中を10年間も滞留しないようにすること、すなわち、化学的安定性をやや落とすことであった¹⁾。分子中に水素原子を含むフロン（HCFC：ハイドロクロロフルオロカーボン）類は比較的寿命は短いと言われている。例えば HCFC-123 の寿命は1.5年と短い。

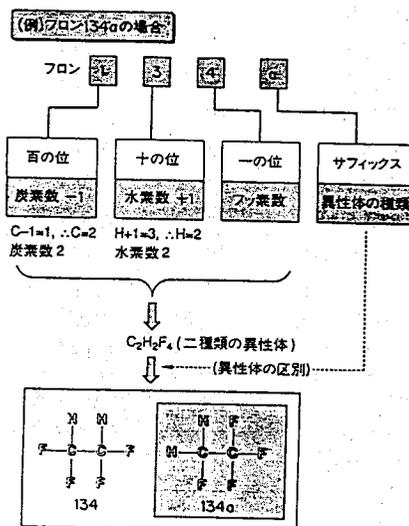
もうひとつの大きな指針は、分子中に塩素原子を持たないようにすること、すなわち上記のようにオゾン層を破壊するのが塩素ラジカルならば、そのようなオゾンとの反応活性種を放出しないフロン（HFC：ハイドロフルオロカーボン）の分子構造を設計することである。

しかしながら調べてみると、単に分子中の塩素を水素なりフッ素なりに置き換えて塩素原子を除くだけでは沸点が下がり過ぎ、樹脂を溶解する能力（相溶性）が低くなる。他方、水素原子の数が多くなると可燃性に転じて来たり、毒性の問題が出てきたりする。代替フロン開発研究のポイントはこれらの指針や特性を相互に考え合わせながら、分子設計を進めることにあった。

フロン類一覧

フロン	化学構造	沸点(°C)
CFC-12	CCl_2F_2	-29.8
HCFC-22	CHClF_2	-40.8
HFC-134a	$\text{CF}_3\text{CH}_2\text{F}$	-26.5
HCFC-142b	CClF_2CH_3	-9.8
HFC-152a	CHF_2CH_3	-25.0
HFC-125	CHF_2CF_3	-48.5
CFC-11	CCl_3F	23.8
HCFC-123	CF_3CHCl_2	27.1
HCFC-141b	CCl_2FCH_3	32
HCFC-124	CF_3CHClF	-12
CFC-113	$\text{CCl}_2\text{FCClF}_2$	47.6
HCFC-225ca	$\text{CF}_3\text{CF}_2\text{CHCl}_2$	51.1
HCFC-225cb	$\text{CClF}_2\text{CF}_2\text{CHClF}$	56.1

注) アミは規制対象フロン。



フロンの命名法(コードナンバーのつけ方)

3. 内外の開発体制

上に述べて来たように、代替フロンの分子設計の基本的な部分はメーカー各社ともほぼ同様の考え方であった。おそらくフロン類以外の化学物質を考えようとした研究者もかなり多数いたと思うが、限られた時間と実験の範囲で具体的な物質を提案することは困難であったものと思う。CFC-11, CFC-12の代替にHCFC-123, -141b, -142b, -124 および HFC-134a などが有望な候補として挙げられ、検討されていることから上記のようなことが明かである。一方、我が国ではフロン類以外の領域からも代替物質を探索しようとする国の研究プロジェクトを既に開始させており、研究成果が期待される場所である。

候補物質の安全性については国際的な協力が必要であった。特に長期間の生物試験を必要とする慢性毒性などの安全性試験については、これを遂行するための国際共同プログラム (PAFT: Program for Alternative Fluorocarbon Toxicity testing) が組まれて、世界中の化学企業10数社がその試験のための多額の費用を出し合っている。代替フロンが環境に与える影響に関する調査を推進する国際協力委員会 (AFEAS: Alternative Fluorocarbon Environmental Acceptability Studies) も発足して、オゾン破壊係数の物理化学的測定技術の進展なども加速されようとしている。

