CICSJ Bulletin Vol.17, No.1, February 1999

このページは「日本化学会・情報化学部会」の責任において運営されています。

目次

◆巻 頭 言

僕の１９９９年問題

お茶の水女子大学理学部　細矢治夫
hosoya@is.ocha.ac.jp

「西暦2000年の１月１日より早い1999年９月９日にもコンピュータは暴走したり、止まったりするかも知れない。」「実は危ないんだ。」といううわさが広まり始めている。情報科学に関わる仕事をしている者にとって、これを対岸の火として手をこまねいていることは許されない、などという格好良い外向けの顔はどこかに置いて、ともかく新年明けましておめでとうございます。今年は平成も１１年と相成りました。僕の生まれは昭和の１１年です。

「何を寝ぼけているのか」と言われそうだが、実は20世紀から21世紀への変り目を迎えるこの２、３年、ゾロ目気違いのこの僕は、何かそわそわして落着かないのである。今度の世紀の変わり目は100年前とは違って、キリスト教ではmillenniumの区切りに当たるというのだ（S.G.グールド「暦と数の問題－グールド教授の2000年問題」早川書房）。この言葉には、「千年間」や「千年祭」の他に、「キリストが再び地上に降りて地上を統治する神聖な千年間」という特別な意味がある。しかし僕はカトリックの学校は出たけれど、キリスト教とはそんなに深い付き合いはしていない。1999年とか2000年に対する僕の問題意識は全然別のところにあるのだ。ある人は僕のことをnumerologist（数秘術者）と決めつけてしまうが、それも違う。この言葉は、数字に何か神秘的乃至は宗教的な意味をこじつける人達のことをいう。僕はただの「ゾロ目気違い」に過ぎない。実は、ＣＩＣＳＪの巻頭言の場を借りて、そのゾロ目気違いの密かなたわ言を開陳しようというのである。御用とお急ぎでない人だけ、お目とおつむを拝借。しかし、これはエイプリルフールの馬鹿騒ぎより、はるかに教育的かつ実践的な生産活動なのである。

今年の僕の年賀状のモチーフは次の式である。

$$\text{HAPPY NEW 11'S} = 11 \cdot 1.1/.11/.11 + 11 \cdot 11/.11/1.\dot{1} + 11/1.1/1.\dot{1}$$

右辺の／は，御存じ、ただの割り算記号である。アメリカや中国など学校で÷という記号を教えていない国は多いし、自然科学の教科書や論文でも÷は使われないのでそれに従った。$1.\dot{1}$というのは循環小数で10／9に等しい。循環小数の考え方と記号は義務教育の中学で習っているはずなのに、この「暗号」が読めない、とボヤかれる大学の先生が何人もおられたのには驚いてしまった。そういう人のために助け船を出すと、右辺は1000＋990＋9＝1999となる。11個の11が11個の点の助けを借りて1999になるところを見てもらいたい。

これに比べると、次の計算式は何とも情けなく、格好が悪い。

$$1+2-3+4^3+5^3+6^3+7^3+8^3+9^3+10=1999$$

しかし、もし僕が89才まで生きていたら、

$$1^3+2^3+3^3+4^3+5^3+6^3+7^3+8^3+9^3=2025$$

というきれいな式を年賀状に使えるのだ。しかも、この年は45x45というとても珍しい平方数の年でもある。その前の平方数の年は何と僕の生まれた1936 (=44x44)年。最近買った僕の真っ赤な愛車には1936という「隠れゾロ目」のナンバープレートが着けてある（「ニコリ」1998年12月号）。この車はとてもその時まではもちそうもないが、2025年というのは僕の長生きの第一のターゲットなのだ。なお、1から始まる３乗数の和が平方数になることは、高校生でも知っている。

$$\sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2$$

次に下のローマ数字を読んでもらいたい。

MCMXCIX

ローマ数字は義務教育では教えないが一般常識である。ヨーロッパだけでなく、歴史の浅いアメリカでも、古い建物の礎石や正面の上部に、その建物の建てられた年号がローマ数字で書かれていることがよくある。Mは、先のmillenniumと同じ語源をもっている。Cは centuryの百で、CMと書くと 1000－100＝900 となる。同様にXC＝100－10＝90 である。次いでに、ローマ数字で50 はL，500 はDだが、その語源は知らない。日本では、高額の金額を書くときに下の３桁の数字を省いて千円単位で書き表すことが多い。それと同じように西欧では、何Mと書くことがある。このMを百万と読み違えると悲喜劇が起こる。昔アメリカで車の保険に入るときにこの記号を見てドキドキした経験がある。やはりこういう日常的な常識は持っているにこしたことはない。

三角関数の数列を 足したり、掛けたりするときれいな整数になることが知られている。残念ながら、1999 は素数なので、1000 個近くの項をかき集めなければならないのだが、これを逆に使って、数値計算の桁落ちを調べるテストに使える。僕のパワーマックに載っている MATHEMATICA では、何と小数点以下百桁以上の精度の計算を苦もなくやってくれたのには感激した。

先ず下のsine の２乗の和を計算するとピッタリ1999 になる。

$$\sum_{k=1}^{999}(2\sin\frac{k\pi}{1999})^2=1999$$

この sine を cosine に変えたのでは 1997 になってしまう。もちろん tangent ではどでかい数になってしまう。次に上の式の和の記号を積の記号に変えると、何とこれも 1999 なってしまうのだ。

$$\prod_{k=1}^{999}(2\sin\frac{k\pi}{1999})^2=1999$$

そこでまた、sine を cosine に変えると、今度は 1 になってしまう。しかし、気をとり直して tangent で頑張ると、カッコの中の 2 を除いた式

$$\prod_{k=1}^{999}(\tan\frac{k\pi}{1999})^2=1999$$

が得られる。実はこれらの式は、次のようなより一般的な公式の応用に他ならない。

$$\sum_{k=1}^{[n/2]}(2\sin\frac{k\pi}{n})^2=n \qquad \prod_{k=1}^{[n/2]}(2\sin\frac{k\pi}{n})^2=n \qquad \prod_{k=1}^{[n/2]}(\tan\frac{k\pi}{n})^2=n$$

数学の公式集を丹念に調べると、少し変形した形で載っている（森口、宇田川、一松「数学公式II」、岩波全書、pp. 19－25）し、exp (ix) ＝ cos x + i sin x というオイラーの公式を使えば、比較的簡単に証明することができる。

何だかんだといろいろなことを得意げに書いてきたが、実は今年の元旦は大失敗をしてしまい、いささか気が滅入っているところである。平成11年１月１日の11時11分にＪＲの1110円の切符を買おうと思って勇んで最寄りの駅に行った。11時ちょっと前である。ＪＲには丁度1110円という値段の切符があることは「時刻表」で先刻調査済み。切符の自動販売機に内蔵されている時計がみんな数分ずつ狂っていることも予め知っていたので、10台位並んでいる機械の約半分について130円切符を１枚ずつ買ってチェックした。２分進んでいる機械が１番早く11:11の切符を出してくれることがわかった。あとのは丁度とか１分遅れだ。時間差があまりないので緊張した。定刻の２分ほど前に１番早い機械に張り付いて先ず130円の切符を買った。もちろんオレンジカードを使って。しめしめ、11:11という数字がまぶしい。そこでやおら1110円のボタンを押してみた。ところが何と出て来たのは11時12分ではないか。畜生、一瞬遅かった。それでは２番目に進んでいる機械へ。ところがそこには、もたもたしている家族連れ。その隣の奴も占領されている。元旦のお屠蘇気分のためか、切符を買っている客の誰も彼もがもたもたして見える。やっと１台、時計の針の進みの丁度よい機械が空いたので、また130円を入れてみた。うまく11:11が出て来た。そこで間髪を入れず1110円のボタンを押そうとした。だがこの機械には、そのボタンがないではないか。欠陥自動販売機だった。１台の機械に初めから張り付いて、その１枚だけを狙えばよかったのに、時間差攻撃などという色気を出したばっかりにこの体たらく、とボヤくことしきりの元旦であった。

しかし気を取り直そう。今年は１月11日、11月１日、そして真打ちに11年11月11日があるではないか。だがしかし、その日は米沢で情報化学討論会。悩み多き1999年ではある。

◆特集：巨大分子をめざしたソフトウェア

大規模科学技術計算の現在：
『巨大分子をめざしたソフトウェア』特集の序にかえて

広島大学理学部化学科　　相田　美砂子
maida@ipc.hiroshima-u.ac.jp

CICSJ Bulletin では、最近、大規模科学技術計算に関連したいくつかの特集を組みました。

Vol. 15 No. 2 (1997)　　生体高分子のシミュレーション
Vol. 15 No. 5 (1997)　　生体高分子の量子化学計算
Vol. 16 No. 2 (1998)　　専用計算機

1999 年の No.1 である今号では、現在実際にどなたでも使うことができる、巨大分子の量子化学計算をめざしているソフトウエアについての特集をお送りいたします。

この特集の内容はソフトウエアの宣伝ではありません。そのソフトウェアのめざしている概念や「心意気」のようなものが彷彿とするような内容の記事を、各ソフトウェアの開発者あるいはその内容を良く知っている方、ヘビーユーザーの方に御執筆をお願いいたしました。最近、ほとんどすべてのソフトウェアは、お金を出して買ったり契約をしたりしないと使えなくなりました。これもこのような時代としては当然のこと、とは思いますが、お金を出して買ったソフトウェアは、「ブラックボックス」になってしまいがちです。ただインプットしてアウトプットを得るための「道具」という認識があればまだ良い方で、それどころか、単なるプログラム名称を示すアルファベットの組み合わせでしかない、という認識しかない場合すらあります。ソフトウェアにはそれを開発した人々の熱い思いがつまっています。そしてそれをさらに改良していこう、という多くの研究者の熱意が寄せられるべきものです。そのような「思い」が伝わるような特集を企画したつもりです。

1998 年度のノーベル化学賞は、みなさま御存知のとおり、理論化学・計算化学の分野の、John A. Pople 教授と Walter Kohn 教授に授与されました。私達にとってなじみの深い方々がノーベル化学賞を受賞されたことは、とてもうれしいことです。現在の量子化学の分野におけるほとんどすべての研究は、彼等の業績に深く依存しています。彼等の業績の意義はいくら強調しても、しすぎることはありません。しかし、今後さらにこの分野が発展していくためには、彼等のおかげで誰にでも使えるようになったソフトウェアをただ使っているだけではいけない、と思います。「化学」をより深く理解することができるような方法論をめざす気持ちを、化学の基礎研究にたずさわっている研究者だけでなく、この分野のソフトウェアを使っているすべてのユーザーが持つことが必要ではないでしょうか。そのような多くの方々の努力が積み重なっていくことによって、化学の分野における理論・計算化学の重要性がますます増していくことと思います。

豊かな「明日」を築くためには、その種は「現在」まいておかなくてはなりません。量子化学の「夢」を語るためには、地道に「現在」の成果を積み重ねていくことが必要です。今号の特集記事はそのような気持ちをこめてお送りいたします。情報化学部会の部会員のみなさまの専門分野は多岐にわたっています。量子化学を専門とされていない方々にも、今号の特集は、最近の量子化学の流れ、という意味で興味を持っていただけるのではないか、と期待しております。

あいだ　みさこ　　AIDA, Misako

平成 10 年 10 月 1 日付けで国立がんセンターから転任いたしました。国立がんセンターには 19 年間所属していましたので、所属が変わった、ということにまだ慣れないでいます。今年は、新しい環境での確固とした地盤づくりをめざして、研究室を整えていきたい、と思っております。

連絡先　〒739-8526　東広島市鏡山１－３－１
広島大学理学部化学科
TEL: 0824-24-7412FAX: 0824-24-0725

Q-Chem: Quantum Chemistry Software for Large Systems

Peter M.W. Gill

Q-Chem, Inc.
Four Triangle Drive
Export, PA 15632, USA

and

Department of Chemistry
University of Cambridge
Cambridge, CB2 1EW, England

Introduction

The1998 Nobel Prize for Chemistry was awarded jointly to Professors Pople and Kohn for the development of computational methods in quantum chemistry and, in particular, for the development of density-functional theory (DFT). The award acknowledges the enormous strides that these men and their disciples have taken during the last three decades and celebrates the revolution in chemical philosophy and practice that their ideas have wrought.

Quantum chemical calculations are being utilised by a continually increasing range of computational chemists and, facilitated by impressive improvements in computer hardware, they can now be used to solve much more complicated problems than one would have thought possible a few years ago. Furthermore, it seems likely that these trends will continue in the coming years: computers will become faster and cheaper and the algorithms and software of quantum chemistry will be radically transformed as researchers strive to apply first-principles quantum mechanics to problems of biological scope and complexity.

In this article, we will review both the classical and DFT approaches to quantum chemistry and we will discuss the theoretical and implementational problems that have restricted the application of these approaches to small systems. Finally, we will describe some of the progress that the Q-Chem collaborators have recently made toward the solution of these problems.

Wavefunction-based approaches

In 1926, Schrödinger introduced wave mechanics, proposing that any physical system in a stationary state possesses a wavefunction Ψ and, for many years, the goal of classical quantum chemistry was to obtain useful approximate solutions to the time-independent electronic Schrödinger Equation

$$\hat{H}\Psi = E\Psi \tag{1}$$

The resulting models (Hartree-Fock, configuration interaction, coupled-cluster and Møller-Plesset theories) were the standard methods of the '70s and '80s and were widely used, especially by academic researchers. It is useful to express the total electronic energy E as

$$E = E_{\mathrm{T}} + E_{\mathrm{J}} + E_{\mathrm{X}} + E_{\mathrm{C}} \tag{2}$$

where these terms are the kinetic, Coulomb, exchange and correlation energies, respectively. In wavefunction-based approaches, each of these four energies is calculated as a functional of the wavefunction. The Fock exchange functional, for example, is given by

$$E_{\mathrm{X}}^{\mathrm{Fock}} = \frac{1}{2}\sum_{i}\sum_{j}\iint \psi_i(\mathbf{r}_1)\psi_j(\mathbf{r}_1)\frac{1}{r_{12}}\psi_i(\mathbf{r}_2)\psi_j(\mathbf{r}_2)\,d\mathbf{r}_1 d\mathbf{r}_2 \tag{3}$$

Density-based approaches

In the mid-60s, Kohn and his collaborators proved that it is theoretically possible to calculate E as a functional of the electron density $\rho(\mathbf{r})$ and, thereby, to avoid consideration of the wavefunction entirely. This was an important discovery because $\rho(\mathbf{r})$ is a much simpler object than Ψ. In Kohn-Sham theory, the kinetic energy E_{T} is computed from the orbitals (as it is in Hartree-Fock theory) but the remaining terms in (2) are calculated as functionals of $\rho(\mathbf{r})$. The Dirac exchange functional, for example, is given by

$$E_{\mathrm{X}}^{\mathrm{Dirac}} = \frac{3}{2}\left(\frac{3}{4\pi}\right)^{1/3}\int \rho^{4/3}(r)\, d\mathbf{r} \tag{4}$$

Many exchange/correlation density functionals have been developed.

Cost and Accuracy

As of1998, the most sophisticated wavefunction-based methods are more accurate than the most sophisticated density-based methods. Nonetheless, it is believed that roughly 90% of modern quantum chemistry computations use one of the DFT schemes. What is it that makes DFT so attractive?

The answer lies in the cost/accuracy ratios of the various methodologies. The computational cost of the very sophisticated wavefunction-based methods (such as the popular CCSD (T) model) increases extremely quickly as the system grows. In fact, the cost of CCSD (T) scales with the *seventh* power of the size of the system and, not surprisingly, CCSD (T) can therefore be applied only to relatively small molecules. This has been the key weakness of wavefunction-based approaches and even spectacular improvements in computer speed will not significantly improve the situation.

In contrast, standard DFT calculations typically yield errors that are 2–3 times larger than those from CCSD (T) and yet their cost grows only with the *third* power of system size. To a pragmatic scientist, this combination of reasonable accuracy at a modest computational cost is irresistible.

The Q-CHEM Software

Although DFT leapt into prominence in the early '90s and was quickly adopted by the majority of computational chemists, a number of significant problems remained unsolved and, in1993, Q-Chem Inc. was founded to address these. The company is currently owned by Peter Gill (University of Cambridge), Martin Head-Gordon (University of California, Berkeley) and Benny Johnson. The Board of Directors is growing but presently includes Dr Gill, Fritz Schaefer (University of Georgia) and John Stott (Eastside Holdings Ltd).

It was recognised from the onset that there were many reasons why it was best to plan the Q-CHEM package from scratch, rather than basing its structure and design on earlier programs. Modern software engineering practice argued that we should adopt an object-oriented programming language but the retention of FORTRAN was favoured on the grounds of compiler efficiency. As a result, the code is a mixture of high-level C++ and low-level FORTRAN routines. The code was designed to function efficiently on vector and superscalar architectures and, although the initial versions do not run on multiple processors, we were careful to employ algorithms and data structures that are well suited to the parallelisation efforts that we are now pursuing.

The first major problem that the Q-Chem collaboration sought to solve was the so-called Coulomb problem that arises when standard algorithms are used to compute the Coulomb energy

$$E_J = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 \, d\mathbf{r}_2 \tag{5}$$

Since the density $\rho(\mathbf{r})$ is usually represented as a sum of gaussian functions, straightforward methods for the evaluation of (5) involve a double sum over all pairs of gaussians and therefore scale quadratically with system size.

Q-Chem Inc has received funding from the US Department of Energy and the National Science Foundation to conduct research into the design of radically new algorithms that compute E_J with a computational cost that scales only *linearly* with system size. These research efforts have been successful and, in 1994, the Q-Chem collaborators published the first algorithm, the CFMM, that computes the Coulomb energy of a molecular electron density in linear work. Such algorithms are known as *O* (*N*) methods. More recently, the Cambridge team has developed another O (N) Coulomb method, the KWIK algorithm, that can compute exchange, as well as Coulomb, energies.

Q-Chem Inc is also involved in the development of local correlation methods. This research, which is supported by the US Air Force, is attempting to develop modified versions of the sophisticated wavefunction-based approaches (such as Møller - Plesset perturbation theory) whose computational costs grow with system size more slowly than the original versions. The Berkeley group is devising well-defined localisation schemes that will recover a large fraction of the correlation energy to be in a small fraction of the time that a normal calculation would require.

Another of the problems associated with conventional DFT algorithms is their reliance on numerical grids to estimate the difficult integrals that arise when one uses any of the standard density functionals (such as B-LYP, PW91, etc). The Cambridge group is about to commence a major investigation into the development of "gridless" density functionals, that is, ones that can be integrated without the need for a molecular grid. This work will be funded by the National Science Foundation.

Density functional theory, in its original form, was a ground-state paradigm. However, there is great interest, of course, in methods that enable molecular excited states to be studied. The Berkeley group has been conducting ground - breaking research in this area and is also currently working on the development of appropriate extensions of DFT for excited states.

The development of new, more accurate density functionals is one of the most important areas of research in quantum chemistry today and Q-Chem is playing a role in this unfolding story. Two years ago, we introduced the G96 functional which, although much simpler than the popular Becke 88 functional, yields similar model chemistries. More recently, in collaboration with Prof Pople, we have developed the EDF1 functional which is superior to the widely used B3LYP functional when used with a basis set of medium size.

Inorganic chemists, who need to understand and predict the chemistry of metal - containing systems, require quantum chemical software that offers a highly efficient treatment of effective core potentials (ECPs) or pseudopotentials. The Cambridge group is currently developing a new algorithm for the calculation of ECP integrals, using a suitably modified version of the PRISM algorithm for two-electron integrals that they introduced some years ago.

The first parallel version of Q-CHEM is now in use and has shown impressive speed-up efficiencies on SGI Origin computers. Continued development of this aspect of the code is currently being supported by the National Institutes of Health and we expect that future versions of Q-CHEM will be fully able to exploit the enormous potential of modern parallel machines.

Q-Chem Inc. has recently released the Q-CHEM 1.2package and it is expected that Q-CHEM1.3 will be released in the summer of 1999. Important theoretical breakthroughs are being made every few months and the Q-Chem collaboration is building vigorously on the solid foundations that Kohn and Pople have laid.

Literature references, timing benchmarks and other information can be found on the Q-Chem web pages at http://www.q-chem.com

大きな分子を精度よく計算できるONIOM法

基礎化学研究所　松原世明
matubara@ifc.or.jp

近年、計算機のめざましい発達とともに、分子の電子状態計算は精度が高いab initio MO法（非経験的分子軌道法）により盛んに行われるようになったが、依然として時間的制約は大きな問題であり計算できる分子中の原子数は限られる。そのため実際の分子をモデル化することもしばしばであるが、そうすることによって分子の特性や反応の本質が失われる場合も多い。大きな分子でもモデル化することなく実際の分子を高い精度で効率よく計算できるONIOM(our own *n*-layered integrated molecular orbital and molecular mechanics)法を紹介する。

1. はじめに

かさ高い置換基を有する有機金属錯体や有機化合物、あるいは不均一触媒、生体分子などのab initio MO法計算では、近年のめざましい計算機の発達があるとは言え、時間的制約から実際の分子をそのまま取り扱えないことが多く、モデル化という操作をよくする。しかしながら実際の分子のかさ高い置換基がその構造や反応性に深く関与している場合、モデル化した分子では解析できない。そこで大きな分子をモデル化することなく、より精度よく計算するためにONIOM法が開発された。ONIOM法は最初、MO法とMM（分子力場）法を組み合わせて、分子全体のエネルギーを求め構造最適化するIntegrated MO + MM(IMOMM)法として開発された。[1] この方法では、反応中心などの重要部分はab initio MO法、かさ高い置換基などのその他の部分はMM法で取り扱い、平衡構造ばかりでなく遷移状態の構造も効率よく最適化できる。その後、反応中心部分の外側の部分をMM法に限らず電子的効果も考慮したMO法を用いる、つまり、反応中心などの重要部分は高いレベルのMO法、その他の部分は低いレベルのMO法で取り扱う、Integrated MO + MO(IMOMO)法が開発された。[2] そしてこれらの手法は、より大きなシステムをより精度よく計算するために、分子全体を中心部分から外側へ順にn層に分け、n個のMO法とMM法の段階的計算レベルの組み合わせで分子全体のエネルギーを表わし構造最適化するというように一般化された概念の基に、ONIOM法として統一された。[3] 実際には分子を2層あるいは3層に分けるのが現実的であろう。以下にONOM法の原理およびその応用を紹介する。

2. ONIOM法の原理

ONIOM法では、より大きなシステムをより正確に計算するために分子全体をn層に分け、それぞれの層を異なったレベルのMO法とMM法で取り扱う。したがって計算対象となるシステムに応じて任意のMO法あるいはMM法を組み合わせることができる。分子を2層あるいは3層に分けたONIOM2およびONIOM3の概念をFigure 1に示す。分子を3層に分ける場合、反応が起こる中心部分、電子的効果の寄与が大きい中間部分および立体的効果の寄与が大きい外側の部分に分けられる。それぞれの層は中心部分より、高いレベル、中間レベル、低いレベルで計算される。2層に分ける場合は、中心部分は高いレベル、外側は低いレベルで計算される。そこで計算レベルは、例えば、

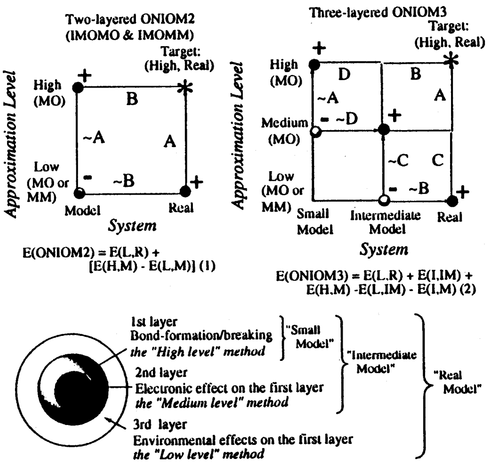


Figure 1. Concept of the ONIOM method.

ONIOM(CCSD(T):DFT:MM)//ONIOM(B3LYP/6-31G*:Charmm)と記述する。分子全体のエネルギーは、ONIOM3 の場合を例にとると Figure 1 の式(2)のように、高いレベル、中間レベル、低いレベルでそれぞれ計算される small model、intermediate model、real のエネルギーの和で表される。ここで intermediate model は small model と、また real は intermediate model と重なっているため、double counting を避けるためその重なった部分のネルギーは差し引く。これが ONIOM 法の根本的原理である。すなわち、Figure 1 に示したように、高い計算レベルでの real (target)のエネルギーを求めるための外挿をしている。構造最適化に必要な分子全体の gradient、hessian やその他の物性値も基本的に同様にして求められる。詳細は文献を参照していただきたい。[1-4] 次に ONIOM 法の応用例を紹介する。

3. ONIOM 法の応用

3.1 $Pt(PR_3)_2$への H_2の酸化的付加反応

$Pt(PR_3)_2$ への H_2 の酸化的付加反応は、基本的かつ重要な有機金属錯体反応の一つとして知られている。これまで ab initio MO 法では、PR_3をそのまま計算することは時間的に困難であるため PH_3に置き換えられて解析され、反応は発熱的で活性化エネルギーは数 kcal/mol であると報告された。[5] しかしながら、R=*i*-Pr, c-C_6H_{11} の場合には反応は進行するが、R=*t*-Bu の場合には反応は進行しないことが実験的に報告されている。[6] このようなホスフィン配位子に依存する反応性については、ホスフィン配位子である PR_3 を PH_3でモデル化すると解析できないが、

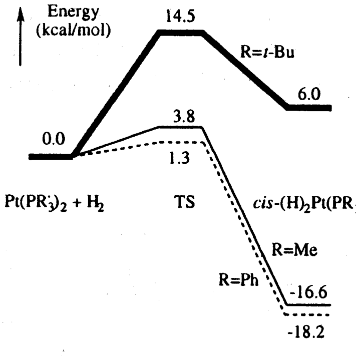


Figure 2. Potential energy surface of $Pt(PR_3)_2$ + H_2 by ONIOM2(MP2:MM3).

ONIOM 法を用いれば実際の分子を時間的制約を受けることなく計算できるため、解析が可能となる。[7] 反応サイトを含む $Pt(PH_3)_2$+H_2は MP2、かさ高いホスフィン置換基は MM3 で取り扱った ONIOM2(MP2:MM3)の計算結果を紹介する。反応のポテンシャルエネルギー面を Figure 2 に示す。R=Me, Ph の場合にはエネルギー障壁は小さく発熱反応、R=*t*-Bu の場合にはエネルギー障壁は大きく吸熱反応であり、実験事実とよく一致している。ONIOM2 の解析結果から、遷移状態構造を Figure 3 に示すように R=*t*-Bu の場合には、かさ高い *t*-Bu 基同士の立体反発のために P-Pt-P 軸が十分に折れ曲がることができない。そのため、H_2と Pt の間の電子の供与、逆供与による相互作用は十分でなく遷移状態は不安定であるため、反応が進行しないことが明らかとなっ

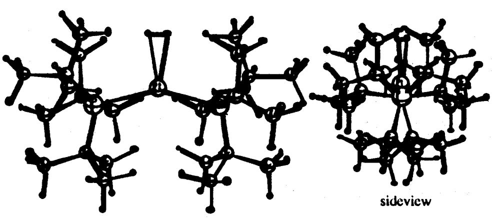


Figure 3. Optimized TS structure on $Pt(P(t\text{-}Bu)_3)_2$ + H_2 by ONIOM2(MP2:MM3).

た。また H_2と *t*-Bu基との立体反発のため、H-Pt-H 面と P-Pt-P 面との二面角は 32 度である。このように ONIOM 法により、R 置換基の遷移状態構造や反応への影響が明らかにでき、その計算時間は full ab initio MO 法と比較して約 1/250 で済む。

3.2 Diels-Alder 反応

ONIOM 法は有機分子や有機反応にも応用され、実験結果とよく一致することが確かめられている。[2,8] 例えば acrylic acid と 2,4-pentadienoic acid の Diels-Alder 反応では、-COOH 官能基の配置によって生成物は４つの異性体を持つが、その選択性は顕著に現れる。ONIOM 法により、反

応中心部分のエチレンおよびブタジエンについては極力高いレベル、その周りの-COOH官能基は適当な計算レベルを選択することでその反応選択性がよく再現されることが示された。[9] ここで重要なことは、計算レベルの組み合わせである。ONIOM3によるacroleinとisopreneあるいは2-*tert*-butyl-1,3-butadieneのDiels-Alder反応の計算例を次に示す。

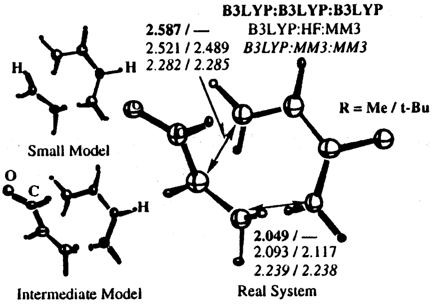


Figure 4. Optimized parameters of TS in the Diels-Alder reaction by ONIOM3 and the model of system.