フロンの寿命とオゾン破壊係数

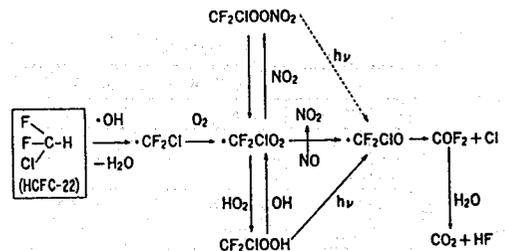
化合物	寿命(年)*	ODP*
CFC-11	52	1.0
CFC-12	101	0.87
CFC-113	79	0.76
CFC-114	197	0.56
CFC-115	393	0.27
HCFC-22	15	0.043
HCFC-123	1.5	0.016
HCFC-124	6.5	0.017
HFC-125	43	0
HFC-134a	24	0
HCFC-141b	6.9	0.081
HCFC-142b	19	0.0455
HFC-152a	2.7	0
HCFC-225ca	1.5 [#]	0.01 [‡]
HCFC-225cb	5.1 [#]	0.04 [‡]

*Scientific Assessment of Stratospheric Ozone: 1989

LLNL 2-D Model [#]Physical Chemical Lab. 測定値

[#]NIST 測定値 [‡]旭硝子推算値

HCFC-22の対流圏での推定分解機構



4. 代替フロン開発の現場

エレクトロニクス産業でプリント基板やシリコンウェーハーなどの洗浄剤として使用されて来ていた CFC-113 は、日本では最も消費量の多いフロンであり、その代替品は上記のような開発方向の他さらにいろいろの制約が加わっていた。沸点が50度前後であること、油脂には溶解力を持ち樹脂には溶解力を示さない性質（選択的溶解性）のあること、洗浄する対象物の隅々まで浸透する（濡れ性または低表面張力）性質のあること、などの要求特性を満たさなければならない。すなわち、CFC-113 にそのままそっくり置き換えうる（drop-in-replacement）物質を追究する必要があった。これまで使用されてきたフロン類は炭素数1個および2個の化合物であったが、この範囲のHCFCまたはHFC にはもはやこれらの条件を満たすものは無かった。これに対し、炭素数3個以上のフロンはこれまでに実用化はされていない。しかしどんな構造の化合物が考えられるのか、しかも合成が容易に出来るものなのか？

ここで当社の研究陣は、合成、触媒、分析、プラントエンジニアリングそして物性評価の各技術を総合化してこの社会的要請の強く、短期間の開発力が求められている課題に積極的に取り組むことを考えて、コンピュータを用いての分子設計技術を開発研究の中に取り込もうと考えた。分子設計技術などと言っても色々な計算手段の組み合わせであって、まとまって一連の体系的な方法論を成すものではない。すなわち、炭素3個以上のHCFC, HFCの構造異性体の数え上げの問題、反応性（平衡反応）の予測計算、樹脂との相溶性（KB 値）の構造-物性相関の問題、オゾン破壊係数（ODP）の予測、麻醉性の予測、臨界点等の圧力-エンタルピー線図の熱物性計算などへコンピュータ化学を応用しようと考えた。

HCFC-225ca および HCFC-225cb の物性

項目	HCFC-225ca	HCFC-225cb	CFC-113
構造式	$\text{CF}_2\text{CF}_2\text{CHCl}_2$	$\text{CClF}_2\text{CF}_2\text{CHClF}$	$\text{CCl}_2\text{FCClF}_2$
分子量	202.94	202.94	187.38
沸点 (°C)	51.1	56.1	47.6
凝固点 (°C)	-94	-97	-35
密度 (g/cm ³) (25°C)	1.55	1.56	1.56
粘度 (cps) (25°C)	0.59	0.61	0.68
表面張力 (dyne/cm)	16.3	17.7	17.3
KB 値	34	30	31
引火性	なし	なし	なし
蒸発速度 (エーテル=100)	101	84	123

化学の分野にコンピュータを活用しようと考えている者の多くは、研究の日常的な進め方を何かしら少しでも変えたいという気持ちがどこかに働いているように思える。合成や触媒の専門家がコンピュータ化学に関心を持つときは、おそらく実験を出来るだけ合理的に進めて効率的な研究をしたいと考えているのだろう。出来るだけ少ない実験量で収率の高い合成反応ルートを見いだしたい。プラント設計でも熱力学計算が重要で、計算出来にくいところは理論的に予測計算したい。

分析化学の分野でも同じである。分光スペクトルなどのデータ処理と解析技術にコンピュータを取り入れたいと考えたのも、そのデータ量と解析の煩雑な手順のことを考えると、日常性の変化を求める動機がはたらいたことはごく自然なことであった。ともかく、日常的な研究開発の現場にコンピュータがどの程度活用出来るのかを試すことはチャレンジ精神に富んでいることだった。それだけに、実行の場面ではそのチャレンジ精神を受け入れてくれる研究現場の風土、あるいはスタッフのスキルが大切だったように思える。

5. コンピュータの活躍場面

5.1 データベースとスペクトル解析

コンピュータ利用の最初のきっかけは、試作成品の分析による構造同定問題から始まった。代替フロンの候補がまだ絞り込めない段階にも、フロンの異性体を片端から造ってみる実験はどんどん進んでいた。当然のように構造確認のためにも分析は精力的に進められ、合成の目的化合物はもとより反応条件や触媒の違いによって副生してくる化合物の構造を解析同定することが研究の進展に大きなキーとなる。主に MS, ^{19}F -NMR が中心となる分析であったが、これらのデータベースが次第に大きくなって行くと共に、スペクトルデータから構造式を迅速に割り出すためには、データベース検索システムを利用するだけでは追いつかない。つまり、これまでに解析したことのない未知化合物にはデータベース検索は無力なので、MS スペクトルに対してはフラグメントイオンピークの組成式 (H, C, Cl, F の個数の組み合わせ) を計算するプログラム、 ^{19}F -NMR スペクトルに対しては化学シフトの推定値計算のプログラム、などいわゆる解析用のパソコン用プログラムが役に立ちはじめ、次第に強力な武器になってきた。

5.2 構造異性体の列挙

フロン類の合成実験は永年続けられてきていたことであったが、改めて代替品を探索することを主眼としてこれまでの実験を整理してみると、いったいぜんたいフロン類とはどこまでの広がりを持っているのであろうか?、どこまで追いかければ全部試したことになるか? すなわち、構造異性体の範囲の全体を知りたいと言う欲求が自然に起こってきた。合成を計画する者もそれを分析する者も、全体像の鳥瞰図があれば仕事はやりやすい。全体から候補を絞り込めば、見落としも避けられようし、ターゲットが明確である。例えば炭素2個のフロン類の構造異性体を書き出しさえすれば (これは手作業でも簡単に行える)、もうすでにその範囲には実用化できそうなものが無いことを誰にでも納得してもらえる。

炭素3個以上のフロン類の鳥瞰図を作るにはコンピュータが使える。

C, H, Cl, F の数に不飽和数を加えた組み合わせ計算で二重結合、三重結合、環状構造、光学異性も含みながら、異性体の数え上げ問題を実行した。世の中には構式組み立て問題の立派なソフトウェアプログラムが有るが、ここでは計算対象の範囲が狭いのでプログラムも簡単で身近に有ったちいさなパソコンで充分だった。今になって振り返ると、鳥瞰図がひとつ出来上がると、次の山が知りたくなり、

次の鳥瞰図を作る。また次の山、次の山と、興奮ぎみにしかし冷静に探索が続いたことを研究チームのメンバーは思い出す。炭素3個の鳥瞰図の中からHCFC-225ca, HCFC-225cbが見つかったことは、筆者も含みチームのメンバーにとって幸せだった。