Figure 4に示したように、small modelはエチレンとブタジエン、intermediate modelではそれに-COH官能基が含まれ、real systemでは残りのMeあるいは*t*-Bu基が含まれる。最適化構造を見てみると、遷移状態におけるお互いの距離は、ONIOM3(HB3LYP:MM3:MM3)よりもONIOM3(HB3LYP:HF:MM3)の方がずっとfull B3LYPに近い値を与えていることが分かる。このことは、この反応システムの場合アルキル基はMM法でよく表現できるが、-COHのような官能基には電子的効果を考慮した計算レベルが必要であることを示している。Table 1に示した反応の活性化エネルギーを見ると、アルキル基はMM3でよいが、-COH基には少なくともHF以上のレベルが必要であるあることが分かる。例えば、ONIOM3(CCSD(T):MM3: MM3)よりもONIOM3(CCSD(T):MP2:MM3)の方がfull B3LYPにより近い値を与える。このようにシステムの分割と計算レベルの選択は、計算精度に顕著に影響し、非常に重要である。

Table 1. Activation barrier (in kcal/mol) for the Diels-Alder reaction by ONIOM3.

High:Med:Low	2-methyl	2-*t*-butyl
B3LYP:B3LYP:B3LYP	19.2	16.4
B3LYP:B3LYP:HF	19.1	17.1
B3LYP:B3LYP:MM3	18.3	15.4
B3LYP:HF:HF	18.6	16.6
B3LYP:HF:MM3	17.8	15.0
B3LYP:MM3:MM3	21.9	18.6
HF:HF:HF	41.1	39.7
MM3:MM3:MM3	42.1	39.9
CCSD(T):MP2:MP2	19.6	16.9
CCSD(T):MP2:HF	19.9	18.9
CCSD(T):MP2:MM3	19.0	17.2
CCSD(T):MM3:MM3	24.7	22.6
MP2:MP2:MP2	12.7	10.1

3.3 有機分子の励起エネルギー

またONIOM法は有機分子の励起エネルギー計算にも応用できる。環状アルケンやエノンなどのn→π*あるいはπ→π*遷移エネルギーは、遷移に関連する二重結合を含む部分を高いレベルのMO、その他の部分を低いレベルのMOで取り扱うことにより、実験値とよく一致すること

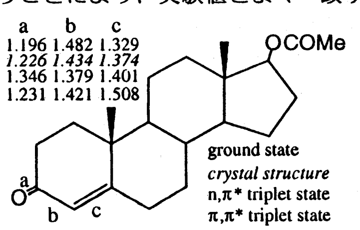


Figure 5. Optimized parameters of testosterone derivative by ONIOM2 (HF:MM3).

Table 2. Excitation enrgies (in kcal/mol) of testosterone derivative by ONIOM2.

	n,π*	π,π*
HF	42.3	44.0
HF:MM3	38.2	41.5
MP2:MM3	87.6	82.7
MP3:MM3	75.9	74.5
MP4:MM3	76.6	74.2
CCSD(T):MM3	73.9	73.0
G2MS:MM3	80.1	78.5
MP2:HF	88.9	85.1
MP3:HF	77.2	76.9
MP4:HF	78.0	76.6
CCSD(T):HF	75.3	74.4
G2MS:HF	81.4	80.9
Experiment	74.5	73.4

とが示された。[10] ONIOM2(HF:MM3)による testosterone 誘導体 ($C_{21}O_3H_{29}$) の基底状態および n→π*あるいは π→π*遷移による三重項励起状態の最適化構造パラメータをFigure 5に示す。4-methyl cyclohexenone をモデルとし、その部分はHFレベル、その他の部分はMM3で取り扱っている。励起状態ではC=C、C=Oの結合距離は伸び、その間のC-Cの結合距離は縮んでいるのが分かる。色々な計算レベルの組み合わせでの励起エネルギーをTable 2に示す。反応中心部分は高いレベルが必要であるが、modelである反応中心部分は周りの電子的効果を充分含むよう大きくしているので、その他の部分はMM3でも良く記述できている。例えば、ONIOM2(CCSD(T):MM3)はONIOM2(CCSD(T):HF)同様、実験値に非常に近い値を与えている。これらの構造最適化を含めた計算時間は40時間と非常に現実的であり、ONIOM法はこのような大きな分子を効率よくかつ精度よく計算できる。

4. 終わりに

ここで紹介したONIOM法の応用はごく一部であるが、その他の有機金属錯体反応や、触媒反応などにも応用されている。[11] また、ONIOM法はすでにGaussian98に組み込まれている。大きな分子を時間的な制約をうけることなく高精度に計算できるONIOM法は、かさ高い置換基をもつ有機化合物や有機金属錯体、不均一触媒、生体分子などへの幅広い応用が期待され、今後の巨大分子の理論的解析に新たな展望を与えるもである。

参考文献

1. F. Maseras, K. Morokuma, *J. Comput. Chem.*, **16**, 1170 (1995).
2. S. Humbel, S. Sieber, K. Morokuma, *J. Chem. Phys.*, **105**, 1959 (1996).
3. M. Svensson, S. Humbel, R. D. J. Frowsw, T. Matsubara, S. Sieber, and K. Morokuma, *J. Phys. Chem.*, **100**, 19357 (1996).
4. (a) S. Dapprich, I. Komaromi, K. S. Byun, K. Morokuma, *J. Mol. Struct. (Theochem)*, submitted. (b) I. Komaromi, S. Dapprich, K. S. Byun, K. Morokuma and M. J. Frisch, *J. Mol. Struct. (Theochem)*, to be submitted. (c) R. D. J. Froese and K. Morokuma, *Encyclopedia of Computational Chemistry*, Ed. by P. v. R. Schleyer, Wiley (1998), in press.
5. (a) K. Kitaura, S. Obara, K. Morokuma, *J. Am. Chem. Soc.*, **103**, 2891 (1981). (b) Obara, S.; Kitaura, K.; Morokuma, K. *J. Am. Chem. Soc.*, **106**, 7482 (1984). (c) Noell, J. O.; Hay, P. J. *J. Am. Chem. Soc.*, **104**, 4578 (1982). (d) Low, J. J.; Goddard III, W. A. *J. Am. Chem. Soc.*, **106**, 6928 (1984). (e) Low, J. J.; Goddard III, W. A. *Organimettalics*, **5**, 609 (1986). (f) Kitaura, K.; Obara, S.; Morokuma, K. Chem. Phys. Lett., **77**, 452 (1981).
6. (a) Yoshida, T.; Otsuka, S. *J. Am. Chem. Soc.*, **99**, 2134 (1977). (b) Fornies, J.; Green, M.; Spencer, J. L.; Stone, F. G. A. *J. Chem. Soc. Dalton Trans.*, 1006 (1977). (c) Paonessa, R. S.; Trogler, W. C. *J. Am. Chem. Soc.*, **104**, 1138 (1982). (d) Packett, D. L.; Jensen, C. M.; Cowan, R. L.; Strouse, C. E.; Trogler, W. C. *Inorg. Chem.*, **24**, 3578 (1985). (e) Yoshida, T.; Yamagata, T.; Tulip, T. H.; Ibers, J. A.; Otsuka, S. *J. Am. Chem. Soc.*, **100**, 2063 (1978).
7. T. Matsubara, F. Maseras, N. Koga, K. Morokuma, *J. Phys. Chem.*, **100**, 2573 (1996).
8. (a) T. Matsubara, S. Sieber, K. Morokuma, *Int. J. Quantum Chem.*, **60**, 1101 (1996). (b) M. Svensson, S. Humbel and K. Morokuma, *J. Chem. Phys.*, **105**, 3654 (1996). (c) R. D. J. Froese, J. M. Coxon, S. C. West, and K. Morokuma, *J. Org. Chem.*, **62**, 6991 (1997).
9. R. D. J. Froese, S. Humbel, M. Svensson, K. Morokuma, *J. Phys. Chem.*, **101**, 227 (1997).
10. R. D. J. Froese, K. Morokuma, *Chem. Phys. Lett.*, **263**, 393 (1996).
11. (a) Y. Wakatsuki, N. Koga, H. Werner, K. Morokuma, *J. Am. Chem. Soc.*, **119**, 360 (1997). (b) R. D. J. Froese, D. G. Musaev, and K. Morokuma, *J. Am. Chem. Soc.*, **120**, 1581 (1998). (c) D. G. Musaev, R. D. J. Froese, K. Morokuma, *Organometallics*, **17**, 1850 (1998).

まつばら　としあき　MATSUBARA, Toshiaki
[専門] 理論化学　[連絡先] 〒606-8103 京都市左京区高野西開町 34-4 基礎化学研究所 電話代表075-711-7708

MOZYMEの蛋白質科学への利用法を探る

東京工業大学生命理工学部　櫻井　実
msakurai@bio.titech.ac.jp

1996年、MOPACの開発者として有名なJ. J. P. Stewartにより極めて斬新な発想に基づく分子軌道計算の新しい手法MOZYMEが発表された[1)]。MOZYMEは蛋白質や一般の高分子の全原子を含む半経験的分子軌道計算を可能にする。筆者らは、この２年あまりMOZYMEの販売元である富士通（株）の和田、鮫島両氏の協力を得て、このプログラムをいかに生体分子の研究に役立たせるか模索してきた。本稿では、筆者らの経験談をいくつかのデータを交えて紹介する。

1. MOZYMEとは？

MOZYMEという名前を初めて聞いた読者は、これはいったい何のことかと思うだろう。発音してみると何となくEnzymeに似た響きがあるので、酵素を扱う何かのソフトウェアだろうか、ぐらいは想像がつくかもしれない。まさにその通りである。これはMolecular OrbitalのMOとEnzymeの合成語であり、元々酵素の分子軌道計算を狙って開発されたソフトウェアである。ただし、こう書くと多少誤解を招く恐れもあるので、もう少し説明させていただく。分子軌道計算では、入力データとしてその分子を構成するすべての原子の位置座標をinputするわけであるが、プログラムの側からみればその分子が蛋白質（酵素）であるのかDNAであるのか、あるいは合成高分子であるのか制限はない。したがって、MOZYMEを使えば蛋白質に限らず、一般の巨大分子の分子軌道計算が実行できると考えてよい。要するに、MOZYMEは本号の特集である一群の”巨大分子の計算をめざしたソフトウェア”の一つである。

それでは、MOZYMEではどれくらいの原子数の分子まで適用可能なのであろうか。この問題に正確に答えるには、もちろん使用するコンピュータ、許容される計算時間、計算の種類（1SCFか構造最適化をやるか否か）などを指定しなければならない。ここではわれわれの計算環境の基での大雑把な見積もりを示すことにしよう。われわれは東工大計算センターのDEC Alpha Server (450MHz)を使用しているが、MOZYME計算はすべて1CPUで実行している。これまで行った最大の蛋白質はバクテリオロドプシン（約２５０残基、原子数３０００弱）であるが、1SCFならばおよそ？時間程度で終了する。構造最適化の時間は初期構造に依存するので一概には言えないが、われわれはこの蛋白質の全水素原子の位置を決めるのに２１日CPU時間の演算をやったことがある。通常の研究では、このあたりが許容演算時間の上限であろう。とすると、蛋白質全原子の位置を最適化するというのはまだ現実的ではなく、最適化する部分を機能上重要なサイトやドメインに絞るとか二面角のみを動かすとかの工夫が必要であろう。バクテリオロドプシンをもう一度例に引くと、発色団レチナールを取り囲む機能上重要なアミノ酸２１残基の構造（側鎖の配向と水素原子の位置）を最適化するのに、およそ２４時間を要した。以上よりわれわれは、MOZYMEの守備範囲はおよそ５００原子ぐらいから高々数千原子の分子までと考えている。下限を設けた理由は、MOZYMEでは小さな分子を扱えないという意味ではなく、それ以下は従来のMOPACで十分に対応できるのではないかと考えるからである。

MOZYMEアルゴリズムの特徴を一言で言えば、”Distance Cutoffを用いた局在化分子軌道法の利用”ということになろう。分子軌道法に疎い読者には何のことかわからないかもしれないが、要するに、大きなサイズの問題を解くのに一度に全体を処理することをせずに、まずサブスペース毎に厳密に解いておいて、後でサブスペース間のつじつま合わせをする、というような意味に解釈してほしい。Distance Cutoffは厳密に扱う範囲を決める。結果として、MOZYMEでは（1SCFの）演算時間は原子数Nの高々１乗程度に抑えられている。また、通常の分子軌道計算では大行列の対角化を実行する

必要があるが、MOZYMEではこの部分が局在化軌道に対するannihilationという操作に置き換えられている。そのためメモリ容量も節約できる。アルゴリズムの詳細に興味をお持ちの読者は、文献1) を参照されたい。

2. MOZYMEの蛋白質科学への応用

上で述べたように、MOZYMEの登場により半経験的分子軌道計算という制限はつくが、蛋白質の電子構造を低分子と同等な近似レベルで研究することが可能となった。これは理論化学の視点からみれば大きな進歩であり、実に魅力的であるように思える。しかしこのことが蛋白質の構造や機能の分子論的理解にどのような寄与をするのか、最初から誰の目にも明かだったわけではない（実を言うと未だに不明な点も多い）。例えば、MOZYME計算を行えば蛋白質全原子の電荷が評価できる。おそらくそのとき得られる各アミノ酸原子上の電荷は、そのアミノ酸1分子あるいはそのアミノ酸を含む蛋白断片に対して得られる値とはわずかに異なるに違いない。そのとき純粋な理論家ならば、やっとわれわれ人類は本当の蛋白の電荷分布を知ることができたと主張するかもしれない。しかし、蛋白質屋は、その差が本当に機能上意味があるのだろうかという疑問を抱くであろう。このことは蛋白質科学の分野で盛んに使われている古典的分子動力学計算を考えてみればわかる。この場合原子電荷はモデル化合物の計算値から引用するが、普通はそれで不都合は生じていない（ようにみえる）からである。筆者の知る限り、現時点でMOZYMEを使った蛋白質の研究として公表されているものは、われわれのものを除いてほとんど皆無でないかと思う[2]。われわれの研究も未だ模索段階であり、読者に十分なデータをお見せすることはできないが、全原子を取り込んだ計算をしてみてはじめてわかった蛋白質の真の姿の一端を以下で紹介しよう。

2.1 MOZYMEデータの一次的利用 -蛋白質の電荷分布-

上で電荷分布のことを述べたので、まずこの点から始めよう。われわれは、予てよりロドプシンやバクテリオロドプシンをはじめとするレチナール蛋白質の発色機構をぜひ理論的に解明したいと考えてきた。レチナールという分子は、ビタミンAの誘導体であるが有機溶媒に溶かすと380 nm付近に吸収極大を示す（プロトン化シッフ塩基（図1）の場合は440 nm）。ところが、蛋白質に結合すると吸収極大は長波長側へ大きくシフトする。視細胞中に存在し明暗視を司るロドプシンでは、吸収極大は500 nmである。このシフトのお陰で脊椎動物は、いわゆる可視光域の光を感じることができるわけである。レチナール蛋白質の吸収極大調節機構に関しては諸説あるが、中でも蛋白質の作る静電場がレチナールの共役二重結合系のπ電子を非局在化する役割を果たしている、という仮説が有力であった。そこでわれわれはその真偽を確かめるため、バクテリオロドプシンのレチナール結合ポケットの静電ポテンシャルをMOZYMEによる電荷情報を使って調べてみた[3]。構造はPDBに登録されている2BRDを使った。ただし、水素原子の位置は1で述べたように21日かけて最適化した。まず、全原子を含む計算からすべての原子上の電荷を得た後レチナール部分の電荷を取り除き、残った蛋白部分の電荷がレチナール結合部位に作る静電ポテンシャルを図示した（図2）。この図では少々わかりにくいかもしれないが、イオノン環側のポテンシャルが比較的高く、シッフ塩基結合に向かうにしたがって徐々に低くなる傾向が見い出された。つまり、共役鎖に沿って緩やかなポテンシャル勾配が生じているのである。π電子を非局在化させるためにはシッフ塩基窒素上のプロトンをイオノン環に向けて引っぱる場があると都合がよいが、図2の結果は全く逆である。すなわち、蛋白の作る静電場は、期待に反して発色団の吸収を短波長シフトさせるはずであるという結果が

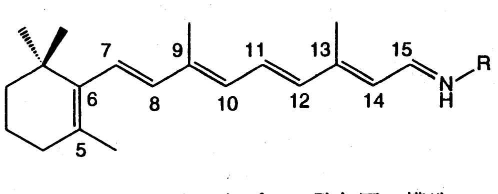


図1 バクテリオロドプシン発色団の構造
all-trans型レチナールがLys216の側鎖にプロトン化シッフ塩基結合している。

出てしまった。実際、蛋白の静電ポテンシャルをあらわに考慮して吸収波長をINDO/Sという方法で計算してみると、60 nm近くも短波長側へシフトしてしまったのである。これはいったいどういうことなのか？元々の仮説が間違っていたのか、それとも、MOZYMEの計算がおかしいのか？結論を言えば、前者である。残念ながらMOZYMEの電荷が完全に正しいことを証明する手だては今のところ無いので、われわれはこの結論を導くためさらに四苦八苦する羽目になるのだが、新しい方法論はこのような状況から生まれるものである。これはMOZYME計算の間接的効用といってよい。その点については2.3節で述べることにして、ここでは図3を見ていただこう。この図は、レチナールが結合していない蛋白質部分のみのMOZYME計算から得られた静電ポテンシャルと図2の差をとったものである。つまり、この図はレチナール（プロトン化シッフ塩基）が結合することによって起こる蛋白質部分の電子分極を反映している（蛋白部分の構造は図2と同じである）。これを見ると大きいところでは15％程度の変化を生じている。したがって、予想以上に蛋白の電子は動きやすいものではないかということがわかる。前にMOZYMEによる電荷計算は、蛋白質科学にとって意味のあるものなのかどうかという問題を提起したが、この結果によればyesという答えが返ってきそうである。実は、上のバクテリオロドプシンの吸収波長はこの蛋白質の電子分極を考慮してはじめて説明できることがわかってきた。レチナールのプロトン化シッフ塩基は、基底状態と励起状態（B_1状態）で大きく電子分布を変える。励起状態では、双極子モーメントが非常に大きくなることが実験的にも証明されている。もし、このようなレチナールの電子分布変化に追随して蛋白部分も電子分極すれば、励起状態が安定化されることになるので吸収極大は長波長シフトするはずである。上で述べた図2に基づく吸収波長計算では、基底状態で最適化された蛋白の電荷をそのまま励起状態の計算に使っていたので、励起状態で本来起こり得る安定化が考慮されていなかったのである。蛋白質の電子分極能は、巨視的物理量に置き換えるとすれば屈折率の2乗に対応する。われわれは、バクテリオロドプシンのレチナール結合ポケットの有効屈折率を、Lorentz-Lorentzの式と立体構造（2BRD）より約1.51と見積もった。この結果とわれわれが独自に開発した溶媒効果理論を使って、この蛋白質の吸収極大を計算したところ実験値を見事に再現した。この仕事はMOZYMEとは直接関係無いので詳細は省くが、興味ある方は文献4,5)を参照していただきたい。

以上から得られる”教訓”は、蛋白質の原子電荷を固定してしまうと問題によっては不合理な結果が生じる、ということである。上では励起状態が関与していたが、基底状態の問題でも次のような場

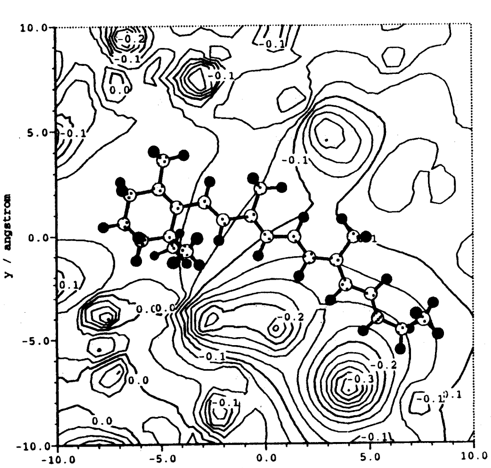


図2 レチナール結合部位の静電ポテンシャル

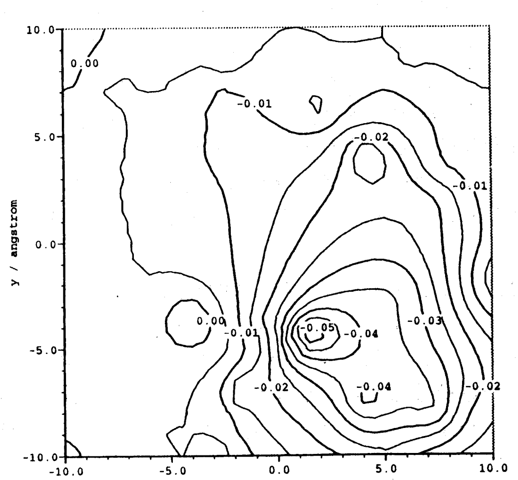


図3 蛋白の電子分極による静電ポテンシャル

合注意を要する。例えば蛋白質中における解離性アミノ酸残基のpKaを評価する場合、解離後の蛋白の電荷は解離前と有為に異なる可能性がある。

2.2 MOZYMEデータの一次的利用 -蛋白質のHOMO、LUMO-

1節で述べたようにMOZYMEでは基本的に局在化分子軌道を用いているため、デフォルトの計算で得られるSCF軌道は通常の分子軌道計算で得られるような正準軌道ではない。正準軌道に移行するためには一度ユニタリ変換を行う必要がある。MOZYMEにはこの機能がついているが、バクテリオロドプシン程度になると4GB弱のメモリーを必要とし、そう安々とできる計算ではない。しかし、われわれは蛋白質のHOMO、LUMOがどのように形成されているのかぜひ見たいと思い、いくつかの蛋白質についてこの計算をやり遂げ、これらの軌道を画像化するのに成功した。一応もっともらしい研究目的をつければ以下の通りになる。

周知の通り蛋白質や酵素は、アミノ酸の一次配列で決まる固有の高次構造をもっている。固有の構造が必要な理由は、一つに、酵素であれば特定の基質を立体選択的に結合するためであり、もう一つは、官能基を有するアミノ酸側鎖を触媒活性の発現に都合のいいように基質結合部位に配置させるためである。しかしながら、これらはX線回折実験等の結果に基づいた旧来の構造生物学的な見解であり、酵素を単に積み木細工に見立てた解釈にすぎない。実際の触媒反応過程を考えてみると、酵素は基質と反応しそれ自身の結合の開裂・生成の後、最終的には始状態に戻る。すなわち、酵素それ自身も自己完結的な反応分子として働いているわけである。一般に、化学反応では、反応分子のフロンティア電子軌道が反応機構や経路を決めるのに本質的な役割を果たしている。酵素においてもこの原理は当然成立するはずであるが、この場合フロンティア軌道は主として活性部位を形成する原子軌道により局所的に形成されると考えるのが自然である。一方、アミノ酸配列が与えられれば高次構造は決まるので、電子軌道も一義的に決まるはずである。したがって、以下の作業仮説が導かれる。

《作業仮説》分子論的視点からみれば、蛋白質・酵素の高次構造はフロンティア軌道を活性中心に閉じ込めるという役割をもっており、さらに言えば、分子進化はこの原理を満たすべく起こったのではないかと予想される。

この仮説の真偽を確かめリゾチーム、バクテリオロドプシンについてのHOMO、LUMO軌道を画像化した。これらの図はモノクロで提示しても意味が無いのでここでは割愛するが、おおよそ次のことが判明している。リゾチームの場合、活性部位に存在するAsp101、Asp52の側鎖領域にHOMOが出現し、その軌道エネルギーは他の占有軌道に比べ1eV以上高かった。これらの残基は、基質との水素結合および反応中間体の安定化に本質的な役割を果たすことが知られており、この酵素について上の仮説の妥当性が示された。バクテリオロドプシンについても、同様に活性中心に存在する発色団レチナールとMet118残基上に、それぞれLUMO、HOMOが局在することが判明した。この事実は、発色団の光化学的性質（吸収極大含む）がMet118により左右されるという実験事実に一致する。今後より多くの酵素について計算を進めるつもりであるが、このような計算から蛋白質の立体構造デザインに関する新しいコンセプトが得られると期待している。

2.3 MOZYMEデータの二次的利用

われわれは、2.1節で述べたバクテリオロドプシンの発色機構の解釈が妥当であるかどうかを判断するため、以下のような新しい計算手法の構築まで行き着いた。この計算法では、MOZYME計算からの電荷分布情報は表にでないが二次的に使われている。ここではMOZYMEがどのように使われているか知っていただければ十分なので、概念的な説明を図4にしたがって行う。

a) 蛋白質全体を量子力学で扱うregion Iと古典的に扱うregion IIに分割する。bRの場合、前者を発色団（プロトン化レチナールシッフ塩基）、後者をその周囲の蛋白質とする。

b) 基底状態は、region Iの波動関数$\Psi_0{}^I$とregion IIに属するすべての原子のatomic charge $\mathbf{Q^{II}}$を用いて記述される。$\mathbf{Q^{II}}$の値は、予めregion IとIIを含む全体系に対する分子軌道計算（MOZYME法）により求めておく。

c) i 番目の励起状態は、region Iの波動関数$\Psi_i{}^I$とregion IIのatomic charge $\mathbf{Q^{II} + q^{II,\mathit{i}}}$により記述される。すなわち、region Iが光を吸収する（$\Psi_0{}^I \rightarrow \Psi_i{}^I$）とそれに伴いregion IIに電荷が誘起される（$\mathbf{Q^{II} \rightarrow Q^{II} + q^{II,\mathit{i}}}$）。ここではFranck-Condon状態を想定しているので、 $\mathbf{q^{II,\mathit{i}}}$としては電子分極の寄与のみを考える。この誘起電荷はregion Iへ反作用場を跳ね返し、$\Psi_i{}^I$と相互作用するので、 $\mathbf{q^{II,\mathit{i}}}$と$\Psi_i{}^I$は互いにself-consistentになるように評価する。われわれの理論では、これらの量が一電子励起配置間相互作用計算の中で求められるように定式化されている。詳細は省くが、$\mathbf{q^{II,\mathit{i}}}$ベクトルはregion II 中の全原子数に等しい次元をもち、その成分は各原子上に誘起される電荷である。この値を評価するためには、regionIIに対する別のモデル化が必要であるがここでは割愛する。

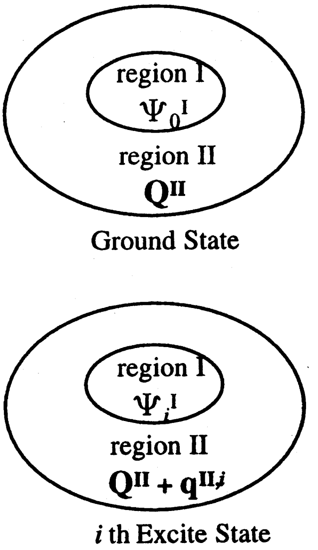


図4 理論の概念的説明

要するに、基底状態の電荷分布についてはMOZYMEから得られたものを使い、発色団の励起に伴う蛋白質部分の電子分極は別の近似を使って補正しようというものである。こうして得られたバクテリオロドプシンの吸収極大は実験値によく一致した。この方法は様々な色素蛋白質の発色機構の解明へ応用できると思っている。

文献

1) Stewart, J. J. P., J. Comp. Chem., **58**, 133-146 (1996).

2) MOZYMEとは別に、 巨大分子の計算法としてdivide-and-conquer methodというのがある。Merzらはこれを用いて蛋白質と水の間の電荷移動を計算している。興味ある読者は次の文献を参照されたい。Nadig, G; Van Zant, L. C.; Dixon, S. L.; Merz, Jr., K. M., J. Am. Chem. Soc., 120 , 5593-5594 (1998).

3) Houjou, H.; Koyama, K.; Wada, M.; Sameshima, K; Inoue, Y.; Sakurai, M., Chem. Phys. Lett., 294 , 162-166 (1998).

4) Houjou, H.; Sakurai, M.; Inoue, Y., J. Chem. Phys., 107, 5652-5660 (1997).

5) Houjou, H.; Inoue, Y.; Sakurai, M., J. Am. Chem. Soc., 120, 4459-4470 (1998).

さくらい　みのる　SAKURAI, Minoru

研究テーマ：レチナール蛋白質の量子化学、溶媒効果理論の構築と生体系への応用、生体分子の計算機シミュレーション、細胞のストレス耐性と糖の役割（特に、トレハロースという糖に興味をもっている。このテーマではNMR、IR、熱分析等の実験も行っています）

連絡先　〒226-8501　横浜市緑区長津田町4259　東京工業大学生命理工学部　tel: 045-924-5795

◆部会記事

第10回ケモメトリックス・ワークショップ開催報告

キッコーマン（株）研究本部　相島鐵郎
LDH00052@nifty.ne.jp

スウェーデン、Umeå 大学の Svante Wold 教授はケモメトリックスの創始者で、数年前に亡くなった PLS 手法を開発した統計学者 Herman Wold の息子である。彼自身 PLS 手法を自然科学分野に広めると共に、線形判別分析と共に多用されるパターン認識手法、SIMCA の生みの親でもある。その高名ぶりから大御所を予想したが、実際にお会いしてみると、世界中を飛び歩く 50 半ばの精力的な研究者であった。また、欧米に多い成功した科学者の例にもれず、数学を専門とする Nauna 夫人と共に創立した"UMETRI"というソフトウェア会社の社長も兼ね、アメリカ支社とスウェーデンで半年ずつを過ごしながら化学会社のコンサルタントも勤めるという多才、多忙ぶりである。今回は本ワークショップのためにわざわざ来日され、11 月 27 日（金）10～17 時、日本化学会会館 5 階会議室において、「多変量モデル化と検量法における最近の発展：直交化信号補正法と wavelet 圧縮法の分析と生産工程への利用 」について下記内容で講演された。

1. 研究と生産工程におけるケモメトリックス
   - 複雑な系とデータの爆発的増加
   - 問題の定型化 - 目的、測定法
   - 計画的な実験法 - 実験計画法と因子
   - 複雑な系の測定 - 多変量解析と試料収集
2. 最新手法
   - より優れた前処理 - 直交化信号補正法(OSC)
   - 情報損失のないデータ圧縮 - wavelet 法
3. 分析への応用
   - 多変量検量法
   - 信号行列データの整理、OSC、wavelet、スペクトル圧縮
4. 製造工程の監視と問題の初期発見
   - PCA と PLS を利用したデータ行列からの情報豊かな少数要約変数の導出
   - 連続工程とバッチ工程
5. 要約と結論
   - ケモメトリックス - デザインと多変量解析の融合

約 50 名の出席者のほとんどは食品、化学、製薬、測定機器など民間企業からの研究者で、今回も大学関係者の参加は少なかった。Wold 教授も講演の中でケモメトリックスの普及動向にふれ、民間に比べて大学関係者が余り興味を示さない傾向が強いのは万国共通であると述べた。これは問題を単純化してアプローチすることが可能な大学の研究に比べ、複雑な問題をそのまま扱わざるを得ない企業においては、必然的にケモメトリックス手法を利用せざるを得ないから、とのことである。日本また然りということになる。

後半には、Nauna 夫人がソフトウェア紹介をかねて、情報損失の少ないデータ圧縮法、多変量検量法、パターン認識手法及び応答曲面法による最適化法について、具体的なデータを使用しながら分かり易くデモンストレーションした。講演終了後、ほぼ出席者全員が参加した 6 階会議室における約 2 時間の懇親会は、参加者どうしの意見交換はもとより、アルコールの効果もあり演者に対して様々な質問も発せられ非常に有意義なものであった。また、講演者が Wold 教授とのことで、夫妻を囲んでの記念撮影まで行われる過熱ぶりであった。3 日後の月曜日にはパリでセミナーを開くとのことで、翌 28 日には離日された。

あいしま　てつお　AISHIMA, Tetsuo
連絡先　〒278-0037 千葉県野田市野田 399 キッコーマン（株）研究本部　電話 0471 23 5980

第３回分析化学のためのケモメトリックス討論会及び初心者のための分析化学に利用されるケモメトリックス講習会

主催　日本化学会情報化学部会
共催　日本分析化学会中部支部・ＣＡＣフォーラム

会期　４月２２日（木）～２３日（金）10時～16時30分
会場　名古屋市工業研究所〔名古屋市熱田区六番三丁目4-41、電話：052-661-3161 交通：地下鉄名城線「六番町」駅下車、徒歩３分〕

１日目：４月２２日（木）初心者のための分析化学に利用されるケモメトリックス講習会内容

1.定性分析に利用される判別分析、ＫＮＮ、ＳＩＭＣＡ、クラスター分析、主成分分析
2.定量分析に利用される重回帰分析、主成分分析、ＰＣＲ、ＰＬＳ
3.ニューラルネットワーク
4.実習（Ｃｈｅｍｉｓｈ　・　市販ソフト）

２日目：４月２３日（金）第３回分析化学のためのケモメトリックス討論会プログラム

1.多変量解析法を応用したポリエステルー混紡糸の混紡率の推定（愛知県警科捜研）○奥山修司・三井利幸
2.多変量解析法による顕微FT-IRデータからの無煙火薬の組成分析（愛知県警科捜研）○肥田宗政・三井利幸
3.ニューラルネットワークによるプラスチックの判別（物質研）田辺和俊
4.トレハロースー水素における特異的相互作用のＰＣＡ法による検討（日本分光）○大久保優晴・赤尾賢一・眞砂央
5.［依頼講演］主成分分析法による微量物質検知（神戸薬大）長谷川健
6.FT-Raman分光法を用いた木材各種組織比率の非破壊的定量（王子製紙森林資源研）○小名俊博・園田哲也・伊藤一弥・柴田勝
7.［依頼講演］Quality of Rice and Chemometrics(Queensland University of Technology)Serge Kokot
8.３Ｄ－ＱＳＡＲにおけるKohonenmapの応用と３way　ＰＬＳによる解析（豊橋技科大）○松岡成郎・荒川正幹・船津公人
9.近赤外－ケモメトリックス及びＦＴラマン－ケモメトリックス法によるポリマーブレンドの非破壊定量分析大阪府警科捜研１・大阪市工研２・関学大理３）○松下晃１・松川公洋２・井上弘２・南幸男１・尾崎幸洋３
10.近赤外スペクトルの解析におけるケモメトリックスと二次元相関分光法の関係（関西学院大学）○尾崎幸洋・村山幸市
11.［依頼講演］ＰＹ－ＧＣを用いたポリマーキャラクタリゼーションにおける未知ポリマー高速検索アルゴリズムの検討（フロンティアラボ１・名大院工２）○渡辺忠一１・大谷肇２・柘植新２
12.主成分分析法による定量分析の改良（愛知県警科捜研１・ジーエルサイエンス２）○三井利幸１・高木肇２・山地武広２
13.［依頼講演］市販点眼剤中香料成分の分析と統計手法の応用（ライオン（株））○判野有子・小池茂行

申込方法：氏名、所属、住所、電話番号、いずれに参加するかを記入の上、官製ハガキ、FAXまたはE-mailにてお申し込み下さい。
申込締切日：3月20日
講習会参加費：主催、共催学会会員6,000円、学生4,000円、非会員8,000円を当日受付にてお支払い下さい。
討論会参加登録費：主催、共催学会会員3,000円、学生1,000円、非会員4,000円を当日受付にてお支払い下さい。
申込先：〒460-0001　名古屋市中区三の丸２－１－１　愛知県警察本部科学捜査研究所三井利幸
［電話052-951-1611（内線3791,FAX 052-953-9629,E-mail:tmit@ma3.justnet.ne.jp］

◆関連行事

第76春季年会(1999)プログラム抜粋[会場別]

G5 会場
20号館207番教室

化学情報・計算機化学

(分子シミュレーション)

3月28日午前

座長 田辺 和俊(10:00～10:50)

1G507 新しい応用分子動力学計算プログラムの開発と仮想現実感環境の構築(東北大院工)○三浦隆治・稲葉祐策・小野津崇之・高見誠一・久保百司・宮本 明

1G508 モンテカルロ法，および分子動力学法による膜分離シミュレーション(東北大院工)○小林泰則・水上浩一・高見誠一・久保百司.・宮本 明

1G509 水溶液中におけるグリシン双性イオンの構造最適化−自由エネルギー勾配法−(基礎化研)○片岡 健・奥山直人・長岡正隆・山邊時雄

1G510 GCMC法によるQCM匂いセンサの応答予測(東工大)中村賢一・中本高道・森泉豊栄

1G511 シリコン酸化膜の応力特性の計算化学的検討(東北大院工・金材研)○山田有場・杉左近潔・遠藤 明・Cheettu Ammal Salai・高見誠一・久保百司・宮本 明・板倉明子・北島正弘

(化学情報・ケモメトリックス)

3月28日午後

座長 三上 益弘(13:10～14:20)

1G526 MoldaNet-Secure Signed Applet と Java 3D技術を用いたネットワークに分散された分子モデリングプログラムの開発(広島大理・インペリアルカレッジ)○吉田 弘・Rzepa H.S.・Tonge A.P.

1G527 主成分分析法による定量分析の改良(愛知県警科捜研)○三井利幸・肥田宗政

1G528 多変量解析法による顕微FT-IRデータからの無煙火薬の組成分析(愛知県警科捜研)○肥田宗政・三井利幸

1G529 有機合成設計システムAIPHOSの開発(18)−クライアント−サーバーシステム化の検討(富士フイルム足柄研)○竹内雅志・船津公人

1G530 ニューラルネットワークによる代替フロン探索(物質研)○田辺和俊・松本高利・長嶋雲兵・吉本昭二・上坂博亨・山本博志

1G531 ニューラルネットワークによる発ガン性予測(物質研)○松本高利・田辺和俊

1G532 ニューラルネットワークによる生分解性プラスチックの判別(物質研)○田辺和俊・松本高利・天野敏男・上坂博亨

座長 有園 尚(14:20～14:50)

1G533 **依頼講演** コンビナトリアルライブラリーデザインにおけるデータベースの利用(日本MDL)岸本隆太郎

(計算化学・理論化学)

座長 高田 章(15:00～15:30)

1G537 **依頼講演** エレクトロニクス材料の原子レベル設計(東北大院工)久保百司

座長 船津 公人(15:30～16:50)

1G540 アルミニウム化学気相成長におけるジメチルアルミニウムハイドライドの水素修飾シリコン表面への吸着に関する理論的研究(東大工)○松脇崇晃・中島 徹・山下晃一

1G541 ジメチルエーテルとOHラジカルの気相反応に関する量子化学計算(東大工)○本村和也・高橋崇宏・山下晃一

1G542 計算化学による振動スペクトルの予測(3)-非経験的分子軌道法による塩素化メタン分子の溶媒効果を含めた赤外スペクトルの計算(物質研)○田辺和俊・松本高利

1G543 電極表面での吸着構造と電荷移動に関する理論的研究(東大工)○亀谷和久・大脇 創・山下晃一

1G544 遷移状態理論によるフッ素化メタン，エタンとOHラジカルとの反応速度の計算(物質研)○杉江正昭・内丸忠文

1G545 アセチレンの超励起状態のダイナミクスに関する理論的研究(慶大理工)○安池智一・藪下 聡

1G546 分子とゼオライト骨格の相互作用−反応指数研究(東北工研)○Chatterjee Abhijit・岩崎孝志・蛯名武雄

1G547 エネルギー選別光解離反応を誘起する制御レーザー場の設計(東北大院理)○黒澤由佳・大槻幸義・藤村勇一

P 会場
13号館体育館

化学情報・計算機化学

3月28日午前

(11:00～12:30)

1PA205 ネットワークを利用した分子構造推定システムSEoNへの候補構造検定システムTUTMASSの付加(豊橋技科大)○梅尾裕美・浦田雅裕・船津公人

1PA206 大規模分子動力学計算のための高速化技術の研究(物質研)○川田正晃・三上益弘

1PA207 水溶液中における5位置換ウラシル誘導体の挙動：互変異性体間の平衡，およびアデニンと形成する水素結合の強さに対する置換基効果(物質研・東工大生命理工・東工大工)○川原俊一・和田猛・川内 進・内丸忠文・関根光雄

1PA208 $(C_5H_5N)_2(CH_2)_n(n=3,4,5)$の異性化反応における理論的研究(三重大工)○塩田昌宏・小関史朗

1PA209 ボランを触媒とするアルデヒドとシリルケテンアセタールとのアルドール縮合に関する量子化学的研究(東海大理)○石川 滋・渡邊幹夫

1PA210 分子軌道法によるベンゼン−置換ベンゼン間相互作用の計算—ab initio法による検討(青山学院大理工)○飯田貴広・遠藤 忠・伊藤眞人

1PA211 He_2^+とC_6F_6との中和反応によるHe_2*の生成過程の理論的研究(九大総理工・九大機能研)○織田恵理香・茂木孝一・辻 正治・酒井嘉子・中川清子・下川利成

1PA212 非古典的金属カルボニルカチオン錯体のカルボニル化触媒反応に関する理論的研究(九大総理工・九大機能研)○茂木孝一・園田高明・酒井嘉子

1PA213 ケモメトリックス手法を用いた材料開発支援プログラムの開発(豊橋技科大)荒川正幹○棚田東作・西村竜一・船津公人

部会役員会報告

第5回幹事会　兼　CICSJ Bulletin 編集委員会

日　時：1998年11月9日（月）15時

場　所：東京大学教養部16号館1階126号室

出席者：吉田元二・時田澄男・三戸邦郎・相田美砂子・高田章
木村美実子・阿部英次・高木達也・清水弘一

議　題：第2回役員会議案について
CICSJ Bulletin発行状況について
・vol.16 No.6 特集「ビブリオメトリックス」
・vol.17 No.1 特集「巨大分子をめざしたソフトウェア（仮称）」
第11回情報化学講習会の報告

第2回役員会

日　時：1998年11月9日（月）18時

場　所：こまばエミナース孔雀の間

出席者：吉田元二・時田澄男・三戸邦郎・相田美砂子・高田章・
木村美実子・阿部英次・伊藤眞人・松岡俊和・米田茂隆・
工藤喜弘・高木達也・細矢治夫・清水弘一

議　題：第21回情報化学討論会について
世話人の米田茂隆氏より最終中間報告
22回、23回の開催地に関する議論が行われた
平成11年度予算案について
原案どおり、予算案が承認された
平成11年度役員について
部会総会議事について
各部より活動報告を行う
ポスター賞投票について

部会総会報告

開催日：1998年11月11日（水）12時

場　所：こまばエミナースダイアモンドルーム

議　題：1998年度事業中間報告

総務より

WWWサーバー移転のお知らせ・その他

編集より

CICSJ Bulletin の発行状況

企画より

情報化学講習会の報告

第6回幹事会　兼　CICSJ Bulletin 編集委員会

開催日：1999年1月7日（木）15時

場　所：日本化学会会館

出席者：吉田元二・時田澄男・三戸邦郎・相田美砂子・

高田章・木村美実子・米田茂隆・高木達也

議　題：新年度役員について

情報化学討論会の予算報告

世話人の米田茂隆氏より最終報告

CICSJ Bulletin の発行状況について

vol.17 No.1 特集「巨大分子をめざしたソフトウェア」

vol.17 No.2 特集「いま情報化学を考える～

アンケート結果を踏まえて～」

vol.17 No.3 「名簿特集号」

アンケートの結果報告

１２～１月部会員異動

［入会］

（個人正部会員　　５名）

冨永　知之（日本ケミファ）　　友田　修司（東京大）

杉山　孝雄（埼玉大）　　西村　竜一（三井化学）

上野　崇　（フルヤ金属）

（学生部会員　　２名）

棚田　東作（豊橋技術科学大）　　廣松　康一（化学品検査協会）

（順不同・敬称略）

前号記事について

前号のCICSJ Bulletin Vol.16 No.6 について、特集記事「ビブリオメトリックスをやってみませんか」の著者、図書館情報大学　小野寺夏生様より、記事上新たに判明した箇所について補足説明がありました。

４ページ右下の、以下の記述について、

「この期間のCAの全抄録件数は19万件あまりなので、Lotkaはその10%弱を抽出したと思われるが、サンプリングの仕方は明らかではない。」

を、次の通りとします。

「この期間のCAの全抄録件数は19万件あまりであるが、Lotkaは累積著者索引を利用して姓がＡまたはＢで始まる著者の論文だけを調査した。」

とし、次ページの表１、文献主題範囲の「？」を「A,Bで始まる著者」とします。

情報化学部会総会のお知らせ

期　日：３月２８日（日）　14:50〜15:00

場　所：日本化学会年会Ｇ５会場

（神奈川大学横浜キャンパス　２０号館２０７教室）

（化学情報・計算機化学部門　休憩時間）

議　題：活動報告、新年度活動予定、新役員等

第22回情報化学討論会のお知らせ

会　期：平成11年11月10日（水）〜12日（金）

場　所：山形県　米沢市内

（詳細は次号以降に掲載します）

◆部会行事

第3回分析化学のためのケモメトリックス討論会及び初心者のための分析化学に利用されるケモメトリックス講習会

主催　日本化学会情報化学部会
共催　日本分析化学会中部支部・CACフォーラム

会期　4月22日（木）～23日（金）10時～16時30分
会場　名古屋市工業研究所〔名古屋市熱田区六番三丁目4-41、電話：052-661-3161 交通：地下鉄名城線「六番町」駅下車、徒歩3分〕

1日目：4月22日（木）初心者のための分析化学に利用されるケモメトリックス講習会内容

1.定性分析に利用される判別分析、KNN、SIMCA、クラスター分析、主成分分析
2.定量分析に利用される重回帰分析、主成分分析、PCR、PLS
3.ニューラルネットワーク
4.実習（Chemish ・ 市販ソフト）

2日目：4月23日（金）第3回分析化学のためのケモメトリックス討論会プログラム

1.多変量解析法を応用したポリエステルー混紡糸の混紡率の推定（愛知県警科捜研）○奥山修司・三井利幸
2.多変量解析法による顕微FT-IRデータからの無煙火薬の組成分析（愛知県警科捜研）○肥田宗政・三井利幸
3.ニューラルネットワークによるプラスチックの判別（物質研）田辺和俊
4.トレハロースー水素における特異的相互作用のPCA法による 検討（日本分光）○大久保優晴・赤尾賢一・眞砂央
5.［依頼講演］主成分分析法による微量物質検知（神戸薬大）長谷川健
6.FT-Raman分光法を用いた木材各種組織比率の非破壊的定量（王子製紙森林資源研）○小名俊博・園田哲也・伊藤一弥・柴田勝
7.［依頼講演］Quality of Rice and Chemometrics(Queensland University of Technology)Serge Kokot
8.3D－QSARにおけるKohonenmapの応用と3way PLSによる解析（豊橋技科大）○松岡成郎・荒川正幹・船津公人
9.近赤外ーケモメトリックス及びFTラマンーケモメトリックス法 によるポリマーブレンドの非破壊定量分析大阪府警科捜研1・大阪市工研2・関学大理3）○松下晃1・松川公洋2・井上弘2・南幸男1・尾崎幸洋3
10.近赤外スペクトルの解析におけるケモメトリックスと二次元相関分光法の関係（関西学院大学）○尾崎幸洋・村山幸市
11.［依頼講演］PY－GCを用いたポリマーキャラクタリゼーションにおける未知ポリマー高速検索アルゴリズムの検討（フロンティアラボ1・名大院工2）○渡辺忠一1・大谷肇2・柘植新2
12.主成分分析法による定量分析の改良（愛知県警科捜研1・ジーエルサイエンス2）○三井利幸1・高木肇2・山地武広2
13.［依頼講演］市販点眼剤中香料成分の分析と統計手法の応用（ライオン（株））○判野有子・小池茂行

申　込　方　法：氏名、所属、住所、電話番号、いずれに参加するかを記入の上、官製ハガキ、FAXまたはE-mailにてお申し込み下さい。
申込締切日：3月20日
講習会参加費：主催、共催学会会員6,000円、学生4,000円、非会員8,000円を当日受付にてお支払い下さい。
討論会参加登録費：主催、共催学会会員3,000円、学生1,000円、非会員4,000円を当日受付にてお支払い下さい。
申　込　先：〒460-0001　名古屋市中区三の丸2－1－1　愛知県警察本部科学捜査研究所三井利幸〔電話052-951-1611（内線3791,FAX 052-953-9629,E-mail:tmit@ma3.justnet.ne.jp〕

◆関連行事

第76春季年会(1999)プログラム抜粋[会場別]

G5 会場
20号館207番教室

化学情報・計算機化学

(分子シミュレーション)

3月28日午前

座長　田辺　和俊(10:00～10:50)

1G507　新しい応用分子動力学計算プログラムの開発と仮想現実感環境の構築(東北大院工)○三浦隆治・稲葉祐策・小野津崇之・高見誠一・久保百司・宮本　明

1G508　モンテカルロ法，および分子動力学法による膜分離シミュレーション(東北大院工)○小林泰則・水上浩一・高見誠一・久保百司．・宮本　明

1G509　水溶液中におけるグリシン双性イオンの構造最適化–自由エネルギー勾配法–(基礎化研)○片岡　健・奥山直人・長岡正隆・山邊時雄

1G510　GCMC法によるQCM匂いセンサの応答予測(東工大)中村賢一・中本高道・森泉豊栄

1G511　シリコン酸化膜の応力特性の計算化学的検討(東北大院工・金材研)○山田有場・杉左近潔・遠藤　明・Cheettu Ammal Salai・高見誠一・久保百司・宮本　明・板倉明子・北島正弘

(化学情報・ケモメトリックス)

3月28日午後

座長　三上　益弘(13:10～14:20)

1G526　MoldaNet–Secure Signed AppletとJava 3D技術を用いたネットワークに分散された分子モデリングプログラムの開発(広島大理・インペリアルカレッジ)○吉田　弘・Rzepa H.S.・Tonge A.P.

1G527　主成分分析法による定量分析の改良(愛知県警科捜研)○三井利幸・肥田宗政

1G528　多変量解析法による顕微FT–IRデータからの無煙火薬の組成分析(愛知県警科捜研)○肥田宗政・三井利幸

1G529　有機合成設計システムAIPHOSの開発(18)–クライアント–サーバーシステム化の検討(富士フイルム足柄研)○竹内雅志・船津公人

1G530　ニューラルネットワークによる代替フロン探索(物質研)○田辺和俊・松本高利・長嶋雲兵・吉本昭二・上坂博亨・山本博志

1G531　ニューラルネットワークによる発ガン性予測(物質研)○松本高利・田辺和俊

1G532　ニューラルネットワークによる生分解性プラスチックの判別(物質研)○田辺和俊・松本高利・天野敏男・上坂博亨

座長　有園　尚(14:20～14:50)

1G533　**依頼講演**　コンビナトリアルライブラリーデザインにおけるデータベースの利用(日本MDL)岸本隆太郎

(計算化学・理論化学)

座長　高田　章(15:00～15:30)

1G537　**依頼講演**　エレクトロニクス材料の原子レベル設計(東北大院工)久保百司

座長　船津　公人(15:30～16:50)

1G540　アルミニウム化学気相成長におけるジメチルアルミニウムハイドライドの水素修飾シリコン表面への吸着に関する理論的研究(東大工)○松脇崇晃・中島　徹・山下晃一

1G541　ジメチルエーテルとOHラジカルの気相反応に関する量子化学計算(東大工)○本村和也・高橋崇宏・山下晃一

1G542　計算化学による振動スペクトルの予測(3)–非経験的分子軌道法による塩素化メタン分子の溶媒効果を含めた赤外スペクトルの計算(物質研)○田辺和俊・松本高利

1G543　電極表面での吸着構造と電荷移動に関する理論的研究(東大工)○亀谷和久・大脇　創・山下晃一

1G544　遷移状態理論によるフッ素化メタン，エタンとOHラジカルとの反応速度の計算(物質研)○杉江正昭・内丸忠文

1G545　アセチレンの超励起状態のダイナミクスに関する理論的研究(慶大理工)○安池智一・藪下　聡

1G546　分子とゼオライト骨格の相互作用–反応指数研究(東北工研)○Chatterjee Abhijit・岩崎孝志・蛯名武雄

1G547　エネルギー選別光解離反応を誘起する制御レーザー場の設計(東北大院理)○黒澤由佳・大槻幸義・藤村勇一

P 会場
13号館体育館

化学情報・計算機化学

3月28日午前

(11:00～12:30)

1PA205　ネットワークを利用した分子構造推定システムSEoNへの候補構造検定システムTUTMASSの付加(豊橋技科大)○梅尾裕美・浦田雅裕・船津公人

1PA206　大規模分子動力学計算のための高速化技術の研究(物質研)○川田正晃・三上益弘

1PA207　水溶液中における5位置換ウラシル誘導体の挙動：互変異性体間の平衡，およびアデニンと形成する水素結合の強さに対する置換基効果(物質研・東工大生命理工・東工大工)○川原俊一・和田猛・川内　進・内丸忠文・関根光雄

1PA208　$(C_5H_5N)_2(CH_2)_n$(n=3,4,5)の異性化反応における理論的研究(三重大工)○塩田昌宏・小関史朗

1PA209　ボランを触媒とするアルデヒドとシリルケテンアセタールとのアルドール縮合に関する量子化学的研究(東海大理)○石川　滋・渡邊幹夫

1PA210　分子軌道法によるベンゼン–置換ベンゼン間相互作用の計算—ab initio法による検討(青山学院大理工)○飯田貴広・遠藤　忠・伊藤眞人

1PA211　He_2^+とC_6F_6との中和反応によるHe_2*の生成過程の理論的研究(九大総理工・九大機能研)○織田恵理香・茂木孝一・辻　正治・酒井嘉子・中川清子・下川利成

1PA212　非古典的金属カルボニルカチオン錯体のカルボニル化触媒反応に関する理論的研究(九大総理工・九大機能研)○茂木孝一・園田高明・酒井嘉子

1PA213　ケモメトリックス手法を用いた材料開発支援プログラムの開発(豊橋技科大)荒川正幹○棚田東作・西村竜一・船津公人

[参考]　アンケート内容

（1）あなたは下記のいずれに属しますか？該当する項目に○を付けてください。

A．大学（現、元）（教官、学生）　B．国公立研究所　C．企業

D．その他（　　　　）

（2）現在の情報化学部会は主として「化学情報・計算機化学」に分類される研究分野が中心となっています。

以下の小分類の中から自分の専門、興味のある分野に○を付けてください。複数選択も可です。また今後加えることを希望する分類（キーワードや通称でも可）があれば自由に書いてください。

（小分類）	（自分の専門）	（興味のある分野）
A．化学情報学	（　）	（　）
B．理論化学・計算化学	（　）	（　）
C．ケモメトリックス	（　）	（　）
D．コンビナトリアルケミストリー	（　）	（　）
E．数理化学	（　）	（　）

F．今後加えることを希望する分類あるいはキーワード

（　　　　　　　　　　　　　　　　　　　　　　）

部会の主な活動としては、化学会年会への参加以外に、ＣＩＣＳＪ発行（年６回）、情報化学討論会（年１回）、講習会（随時）、ワークショップ（随時）、ホームページによる情報提供等があります。

（3）講習会は年１回を目標に最新の研究動向をわかりやすく会員の皆様に知っていただくために開催しています。今後開催を希望されるテーマを書いてください。（会員の皆様が中心となって開催される講演会、勉強会等のワークショップ開催への援助も可能です。上記の部会連絡先へ気軽にお問い合わせください。）

（4）情報化学討論会は会員の皆様が自由に研究成果の発表、討論ができる交流の場となっております。現行討論会への要望、他の学会・研究会との交流への意見等ありましたら自由にお書きください。