F \ H	0	1	2	3	4	5	6	7	8
0	C ₃ O ₃ 1	C ₃ O ₂ F 1	C ₃ O ₂ F ₂ 4	C ₃ O ₂ F ₃ 5	C ₃ O ₂ F ₄ 6	C ₃ O ₂ F ₅ 5	C ₃ O ₂ F ₆ 4	C ₃ O ₂ F ₇ 2	C ₃ F ₃ 1
1	C ₃ HO ₂ 2	C ₃ HO ₂ F 5	C ₃ HO ₂ F ₂ 10	C ₃ HO ₂ F ₃ 13	C ₃ HO ₂ F ₄ 13	C ₃ HO ₂ F ₅ 10	C ₃ HO ₂ F ₆ 5	C ₃ HO ₂ F ₇ 2	
2	C ₃ H ₂ O ₂ 4	C ₃ H ₂ O ₂ F 10	C ₃ H ₂ O ₂ F ₂ 19	C ₃ H ₂ O ₂ F ₃ 22	C ₃ H ₂ O ₂ F ₄ 19	C ₃ H ₂ O ₂ F ₅ 10	C ₃ H ₂ F ₄ 4		
3	C ₃ H ₂ O ₂ 5	C ₃ H ₂ O ₂ F 13	C ₃ F ₂ O ₂ F ₂ 22	C ₃ H ₂ O ₂ F ₃ 22	C ₃ H ₂ CF ₄ 13	C ₃ H ₂ F ₅ 5			
4	C ₃ H ₂ O ₂ 6	C ₃ H ₂ O ₂ F 13	C ₃ H ₂ O ₂ F ₂ 19	C ₃ H ₂ CF ₃ 13	C ₃ H ₂ F ₄ 6				
5	C ₃ H ₂ O ₂ 5	C ₃ H ₂ O ₂ F 10	C ₃ H ₂ CF ₂ 10	C ₃ H ₂ F ₃ 5					
6	C ₃ H ₂ O ₂ 4	C ₃ H ₂ CF 5	C ₃ H ₂ F ₂ 4						
7	C ₃ H ₂ O 2	C ₃ H ₂ F 2							
8	C ₃ H ₂ 1								

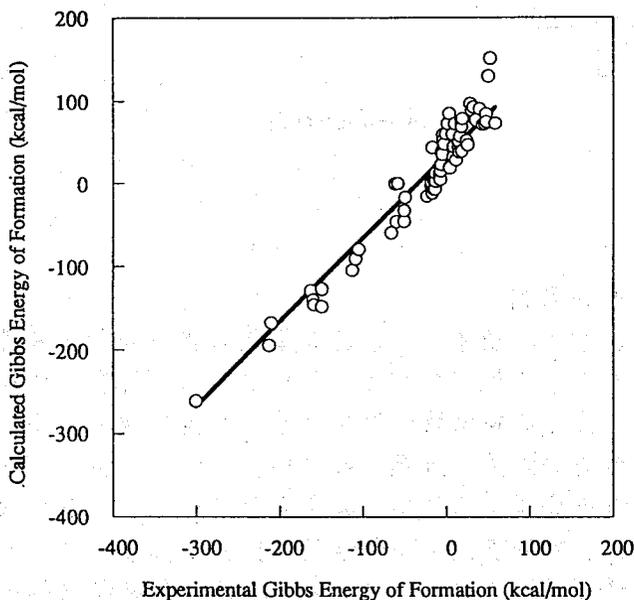
炭素数3個のフロン類構造異性体数

5.3 分子軌道法と反応設計

フロン類の特性は、水素、塩素、フッ素の個数の組み合わせと異性体構造によってそれぞれ異なる。これらの機能の差は、特にフッ素の性質に依存しているものと思われる。すなわち大きな電気陰性度を持ち、水素に次いで小さい原子であり、炭素-フッ素間の結合力が強く、分極率は低い。これらの性質が他の炭化水素系化合物に比べて特異性を発現する理由と考えられている。これに水素原子が付くか付かないかによって、熱的安定性や化学的安定性が異なってくる。塩素原子は洗浄力に何らかの関係を示している。これらの特徴は代替品候補の分子軌道からも充分考察できるので、最安定構造の原子間距離、結合角、双極子モーメントなどの計算値は参考になった。しかしながら、合成反応の経路を探索する研究者が最も期待したのは、気相反応のGibbsの生成エネルギーを分子軌道法で計算予測することであった。

従来のフロンの製造反応は、原料の塩素化合物を触媒存在下にHFでフッ素化するものであったが、代替フロンの場合にはフッ素化だけでなく、水素化、塩素化及び異性化などを選択的に精緻にコントロールする技術が必要になる。多くの選択肢を持つ合成経路を考える上で、素反応の生成Gibbsエネルギーの値が分かれば、温度条件の選択や平衡反応の反応率、および反応速度の推算などが出来、その反応の有効性を判定することで選択肢の中から大きな枝刈りができる。

言うまでもなくGibbsエネルギーはNBSやJANAFの熱力学データ集から調べることが出来るが、ここで取り扱おうとしている化合物の実験データの全てが得られる訳ではない。そこでGibbsエネルギーはエンタルピーとエントロピーから求められることに着目し、エンタルピー、エントロピーの両項ともに化合物の項から構成元素の項を差し引く形で定義しておき、化合物の項は分子軌道法で構造最適化を行った振動、回転、並進の分配関数から、また構成元素の項は元素単体の同じく分配関数を H_2, Cl_2, F_2 などに対してあらかじめ計算したものを用いた。多数の素反応を計算しなければならないから、半経験的分子軌道法で我慢することとしMopacの中に元素単体の分配関数は内臓させ、MNDO, AM1, PM3の計算精度の比較を行った。どの計算法を用いても実験値と計算値との間に高い相関 ($r=0.96$; $n=74$ 化合物) を示した²⁾ ので、研究者に強い信頼を与えた。コンピュータによる推算値が合成化学担当者に認められる為には、計算値と実測値との間に高い相関の有ることを十分に証明することの重要性を痛感した。他方、或る素反応が選択的塩素化反応か否かの判定にフロンティア軌道理論が適用され、実験事実も予測計算を支持したことは印象に残った。



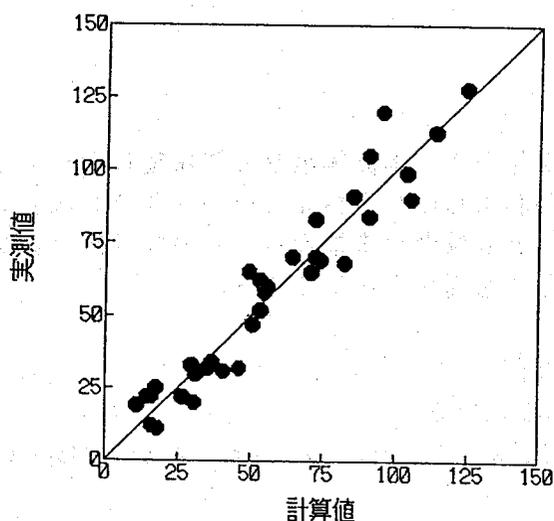
MNDO calculated and experimental Gibbs energy of formation for all compounds studied.

5.4 構造物性相関 (QSPR)

QSPRは色々な場面で利用された。フロン類の沸点の予測計算はかなり精度の高い経験式が見い出されていて、構造と物理化学的性質との間にかなり捕えやすい相関があるものと予想していた。そこで、理論的に解決すべき課題と統計的あるいは経験的に解決すべき課題とを検討、区別してそれぞれ異なるアプローチを取ることにした。麻醉性はQSPRの応用を試みた例であるが、分子設計の段階で候補化合物を絞り込むのに効果を示したのは、溶解性に関するQSARであった。