（5）現在情報化学部会のホームページがご覧になれます。（http://www.nimc.go.jp/cicsj/）

ホームページをご覧になった事がありますか？○を付けてください。

（Yes：ホームページ有用）　（Yes：ホームページ不要）　（No）

今後会員の皆様に役立つ情報を増やしていこうと考えています。下記の質問にお答えください。

A．部会のホームページで役にたったものは？（　　　　　　　　　　）

B．部会のホームページであると有用なものは？（　　　　　　　　　　）

Ｃ．普段利用しているホームページでお勧めは？（　　　　　　　　　　　　　　　）

（　　　　　　　　　　　　　　　）

新たに部会員に有用なホームページを発見された場合にはメールにてアドレスをお送りいただければ助かります。関連サイトのページに掲載させていただきます。

（６）ＣＩＣＳＪ（会誌）は定期的に会員の皆様に情報を提供する貴重な情報媒体になっています。該当する項目に○を付けてください。

Ａ．会誌を読みますか？　（いつも）　（時々）　（ほとんど見ない）

Ｂ．目を通す記事はどれですか？複数選択可。

（特集記事＝本文）　（部会記事）　（部会行事）　（関連行事）　（編集後記）

Ｃ．現在の発行回数（年６回）は？　（多い）　（適当）　（少ない）

Ｄ．今後特集を希望するテーマ、執筆者への要望、新規に加えたいコーナー等ありましたら自由に書いてください。

（７）上記の現行行事以外で新たに希望される活動や部会へのご意見があれば自由に書いてください。

CICSJ Bulletin Vol.17, No.2, April 1999

このページは「日本化学会・情報化学部会」の責任において運営されています。

目次

◆巻 頭 言

アンケート

部会長　吉田　元二

motoy@zoom.cse.nagoya-u.ac.jp

本号は、昨年の秋に実施した情報化学部会のアンケートの集計と部会員の方の部会への要望、提案等の特集号となりました。最近、「情報化学部会は何を目的としているのですか」とか「もう情報化学部会の存在意義はなくなったのでは」といった声を耳にすることが多くなりました。私自身はこれからますます部会は重要になってくると考えておりますので、所感を述べさせて戴きます。

CICSJ や情報化学討論会の講演要旨集を見てみますと、その内容に時代の変遷を感じざるを得ません。昭和４０年の前半は情報管理、命名法の話題が多くみられます。丁度この頃は、企業においてもバイナリ形式で IBM カードや磁気テープに記録された化合物や特許の情報を購入し、それに自社のデータを追加して情報検索システムを作る事が非常に重要であった時代です。従って如何に効率よく検索するか、構造情報を如何に文字またはバイナリで表現するか、あるいはスペクトルデータを如何に集積するかといったことが研究の中心でした。昭和４０年代の後半になりますと、定量的構造活性相関を用いた薬物設計に関心が集まり、その関連の論文が多くなってきました。昭和５０年代前半には、分子力学法による有機化合物の構造計算に関心が寄せられ、さらに計算機の高速化が進み、分子軌道計算が現実的になるにつれてその関連の論文が増えてきました。そして、昭和６０年代に入るとそれらの技術を統合化させた分子設計システムの構築に関する論文が増えてまいりました。そして、現在は今までの研究成果を駆使して現実の分子設計、反応設計あるいは材料設計に役に立てるにはどうしたら良いかが最大の関心事になってきているように思います。このように、情報化学部会は常にその時代に要求される分野を反映して進歩してきているのです。確かに理論化学についての専門の討論会があり、それはそれで非常に大切なことであると思っています。しかし、実際の研究開発に計算機を活用して役立てるためには、色々な分野の知識が必要となるのです。そのためにこそこの部会があると思っています。実際の研究開発は、プロブレムオリエンテッドなのです。解決しなければならない課題がまずあるのです。それに応えるための計算科学部隊であるわれわれがあまりにもメソッドオリエンテッドあるいはプログラムオリエンテッドではないかいう危惧の念を持っております。われわれもプロブレムオリエンテッドに徹し、解決すべき課題に対して手持ちのプログラムだけで対処するのではなく（そんなことがいつも可能なはずがありません）、その課題の化学を、物理をしっかりと考え、必要ならば理論を構築し、プログラムを開発するぐらいの意欲こそが必要だと思っております。そして、そこに情報化学部会の次の発展が約束されるのではないでしょうか。会員の皆様方の忌憚のないご意見をお待ちしております。

◆特集part 1：部会員からの声

情報化学ってどんな学問？

創価大学工学部　伊藤眞人
itomasa@t.soka.ac.jp

情報化学は拡大する？

今から20年ほど前、まだ大学院生だった頃、個人的に面白いと思った文献を、専門の有機化学に限らず、分野別に分類してリストにしようとしたことがある（生来の怠け癖と移り気のせいですぐに挫折した）。その時用意した分類項目の一つに、「情報・計算」という項目があった。「情報化学」に対する当時の化学分野での一般的な認識はこの2点、端的に言うとChemical AbstractsのTSS（on-lineという言葉はまだ一般的ではなかった）検索と大型計算機を使う分子軌道計算にしかなかったのではないかと思う。今回のアンケート結果の分野を表すキーワードとそれが包含している内容をみると、情報化学のカバーする分野はずいぶん広がったものだとつくづく感じる。

そして、「今後加えることを希望する分類など」の項にある分野が、本当にそのまま付け加わるとしたら、20年後にはどれほどの広がりを見せるのだろうか。化学会の下にある一部会でありながら、親学会の枠組みから大きく離れた学問領域にまで広がってしまう。現在でさえ部会のidentityが問題になっているというのに．．．

それはともかく、「情報化学とは」という命題に対して、近過去から近未来まで成立しそうな一つの作業仮説を提示してみると、次のようになる。

「化学とその応用分野での、(1) 化学情報の綜合的な活用、(2) 計算機とソフトウエアの活用、(3) 情報科学の手法の活用」に関する学問

遺漏を残さぬようにと考えたので、風呂敷をやや広げすぎたきらいがあるが、考えを進める上で一つの足がかりにはなるのではないか。

情報化学＝計算機利用化学？

さて、(2) はもとより、(1)や(3)のカテゴリーにあてはまる研究・開発活動を支えているのは言うまでもなく計算機である。また、上に述べた分野の拡大を可能にしたのが計算機の発達であることは言うまでもない。となると、「情報化学＝計算機利用化学」という式を立ててみたい誘惑に駆られる。

上はスパコンから下はパソコンまで、計算機を研究・開発のために利用している化学者・技術者の数は、この20年間でいったい何倍に増えたのだろうか。計算機がなくては研究・開発そのものが成り立たない分野も研究者も著しく増えている．．．はずである。

ところが、この増加がそのまま情報化学の専門家の増加、あるいは部会員の増加につながっているかというと、どうやら、必ずしもそうではないらしい。もし、「情報化学＝計算機利用化学」という式が成り立つとしたら、これは非常に不可解なことである。この問題を考えるには、どのような人達が部会に加わっているかではなく、どのような人達が「加わっていないか」に着目したほうがよさそうである。

部会に関心のない計算機利用者がいる？

振り返ってみると、化学関係の研究に不可欠な道具として計算機を活用しているが、部会あるいは情報化学とは関係ないと考えている化学者は昔からいたし、今でも相当数いるのではないか。そうでなければ、たとえば、日本化学プログラム交換機構（JCPE）の個人会員の数のほうが、情報化学部会の部会員の数よりもはるかに多いという実態を説明することはできないだろう。それでは、これはどのような人達だろうか？

たとえば、紫外線からマイクロ波に至る分光学やX線結晶解析などの構造化学の分野の研究者が考えられる。構造とスペクトルの関係を明らかにする上で、計算機を用いる理論計算は昔から不可欠であり、初期の計算機の進歩と歩調を合わせて発展してきたとも言っていい。特に初期には解析のためのプログラムをも自作していたはずである。それでも、彼らの中には分光学、構造化学を専門としており、情報化学とは関係ないと考えている研究者が多いのではないか。なるほど彼らにとっては重要なのは分子構造という実験「事実」であり、計算機やプログラムは「事実」を知るための手段・道具に過ぎないのかも知れない。

計算機と共に進歩してきたもう一つの分野が理論化学・計算化学である。こちらは計算結果そのものが研究の成果であるという点では、上述した実験化学者達よりも計算機やプログラムとの関わりは深いはずである。それでも、この分野の専門家の大部分が、情報化学と関係があると考えてい

るわけではなさそうである。彼らの関心は専門分野の化学「理論」の構築と検証にあって、計算機やプログラムは彼らにとってもこれを実現するための手段であり、道具に過ぎないのかも知れない。

これらの分野に限らず、研究・開発の主要な手段として計算機を使っていながら、「情報化学とは関係ない」と考えている人達は、「情報化学」のカテゴリーを先の作業仮説の(1)と(3)だけであり、自分たちの活動が該当する(2)は「情報化学ではない」と考えているのかも知れない。しかしそれだけだろうか？研究等を行う上での、計算機やプログラムに対する位置づけが、情報化学部会に関わっている人たちとは違っているという可能性はないだろうか？いいかえれば、情報化学部会に関わっている人たちにとっては、研究・開発等にたずさわる上で、計算機やプログラムは手段・道具以上の存在なのだろうか？これに対しては部会の関係者のほとんどが「No」と答えるだろう。それではどこに違いがあるのだろうか。

事実、理論、そして．．．．？

「事実」と「理論」は科学の「両輪」のように考えられ、引き合いに出される。化学でも例外ではない。この２つがあまりにも強調されるので、これだけが化学の研究対象であると信じている人も少なくあるまい。

「応用科学」の分野になると事情は一変する。「工学」を例に取ろう。工学では、「予測」と「設計」の２つの手法が重要な要素をなしている。批判を恐れずに一言で表すとすれば、「工学とは『方法論』の学問である」と言えるかも知れない。また、部会と関係の深い「構造活性相関」は、「薬学」分野の重要な「方法論」の一つであると位置づけることができる。

ところで、「方法論」は決して応用科学だけの専売特許ではない。化学の分野には「方法論」を研究対象としている分野が昔からある。「分析化学」である。分析法の進歩が化学の進歩に果たしてきた役割は歴史的に見ても非常に大きい。また、「事実」と「理論」を対象とする化学の他の分野でも、その発展に対して「方法論」の進歩が果してきた役割は決して小さくない。それにもかかわらず、研究・開発の目標と成果が先に述べた「事実」とも「理論」とも異なるために、「両輪」の信奉者によってその重要性はともすればなおざりにされがちである。

情報化学に話を戻そう。先に述べた情報化学に関する作業仮説で挙げた3つのカテゴリーのいずれも、研究・開発の目標は新しい「方法論」の確立にあるのではないだろうか。3つとも「活用」で止めていたが、これは「活用法」としたほうがいいのではないか。批判を覚悟でさらに単純化すれば「情報化学とは、化学とその応用分野での、計算機利用の方法論の（確立と活用に関する）学問である」といえるのではあるまいか。「方法論」という観点を念頭に置いて、アンケートの結果をご覧になってみてはいかがだろうか。

情報化学部会の今後は？

将来、化学のあらゆる分野で計算機と化学計算プログラムの利用が当り前になれば、「計算化学」も「情報化学」も各分野に吸収されて存在意義を失うと考える人もいるかも知れない。しかし、おそらくこれは正しくないだろう。化学の研究の対象と目標が「事実」「理論」「方法論」の３つにあり、情報化学が計算機利用の「方法論」をその守備範囲に含めている限り、計算機利用の拡大は部会の守備範囲を広げることはあっても奪うことはないだろう。

部会の今後の活動を考えていく上では、基礎から応用までさまざまの視点から「方法論」を意識し、関連する他の分野の人々にその重要性を理解してもらうような働きかけを行うことが、会員層を広げ、活動をさらに活性化するのに有効ではないだろうか。

私にとって情報化学は上述した広い守備範囲で考えても専門外であり、その成果の一端を使わせて戴きたいとは思っているものの、部会では「生涯一会員」であると思っていた。それが、どうした経緯か２年間も役員の末席を暖めてしまった。討論会に参加し、ニュースレターを拝見し、さらに本号のアンケートを拝見して感じたことを、頼まれたのを機会に思いつくままに書いてみた。部会の諸兄にとって何らかの役に立つ内容になっていれば幸いであるが、自明のことばかり、あるいは見当はずれなことばかりではないかと恐れる。お世話になった諸兄に対する恩返しになったかどうかわからないが、これをもって2年間の活動のまとめとさせて戴きたい。

情報化学・計算化学・理論化学

分子科学研究所　岩田末廣(iwata@ims.ac.jp)

多分私は情報化学部会が創立されてからの会員と記憶しているが、部会員としては送られてくる会誌を読むだけで情報化学討論会にも 1 度ぐらいしか出席したことがなく、部会の運営にも何ら寄与してこなかった。たまたま E-mail で送りつけられたアンケートに答えたために、作文をさせられる羽目になってしまったようである。従って、情報化学部会がどんな活躍をすべきかを論ずる資格はない。井口洋夫先生が「化学と工業 2 月号」で述べられているように化学で取り扱う物質の数は膨大でありその量は文字通り日毎に増えている。これらの物質に関する情報の中から必要な情報を探し利用することは単純ではない故に「情報化学」と言う研究分野が、電子計算機の学術利用が始まってすぐに開拓されたのであろう。「情報物理学」はとてもありそうもないが、「情報生物学」「情報薬学」などはあっても不思議ない。実際、遺伝子配列情報の収集・利用は巨大な資金が投入されている大きな研究分野となっている。私が博士課程を終了した頃から電子計算機の学術利用が我が国でも容易になったこともあって化学科の同級生 22 人(1962 年学部卒、博士課程進学は 13 人だったと思うが)のうち二人が情報化学に移りその後情報科学の専門家になってそこで教授に就任している。同級生の中で彼らが最初に教授になったと記憶している。情報科学の初期の発展に、情報化学は強く関与してきたと云える。

しかしながら、Mulliken が 1966 年のノーベル賞講演で「計算機から化学情報を得る」と述べたとき、彼が意図していたのは「情報化学」的な意味ではなく、理論に基づいた計算を実行して新しい知識を取り出すことであった。私自身の計算機との関わりも「新しい化学情報を作り出す」立場からであったし、これからもそのつもりである。だが、情報化学の重要性を軽く見ているわけではないつもりである。現代産業活動の中における三次産業の重要性にたとえることもできよう。

「情報化学」中で、良質なデータベースの作成方法の研究と実際にデータベースを作り上げ、さらに公開する作業は特別な位置を占めていると思う。私自身は、文献データベースである QCLDB(Quantum Chemistry Literature Data Base)の計画段階から参加してきたので、これら一連の作業には研究時間と人力が、従って資力が要求されることを知っている。QCLDB は当時北大教授であった大野公男先生を中心にした QCDB(Quantum Chemistry Data Base)研究会が設計し、作成を開始した量子化学関連の文献のデータベースである。すでに 20 年余にわたってデータは集積されていて、世界中の研究者によって広く利用されている[注]。現在でもこの作業は新しい世代に引き継がれて続いており、新しい計算機環境に見合ったデータベース作成方法と公開方法の研究も進められている。　残念ながら、これらの作業に対する評価は必ずしも高くないことも知っている。

化学研究のために計算機を駆使することは、現在では学会講演、論文作成にいたるまで不可欠になっている。電子メールによる通信と www による情報交換は我々の日常の一部になっている。化学教育に computer literacy の必要性を語っていたのはついこの前のことであった。私は日常使っている計算機を「情報機器」と呼ぶのにはまだ抵抗がある。現在でも計算機がもっとも威力を発揮するのは理論とモデルに基づいた理論計算である。1998 年のノーベル化学賞が量子化学計算に巨大な功績を残し

た Pople と Kohn に与えられたのは時機を得たというよりようやくというのが実感である。今化学研究で化学計算が果たしている役割を否定することができなくなっている。今、計算化学で行っている計算は、量子力学・古典力学・統計力学などの理論に基づいている。学部と大学院の教育でいっそうこれら理論化学の教育が重要になっている。それには理論化学の研究室を化学関連学科と専攻に増やさなければならない。

情報化学でも計算化学でも今多くの計算機支援化学(computer aided chemistry)ソフトウエアが市販されている。これらの多くは、使いやすいようにいわゆる man-machine interface に工夫が施されている。しかしながらこれらのソフトウエアを使って化学情報を得て、あたらしい情報に転換させるには、ソフトウエアが基づいている理論の基礎を把握していなければならない。情報化学・計算化学・理論化学にわたる広い意味の計算機支援化学(computer aided chemistry)の研究者を意図的に養成する事が急務だと思う。企業においても、企業内教育だけでなく、博士研究員を経た経験ある研究者の採用をぜひ考えてほしいものである。

この数年の間に、いろいろな形で博士研究員(ポストドク)が急速に増えている。「博士研究員の 2000 年問題」と冗談交じりに語られている(日本物理学会誌 2 月号)ように、昔あった over-doctor の年長版 over-postdoc が出現するのは時間の問題である。アンケートの回答にもあったが、情報部会誌で求職・求人欄をもうけるのは一案であろう。比較的限られた分野であるので、適切な職・人材を見つけるきっかけになるのではないだろうか。

[注]QCLDB は現在次の方法で公開されている。1)分子科学研究所計算機センターの利用者は on-line コマンドで検索できる。2)Elsevier 社が出版する J.Molecular Structure (THEOCHEM)の特別号として毎年 1 回出版される。3)化学情報協会を通じて検索ソフトウエアとデータを購入し、サイトライセンスを得た後、各サイト内で公開する。主要大学計算機センターにもインストールされているはずであるが最新のものになっていない可能性が高い。4)新しい試みとして web(http://qcldb.ims.ac.jp/index.html)でも試験公開を始めている。1 年ごとの登録制で利用課題とパスワードを取得して利用する。ただしデータは最新のものにはしていない。

情報化学討論会をしまりのあるものにしていただきたい

金谷重彦（山形大学・工学部・電子情報工学科・生体情報講座）

情報化学討論会に対する私見を以下に述べさせていただきます．情報化学討論会が最先端の議論の場となるように改革願えれば幸いです．

（1）情報化学討論会をしまりのあるものにしていただきたい．

情報化学討論会の発表，特に要旨集において，報告されている内容が情報化学という立場からみて新規なものかはっきりしない発表が多数見受けられます．第２０回の討論会要旨集をみると，２０件の一般口頭発表のうち文献の引用の無いものが９件，文献引用があるもののうち研究の背景，情報化学もしくは化学における意義，先端研究等との比較がないものが５件あった．４ページも費やして書かれている要旨集からこの討論会で議論されていることに本当に意味があるのか，あるいは，要旨集の情報は質的にどういったレベルのものなのか不明であります．進行中の研究ではなく，<u>むしろ審査つき雑誌に掲載されるに値するものを口頭発表で発表していただきたい!</u>．そのために，<u>発表に対する事前審査をしていただきたい．</u>背景もはっきりとせず化学に関わる対象についてとりあえず解析をしたとか，.単にデータを集めてデータベースをつくっているといった発表で情報化学討論会を埋めるのでは，ますます情報化学討論会をしまりのないものにすると思います．

（2）口頭発表とポスター発表の違いを明確にしていただきたい．

口頭発表は，審査を通った発表の場として位置付け，ポスター発表は，進行中の研究の発表の場にしていただきたい．

（3）国際的シンポジウムにしていただきたい．

毎年行われている討論会を，数年に一度のペースでもよいから，外国の研究者を招いて議論できるシンポジウムにしていただきたい．外国の研究者を招くということは，世界でどのような研究がなされているかを知る機会を得ることができると同時に，<u>日本の研究をアピールする場にもなる．</u>このような国際的立場での情報化学の発展に貢献する場をつくっていただきたい．

（4）情報化学研究のレビューをまとめていただきたい．

２０数回の討論会を通して，なにが解けていて，なにが解けていないテーマなのか，について情報化学討論会に長く携わっておられる研究者により<u>レビューにまとめていただきたい．</u>２０数回の討論会でなされた研究発表がどの程度公表論文という形式にまで展開されているかについても，同時にレビューにまとめていただきたい．このレビューは，今後の情報化学討論会の発表に対して，以前になされた研究で解けなかったものが実現できた，あるいは，すでになされている研究で発表に値しないといった判断に役立つと期待できます．

以上

CISCJ アンケート結果を読んで ― 感想と要望 ―

住友化学工業(株)有機合成研究所　河合　隆
kawait2@sc.sumitomo-chem.co.jp

アンケート結果（特に「今後加えることを希望する分類等」）を読んでまず感じたことは、「情報」化学という分野が実に膨大な分野であるということである。かつては情報化学というと、計算化学(Computational Chemistry)のことを指していたように思う。しかし、現在では計算化学は勿論のこと、情報検索・化学教育から WWW(World Wide Web)に亘るまで、およそ情報と化学がクロスオーバーするものなら何でも情報化学で取り扱う話題となっているようである。このような状況下で次の特集で何を読むことができるかということは、一読者としては次に何が出てくるか、単純に楽しみにしていればいいが、その一方編集に携わる方々にとっては何を次の特集に取り上げるかに頭を悩ませておられることと思う。様々な話題を取り上げて下さることに感謝したい。

本誌で扱う特集のテーマとしては、今後も計算化学が主体となるものと思うが、個人的にはできるだけ広範囲の分野を取り上げていただきたいと思う（贅沢な望みで申し訳ありません）。この点で、前々号で取り上げられていた書誌学に関する特集は、筆者にとっては全く畑違いの話題であるにも拘らず、非常に興味を持って読むことのできた特集であった。

講習会に関しては、社会人として多様な分野の最先端に触れることのできる格好の機会であり、今後もいろいろな講習会を開催していただきたいと思う。筆者は合成反応設計プログラムの開発に携わるものであるが、社内ではコンピュータを使っていれば、分子軌道法からケモメトリックスまで全てに通暁しているものと思われることが多く、閉口することがある（それどころか、パソコンのアプリケーションの問題についても相談されることがしばしばである）。このような筆者にとって、本会で企画される講習会は、非常にありがたいものである。他の企業の研究所の方でも同じように感じておられる方は少なくないのではないかと思っている。

情報化学討論会については、筆者自身一度も参加したことがないので、具体的にあれこれ論じることはできない。しかし、アンケートの結果を読んで同感したのは、1) 情報化学の間口が非常に広くなってきている現在、参加者が必ずしも討論項目に関して通暁してはいないということ、2) あまり十分とは言えない時間で行われた討論を深めずにそのまま終えてしまうのはもったいないという２点である。1)の点に関しては、アンケートに書かれているような「パネル討論」や「専門的な知識がなくても気軽に参加できるテーマ」などを取り上げることは有効であるように思われる。また、2)の点に関しては、掲示板やメーリングリストを討論会の後に開設することによって、ディスカッションを継続し、深めていくことができるのではないかと思う。

また、情報化学討論会は構造活性相関シンポジウムと併せて開催されるのが通例であるが、できれば午前と午後に分けるなど、ずらした方が相互に聞きたい講演を聞くことができるというメリットがあるのではないかと思う。

最後に筆者の意見を述べる場を提供してくださった日本化学会情報化学部会の編集委員の方に感謝します。

情報化学部会の活性化のための提言

物質工学工業技術研究所　田辺　和俊　ktanabe@home.nimc.go.jp

今回の特集号で部会の過去、現在、今後を考える動機となったのは、多分、会員数の低下が最大の要因ではないかと思います。大学は学生数が減少し、当部会に限らず、どこの学会も会員の確保に大変なようです。一方、企業では利益に直結しにくいこの分野では最近の不景気のためにどこの企業でも研究者が削減され、今後は外注さえ起こりかねない状況です。特に、当部会は会員の高齢化の問題を抱えているのではないでしょうか。過去の会員の年齢別数の変化を調べるのは困難でしょうが、現在の会員の年齢構成はどうなっているのでしょうか。

当部会がカバーする分野は、情報化学（データベース、ケモメトリックスなど）と計算化学の２つに大別できると思いますが、後者については大学に講座や研究室が存在するので、会員数の低下はあまり心配ないと思われます。それに対して、前者の教育を行っている大学の講座が国内ではきわめて僅かです。そのため、この分野に参入する若手の研究者が少なく、情報化学討論会でも情報化学のセッションではロートルのサロンになりつつあるように思います。しかし、この状況は日本だけでなく米国でも同様のようで、私が数年前に出席した米国化学会の年会でも、後者に相当するComputers in Chemistry (COMP)のセッションは大盛況なのに、前者のChemical Information (CINF)のセッションはロートルばかりで寂しかったのを思い出します。

しかし、このような低迷状況に手を拱いてばかりでは、凋落傾向に歯止めがかかりませんので、何とか活性化対策を講ずる必要があります。積極的な提言をしてみようと思います。

アンケートではセミナーや講習会の開催を望む声もあるようですが、昨今は人が集まる企画を立てるのは大変です。

部会の年齢構成を考えると、ここは若手の部会員が増えるような企画を考えるべきではないでしょうか。そのためには今何が一番必要かを考えてみると、若手の研究者、特に大学院生が博士号を取るための論文発表できる場を設けることが最も必要ではないかと思います。そのためには投稿論文を掲載してくれる日本語の論文誌があってほしいと思います。

計算化学の論文は国内でも日本化学会誌など化学会の雑誌で受理されそうです。しかし、情報化学分野の論文を発表する雑誌となると国内にはなさそうです。英語で論文が書ければJournal of Chemical Information and Computer SciencesやComputers and Chemistryなど海外の雑誌に投稿すればよいのですが、やはり若手の研究者が日本語で気軽に投稿できる雑誌がほしいと思います。

そこで、そのために、当部会で論文誌を発行することを提案します。勿論、最初は投稿論文が集まるかという心配があるでしょう。しかし、そこは部会員に半ば強制的に論文投稿を依頼してもよいのではないでしょうか。

論文誌ができて、投稿論文が掲載されるようになると、大学院生などの若手の会員が増えて、当部会も活性化すると期待できるのではないでしょうか。

情報化学部会の活性化のための提言

物質工学工業技術研究所　田辺　和俊　ktanabe@home.nimc.go.jp

今回の特集号で部会の過去、現在、今後を考える動機となったのは、多分、会員数の低下が最大の要因ではないかと思います。大学は学生数が減少し、当部会に限らず、どこの学会も会員の確保に大変なようです。一方、企業では利益に直結しにくいこの分野では最近の不景気のためにどこの企業でも研究者が削減され、今後は外注さえ起こりかねない状況です。特に、当部会は会員の高齢化の問題を抱えているのではないでしょうか。過去の会員の年齢別数の変化を調べるのは困難でしょうが、現在の会員の年齢構成はどうなっているのでしょうか。

当部会がカバーする分野は、情報化学（データベース、ケモメトリックスなど）と計算化学の２つに大別できると思いますが、後者については大学に講座や研究室が存在するので、会員数の低下はあまり心配ないと思われます。それに対して、前者の教育を行っている大学の講座が国内ではきわめて僅かです。そのため、この分野に参入する若手の研究者が少なく、情報化学討論会でも情報化学のセッションではロートルのサロンになりつつあるように思います。しかし、この状況は日本だけでなく米国でも同様のようで、私が数年前に出席した米国化学会の年会でも、後者に相当するComputers in Chemistry (COMP)のセッションは大盛況なのに、前者のChemical Information (CINF)のセッションはロートルばかりで寂しかったのを思い出します。

しかし、このような低迷状況に手を拱いてばかりでは、凋落傾向に歯止めがかかりませんので、何とか活性化対策を講ずる必要があります。積極的な提言をしてみようと思います。

アンケートではセミナーや講習会の開催を望む声もあるようですが、昨今は人が集まる企画を立てるのは大変です。

部会の年齢構成を考えると、ここは若手の部会員が増えるような企画を考えるべきではないでしょうか。そのためには今何が一番必要かを考えてみると、若手の研究者、特に大学院生が博士号を取るための論文発表できる場を設けることが最も必要ではないかと思います。そのためには投稿論文を掲載してくれる日本語の論文誌があってほしいと思います。

計算化学の論文は国内でも日本化学会誌など化学会の雑誌で受理されそうです。しかし、情報化学分野の論文を発表する雑誌となると国内にはなさそうです。英語で論文が書ければJournal of Chemical Information and Computer SciencesやComputers and Chemistryなど海外の雑誌に投稿すればよいのですが、やはり若手の研究者が日本語で気軽に投稿できる雑誌がほしいと思います。

そこで、そのために、当部会で論文誌を発行することを提案します。勿論、最初は投稿論文が集まるかという心配があるでしょう。しかし、そこは部会員に半ば強制的に論文投稿を依頼してもよいのではないでしょうか。