洗淨用の代替フロンに要求される特性の中で、油脂には溶解力を持ち樹脂には溶解力を示さない性質が重要であることは前にも述べた。この性質の指標にKB値が一般に用いられていた。これは検体にカウリプタノール樹脂と呼ばれる天然の樹脂を滴下してその溶液の曇点の濃度で表わす測定方法である。天然物であるために種々の分子構造を持つポリマーおよびオリゴマーからなる樹脂である。従ってこの指標の発現機構がわからない。しかしフロン類の範囲に限って測定値を比較するのであれば、発現機構は同じと考えられる。フロンの異性体構造間の違いを表現出来れば良いはずである。KB値既知の約30種のフロン類について実測値と計算値の相関係数0.96を与えられた相関式を、試行錯誤の末に得た。この式を用いて候補化合物のKB値を予測した。炭素数3の候補化合物を絞り込んで、実際に50種の化合物の合成と精製が行われたが、それらの実測値は予測値に良く一致した。明らかに構造式と溶解性の機能発現との間に相関があると思われるが、相手が天然物だけに追及は行わなかった。厳密な実験はHCFC-225を絞り込んだ後に、種々の合成樹脂に対する溶解性の試験で精力的な調査が行われた。

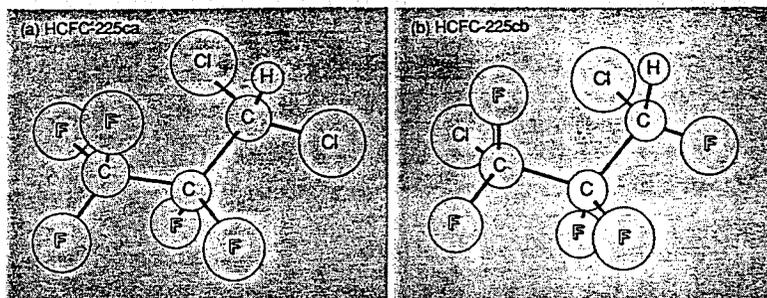
サンプル数 = 34 相関係数 = 0.962
 誤差標準偏差 = 8.9 誤差平均 = 6.8



6. おわりに — 実験と計算は車の両輪

コンピュータを利用した反応性の予測や物性の推算を行いながら、その結果に刺激されて実験が行われ、そこに分析があり、その結果が次の実験と計算の必要性を感じさせるといふ循環作用を生んで行った。ひとつの山の頂に達すると、さらに次の山がはっきりと見えてくるようであった。その過程で、炭素3個のフロンでは中央の炭素がジフルオロメチレンの場合に化学的熱的に安定性を保つことがはっきりしてきた¹⁾。このことは永くフッ素化学にたずさわって来ていた研究者の認識の底にはんやりと存在していたことであつたが、上に述べたような

過程の中で明確に定量性を増して来たのである。光学異性体まで区別して数えたとき366種に達する炭素3個の異性体群の中から、50種に絞り、実際に合成して物性のスクリーニングを行った。ひとつを合成する度に、反応性の計算の確認が行えたとし、KB値の実測データが増えて行った。その中にHCFC-225の2種の異性体が存在し、両者ともにCFC-113の物性に近い物質であった。環境への影響も寿命の測定やODPの推算を通してオゾン層の破壊には極めて影響の少ない物質と考えられている。また生物や人体に対する影響の点も上記のようにPAFTで試験が続けられている。



以上、不十分だったと思うが、本誌編集長のご希望に沿うように研究開発におけるコンピュータ化学の利用の流れを紹介したつもりである。もとより本文中の計算、実験、分析のいずれも筆者が直接行ったものではない。プロジェクトチームの諸氏にこの場を借りて感謝する次第である。

引用文献

- 1) 森川, 化学, 45(2) p89 (1990).
- 2) T. Miyazima, S. Morikawa, Reports Res. Lab. Asahi Glass Co., Ltd., Vol.41, No. 2 (1991)

文 献 紹 介

[コンテンツ・リスト]

JOURNAL OF CHEMICAL EDUCATION

Jan. 1993, VOL.70, NO.1

Journal of Chemical Education : New Issues for 1993. edited by James P. Birk, p.62.

Computer Series:145

LIMSport II : Use of the Interfaced Balance for Pressure Measurements,
Streamlined Synthesis, And Titrations, Ed Vitz, p.63.

A Direct Approach for Solving Simultaneous Equations: Combining a Basic
Understanding of Equilibria with the Convenience and Power of a Spreadsheet,
D.R.Carter, M.S.Frye, and W.A.Mattson, p.63.

Feb. 1993, VOL.70, NO.2

Computer Series:146

Computer-Assited Learning: IRExpert, an Infrared Interpretation Assistant,
D.Cabrol, J.P.Rabine, D.Ricard, M.Rouillard, and T.P.Forrest. p.120.

JOURNAL OF CHEMICAL INFORMATION

AND COMPUTER SCIENCES

□ VOL.33, March / April 1993

184 Problems and Solutions in the Development of Hard and Soft Versions of a Chemical Reaction Hazards Database
Peter E. Barker, Leslie Bretherick, and Frank A. A. Kay*

193 About Metrics of Bibliometrics
Milan Kunz

CHEMICAL COMPUTATION

197 Graph Automorphism Perception Algorithms in Computer-Enhanced Structure Elucidation
*Marko Razinger, K. Balasubramanian, and Morton E. Munk**

202 Neural Network Approach to Structural Feature Recognition from Infrared Spectra
D. Ricard, C. Cachet, D. Cabrol-Bass, and T. P. Forrest*

211 List Operations on Chemical Graphs. 3. Development of Vertex and Edge Models for Fitting Retention Index Data
*Ch. Duwenbeck and P. Zinn**

220 List Operations on Chemical Graphs. 4. Using Edge Models for Prediction of Retention Index Data
*Ch. Duwenbeck and P. Zinn**

231 The Walk ID Number Revisited
*Wolfgang R. Müller, Klaus Szymanski, Jan V. Knop, Zlatko Mihalić, and Nenad Trinajstić**

234 "SpectraGraph" and "SpectraSort": Mass Spectral Display and Interpretation Software for the Macintosh
Paul R. West and Andrew J. Mort*

240 Topological Organic Chemistry. 6. Graph Theory and Molecular Topological Indices of Cycloalkanes
Harry P. Schultz and Tor P. Schultz*

245 Sequential Assignment of 2D-NMR Spectra of Proteins Using Genetic Algorithms
R. Wehrens, C. Lucasius, L. Buydens, and G. Kateman*

MOLECULAR MODELING

252 Implicit Runge-Kutta Method for Molecular Dynamics Integration
Dušanka Janežič and Bojan Orel*

258 Topological Approach of ¹³C NMR Spectral Simulation: Application to Fuzzy Substructures
Annick Panaye, Jean-Pierre Doucet, and Bo Tao Fan*

266 ANALOGS: A Computer Program for the Design of Multivariate Sets of Analog Compounds
Mario Marsili and Heinz Saller*

270 Extending Crystallographic Information with Semiempirical Quantum Mechanics and Molecular Mechanics: A Case of Aspartic Proteinases
Amiram Goldblum, Anwar Rayan, Amit Fliess, and Meir Glick*

275 Correlation of Graph-Theoretical Parameters with Biological Activity
Gordon G. Cash and Joseph J. Breen*

編集後記

「薫風の」という形容がぴったりの季節になりました。部会員の皆様は各方面に御活躍のことと存じます。Vol.11, No.3号をお届けします。お陰様で充実した誌面を作ることが出来ました。執筆いただいた諸先生には改めて御礼申し上げます。今回の特集は今年2月に開催された第4回ケモメトリックス・ワークショップの講演者の方々に、当日の話題を中心として御執筆いただいたものです。今後もこれまでのような独自の企画と合わせて、今回のようなワークショップや討論会と連携した特集も続けていきたいと考えています。御意見、御感想をお待ちします。

本誌を部会員相互のコミュニケーションの場として更に機能させていきたいと願っています。本誌に対する御意見、御感想はもとより部会活動全般に対する御提言等をお寄せいただきたいと思います。頂戴した御意見は「声」ないしは「提言」欄を設けて掲載させていただきます。事務局宛お送り下さい。

(編集代表 四居)