論文誌ができて、投稿論文が掲載されるようになると、大学院生などの若手の会員が増えて、当部会も活性化すると期待できるのではないでしょうか。

「情報化学」雑感

徳島大学薬学部　中馬　寛
hchuman@ph. tokushima-u. ac. jp

科学・技術の発展を振り返ってみたとき，今世紀最後の大きな出来事は情報科学・技術の出現，その飛躍的発展，一般社会へ与えた種々の影響である．情報科学・技術を抜きには次世紀を展望することはできないゆえんである．また学問領域はますますボーダーレスになりつつある．「アンケート」の中でも，今後の情報化学部会の活動に期待することとして，情報科学を基盤とする幅広い分野からのトッピクス，他の学会との交流，従来の情報化学分野を超えた対象への取り組みなどが挙げられている．化学自身も物理，生物との垣根を越えたところから大きな収穫が得られてきている．そうなると情報化学も従来の情報化学分野を超えた対象へとますます目が向けられ，他分野の交流も当然のごとく要求されてくることとなる．一方，化学は物質文明の基盤を作り上げ多くの恩恵と同時に負の遺産も作り出してきたように，化学の化学たるゆえんは，「もの」の変換，すなわち「もの」を作り出すことにあろう．ここで「今後の情報化学とは？」と大それたことを述べる気も資質もないが，今後，情報化学は「もの」の変換という化学の中心的命題を含みながら，情報化学の上位概念の情報科学である枠内で発展していくことが予想されよう．その中には，「アンケート」の計算化学，理論化学や化学情報学，ケモメトリックス，数理化学，コンビナトリアルケミストリーのような分野と新しい対象分野や方法論が含まれてこよう．皆様の御尽力で，年に一度の情報化学討論会も幅広い情報化学的な演題が集まり，様々な専門分野の研究者が様々な観点から議論できる貴重な場となっている．現状の討論会での領域から，より広く「もの」を基盤とする化学としての情報関連領域を包含するまで発展し，情報化学分野の中核的な学会となっていくことを今後の情報化学部会活動に期待する．

乱暴にも化学は「もの」の変換と述べたが，最近感じていることは「もの」のもつ別の側面である．薬を「もの」としてとらえれば，例えば定量的構造活性相関解析も創薬の一方法と位置づけられ，さらにそれは情報化学と密接な関係があることは周知である．さて実際の創薬となると，創薬研究者と同時に医療の現場にいる医師，看護婦，薬剤師そして患者の存在を忘れては決して創薬を語れないこととなる．当然のこととして創薬という言葉には本来初めからこのような全体的視野を内包している．今後必要となるのは創薬研究と医療現場との双方の情報のフィードバックであり，ここで情報科学と技術が大きな寄与をなすことが期待されている．事実，このような意味での新しい創薬科学がファーマコインフォマテイクッスやファルマインフォマテイクッスの名前で提唱されてきている．薬以外にも「もの」のもつ情報の流れが大きな影響を及ぼす例は多々あろう．

情報化学は「もの」が原点と述べたが，一般に「もの」のもつ情報とその流れの処理，加工，高次情報への変換などは今後ますます重要となるであろう．そしてその中から新たな情報化学の分野が生まれてくるかもしれない．

こんな具合に情報化学を考えてみると，いつもいろいろなことが浮かんできてまとまりがつかないぐらい，それは幅広く，ダイナミックなことが改めて分かる．

情報化学という分野は過去の成果を基に，今後我々が構築していくものであり，だからこそ自らの視点で，「情報化学とは？」を考えてみることは，より広い枠組みの中で自らの座標とその方向を多少は明らかにできる可能性がある．こここそが今後の情報化学と情報化学部会活動を考えるときの出発点ではないであろうか？　読者も「アンケート」結果を通して感じることなどをこの紙面に寄稿していただき，さらなる議論ができれば今後の情報化学と情報化学部会の活動にとって大きな一助となろう．

Mosaicな情報化学部会

日本アイ・ビー・エム株式会社 RS/6000 製品事業部 平尾義郎
yhirao@jp.ibm.com, yoshiro@hirao.com

当方 日本IBMでRS/6000というUNIXワークステーションの製品事業部に所属し、技術的支援の仕事を致しております。仕事上、何名かの幹事の方とのお付き合いもあり、そういった縁で、当稿を書くに至った次第です。最初は、計算化学や情報化学の研究者でもない自分がこのようなことを書いて良いのやらと迷いましたが、とにかく自由に忌憚のない意見をとの事ゆえ、かなり無責任かもしれないことを以下に記してみようと思います。

いきなりですが、情報化学って難しいといいますか、皆さんの間で共通するイメージを確立することが容易でないと思います。元々、物理や生物との境界領域も含め、多岐な研究対象が存在している化学に、方法論として「情報」や「計算機」が加わり、理論化学をベースにした計算手法の開発とシミュレーション・ソフトウェアの作成、既存ソフトウェアの応用技術、計算機アーキテクチャー、データベース、グラフィックス、ソフトウェア・ツール、チューニングや並列化などのプログラミング・テクニック、最近ではインターネットの活用、マルチメディア技術、ネットワーク・コンピューティングなどのトピックが入ってきます。更に、メンバーの方には、主に理論的研究に興味ある方もいらっしゃいますし、応用や実務的な観点に興味の方々もいらっしゃいます。アンケート結果での部会への要望を見ましても、わかりやすい解説や動向の説明を求める方もおりますし、研究成果の議論や討論に重きを置きたい方々もいらっしゃるようです。あたかも I*J*K*L の四次元マトリックスのどれかに属している方々の集合体の感があります。

長々と書きましたが、要するに情報化学部会の各メンバーの方々の最大公約数的なものを抽出することができないように思われます。このような状態のなか、情報化学部会はどうあるべきかは大変難しくかつ有益な議論ですが、一つのアイデアを提示させていただくと、小さなワーキング・グループの集合体になり、情報化学部会全体としては、各種ワーキング・グループや他の学会や団体活動に対する Directory Service を行うというのはいかがでしょうか。ホームページのイメージで言いますと、各種ワーキンググループや他の学会や団体活動に対するリンク集になるということです。

化学分野への応用はともかく、計算機アーキテクチャーや動向といった話は、情報化学部会の外部に多くの情報があります。例えば、計算機の将来動向については、情報化学部会の中で一から議論を興すよりも、米国エネルギー省のASCI(Accelerated Strategic Computing Initiative)プロジェクトが一つのレファレンスになると思います。多様な部会員からの各種の要望をすべて一つの器で満たそうとするのではなく、活動の細分化・明確化、及び他の団体が実施していることへの Directory Service を行うと、わかりやすくなるように思います。

以上、冒頭で述べましたように、門外漢ゆえの、かなり無責任かつ身勝手な雑感に恐縮しながらも、同時に、Mosaicから始まったWeb Browserが飛躍的に発展したように、Mosaicな貴部会の発展を切に念じつつ、筆を置くならぬキーボード・タイピングを終えることと致します。

研究者とデータベース作成者の接点として

科学技術振興事業団（ＪＳＴ）
研究基盤情報部　前田　知子
chikako@tokyo.jst.go.jp

情報化学部会は、化学関係のデータベース開発に携わってきた者にとって、事業成果を発表できる場として、また情報化学分野の研究者の方々と直接に交流できる場として、たいへん貴重である。

化学分野の情報を取り扱うためには、例えば化合物の登録方法に関してなど、データベース化のための研究や工夫が必要になる。こうした業務上の成果を、ＪＳＴではこれまでにも、情報化学討論会で発表してきた。情報化学討論会は、討論主題の中に、「化学情報学－情報検索、データベース、コンピュータネットワーク.....」と明記されており、ＪＳＴの業務について発表できる最適な場であると言える。

現在私は、ＪＳＴ高機能データベース開発事業の一環として、高分子データベース "PoLyInfo" の開発を担当している。高分子学会には、情報化学部会に近いものとして、高分子計算機科学研究会があり、この研究会でもデータベースに対する関心は高まりつつあるが、中心となるテーマではない。また、ＪＳＴでは、旧 JICST 時代から毎年、情報科学技術研究集会を主催しており、データベース開発事例などについての発表が行われている。この研究集会でも化学系のデータベースが取り上げられてはいるが、情報化学討論会は、化学系データベース構築の重要さや問題点について一層有意義な議論ができる場である。

データベースに対する利用者の要望が多様化、高度化してきている中、データベースも単なるデータの集積・参照機能を持つだけではなく、計算化学や材料設計で活用できるデータと機能を備えることが求められている。ＪＳＴがこのような新規性のあるデータベースを開発するためには、大学・企業の研究者の方々との協力が不可欠であり、今後は、これまで以上に情報化学部会との関わりが重要になって来ると思う。

◆特集part 2 ：情報化学部会活動に対するアンケート結果の報告

情報化学活動に対するアンケート結果の報告

情報化学部会では部会員の皆様のお考えを今後の活動により一層反映させていくために昨年秋、「情報化学部会活動についてのアンケート」を実施致しました。67名（大学28名、国研等8名、企業31名）の部会員の方々から貴重なご意見をいただきました。ようやく結果がまとまりましので報告させていただきます。皆様のご協力に対し厚く御礼申し上げます。

1　アンケート結果のまとめ

1．1　回答数および回答様式

	大学	国研等	企業	総合	割合(%)
回答数	28	8	31	67	
回答様式					
E-mail	18	4	21	43	64
Fax（討論会持参含む）	7	3	9	19	28
Web	3	1	1	5	7

1．2　専門、興味分野について

	大学	国研等	企業	総合	割合(%)
・専門分野					
A.化学情報学	11	1	6	18	27
B.理論化学・計算化学	15	6	13	34	51
C.ケモメトリックス	5	1	5	11	16
D.コンビナトリアルケミストリー	1	0	3	4	6
E.数理化学	2	0	0	2	3
・興味分野					
A.化学情報学	13	0	17	30	45
B.理論化学・計算化学	17	5	19	41	61
C.ケモメトリックス	7	1	10	18	27
D.コンビナトリアルケミストリー	8	1	8	17	25
E.数理化学	7	0	4	11	16

<u>今後加えることを希望する分野</u>

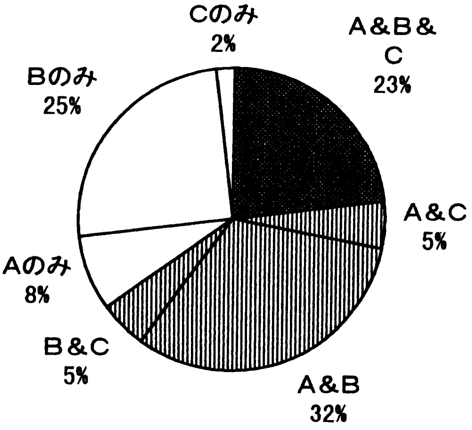
[大学]

・コンピュータシミュレーション、バイオインフォマティックス

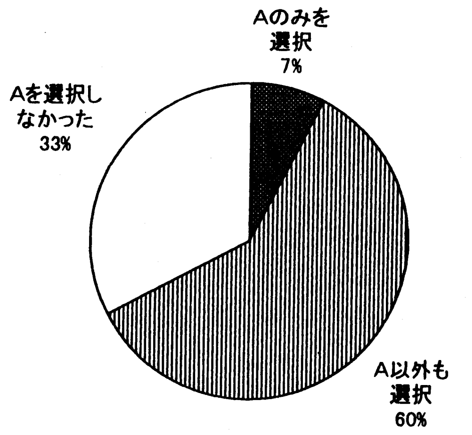
・化学または化学教育への活用

・情報化学（部会名は分野名から削除されたのか）

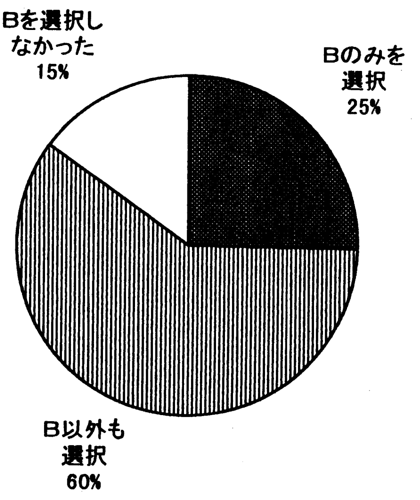
専門あるいは興味のある分野
A: 化学情報学、B: 理論・計算化学、C: ケモメトリックス



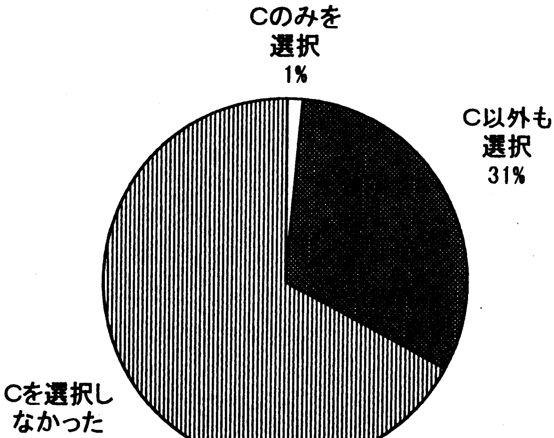
A(化学情報学)について



B(理論・計算化学)について



C(ケモメトリックス)について



・定量的構造活性相関(QSAR)、構造活性相関(SAR)
・化学グラフィックス
・化学情報教育、ネットワーク利用技術
[国研等]
・構造活性相関
・インターネット、グラフィックス
・物理との境界領域
[企業]
・種々のソフトウエア（データベース関連、分光・構造解析ｿﾌﾄ、化学工学、薬等）
・幅広く化学、化学工学を見据えた分類を検討してほしい
・データベース
・粉体、レオロジー、焼結体のシミュレーション
・ＡＩ
・化学材料のデザイン、物性予測
・並列処理、オブジェクト指向
・計算機利用技術（ベクトル、RISC、専用計算機それぞれに対する技術）
・ソフトウェア評価（ユーザーの声をもとに）
・メディシナリー・ケミストリー

１．３　講習会で希望するテーマについて

[大学]
・トポロジー、グラフ理論への応用
・大阪での開催、できれば土・日に
・バイオインフォマティックスの現在、未来
・講習会不要（学会発表充実、「情報化学」の名前の普及優先）
・北海道から参加できない
・情報科学を基盤とする幅広い分野からのトピックス
・化学がどのような形で世の中の役に立っているかを初心者向けに
・データベース関連技術、人工知能、知識ベースの化学への応用
・最近のCAMD(Computer-Aided Molecular Design)の進歩について
[国研等]
・並列処理、ネットワーク技術等を応用した計算化学
[企業]
・分子動力学、巨大分子の分子軌道法
・合成反応設計、反応データベース、CMS(Chemical markup Language)
・各種コンピュータシミュレーション(化学物質の環境動態など)
・インターネットの利用、情報検索
・コンビナトリアル・ケミストリー、データマイニング、遺伝子と化学情報
・合成経路設計システム、物性値推算システムの最新動向

・コンピュータによる反応設計、薬物設計
・ナノ～サブミクロン領域計算へのアプローチ
（量子の情報のマクロへの生かし方、分子計算と流体、構造計算の組み合わせ）
・ケモメトリックス
・化学系出身エンジニアのレベルで活用可能な技術
・粉体、ペースト、セラミックス等のシミュレーション
・コンピケム、ケモメトリックス（大量活性データの解析）
・コンビナトリアル・ケミストリー

1．4　情報化学討論会への要望について

[大学]
・北海道に住んでいるので参加したことが無い
・今後議論したい
・日本機械学会（熱部門：計算力学）、応用物理学会、電気学会、化学工学会などとの交流
・化学ソフトウエア学会との交流
・「情報学」と「理論・計算化学」に分けたので内容が見やすくなった、
後者は春の「計算化学シンポジウム」との棲みわけを検討すべき
・終了後、WWWで討論できるような掲示板を設ける
・パネル討論などがよい、専門的な知識がなくても気軽に参加できるテーマがほしい
・パネル討論希望
・討論会しまりがない、
研究の背景をしっかりと認識し、進めている研究か否かぐらい査読して口頭発表を決める
[国研等]
・開かれた討論会にしてもらいたい
[企業]
・高分子学会の計算化学部門との交流
・高分子学会との合同セミナー、勉強会
・構造活性相関シンポの聴きたい講演と重なることがある、ずらせないか
・パネル討論希望

1．5　部会ホームページについて

	大学	国研等	企業	総合	割合(%)
A-1.ホームページ見たことがある	22	7	14	43	64
A-3.ホームページ見たことがない	5	1	16	22	33
E-1.電子掲示板必要	22	8	21	51	76
必要無い	3	0	5	8	12

ホームページについての意見

B.役に立ったもの

・CICSJ Bulletin　（１１件）

・化学関連サイト　（８件）

・講演会の情報　（７件）

・討論会の情報　（７件）

・その他（ワークショップ、執筆依頼時の部会紹介用HP）

C.今後有用なもの

・各種化学サイト充実　（１２件）－－－海外の化学情報、関連分野の論文誌の内容等

・PDF形式によるCICSJの公開

・解説記事と討論会要旨全文（過去数年分について）

E.電子掲示板の必要性

・（YES）と答えた人の意見：

部会員相互の情報交換に便利

YESだがメールリストの方が便利

会員による情報交換が自由にできるページ（質問、問い合わせ、回答等）が欲しい

会員の関心が多岐にわたるので仕掛けが必要（ex.ジャンルごとに分ける）

・（NO）と答えた人の意見：

発言に対する責任の点で登録制のメールリストの方がよい

１．６　CICSJ Bulletin（会誌）

	大学	国研等	企業	総合	割合(%)
A.発行回数					
A-1.会誌いつも読む	19	6	23	48	72
A-2.会誌ときどき読む	7	1	8	16	24
A-3.会誌ほとんど見ない	2	1	0	3	4
B.目を通す記事					
B-1.特集＝本文	25	7	29	61	91
B-2.部会記事	16	5	17	38	57
B-3.部会行事	21	7	23	51	76
B-4.関連行事	21	5	21	47	70
B-5.編集後記	16	3	14	33	49
C.発行回数					
C-1.多い	3	1	2	6	9
C-2.適当	24	6	27	57	85
C-3.少ない	1	0	2	3	4

CICSJに対する意見

C.発行回数（現在、年６回）：

・（多い）と答えた人の意見

編集委員の負担が大きい、４回ではどうか

季報（４回）でよい

ホームページで十分、会誌を希望しない人には会費を安くしてほしい

D.編集に対する希望：

[大学]

・内外の化学関連学会の情報活動

・トポロジー化学のやさしい解説

・以前に試みられたが解明されていない課題のレビュー

・電子化することだけはやめてほしい

・非線形化学

・ケモメトリックス（環境ホルモン関連）

・読者の投書欄

・環境ホルモンと情報化学

・関連分野の求人・求職情報、一般のパソコンを含むハードやｿﾌﾄの情報

[国研等]

・オンライン出版で十分

・研究の世界的な動向

[企業]

・分子動力学

・実験計画法（物性予測への応用の観点から）、ニューラルネットワーク

・CML(Chemical Markup Language)の動向

・インターネットの利用、情報検索

・どうやって結果をまとめるか、レポートに仕上げるかのコツやノウハウ（TeXや構造式等）

・物性推算システムの最新動向（構造活性相関等）

・関連分野の求人・求職情報

１．７　新たに希望する活動、部会への意見

[大学]

・メーリングリストが動くようになって大変良い

・IJCのように学術論文を募集してWWW上で公開

[国研等]

・若手研究者をふやす行事を考えてほしい

[企業]

・他の部会や関連学会との共催でのシンポジウム

・情報検索を学問という立場で考えることは可能か

・以前の３次元QSARセミナーのようなもの

・会員への特典（化学ソフトウェアの割引制度など）

・類似の学会と統合し、より広い交流を目指す

2　アンケート結果の分析

1 部会員の専門および興味分野について ＜設問（2）に関して＞

「理論化学・計算化学」を専門とする人と、「化学情報学」あるいは「ケモメトリックス」を専門とする人がほぼ半数ずつになっています。またこの部会の特徴として自分の専門以外の分野にも広く興味を持っている人が多いようです。今後は、一つの分野に偏ること無く、また絶えず新しい分野も取り込みながら活動していくことが望まれているようです。

2 講習会への要望 ＜設問（3）に関して＞

アンケートで提案された案を今後の講習会企画の参考にさせていただきます。

3 情報化学討論会について ＜設問（4）に関して＞

構造活性相関だけでなく、他学会との交流が望まれていることがわかります。講習会、ワークショップの合同開催についても検討していきます。

（1）部会員が広く議論・交流出来ること、（2）高いレベルの発表・議論を維持すること、の両面が望まれています。

4 ホームページについて ＜設問（5）に関して＞

部会ホームページを一度も見たことがない人が多いようです（33%）。

他の設問でも情報（他学会の情報も含めて）提供を求める声が強く、コンテンツの充実（各種の化学サイト充実等）が望まれています。

また部会員間の情報交換を盛んにする手段として、電子掲示板機能の設置が効果的と考えられています（76%）。

5 CICSJについて＜設問（6）に関して＞

会員の皆さんが特集記事だけでなく行事や編集後記にも目を通されていることが確認できました。部会内のコミュニケーションをさらに盛んにしていくためには、部会活動の様子をもっとわかりやすいものにしていくことも重要です。

媒体（冊子か電子ファイルか）については意見がわかれました。それぞれ良さがあるようですので現時点でどちらかに統一することは難しいと考えられます。ホームページの充実も含めて検討していきます。

3　今後の活動に向けて

情報化学部会は種々の異なる専門分野を包含するかなり学際的な集団であること、そして部会員の皆さんがお互いの分野に非常に興味を持っていることがわかりました。

部会に期待されていることを大括りにまとめてみると、

１）情報化学に関連する情報の提供及び情報交換できる場の提供、

２）情報化学分野に係る層の拡大、

３）情報化学研究のレベルアップ、

となります。

これらのご期待に応えていくための案として以下の項目につき検討を開始します。

・今後「情報化学」に取り込んでいく新しい領域

・CICSJ に加えて、電子メール、ホームページ（電子掲示板含む）も有効利用した、部会員への情報提供の増加

・他学会（例えば高分子、触媒学会）との交流促進（具体的には共同の講習会・討論会等の企画）

・討論会の内容充実（領域ごとのレビュー講演、研究動向や部会活動に関するパネルディスカッション開催等）

・論文誌発行（研究者の発表機会増加、レベルアップを目指して）

・情報化学全般を一通り学べる基礎講座開催（討論会の機会を利用して開催等）

上記の検討課題につきましては、今後も CICSJ Bulletin、電子メールを通してあるいは情報化学討論会の場で継続的に議論を進めながら実施していく予定です。今回のアンケートのまとめに対する自由なご意見等をお待ちしております。

ご意見の送り先：（社）日本化学会 情報化学部会

E-mail： cicsj@chemistry.or.jp　　　　FAX: (03)3292-6318

◆関連記事

Web における構造作図・検索ツール

科学技術振興事業団　時実　象一
<tokizane@pc.highway.ne.jp>

1.　はじめに

本テーマについては以前CICSJ Bulletin 15(4), 1997 に投稿し[1)]、その中で Daylight – GRINS[2)]、MDL – Chemscape[3)]、CambridgeSoft Corp. の ChemDraw/ChemFinder[4)]、Tripos の WebSketch[5)] について紹介した。

その後約 1 年半を経てこれらのサイトを再訪問してみたところ、各サイトとも現在もアクセス可能であった。要約すると

- Daylight - GRINS 特にバージョンアップはされていないようである。
- Chemscape Chime Ver 2.0 となっているが画面上の変更は無いようである。
- ChemDraw 現在 ChemDraw Plugin としてダウンロード可能で、ここで作図した構造は ChemFinder で検索に使える。
- Tripos Sketch & Fetch はかなり画面が変更され，使いやすくなった感じである。

その後新しい構造作図インタフェースモジュールが開発されているので、今回はこれらについて紹介したい。今回紹介するのは ACD / Structure Drawing Applet (ACD/SDA)、Synopsys Scientific Systems Ltd. が開発した Accord ActiveX Chemistry Control と ChemAxon Ltd. の Marvin Sketch である。これらはインターネットまたはイントラネット上で化学構造検索を行ないたいという需要に答えるものである。これらのうち ACD/SDA は Java アプレットで mol ファイルをサポート、Accord ActiveX Chemistry Control は Active X で SMILES、mol ファイルをサポート，Marvin Sketch は Java アプレットで SMILES、mol ファイルをサポートしている。

Accord ActiveX Chemistry Control はそれ自体には構造作図モジュールは持っておらず、ISIS Draw で作図した構造をコピーして貼りつけ、各種形式に変換し、検索に用いる。

2.　ACD/SDA[6)]

Advanced Chemistry Development, Inc. では ACD / Structure Drawing Applet (ACD/SDA) を無料でダウンロード提供している。これは Java アプレットで Netscape 2.0 以上、Internet Explorer 3.0 以上で動作し、ブラウザ上での構造作図を可能にする。このアプレットのパラメータの指定方法は Example として示されているので、簡単に利用できる。

図 1 は通常の作図画面である。たとえばこの画面を呼び出すには

```
<APPLET CODE="StructureEditorApplet.class"
width=500 height=350>
</APPLET>
```

と HTML 中で指定すればよい。

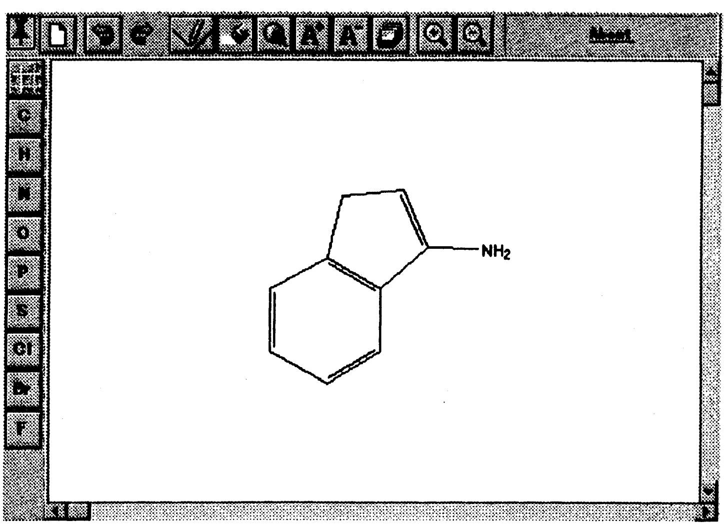


図 1. ACD/SDA の作図画面

をクリックするとテンプレートのウィンドウが

表示される (図 2)。

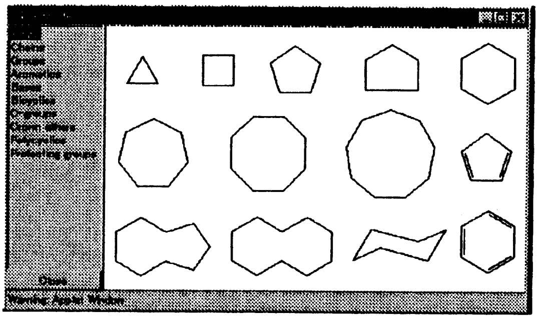


図 2. ACD/SDA の環テンプレート

ここに示したのは環のテンプレートであるが、使いたい環のどこかの結合をクリックすると有効となり、作図画面にもどる。環を縮合するときは最初にクリックした結合が縮合点となる。テンプレート・ファイルは mol ファイルなので自分でカスタマイズできる。

画面左側に主要な元素があるが、をクリックすると周期表が表示される (図 3)。選択した各元素について簡単な説明が表示されるようになっている (上部中央)。

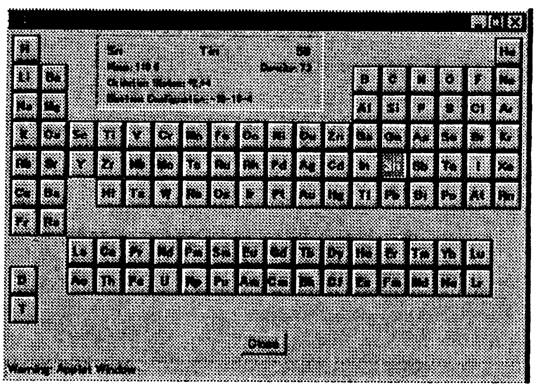


図 3. ACD/SDA の周期表画面

鎖を伸ばすのは適当な原子のうえからドラッグすればよい。結合の上をクリックすると単結合、二重結合、三重結合、単結合という順に変更される。また構造を移動したいときはをクリックしてから構造の周りをドラッグするとその構造が選択されるので、これをつかんで好きな位置までドラッグすればよい。

選択したテンプレートや元素を無効にするには(Normal) をクリックする。は電荷の指定ボタンである。

図 4は一般式構造 (Substructure) の作図画面である。

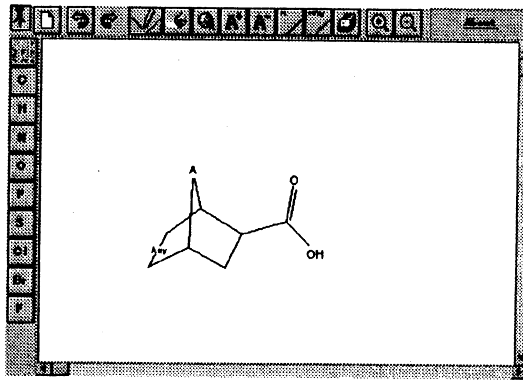


図 4. ACD/SDA で描いた一般式構造

原子を一般化するにはを、結合を一般化するにはを使う。ボタンの右下の三角の部分をクリックするとたとえば原子なら A, Q, HO (Not 水素), L (リスト), NL (Not リスト) などを選択できる。結合の場合は A, Ar (芳香族), S/D (単結合か二重結合), Rn (環), Cn (鎖) が選択できる

作成した構造図は mol ファイルに保存することができる。したがって mol ファイルを受け付ける化学構造データベースなら何でも検索可能となる。

3. Accord ActiveX Chemistry Control[7]

Synopsys Scientific Systems Ltd. が開発した Accord ActiveX Chemistry Control の最終ベータ/体験版が入手できたので試用してみた。これは作図された科学構造から SMILE, mol ファイルなどのデータを生成し，構造データベースにリンクするためのツールであり，作図モジュールは自分では持っていない。

これは ActiveX を使うので Internet Explorer を使う必要がある。IE 3.0 では支障なく動く。作図ソフトとしては ISIS Draw が使える。ChemDraw は今のところサポートしていない。ISIS Draw を使ってたとえば図 5 のように作図

した後 Copy して、Accord ActiveX Chemistry Control の "Paste Chemistry" ボタンをクリックすれば図 6 のように図が貼り付けられる。

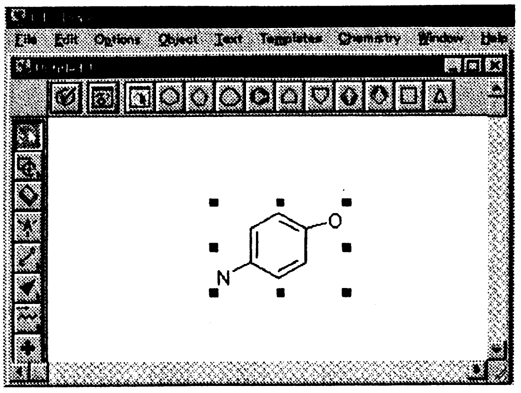


図 5. ISIS Draw による作図

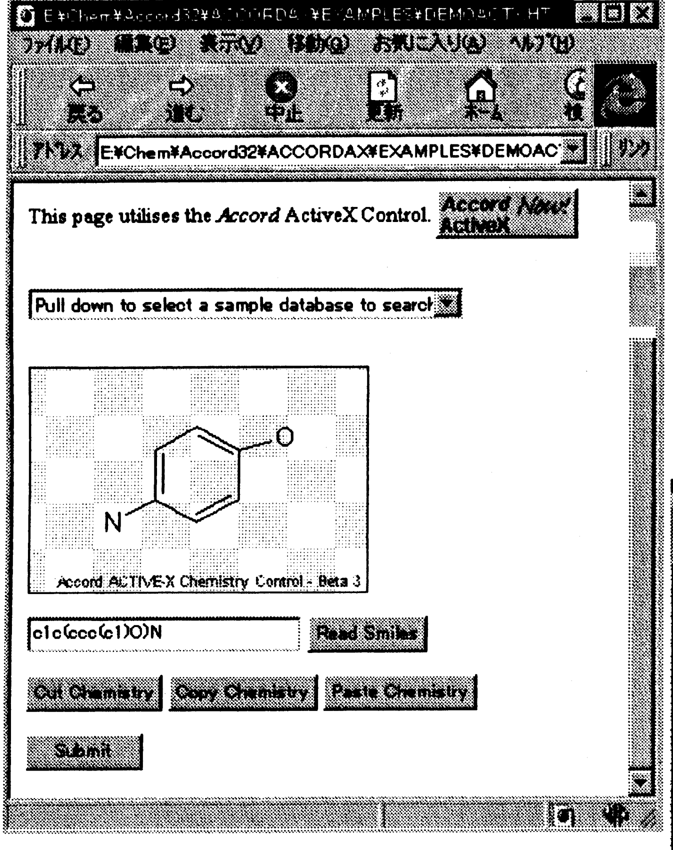


図 6. 図 5 の構造を Paste Chemistry でコピーしたところ

次に "Pull down to select a sample database to search" をクリックして、たとえば ChemFinder Web Database を選択し、下の "Submit" をクリックすると、直ちに ChemFinder に接続して、図 7 のような回答が得られる。以上の操作は極めて簡単である。

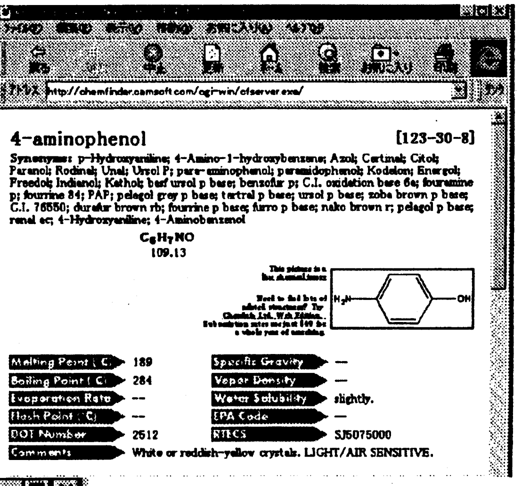


図 7. 図 6 で ChemFinder を選択して、Submit をクリックすると、すぐに ChemFinder から回答が得られる

またこの ActiveX Control はインターネット、イントラネット上の MDL mol ファイル, Rxnfile, Sketchfile, SMILES, SMD or Accord native format の化学構造を読み取ることができる。読み込んだ構造は簡単に 3D 化できる (図 8)。

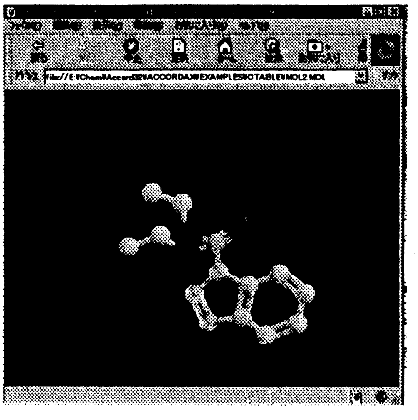


図 8. サンプル・データを 3D 表示したところ

本ソフトは Web 上での化学構造検索その他の化学構造情報処理に非常に有用と思われる。今後他の作図ソフトとの連携が望まれる。

4. Marvin Sketch[8)]

ChemAxon Ltd. の Marvin Sketch という構造作図モジュールが無償でダウンロード可能である (ユーザ登録は必要)。ただしインストールするには Java Runtime Environment が必要で、Windows ユーザなら JDK 1.1.7B または 2.0 を Sun のサイトからダウンロードしてインストールする必要がある。この JDK は DOS ウィンドウを開いて (懐かしの) DOS コマンドで起動しなくてはならない (パスを正しく指定する必要がある)。

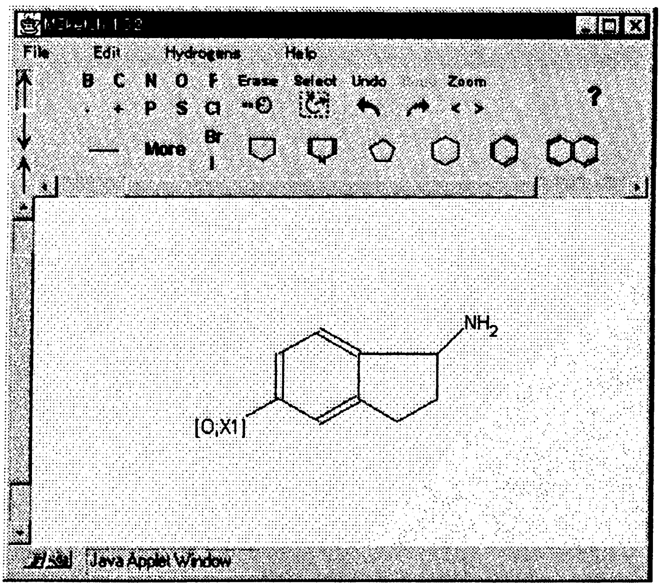


図 9. Marvin Sketch の作図画面

このプラグインはディレクトリにできる index.HTML から起動できる。ここで Sketch a Molecule をクリックするとコンソール画面と図 9 のような画面が開き、ここで簡単に作図できる。

メニューの "More" をクリックすると周期表 (図 10) が表示されるが、ここには一般式原子 (X, R) や立体化学結合も含まれている (上部中央)。

作図した構造についてはコンソール画面で mol ファイルなどに変換できる (図 11)。これらの情報をデータベースに送るためのコマンドはアプレットのパラメータとして用意されている。

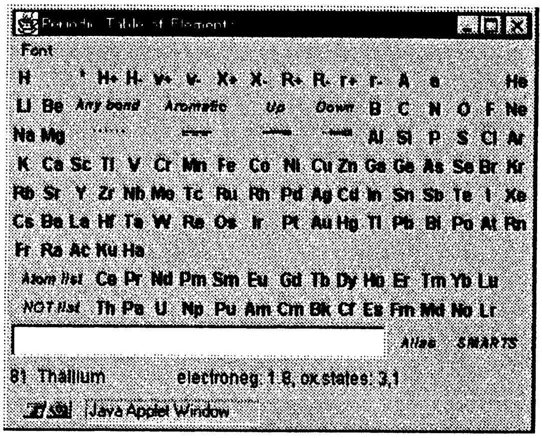


図 10. Marvin Sketch の周期表画面

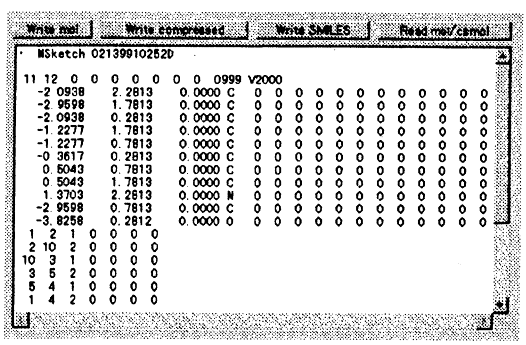


図 11. Marvin Sketch のコンソール画面で

mol ファイルデータを生成したところ

5. 終わりに

なお商業オンライン検索サービスでは STN が 1999 年 4 月に発表する STN on the Web[9] で構造作図プラグインを採用する予定である。これは従来検索インタフェースとして使われてきた STN Express の構造作図モジュールを使用している (図 12)。また ISI の ISI Chemistry Server[10] でもほぼ同じモジュールがプラグインとして使われている。後者のプラグインは同じ Thomson グループの Derwent Information が計画している Derwent Discovery でも使われることになっている。

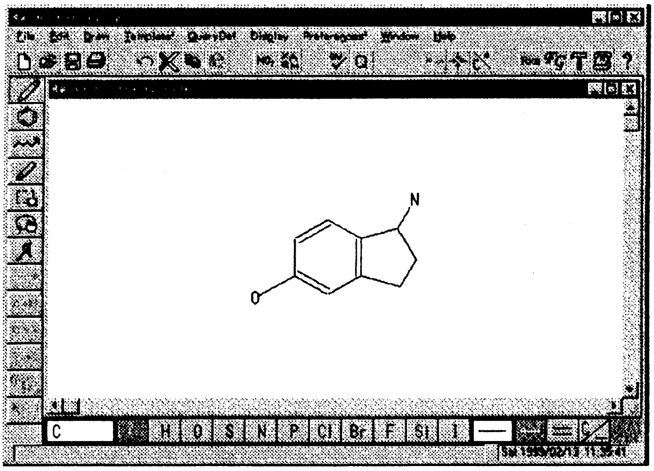


図 12. STN on the Webの構造作図プラグイン

今後とも Web で構造作図と検索をサポートするツールはさまざま出現すると思われるので注目していきたい。いまや化学構造検索システムはこれらのツールを使えば誰でも組み立てられる時代になってきている。

[参考文献]

1) 時実象一、「WWW 上での構造検索手法」、CICSJ Bulletin 15(4), 1997

2) Daylight - GRINS
<http://www.daylight.com/daycgi/testgrins>

3) MDL - Chemscape
http://www.mdli.com/chemscape/index.HTML

4) CambridgeSoft Corp. - ChemDraw/ChemFinder
<http://www.camsoft.com/>

5) Tripos - WebSketch
<http://www.tripos.com/SandF.HTML>

6) ACD/SDA
<http://www.acdlabs.com/products/java/sda/>

7) Accord ActiveX Chemistry Control
<http://www.synopsys.co.uk/products/software/accordax/activex.HTML>

8) Marvin Sketch
<http://www.chemaxon.com/marvin/>

9) STN on the Web <http://stnweb.fiz-karlsruhe.de/>

10) ISI Chemistry Server
<http://www.isinet.com/products/chem/server/server.html>

会員募集

Chemical Structure Association

Chemical Structure Association (CSA) <http://www.chem-structure.org/> は化学構造情報に関する仕事をしている人達のための国際的組織です。現在会長は Janet Ash <ash@euronet.nl > です。

CSA は分子設計、コンビナトリアル化学、化学構造情報処理システム、化学情報データベース、オンライン・システム、インターネットの化学情報、マニュアルと CD-ROM のツールに関心のある会員からなっています。CSA の会員の多くは化学構造情報の世界的な専門家です。

会員には年 3 回ニュースレターが郵便及び Web でパスワード付きで提供されます。また他の機関と協力して国際会議や討論会が開かれ、また CSA は化学構造処理に関する討議の場を提供します。(過去のニュースレターは CSA のページで誰でも見られます)

●1999 年の CSA 共催の会合

- 1999 年 3 月 21-25 日: カリフォルニア州アナハイムで開かれる米国化学会年会において、「インターネット時代のオンライン・サービス」 と題する CSA/CINF シンポジウムを共催。CAS, ACS, Dialog Corporation, Questel/Orbit, ChemWeb, ISI, ChemConnect, TDS, NIST などから発表者がある予定。
- 1999 年 6 月 6-10: オランダ、Noordwijkerhout で 5th International Conference on Chemical Structures を開催。

●会員募集

入会希望の方は時実象一 <tokizane@pc.highway.ne.jp> (〒102-0081 東京都千代田区四番町 5-3 科学技術振興事業団情報事業部 tel: 03-5214-8402, fax: 03-5214-8420) までご連絡下さい。年会費 20 ポンド (円で支払い可能) です。

CICSJ Bulletin Vol.17, No.4, August 1999

このページは「日本化学会・情報化学部会」の責任において運営されています。

目次

◆特集：データマイニング

データマイニング特集を企画するにあたって

大阪大学大学院薬学研究科　高木達也
ttakagi@phs.osaka-u.ac.jp

「現代は情報の時代だ」というような言葉は、もう、聞き飽きるほど聞きました。多分、この文章をお読みの皆様も同様でしょう。では、少なくとも私自身が、そして、ひょっとしたら多くの方々が、情報をいかに活用しているのかと問い直してみると、少々心もとない気がします。…そう言えば、「情報はあふれている。あとはそれをどう使うかだ」という文章も聞き飽きました。しかし、「そうですね、じゃぁ、こう使いましょう」という文章には、なかなかお目にかかれません。もちろん、問題提起は重要なのですが、この件に関する限り、問題提起ははるか昔に終わっています。その意味では、まだ、現代は情報の時代ではないのかもしれません。そろそろ、情報をどう料理するのか、解決法を示さなければいけない時期に来ているといっていいでしょう。

幸い、「データマイニング」という言葉が最近、盛んに聞かれるようになってきました。これら大量の、必ずしも整理されているわけではない生の情報は、それを眺めているだけでは何の実質的な情報にもなり得ないのですが、データマイニングはそれらの「生情報」から、整理、視覚化、抽出、分類…といった作業を行うことによって、有益な情報に加工する方法・過程だと（少なくとも私は）認識しています。

データマイニングで用いられる方法論は多岐にわたっています。ANN や HMM もその一つですが、今回、ご論文を頂きました先生方に解説いただきましたように、まだあまり国内の化学分野で応用がはじめられていない実に多様な手法が、すでに他分野では盛んに用いられつつあります。いずれも金田先生がご論文中で触れておられます、「Ockham's razor」の原則に従い、「いかに単純な手法、前提のみで、より多くの有益な情報を掘り起こせるか」が競われている方法論だと、私は、素人なりに考えています。

少し話は横道にそれますが、ここに出てくる「Ockham」という人は…いまさら解説するのは読者の皆様に失礼かもしれませんが…「William of Ockham（Guillelmus de Ockham）」として有名なフランチェスコ派修道士にて論理学者、神学者（スコラ哲学者）（1286-1347?, 1349?）で、彼自身、同様のことを何度か述べているようですので参照する文献によって原文でさえ少しずつ異なるのですが、そのうちの一つは、'Pluralitas non est ponenda sine neccesitate.' となっています。私はラテン語を解しませんので、Physics FAQ[1] によって、"Entities should not be multiplied unnecessarily."と解しています。原理自身が当該の原理に従って簡潔に記されているという見事な例だと常日頃感じております。

素人が口幅ったいことを申し上げてしまいました。今号では、これから化学分野で重要な役割を担うことが確実であろうデータマイニングを取り上げ、第一線で活躍しておられる先生方に解説記事を書いて頂く事ができました。ご多忙な中、ご論文を執筆頂きました先生方に、この場を借りまして厚くお礼申し上げます。なお、この「企画するにあたって」などという偉そうな文章を書いている私が「企画」したことといえば、「言い出した」ということだけです。私の「言い出し」の犠牲になり、実際に企画の大半の労をおとりくださり、また、各先生方に仲介の労をおとり頂きました、関西学院大学の岡田孝先生にはお礼の言葉もございません。深く感謝の意を表させて頂きますことで、お許し頂けましたら幸いです。

先述のように、また、ご論文中で岡田先生が触れておられますように、国内の化学分野では、decision tree をはじめ、まだ多くのデータマイニング手法がほとんど応用の機会を与えられていないように思われます。この企画が、データマイニングの化学分野への普及の一助にでもなれば、望外の幸せです。

参考文献

1) Chase, S. I.; Weiss, M.; Gibbs, P.; Hillman, C.; Urban, N. *Physics FAQ* http://www.weburbia.com/physics/faq.html

---

たかぎ たつや　　TAKAGI, Tatsuya
連絡先:〒565-0871 大阪府吹田市山田丘 1-6　tel/fax 06-6879-8250
:大阪大学大学院薬学研究科 応用医療薬科学専攻 医療薬物科学講座 医薬情報解析学分野

データマイニング技術の現状

日本アイ・ビー・エム株式会社コンサルティング事業部　塩田 千幸
SHIOTAC@jp.ibm.com

大量のデータから知識を獲得する方法としてデータマイニングが本格的に利用され始めている。データマイニングを成功裏に行うためには、プロセス全体が重要である。プロセスの中心にデータマイニング手法が位置付けられるが、その技術面も広く捉え、利用目的に応じてデータ特性との関係等も総合的に考える必要がある。ここでは、現時点で利用が進んでいるビジネス分野での応用をベースとして、データマイニングの概要、プロセス全般、データおよび技術面の順に要点を整理・紹介する。

1 データマイニングの現状

1.1 データマイニング利用の概況

ビジネスの世界では膨大なデータの蓄積が図られてきた。例えば、大手銀行では１０００万件以上の顧客に関する属性データと日々の取引データを保有している。この種のデータは日常業務のために利用するだけでなく、様々な角度から分析して最大限の価値を引き出そうとする試みが各方面でなされている[3)]。このような潮流の中で大量のデータから知識を獲得する手段としてデータマイニングは一般化しつつある[5)]。既に米国の企業では数年前から数テラバイトのデータウェアハウスを構築し、大量データの分析に取り組んでおり、我が国でも急速に利用が進み始めている。米国では企業は一般に期待効果の15-20%をデータマイニングに投資する傾向が見られるという[13)]。化学分野でのデータマイニングについては、欧米のビジネス系文献を検索した結果によると、利用の方向は製薬とバイオテクノロジー関連が多いようである（例えば参考文献 10)など)。

データマイニング適用については、データに関わる問題、データマイニングの目標設定や結果の評価の問題などがあり、単純にデータに手法を適用して完了という訳にはいかない。データマイニングを成功させるためには、後述するプロセス全般の問題への対処が必要である。

1.2 データマイニングとは

現在使われる意味でのデータマイニングは比較的新しい概念であり、まだ発展段階にあって、定義も確定したとは言い難い。コンピュータ技術の関係者、人工知能・機械学習技術の関係者、ビジネス関係者により考え方が少しずつ異なる[14)]。本稿では複数の手法の利用や試行錯誤を含めた一連の流れが重要であることを強調する意味で、データマイニングを、「大量のデータから探索・分析手法を用いて有用な知識を新たに獲得するプロセス」として扱う。特に、大量のデータを対象にすることがデータマイニングの最大の特徴であるが、それは分析者にとって、価値ある情報が簡単には発見できないことだけでなく、発生頻度の少ない事象も取り扱えること、十分な量の分析用データと検証用データが確保できること、信頼度の高い結果を得るために試行錯誤を繰り返すことができること等も意味する。

データマイニングの分類についても、いくつかの考え方がある[1)]。例えば、未知の情報をデータから抽出するような「知識発見型」と、予見を実証するか反証するための「仮説検証型」の２つのタイプに分ける考え方がある。この分類はデータマイニングのプロセスと関係がある。また、扱う問題の目的による分類では、「予測型」と「記述型」に分けられる。使用法(operation)により、予測モデリング(predictive modeling)、データの層別(database segmentation)、リンク分析(link analysis)、特異性の検出(deviation detection)に分ける考え方もある[4)]。このような分類は手法の分類と重なる。

データマイニング手法はデータ分析手法と見れば統計的手法と接点がある。両者に明確な境界線を引くことは困難であり、それほど意味のあることとも思えないが、応用の観点から見ると、データマイニングの方が、より大量のデータを曖昧な状態で取扱う、という違いがありそうである。

データマイニング・プロセスは、(1)目標設定、(2)前処理、(3)データマイニング手法の適用、(4)後処理、(5)結果の利用（意思決定等）のステップから構成され、主に(2)-(4)が必要に応じて反復される。このプロセス中で手法を適用するステップ（3）は重要ではあるが、プロセス全体に占める時間的・作業負荷的な割合は多くなく、通常は半分にも満たない。特に時間がかかるのは前処理のステップで、データ源からのデータ抽出（目的に対して適当でないデータの除外、分析用データと検証用データへのランダム分割など）、エラーデータの除去や不整合・欠損のチェックを行うデータ整理、分析に適したようなデータ変換、データ統合などの準備作業を含む。これらの作業では人の判断に依存する試行錯誤を繰り返すが、目標の設定・課題の定義などの最初のステップと合わせて、データマイニングの成功のために人間の果たす役割は大きい。

1.3 データマイニング・プロセスの特徴

データマイニング・プロセスには曖昧な要素が含まれがちである。まず、目標設定については、取り組むべき課題や価値ある情報を期待する領域が曖昧な場合が多い。例えば、「ダイレクトメール(DM)の反応率を高める」ためにデータマイニングを行うとすると、反応率を高める対象、DMの時期・内容、他の補完手段との組合せ、「反応率」以外の評価尺度などに曖昧性があり、人が総合的に判断しなければならない。

目標設定後に、必要なデータを準備する段階の問題は次節で触れる。

適切なデータマイニング手法を用いて情報を抽出した後で、その情報に基づき行動を起こし、結果を評価するまでのステップでも、結果から価値を見出す部分に曖昧さがあり、人の判断と裁量に任される度合が高い。例えば、「DMの反応率向上につながるような価値あるルール」を求めている場合、「少数の顧客層に 95%の確度で成り立つルール」と「少し多めの顧客層に 75%の確度で成り立つルール」の何れが価値があるのか等の判断が必要である。その際に確度の高さの評価基準も必要であり、例えば、従来のDMの反応率が 2%程度であれば 20%の確度でも利用上は十分高いと言える場合もあろう。

一連のステップでは、視覚化の技術が関係者の理解を深め適正な判断をする上で有効である。視覚化は、様々な工夫がされたツールがあり [13]、決定木のように手法と一体と見なせる場合や後処理でのプレゼンテーション技法として使われる場合がある。一般に結果を見て判断する関係者が必ずしも分析に詳しくないこと等からシンプルで分かりやすいものが望ましい。ただし、視覚化は高次元データの表現には弱いという限界がある。また、データマイニング手法の選択に当っても「分かりやすさ」は重要なポイントである。例えば、同程度の精度が得られるならば、ニューラルネット（バックプロパゲーション）でなく決定木の方が分かりやすさを理由に選ばれる。ニューラルネットの方が精度が高い場合でも、決定木の誤判別ケースを検討して新たに変数変換を追加すること等で決定木の精度を向上させられれば、決定木を採択することがある。

データマイニングでは目的に応じて複数の手法を組み合わせて使用することが効果的である。例えば、クラスタリングにより個体を均質な層に分類し、その層毎に成り立つルールをアソシエーション分析で抽出する等である。また、統計的分析手法の補完的な利用も有効である。例えば、主成分分析等により多数のデータ項目を整理して絞り込みデータマイニングを行うこともある。

データマイニング・プロセスでは、データの特徴の理解と分析結果の評価のサイクルを満足できる結果が得られるまで反復する。1回のサイクルで終わることは稀である。また、データの更新が進めば分析内容の向上や予測モデルの修正のために、随時、プロセス全体を改めて繰り返す必要がある。

2 データマイニング利用に関する技術的側面

2.1 分析対象データの問題

データマイニングの有効性を決める最大の要因は分析対象となるデータであり、十分な質と量のデータが必要である。データの内容も重要で、適切なデータ変換の方法を選択することの方が適切な手法の適用よりも結果に大きな影響を及ぼすことさえある。

分析対象になるデータは大量に保管されていても、そのままではデータマイニングに適した内容とは限らず、生のデータ利用が必ずしも良いとは言えない。例えば、異常データの混入、データマイニング目的に不適当な個体の混在等の問題がある。データ内容について深い知識を持った人がデータマイニングのプロセスに参画することが重要である。データの問題はデータマイニングのテーマやデータの種類により個別の対応が必要であるが、ここでは、データマイニング手法適用に関わる一般的なポイントを述べる。

a．<u>データ個体</u>　：　データマイニングでは「標本データ」は扱わないのが基本なので、通常はデータの代表性等の標本データに関わる問題はない。ただし、母集団の定義の問題はあり、データベースに登録されている全件データが常に母集団とはいえない。また、全件データをそのまま分析対象とするのが良いとも限らない。分析目的に合った個体のみを抽出した方が良い場合がある。ただし、特異データについての吟味から新しい発見が得られる場合もあることに注意が必要である。通常、オリジナル・データは分析用データと検証用データにランダムに分割するが、その際、データの発生時点の偏りやデータ重複の有無等への配慮は必要である。

b．<u>データの均質性</u>　：　マイニング結果を利用する観点からデータ全体を見ると均質とはいえないことが多く、均質な層に分けた方が望ましい場合がある。単純な分類基準が事前に想定されにくい場合も、クラスタリングによりデータの均質性を確かめることは意味がある。勿論、支障がなければ全体を分割せずに分析すべきである。

c．<u>データ項目</u>　：　横断面データ項目と時系列データ項目があり、項目タイプには連続量をとる数値タイプと離散量をとるカテゴリー・タイプがある。相互に相関関係のある項目も多い。カテゴリー項目には階層構造（大分類・中分類･小分類）のある場合があり、マイニング手法と目的に応じてデータとして利用するレベルを決める。データ項目については、内容の定義が曖昧であったり、とる値が不確実な場合がある。これは、必ずしも分析目的で収集されたのではない既存データを利用しようとすることにも起因する。また、データ項目数は数百以上あることが多いが、全項目を分析対象にしない方が、データマイニング結果を解釈・利用する際に都合のいい場合も多い。例えば、意味の重なる項目を絞り込み、アソシエーション分析で自明のルールが大量抽出されることを回避したり、決定木が複雑になりすぎないようにする。採否判断も含め、データ項目の利用に当っては業務知識が不可欠である。

d．<u>欠測値</u>　：　発生頻度は一概に言えないが、欠測値は基本項目でない属性項目や古いデータに含まれることがある。通常は欠測の発生頻度が多い項目は除くか、不明データのあるケースを分析対象外として除外する。連続量をとる項目の場合は、離散化して欠測値のカテゴリーを設ける場合もある。明確な根拠もなく欠測値を他の値で置き換えることはしないのが通例である。

e．<u>データの正確性</u>　：　本来、正確であるべき項目と誤差があっても仕方がない項目がある。特異値発生･混入の理由に応じて、複数の項目についてデータが過大または過小になっていたり、データ項目間の関係にも歪みが生じていることがある。このような場合は特異値混入の影響を受けにくい項目を分析するか、項目間の関連性に基き妥当な値に補正する等の対策をとる必要がある。

f．<u>データの定常性</u>　：　データ項目とデータ個体について、それぞれ時間的側面を考える必要がある。即ち、時系列データとして扱われるべき項目がある一方で、データ全体に含まれる個体の新旧という問題がある。時系列データ項目のタイプには、等間隔データ、不定間隔データ、発生間隔のデータがあり、よく使われるのは等間隔データである。時間的要素を考える場合、ビジネス・データは経済環境などが常に変化するため基本的に非定常である。従って、時系列データ期間は、マイニング目的（前兆を早目に察知したい等）に必要な期間と定常性が見込まれる期間限界のバランスを考える必要がある。また、マイニング対象個体データ期間も定常性が想定できる範囲に限定すべきである。

g．<u>データ変換</u>　：　通常の統計分析と同様にデータ変換の有効な場合がある。データ変換には、新しい変数の合成、正規化、平滑化、差分・比率計算、カテゴリー化などがある。特に、数値項目について、比率計算とカテゴリー化が重要である。例えば、予測モデルの精度向上や決定木の肥大化防止等のために複数項目の比をとることがある。カテゴリー化は、分析目的と手法に合わせ分析者がデータ分布を見て事前に区切ることが必要な場合もあるが、決定木やクラスタリング等でデータマイニング・ツールがデータから判定した適切な閾値を設定する場合もある。また、カテゴリー項目でも、原データのと

る値が細分化されすぎているので再カテゴリー化した方がよい場合があり、例えば、アソシエーション分析でうまくルールを抽出する上で有効である。

2．2 主要なデータマイニング手法

データマイニングは、機械学習、パターン認識、統計学、視覚化手法、データベース技術などの系統を引いており、データマイニングのための手法と呼ばれるものは多い。しかし、データマイニング手法について明確な定義はなく、必ずしも新しい手法ばかりでもない。 それが最近になって一種の流行ともいえる状況にあるのは、データウェアハウスの導入等で大量のデータを取り扱いやすくなってきたこと、コンピュータの計算能力の向上で手法面の制約が緩和されたこと、蓄積されたデータを活用したいというニーズの高まり、各種ソフトウェア・パッケージが整備され利用可能になったこと等が背景にある。

データマイニング手法は、扱う問題の目的によって、予測のための手法と特徴を発見し記述するための手法とに大別できる。前者の結果は予測精度で評価され、後者の結果は説明力・簡潔さなどで評価される。また、データ項目として離散量と連続量の両方を同時に扱えるように拡張されている手法の使い勝手が良い。

予測のための手法では、データの特定の項目(被説明変数) を複数の他の項目(説明変数) で説明させるモデルを構築する。被説明変数のとる値が離散量の場合はクラス分類(classification)と総称され、決定木(decision tree)とバックプロパゲーションがよく利用される手法である。連続量の場合、ＲＢＦ(Radial Basis Function)が利用され、後の２つは教師付き学習(supervised learning)とされるニューラル・ネッワークを用いた手法で、予測モデルとしてはブラック・ボックスである。また、記憶に基づく推論(MBR:Memory-Based Reasoning)のように、類似例からの距離を計算して近隣性を判定し、クラス分類や予測を行う手法もある[1]。

記述のための手法では、データ項目間に説明変数・被説明変数の区別をつけず、データの内容・構造を理解しやすく簡潔に表現することを目的とする。クラスタリングとアソシエーション分析がよく利用される手法であり、前者には教師なし学習(unsupervised learning)とされるニューラル・ネットワークを用いた手法(Kohonen Feature maps)も含まれる。

また、前処理のための技法として、カテゴリー化の方法(平滑化、クラスタリング等)、データのサンプリングなどがあり参考文献12)に詳しい。

ここでは、手法の全般的な説明は他の文献に譲り[1),12),14)]、比較的使いやすく、ビジネス分野等でよく利用される手法の特徴を利用上の観点から述べる。

ａ．<u>決定木</u> ： 被説明変数のとる値(クラス) を説明変数の組み合わせにより説明するルール群を tree 構造で生成する。決定木の手法では、ノード分割の基準・形態、枝刈り(pruning)の方法などにより異なるアルゴリズムが使用される。一般的な特長として、生成されたルールの分かりやすさと結果の使いやすさ、計算負荷の軽さ、特異データに影響されにくいこと等がある。一方、説明変数間の相関が高い場合に決定木が大きくなること（この問題に対応したアルゴリズムもある[18)]）、最小の決定木が得られる保証がないこと(NP 困難といわれている）[16)]等に注意が必要である。決定木の枝の分岐は各ノード単位に最適判定を繰り返して行うため、ツリー全体として最適分岐の形状である保証はない。また、説明変数全体を総合的に組み合わせて扱わないとクラス判別が困難な場合は、一般にバックプロパゲーションの方が望ましい。

ｂ．<u>バックプロパゲーション</u> ： 取り扱える問題の範囲が広く、変数間の関係が高度に非線形や複雑な場合でも良い結果を出せることが特長である。一方、計算の負荷が重いこと、内容がブラックボックスであることに注意が必要である。非常に多数で雑多な不正利用パターンを見分けて被害を最小に抑えることだけが重視されるような「カード不正利用検出」がビジネス分野での典型的な適用例である。

ｃ．<u>クラスタリング</u> ： データ全体を異質な個体群の集合とみなして、できるだけ均質なクラスターに分割するための手法であり、従来からクラスター分析として統計的手法がよく利用されている。クラスターを求めるためには、個体間の類似性の尺度とクラスター間の距離尺度が必要である。前者には、

個体をデータ空間内の点として、２点間の距離、２つのベクトルの角度(sin)、カテゴリー・データの場合の共通特性の数などがある。後者は、最も近い個体間、最も遠い個体間、クラスター中心間などで定義できる。クラスタリングを予測モデリングなどに先立つ予備的な分析として利用することも多い。データマイニング手法としては、連続量をとる変数と離散量をとる変数を同時に扱えること、クラスター数を事前に指定しなくてもよいこと、非常に大規模なデータを高速に処理できること等が望まれ、実用化されている(例えば参考文献8)。

ｄ．アソシエーション分析　：　同時に生起する傾向のある事象群とそれらの関係を発見するために用いられる。カテゴリー・データを対象に、全項目を組み合わせた『多重クロス集計結果』から一定の条件を満たす項目の組み合わせをルールとして抽出する（図２．３参照)。データ項目は全て離散量でなければならない。特徴として、結果がルールの形で表現されるので理解しやすいこと、ルール抽出の条件設定によっては無意味なルールが多数抽出されること、発生頻度の少ないカテゴリーのルールは出にくいこと等がある。最後の問題に関しては、カテゴリーをまとめて中分類、大分類などを設け発生頻度を揃える方法(taxonomy)がある。ルール抽出条件を変えて試行錯誤を繰り返す必要性が高い。

２．３　結果の有効性の評価

データマイニングの有効性は設定された目標(例えば、データを理解すること、ＤＭの反応率を上げること等）の達成度で評価できる。逆に達成度が評価できるような目標を設定しておく必要もある。一般的な有効性の指標としては、定量的なものでは、確実性（予測精度など)、簡潔性（ＭＤＬなど)、有用度（金額換算したコスト節約額など）等があり、定性的なものでは、新規性、明瞭性などがある。

予測のための手法については、予測対象が離散量であればエラー率（分析用データ、検証用データについての）による精度評価が一般的であり、連続量の予測であれば推定値と実際の値の差に基づく指標（平均二乗誤差など）で評価する。ただし、精度評価ではoverfittingに注意が必要である。例えば、決定木の手法では分析データから求められた末端ノード数が多いほど判定の精度は高くなるが、適正なノード数を越えているとoverfittingの問題が生じる。分析データとは別の検証データに決定木を適用して精度低下の程度を調べて適正な大きさまで決定木の枝刈りをすることが重要である(図２．１）[1]。

図２．１　分析データと検証データでの精度差

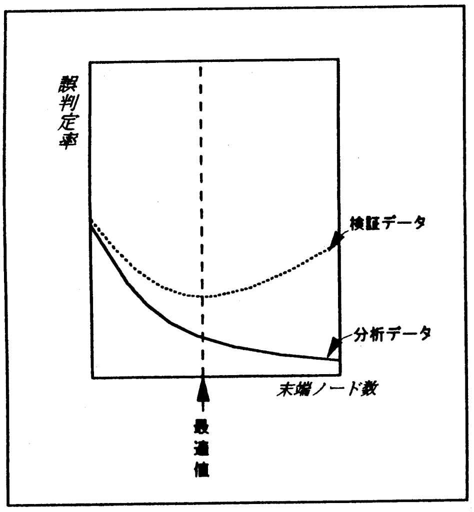
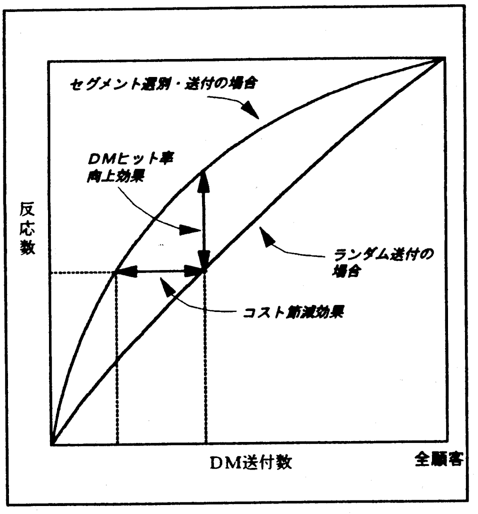


図２．２　データマイニング効果の評価例



記述のための手法の評価には、ＭＤＬ(Minimum Description Length)が使える。ＭＤＬは、データを表現するモデルを、モデル自身も含めたデータの符号化に要する符号長で評価する[17]。
例えば、クラスタリングによってデータを何個のクラスターに分類するのが適当かを判定する指標として参照できる[15]。また、決定木の最良サイズ判定にもＭＤＬ基準が使える[10]。一方、アソシエーショ

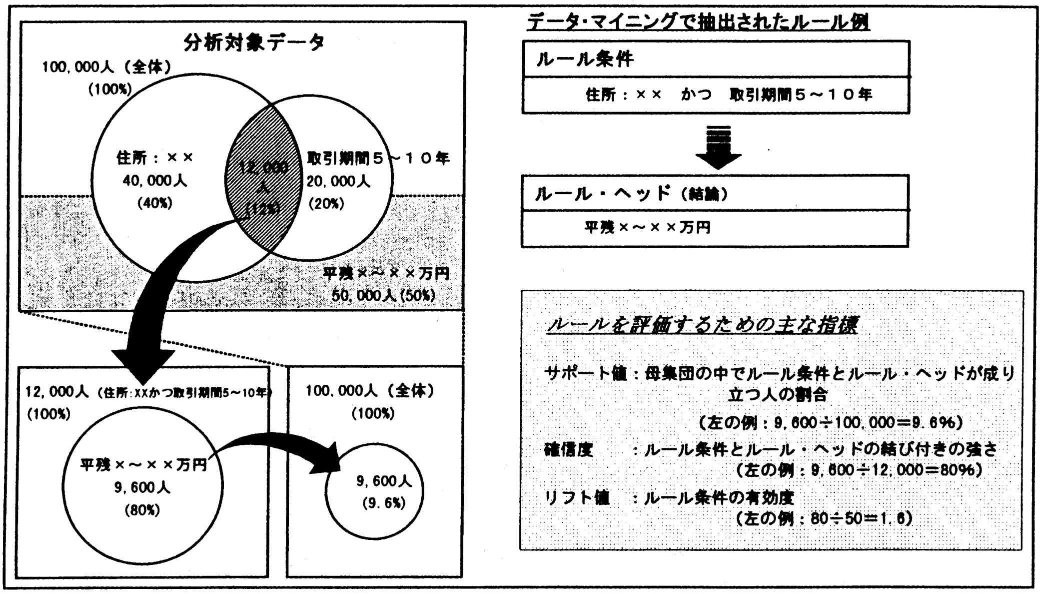
ン分析によりデータを記述した結果には不適当である。

マイニング結果を総合的に評価する指標として、リフト値(Lift)が役に立つ。i 番目のセグメントについてのリフト値は次式で計算できる。

$$(\mathrm{Lift})_i = \frac{(\text{セグメント内レスポンス数}(Ci)/\text{セグメント人数}(Ni))}{(\text{全体レスポンス数}(C)/\text{全体人数}(N))}$$

ここで、セグメントは、例えば、決定木で得られた末端ノードを示し、レスポンス数は特定のクラスに帰属する人数（例えばDM反応者数）を示す。リフト値が１を超えて大きいほど、そのセグメントはクラスの特性を明確に持ち、０に近いほど逆の特性を持つ。降順にソートしたリフト値（累積リフト値）を並べてリフト・チャート(Ｘ軸：ΣNi/N, Ｙ軸：(Lift)i）に表すことにより被説明変数(例えば、ＤＭに対する反応有無）に関する集中度が分かるので、特にクラス分類の手法の評価に効果的である。リフト値は１を境に大小の相対評価はできるが、データ件数により上限が変動するため、Ｙ軸を(ΣCi/C)とした図を使うことが多い（図２．２)。同様の意味合いでアソシエーション分析でもリフト値が評価指標として利用できる（図２．３)。ただし、アソシエーション分析では累積値を利用することはない。その他の手法でもセグメントが決められればリフト値を計算でき、例えば予測のための手法間で結果を

図２．３ アソシエーション分析から得られたルールの有用度の評価指標



比較できる。

最後に、結果の経済的価値を評価する例として「ローン与信審査モデリング」を取り上げる。この場合、信用リスクと収益機会のバランスをとりながらビジネスを推進するために、「ローン与信審査の精度を上げる」ことが目的である。データマイニング手法として決定木を利用して通常客と不良客をクラス分類すると、分析結果の精度評価にはエラー率が使える。しかし、経済的価値の視点で見ると通常客を不良客と見誤るコストより不良客を通常客と見誤るコストの方が高いため、単純な精度だけでなくコスト（価値）で重み付けられた精度を見る必要がある。従って、エラー率精度表(Confusion matrix)に同じ枠組みの価値ウェイト表(Value matrix)を組み合わせ、要素毎の積和で総合価値（コスト）を求め、分析ケース間の比較をする。

3 今後の展望

今後データマイニング利用が進むと予想される新しい応用分野に、テキスト・マイニングがある。テ

キスト・マイニングは、膨大な量のテキスト・データを対象に、キーとなる情報の抽出、主題別に文書の再編成、文書群から主要なテーマの発見、関連文書の探索などを行う[7]。そこでは、カテゴリー化、クラスタリング、検索などの技術・手法が用いられ、特許情報の分析、Ｅメールの応対処理、アンケートのフリーコメント分析など様々な領域での利用が考えられる。

今後、データマイニングはビジネス領域だけでなく更に一般的に利用される傾向が強まり、ツールの整備が進むことが期待される。ツールの進化の方向性としては、各種のデータマイニング手法群と統計分析手法等が統合されるものと、特定の応用分野に特化したソフトウェアの中にデータマイニングが一機能として組み込まれるものが考えられる。さらに先進的な理論・アルゴリズムも利用分野の課題解決に適した形で応用可能になることも期待される[6),9)]。

参考文献

1)Berry,M.and Linoff,G. Data Mining Techniques. John Wiley, 1997

2)Bigus, J. Data Mining with Neural Networks. McGraw-Hill, 1996
(社会調査研究所, 日本ｱｲ･ﾋﾞｰ･ｴﾑ訳：ﾆｭｰﾗﾙ･ﾈｯﾄﾜｰｸによるﾃﾞｰﾀﾏｲﾆﾝｸﾞ, 日経BP社,1997)

3)Brachman,R.,Khabaza,T.,Kloesgen,W.,Shapiro,G. and Simoudis,E.
Mining Business Databases, Communications of the ACM Vol.39,No.11,42-48(1996)

4)Cabena,P. et al. Discovering Data Mining. Prentice Hall, 1998

5)Fayyad,U.,Haussler,D. and Stolorz,P. Mining Scientific Data.
Communications of the ACM Vol.39,No.11(1996)

6)Fayyad,U. et al.(Eds.) Advances in Knowledge Discovery and Data Mining. MIT Press,1996

7)Gotthard,W.,Seiffert,R. and England,L. Text Mining, Research Report, IBM, 1997

8)IBM. IBM Intelligent Miner for AIX User's Guide, IBM, 1996

9)Komorowski,J.,Zytkow,J.(Eds.) Principles of Data Mining and Knowledge Discovery. Springer,1997

10)Mehta,M., Agrawal,R. and Rissanen,J. (1996): SLIQ:A Fast Scalable Classifier for
Data Mining, Proc.of the Fifth International Conference on Extending
Database Technology

11)Moran,N. A crucial role for IT in drug research. Financial Times Surveys. Jan.13,1999,Page9

12)Weiss,S.and Indurkhya,N. Predictive Data Mining. Morgan Kaufmann, 1998

13)Westphal,C.and Blaxton,T. Data Mining Solutions. John Wiley, 1998

14)塩田千幸. データマイニングの手法と実際, 計算機統計学, Vol.10,No.2,127-144(1997)

15)鈴木秀男, 圓川隆夫. ＭＤＬ基準によるデータの分類, 応用統計学, Vol.22,No.3,117-132(1993)

16)福田剛志, 森本康彦, 森下真一, 徳山豪. データマイニングの最新動向, 情報処理,
Vol.37,No.7,597-603(1996)

17)松嶋敏泰. 統計モデル選択の概要, オペレーションズ・リサーチ, Vol.41,No.7,369-374(1996)

18)森本康彦, 福田剛志, 森下真一, 徳山豪. 領域及び区間分割を用いた決定木の作成,
情報学基礎研究会報告, 1997.8.

しおた ちゆき SHIOTA, Chiyuki

最近は金融機関向けにデータマイニングを使ったコンサルティング・サービスを提供しています。
連絡先 〒103-8510 東京都中央区日本橋箱崎町19-21 日本ｱｲ･ﾋﾞｰ･ｴﾑ(株)金融ｻｰﾋﾞｽﾏﾈｼﾞﾒﾝﾄ･ｺﾝｻﾙﾃｨﾝｸﾞ
電話直通 03-3808-5367 FAX 03-3664-4768

決定木学習：実用期を迎えたノンパラメトリック統計手法

同志社大学大学院　金田 重郎 [8]
skaneda@mail.doshisha.ac.jp

決定木は、対象物の属性について質問を積み重ねてゆくことにより、最終的にある種の判断を実現する知識表現である。専門家の判断事例を蓄積し、この事例から自動的に決定木を生成する決定木学習手法の発達により、近年、SPSS 等、多くの統計解析ツールにパッケージとして組み込まれるようになった。決定木は、線形モデルを用いた従来の統計手法とは異なり、非線形の知識表現であり、知識記述能力が高い。本稿では、決定木学習の代表的な手法である ID3,CART について紹介して、決定木学習の特長と適用範囲について紹介する。

1　はじめに

決定木モデル (樹形モデル) は、対象物の属性について質問を積み重ねてゆくことにより、最終的にある種の判断を実現する知識の表現方法である。古くから使われていた表現であるが、近年、ID3,CART 等の、事例から自動的に決定木を導く手法 (機械学習アルゴリズム) が種々開発されたことにより、顧客データ分析等の多くの応用への適用が活発化している。

決定木は、従来の重回帰分析等のパラメトリックな統計手法とは異なり、事前の事例分布や誤差に何らの分布も仮定しない、ノンパラメトリックな統計手法である。有力な統計解析パッケージである S-PLUS には以前から搭載されていた[1]が、数学的基盤が弱いと見られ、統計利用分野では、さほど重要視されてはいなかった。

これに対して、人工知能 (AI) 分野では、機械学習の重要な 1 分野として、この 10 年程の間、盛んに研究され、実用的な手法が種々開発されている。代表的手法に ID3 がある。

ポイントカード、電子商取引 (EC) の進展により、大量のデータが蓄積され初めている状況から、最近では、決定木は、SPSS 等、多くの統計解析ツールに組み込まれている。線形モデルを用いた従来の統計手法とは異なり、非線形の知識表現であり、知識記述能力が高い。

本稿では、ID3 に焦点を絞りながら、決定木学習手法の原理・利点について紹介したい。

[1]S-PLUS では、樹形モデルと呼んでいる。

2　機械学習と決定木

2.1　機械学習

機械学習とは、事例からその背後にある一般的法則を帰納する技術を言う。決定木学習はその一分野であり、現実の社会に存在する事例 s_j を以下の形で表現する。

$$s_j = (a_{1,j}, a_{2,j} \cdots a_{i,j} \cdots a_{n-1,j}, a_{n,j} : C_j) \quad (1)$$

ここで、j は事例番号であり、事例の個数を M とするとき、$j = 1, 2 \cdots M$ である。$a_{i,j}$ は属性値である。**属性**とは、例えば、判断すべき対象が病名であり、事例が個々の患者であるとしたとき、患者個々の性別、咳の有無、発熱の有無、体重等が属性であり、その値を**属性値**と呼ぶ。ここでは、属性値に、1 から n までの番号を付した。n は、属性の個数である。C_j は事例 s_j が有するクラスであり、例えば、患者 s_j に対する病名である。これら属性からクラスを決定する知識を自動的に獲得することが、決定木学習の任務となる。即ち、病気の診断であれば、病名未知の患者を診断する知識を決定木として診断事例から得ることになる[2]。

機械学習とは、**M 個の事例（学習事例）が与えられた時に、クラスを推定する知識を自動的に生成する**ことである。但し、少ない事例からでも、高精度の知識が得られることが望ましい。

[2]統計分野では、クラスを被説明変数、属性を説明変数と呼ぶ。この呼び名の違いは習慣的なものであるが、属性値が数値ではないことを強調するために、本稿では、「属性」を利用する。

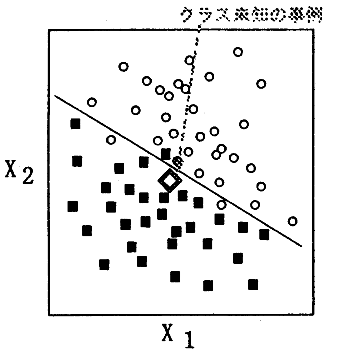


図 1: 線形分離の例

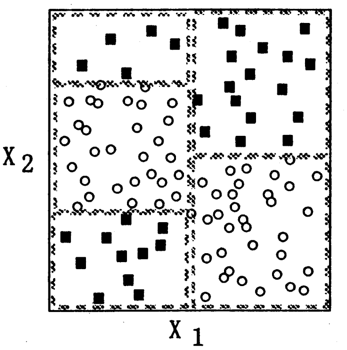


図 2: 線形非分離の例

2.2 線形分離

図1は、2個の属性,x_1,x_2 を持つ事例を2次元的に配置したものである。事例の持つクラスを○と■で表現している。例えば、過去の患者が、その症状により配置され[3]、事例対応のクラスである○または■が2種類の病名に相当する。

前述の目標「クラスが未知の事例に対して、高い精度でそのクラスを推定する知識を得る」について考えてみる。例えば、図1中の◇のように、属性値は分かっているが、クラスが未知の事例が得られたとする。病気の症状は問診できる。◇が○であるか■であるかを判断する必要がある。この場合、図1に示したように、領域を2分する直線[4]をあらかじめ引いておけば、直線のどちら側に入るかで、○か■かを判断できる。図1では、◇は■と推定できる。この直線による領域分割が、「高い精度でそのクラスを推定する知識」である。

図1のように、空間を直線で分割するものを線形モデルと呼ぶ。重回帰分析や判別分析等の統計解析手法は、基本的には線形モデルを用いている。数学的に扱いやすく、図1の様に、線形分離できない事例があったとしても、誤差を最小化すること (残差平方和の最小化) を線形代数の範囲で保証できる。

2.3 線形非分離

しかし、直線による分割 (線形モデル) には、大きな問題点がある。図2のような場合はどうしたらよいのであろうか？この例では、1本の直線で領域を分割することはできない。この場合 (線形非分離) には、図2で示すように、複数の領域に分割することが望ましい。

一方、統計分野では、**「オッカムの剃刀」**と呼ばれる大原則があり、「クラスを推定する知識」は、できるだけシンプルであるほうが、将来の事例に対する推定性能が良い。従って、機械学習の課題は以下に集約される。

【目標】:多次元空間上に分布する事例 (これを**学習事例**と呼ぶ) を、最小の個数の多次元直方体 (これを以下、キューブと呼ぶ) でカバーし、キューブ毎にクラスを1個定める。キューブ内の事例と、キューブのクラスは一致しているものとする。

上記目標に対して、以下のことが分っている[5]。

【効率的計算手法の非存在:】最小個数となるキューブを求める「効率的」方法は存在しない。ここで、効率的とは、事例数を M とするとき、カバーを求める時間が、M の2乗や3乗といった指数関係を遙かに越えることを言う。現実的なサイズの問題では、解を見つけることは不可能である。

即ち、図2のような、最小個数のキューブによるカバーを求めることはできない。であるからこそ、従来の統計解析では、空間分割能力が劣ることを承知の上で、線形モデルに限定し、その代償として、数学的厳密性を得ていたのである。

2.4 決定木とは何か？

以上のような背景から、非線形のモデルを用いた推定手法として、研究されてきたのが、決定木

[3]実際の病名の診断では、属性は多数あるので、多次元空間上の点となるが、ここでは、説明のために、2次元とした。

[4]多次元空間上の超平面

[5]現実には、属性値に誤り等があり、図1のように、他キューブに入ってしまう学習事例 (***ノイズ事例***) が存在することが多い。ノイズ事例も含めて全事例を厳密に分類すると、キューブ数が多くなり、未知事例への推計性能が低下する。ここでは、ノイズ事例はないとしておく。

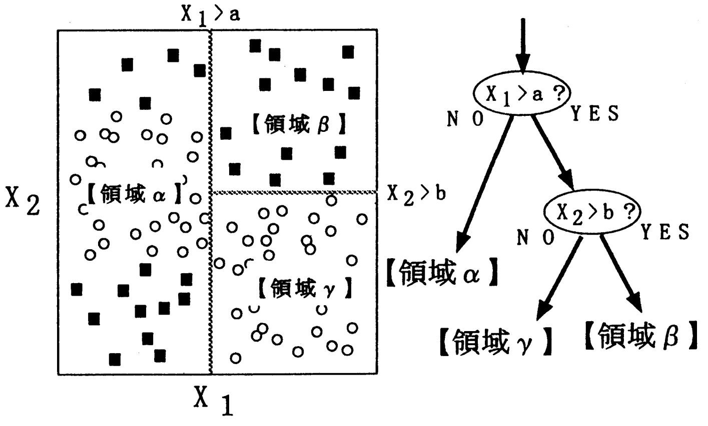


図 3: 決定木の原理

である。図3を用いて、その考え方を示す。

決定木では、「分割統治 (Divide and Conqure)」により、判別知識を生成する。即ち、学習事例が分布する空間に対して、まず、全体を眺めて、ある属性により大きく分ける。図3の例でみると、まず、属性 X_1 に着目して、それが a よりも大きいか小さいかで、全体の領域を左右に2分割する[6]。

次に分割後の右側領域に注目する。上半分と下半分で、クラスが分かれている。そこで、中央付近 ($X_2 = b$) で、2領域に分ける。これにより、少なくとも、右半分の領域では、学習事例のクラスは、きれいにそろう。これらの領域 (キューブ) に未知事例が飛び込んで来た場合には、その未知事例のクラスは、右の上半分は、「■」、下半分は、「○」として良い。図3の左半分については分割を続ける必要がある。

以上をグラフ的に表現したものが、図3右側の決定木である。分割統治で領域を分割してゆき、分割された領域中のクラスが一意になった時点で、その枝については、分割を止める。枝別れの部分を「**ノード**」とよび、クラスが一意となってキューブが完成したものを、「**リーフ**」と呼ぶ。このような構造全体を「**ツリー (木)**」と呼ぶ。

具体的には、領域 β については「■」、領域 γ については「○」とする。左側の枝についても、同様の処理をすることにより、決定木が完成する。キューブが求まれば、学習事例はすてる。図3の右側に示したような決定木のみが保存され、今後生じてくる事例のクラス判定に供される。

新しい (クラスが未知の) 事例が到着した場合には、この決定木の上から、その事例を流して、属性値に対するテストを行う。例えば、図3において、新しいクラス未知の事例 F を考える。この事例 F では、X_1 の値は a より大きな A であり、X_2 の値も b より大きな B であったとする。決定木の最上部から、事例 F を流す。最初に、X_1 について、テストが行われ、その結果 F は、右の枝に流れてゆく。次に、X_2 がテストされ、更に右下の枝へと流れてゆく。このノード (領域 β) では、学習事例から、すでにクラスが「■」であるとされているので、この未知事例 F のクラスは「■」となる。

2.5 決定木学習アルゴリズム

複数の属性と、唯一のクラスを持つ事例を考え、その事例の集合を T とする。T は**学習事例**と呼ばれる。この時、事例の持つクラスは、$\{C_1, C_2, C_3 \cdots C_j \cdots C_n\}$ とする。n は、クラスの個数である。決定木作成のプロセスは以下のようなものである。

*STEP*1 : T がすべて同一のクラス C_j であるならば、T に対する決定木はリーフであり、そのクラスは C_j である。

*STEP*2 : T が全く事例を含まない空である場合には、上記 *STEP*1 と同様にリーフとするが、そのクラスは、事例からは決めれない[7]。

*STEP*3 : T が複数のクラスを有する場合には、ある属性を選択して、その属性が持ちうる属性値により、事例 T を分割する。そして、

[6] どうやって、分割に利用する属性と属性値とを定めるかについては、後述する

[7] この決め方は、ツールに依存する。後述の C4.5 では、ひとつ上流側のノードへと流れて来た事例 T のクラスのなかで、最頻クラス。

(高い、黒色、茶：-)
(高い、黒色、青：-)
(低い、黒色、青：-)
(高い、ブロンド、茶：-)
(低い、ブロンド、茶：-)
(低い、ブロンド、青：+)
(高い、赤色、青：+)
(高い、ブロンド、青：+)

図 4: 学習用データ

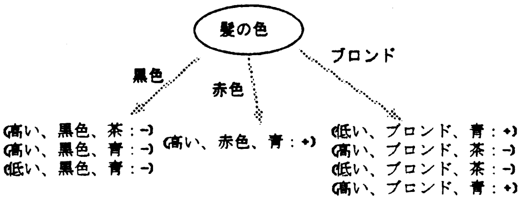


図 5: 髪の毛による分割

分割された事例集合の各々について、上記 $STEP1 \sim STEP3$ を再帰的に実行する。

3 決定木学習アルゴリズム ID3

上記の決定木学習で、分岐させる属性の決定方法が問題となる。生成された決定木の性能に大きな影響を与える。代表的な決定木作成アルゴリズム ID3[3] のエントロピーを利用した属性選択法について紹介する。

まず、学習事例として、図 4 に示す 8 個を考える。各事例は、「背の高さ」「髪の毛の色」「目の色」の 3 個の属性と、2 値のクラスから構成される。属性値は、背の高さが 2 通り、髪の色が 3 通り、目の色が 2 通りなので、考え得る事例の種類は $2 \times 3 \times 2 = 12$ 通りであるので、4 種類の事例は未収集である。未知事例を含めた、全事例のクラスを判定する決定木を生成する。

Quinlan の ID3 では、上記のような分割の良さを評価する手段として、情報エントロピーを用いる。情報エントロピーは、情報の乱雑性を評価するために用いる量であり、クラス数が 2 の場合には、「+」「-」のクラスの発生確率を p_+, p_- とする時、以下の式で与えられる。但し、log は log_2 であり、エントロピーの単位は bit である。

$$E = -p_+ \cdot log(p_+) - p_- \cdot log(p_-) \quad (2)$$

従って、図 4 のエントロピーは

$$E = -(\frac{3}{8}) \cdot log(\frac{3}{8}) - (\frac{5}{8}) \cdot log(\frac{5}{8}) = 0.954bit \quad (3)$$

となる。

次に、髪の色で、この事例を分けてみる。図 5 のように、事例は 3 個の集団に分離する。ここで、髪の色が黒と赤の場合には、クラスがそろっており、前述の $STEP1$ から、これがリーフとなる。一方、髪の色がブロンドの場合には、クラスが分かれており、引き続き、髪の色以外の属性での、分割を繰り返す必要がある。

髪の毛による分割以前と以後で、エントロピーはどう変化するかを考える。式 (2) の定義から、この場合、髪の色が黒、または赤の場合にはエントロピーはゼロである。エントロピーゼロとは、まったく乱雑性がない状態である。ブロンドの場合には、「+」と「-」が半々なので、エントロピーは、以下のようになる。

$$E = -(\frac{2}{4}) \cdot log(\frac{2}{4}) - (\frac{2}{4}) \cdot log(\frac{2}{4}) = 1bit \quad (4)$$

そして、分割後のエントロピーの期待値は、それぞれの枝に流れている事例数を考慮して、以下のように求める。

$$(\frac{3}{8}) \cdot 0 + (\frac{1}{8}) \cdot 0 + (\frac{4}{8}) \cdot 1 = 0.5bit \quad (5)$$

従って、分割前 (式 3) と分割後 (式 5) のエントロピーの差は、

$$0.954 - 0.5 = 0.454 \quad (6)$$

となる。

以上から、「髪の色」を属性として選択した時のエントロピーのゲインは、0.454bit である。同様にして、他の属性 (背の高さ、目の色) のエントロピーゲインは、(1) 背の高さ：0.003 bit、(2) 髪の色:0.454bit、(3) 目の色：0.347bit、である。

髪の色による分割は、エントロピーゲインが一番大きい属性であったことが分かる。エントロピーゲインが大きい事は、事例集合の乱雑性が大きく減少することを意味する。即ち、図 2 の例のように、うまく事例をカバーする可能性が大きいことになる。このようにして、ID3 では、エントロピー削減が一番大きな属性をつぎつぎと再帰的に利用しながら、枝を伸ばして、決定木を構築する。

決定木の目標は、できるだけ少ないノード（属性値の質問）数で、決定木を構成することである。これは、言い換えると、図 2 のように、多次元空間をできるだけ少ないカバーで覆うことを意味している。最小ノード数の決定木を構成する効率的な方法は本質的に存在しない。しかし、情報エントロピーに基づく方法は、簡単な方法ながら、優れた近似解を導くことが知られている。

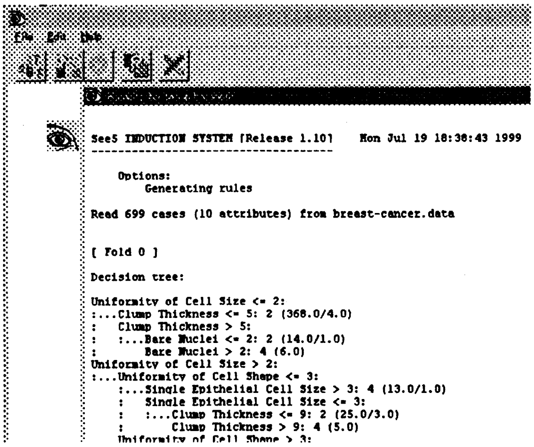


図 6: See5

但し、上記の原理的手法だけでは、現実の問題を扱うことは困難であり、種々の改良が加えられた。問題点と対処策を以下に示す。

過学習の防止：現実の問題では、測定誤差等により、別クラス領域に、事例が入り込むことがある。この場合、事例集合のクラスが一意になるまで枝の展開を繰り返すと、決定木が過度に大きくなり、推計性能が低下する。これを「**過学習**」と呼ぶ。この問題を解決するため、ID3 改良版の C4.5[4] では、決定木の成長を途中で止めたり、あるいは、一度成長した決定木を刈り込む機能を持っている。これを、「**プルーニング**」と呼ぶ。

過度に平坦な木の防止：上述のエントロピーに基づく属性選択は、「平たく横に幅広い木」を作りがちである。この問題を回避するため、C4.5 では、*split info* と呼ばれる評価尺度を利用している。

Missing Value への対処：現実の事例では、ある属性値が何らかの理由で測定されておらず、属性値が不明の場合がある。これを「**Missnig Value**」と呼ぶ。C4.5 では、この種の事例に対しても対処可能としている。

数値属性の許容：以上の説明では、属性値はあくまで飛び飛びのシンボル値であり、その順序には意味がない (名義尺度) とした。しかし、現実の応用では、数値属性がしばしば現れる。この様な属性に対して、C4.5 では、自動的にあるスレショールド（閾値）を求めて、そのスレショールドより大きいか否かを尋ねる質問を生成する。これにより、数値属性と名義尺度属性とが混ざった事例を許容する。

尚、C4.5 では、決定木形式のみではなく、ルール形式の知識も生成するようになっている。ID3 は、人工知能分野では極めて有名な手法であり、これを利用した研究や改良版が多い。その最大の特徴は、属性を原則として名義尺度としていること、そして、枝分かれが 3 以上の複数分岐（マルチスプリット）である点である。後述の CART では、マルチスプリットは実現できず、名義尺度が多い応用は得意とは言い難い。

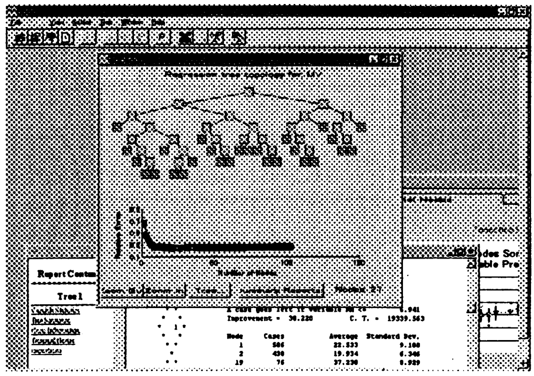


図 7: Salford 社 CART

4　2 進木学習アルゴリズム CART

ID3 とは異なり、統計家に好感を持って迎えられている決定木学習アルゴリズムに CART がある。CART は、属性値 (場合によってはクラスも) が連続値であることを基本とする。分岐は 2 分岐に限定される。このため、「2 進木」の名前がある。CART では、クラスが名義尺度の場合には、「分類 2 進木」とよび、クラスが連続値の場合には、「回帰 2 進木」と呼ぶ。

CART も、分割統治により、領域を分割して、決定木を作成するところでは、ID3 と同一である。しかし、分岐すべき属性（分岐変数）の選択には、エントロピーではなく、GINI インデックスを利用している。例えば、あるノード t において、事例集合は、C 個のクラスに分散しているとする。この場合、各クラス毎の事例が、事例全体の中に生起する確率を $p(j|t)$ で表すとき、GINI インデックスは以下の式で表される[8]。

$$GINI = 1 - \sum_{j}^{C} p(j|t)^2 \qquad (7)$$

[8] 例えば、クラス 2 の場合について考えてみると、クラスが一方に偏っていれば、この GINI 分散指標も 0 となり、逆に等確率にバランスしていると最初値になる。このことから、この GINI インデックスとエントロピーの両者間には、大差はない。

ID3との最大の違いは、名義尺度の属性に対する振る舞いである。ID3では、属性値が3種類以上存在しても、決定木の展開に支障はない。しかし、CARTでは、もともと2分岐しか考えていない。このため、CARTでは、名義尺度の変数であることを陽に宣言すると、(属性値がn種類あるとして)n個を2グループに分割するすべての組み合わせを調べる。nが大きくなると、組み合わせは膨大となる。このことからも、CARTでは、従来の統計解析と同様に、属性としては、連続値（数値）を強く意識している。尚、CARTでもMissing Value、過学習に対する対策はID3と同様に盛り込まれている。

CARTのID3と比較した利点の一つは、クラスが連続値でもよいことである。これを、回帰2進木と呼ぶ。回帰2進木では、群間平方和を最大化(群内平方和を最小）することを評価基準として、木を展開するが詳細は文献[4]に譲る。但し、頑健性（ロバストネス）のある評価値として、2乗和ではなく、偏差(誤差の絶対値)を用いることも可能としている。また、CARTが用いているGINI分散指標では、2分岐した一方に流れる事例が極端に少ない選択が有利となる欠点が知られており、この問題に対処するため、Twoing Ruleと呼ばれる、改良された指標も準備されている。

5 決定木の適用分野

5.1 商用パッケージ

ID3及びCARTは、ソフトウエアパッケージとして市販されている。図6は、ID3の商用版であるRulequest社[6]のSee5の動作画面である。但し、C4.5は、Quinlanの著書[3]に付属しており、研究目的であれば無料で利用できる。C5.0及びSee5は、C4.5の機能強化版である。

一方、CARTは、Salford社[5]より市販されている。動作例を図7に示す。CARTでは、種々のデータにより決定木を分析できる。ソフトウエア製品としては、前述のSee5より完成度が高い[9]。

5.2 決定木の適用分野

決定木の適用範囲については、研究を待たなければならない点が多い。しかし、一般的には、以下が言われている。

可読性の高さ： 名義尺度を利用でき、しかも、得られた知識が人間に可読である。従来の重回帰分析では、ダミー変数とよばれる「0」「1」の2値を取る属性を、各シンボル値に対して割り付けていたので、直感的理解には向かない。

部分構造抽出：属性Aがある値では属性Bが影響し、属性Aが別の値となるとBは無関係でCが影響する場合でも、決定木は知識を表現できる。これに対して、重回帰分析では、属性の部分集合を取り出すので、細かい部分構造を解析できない。

大きな決定木の出現：多数の属性の値を決めない限りクラスが決まらない場合については、注意が必要である。この種の問題を決定木記述すると、決定木のどの分岐も、すべての属性を通るまで枝を伸ばすので、決定木が巨大となる。

学習時間の短さ：決定木学習手法は、ニューラルネットのような他の学習手法に比べて、処理が高速である。このため、巨大データ解析や、ブートストラップ手法[7]に適している。

6 おわりに

決定木学習手法の研究は、手法自体の研究よりも、実データにより評価し、何が本当に求められているかを検証してみるべき時期にあるように著者は感じている。何より、名義尺度が使えること、得られた知識が人間に可読であること、そして、非線形モデルであることは、従来の重回帰分析には見られない大きな特長である。これら特長を生かして、決定木の適用範囲が、化学・社会学等、多岐に広がってゆくことを念ずるものである。

参考文献

[1] 秋葉他,「例からの学習技術の応用に向けて（1及び2）情報処理,Vol.39,No.2-3.

[2] L.Breiman, et. al,"Classificaion and Regression Trees", Wadsworth,Inc.,1984.

[3] Quinlan 著、古川訳,「AIによるデータ解析」トッパン、1995.

[4] 大滝他,「応用2進木解析法」日科技連、1998.

[5] http://www.salford-systems.com/

[6] http://www.rulequest.com/

[7] B.Efron et. al,"A Introduction to the Bootstrap",Chapmann & Hall,1993.

[8] 金田重郎 http://www1.doshisha.ac.jp/~skaneda

[9]尚、S-PLUSだけではなく、主要な統計解析ツールには、近年、決定木ツールが追加されている。これらの詳細は販売元に問い合わせされたい。

統計的データマイニング法の適用とその解釈

塩野義製薬(株)解析センター　渡辺秀章・田崎武信
hideaki.watanabe@shionogi.co.jp・takenobu.tasaki@shionogi.co.jp

統計学の分野で提案された方法でデータマイニングにも有用であると考えられる３種類の方法を紹介する．一つの事例の解析を通してそれらの方法の特徴を議論する．

１．はじめに

コンピュータ時代の黎明期から10数年前まで，多変量データ解析の花形は回帰分析，判別分析，因子分析，クラスター分析などで代表される多変量解析法であった．いまになって「古典的」と形容されることもあるそれらの多変量解析法には，変数の線形結合を基礎にする，全所的な基準を最適化する，正規性を仮定する，階層的なアルゴリズムを基礎にする，といった性格がある．しかしながら，このような性格は現代社会で得られる膨大で，かつ複雑なデータの解析には馴染まないと見なされつつある．とくに，それらの方法を適用して得られる結果の解釈は方法の中身をある程度知らなければ難しいこと(性格や限界を把握していなければ正しい結論は得られないと繰り返し警告されていること)から，古典的な多変量解析法に対する期待は薄れつつある．このような古典的な多変量解析法に代わり，コンピュータの飛躍的な発展を背景に，機械学習やデータベースの分野から，仮定や前提条件を必要とせず，しかも非線形な構造の検出を狙いにした方法が提案され，データマイニングの方法として脚光を浴びている．それらの方法の代表格がニューラルネットワーク法，知識発見法，連関ルールであり，データを出発点，発見的な考えを基礎にし，方法の中身を知らなくても適用結果を直観的に解釈できることから，種々の分野で導入が試みられている．しかしながら，統計学の分野においても，古典的な多変量解析法からの脱却を意図した研究が行われ，とくに非線形性の問題に挑戦を向けた方法がいくつか提案されてきた．これらの方法はデータマイニングの方法として有望であると考えられる．本小文では，統計学の分野で発展した方法を統計的データマイニング法と呼び，それらのうちから３種類の方法をとりあげて紹介し，Brownの前立腺癌データ[1)]への適用結果の解釈の仕方を示して３方法間の対比を行う．

２．各方法の概略

ここでは以下の３種類の統計的データマイニング法の概略を紹介する：CHAID(CHi- square Automatic Interaction Detector)法[2)]，SPAN(Search Partition ANalysis)法[3), 4)]，MARS (Multivariate Adaptive Regression Spline)法[5)]．これらの方法はいずれも，応答と説明変数がペアになった，いわゆる回帰型のデータに適用される．なお，これらの方法の比較対照として古典的な多変量解析法に位置づけられるロジスティック回帰分析法もとりあげる．最初に３種類の統計的データマイニング法の概略を紹介する．

CHAID法：名前が示すように，説明変数間の交互作用を検出するための方法論として開発された．その手順は説明変数のとり得る値に基づいてデータを階層的に，互いに排反になるようにその部分データに分割することからなる．この方法は他のAIDと名前の付される方法やCART法[6)]とともに階層的反復分割法として知られている．階層的反復分割法の諸法は大雑把にデータを分割する際の基準あるいは分割の停止基準で区別することができる．CHAID法は応答がカテゴリー変数で与えられる場合を主に対象としており，分割基準はカイ２乗検定のp値であり，それが所与の有意水準よりも小さければ分割を継続し，大きければ分割を停止する．

SPAN法：データを２個の部分データに分割する方法である．その際にある分割基準を参照して応答ができるだけ均質になるように，すなわち一方では一様に小さく，もう一方では一様に大きくなるように分割ルールを構成し，最適な分割ルールを探索することである．得られた部分データはいくつかの部分グループで特徴付けられる．ここに，部分グループは説明変数に関する積集合の形式で表現さ

れるが，部分グループは互いに素でなくてもよい，つまり部分グループは互いに重複しあってもよい．いいかえると，SPAN 法は部分データ集合を部分グループの和集合の形式で表現する．その意味でSPAN 法は非階層的な分割方法であるといえる[7]．

MARS 法：上記の AID 法と SPAN 法に共通の難点は座標軸に平行な直線による分割を基礎にすることである．このような分割には解釈を平易にするといった長所があるが，分割する直線について応答が不連続になってしまい，分割点の近傍で適切に解釈することができないといった短所もある．MARS 法は反復分割法のアルゴリズムを基礎にし，スプライン関数を利用することで分割点での不連続性を回避した方法である．データはスプライン関数の積として得られる基底関数の総和によって説明される．MARS 法は階層性を捨てることで説明変数の主効果にも注目している．階層的反復分割法が性格的に交互作用にしか注目していないことと対照的である[8]．

3．各方法の適用結果の比較

前節の統計的データマイニング法と古典的なロジスティック回帰分析法を前立腺癌患者におけるリンパ節転移のデータ[1]に適用し，各方法でデータの特徴がどのように把握されるかを検討する．そのリンパ節転移データでは，前立腺癌患者のリンパ節転移に影響を及ぼすリスク因子を探索するために，前立腺癌患者 53 例を対象に，X線(+，−)，ステイジ(+，−)，グレイド(+，−)，酸性フォスファターゼ(ACP)，年齢，リンパ節転移の有無が観測されていた．ここに，「+」は「−」に比して好ましくない徴候を示す．ACP と年齢は連続値で観測されていたが，ここでは CHAID 法を適用するため，そして他のリスク因子が 2 値で観測されていることから，田崎・後藤[9]に倣い，ともに 60 未満(−)と 60 以上(+)の 2 値データとしてとり扱った．これらの 60 という切断点はこのデータに対する Brown[1]の解析過程で示唆されたものである．なお，リンパ節転移なしとありの例数はそれぞれ 33 例(62.3%)と 23 例(27.3%)であった．

3.1 ロジスティック回帰分析

ロジスティック回帰分析法を適用するにあたり，リンパ節転移の有無を応答，5 個の候補リスク因子を説明変数とした．このとき，リンパ節転移のなしを 0，ありを 1，リスク因子の−を 0，+を 1 とした．ロジスティック回帰分析法の適用結果を表 1 に要約する．

表 1．ロジスティック回帰分析法の適用結果

リスク因子	回帰係数	標準誤差	カイ 2 乗統計量	p 値	オッズ比
切片	-2.8086	−	−	−	−
X線	1.9991	0.8234	5.8953	0.0152	7.3830
ステイジ	1.3867	0.7762	3.1921	0.0740	4.0020
グレイド	0.7083	0.8167	0.7521	0.3858	2.0300
ACP	1.6285	0.7965	4.1809	0.0409	5.0960
年齢	-0.6505	0.7468	0.7587	0.3837	0.5220

表 1 の p 値は回帰係数がゼロに等しいとする帰無仮説を検定したときの結果であり，p 値が小さいリスク因子はリンパ節転移の有無に影響を及ぼしていると解釈できる．表 1 から，X線，ACP，ステイジに対する p 値がいずれも慣習的な有意水準 0.05 に近いこと，そして p 値の最も小さい X線がリンパ節転移の有無に最も強い影響を及ぼし，次いで ACP，ステイジの順に影響を及ぼしていることが示唆される．オッズ比からは，例えば X線において，X線が+である患者のリンパ節転移のリスクは−である患者に比べて約 7.4 倍高いと解釈できる．

このようなロジスティック回帰分析法を適用するにあたっては，モデルの適合度検定を行って回帰モデルの妥当性を確認しておくことが重要である．適合度検定の p 値は 0.4365 であり，モデルは適合していると考えても差し支えない．しかしながら，上記の解析は各リスク因子の主効果のみを評価したものであり，各リスク因子間の交互作用は評価されていない．適合度検定から主効果のみの回帰モデルであっても妥当である結果が得られているが，これが最良の回帰モデルであるとは限らないこと

に注意したい．候補リスク因子間の交互作用を考慮したモデル，例えばX線×ステイジを説明変数に加えた回帰モデルがより望ましいかもしれない．交互作用も含めて最良の回帰モデルを導出することは容易ではなく，試行錯誤を必要とする．ロジスティック回帰分析法は現在でもとくに医学データの解析でよく用いられるが，データの特徴を把握するには馴染みにくいと考えられる．

3.2 CHAID法

CHAID法を適用するにあたり，分割の停止基準である有意水準を0.05とした．この場合にはリンパ節転移と有無とリスク因子の2値カテゴリーの組み合わせに基づいて(部分)データを分類することで2×2分割表を作成できる．この2×2分割表に対するカイ2乗検定を行ったときに最も小さいp値を与えたリスク因子で(部分)データを分割する．ただし，すべてのリスク因子でp値が所与の有意水準より大きければその部分データの分割を停止する．CHAID法の適用結果を図1に示す．

図1の網かけを施した部分は分割の終了した部分データを示している．図1において，最初にデータはX線で2個の部分データに分割され，X線＝＋の部分データはさらにステイジで分割されている．これより，{X線＝＋かつステイジ＝＋}の部分データでのリンパ節転移率は90.0%(9/10)であり，X線とステイジがいずれも＋であれば，リンパ節転移のリスクの高くなることが示唆される．他方，X線＝－の部分データはACP，グレイド，ステイジで順次に分割されており，{X線＝－かつACP＝－}の部分データと{X線＝－かつACP＝＋かつグレイド＝－かつステイジ＝－}の部分データでのリンパ節転移率はそれぞれ10.5%(2/19)と11.1%(1/9)であった．これより，X線とACPがいずれも－であるか，ACPが＋であっても，X線，ステイジ，グレイドがすべて－であればリンパ節転移のリスクが低いことが示唆される．ACPとグレイドはX線が＋である部分データの分割で登場していないのに対し，X線が－である部分データの分割には登場していることに注意したい．これはACPとグレイドの影響の仕方がX線の＋と－とで異なることを意味している．これより，リンパ節転移の有無に関してリスク因子間の交互作用の存在がうかがわれる．

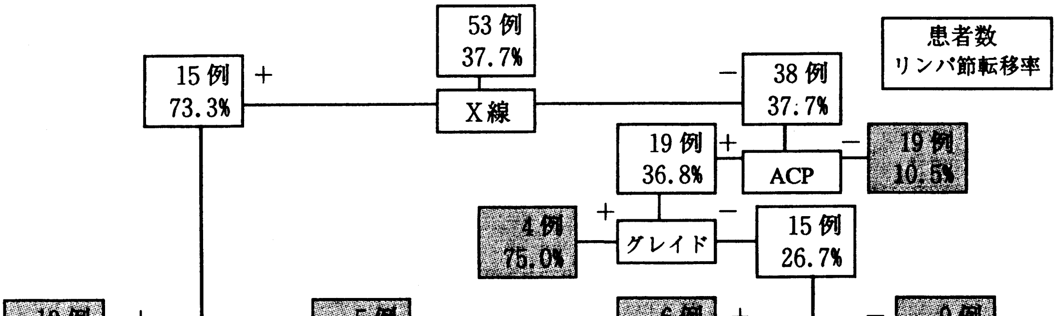


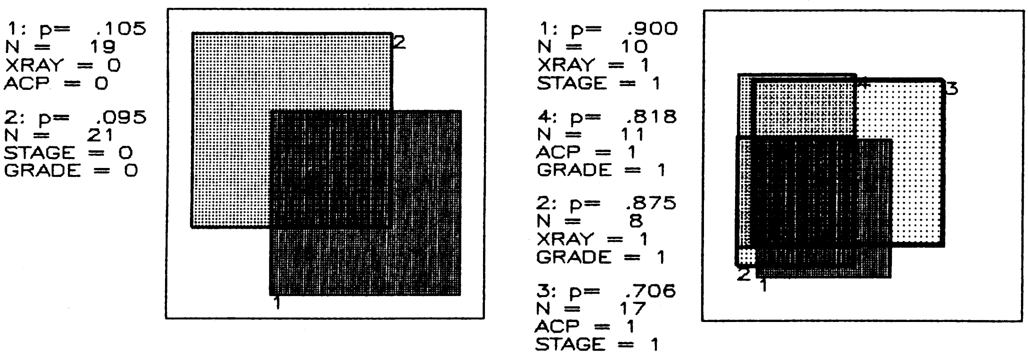
図1．CHAID法の適用結果

3.3 SPAN法

SPAN法を適用するにあたり，データを2分割するための分割基準として，応答が2値であることからカイ2乗統計量を用いた．SPANの適用結果を図2に示す．実際には，SPANの提案者であるMarshallが開発したソフトウェアを利用しており，図2はその出力の一つである．図2の左側はリンパ節転移率の低かった部分データを示しており，ここではそれを*A*で参照することにする．部分データ*A*は2個の矩形で表現されていることがわかる．番号1の付された矩形の説明は*A*の左側部分に与えられており，番号1の矩形は「XRAY=0」，すなわちX線＝－と「ACP=0」，すなわちACP＝－を同時に満たす患者の集合を示している．「N=19」はその集合に属する患者数が19であること，p=0.105はリンパ節転移率が10.5%であることをそれぞれ示している．番号2の矩形も同様である．これらの矩形はSPAN法では「部分グループ」と参照されている．*A*は{X線＝－かつACP＝－}と{ステイジ＝－

(STAGE=0)かつグレイド＝－(GRADE=0)}の2個の部分グループの和集合として表現され，Aに属する患者数は32例であり，リンパ節転移率は12.5%であった．X線＝－かつACP＝－，またはステイジ＝－かつグレイド＝－のいずれかを満たす患者のリンパ節転移率は低いことが示唆される．

図2の右側はリンパ節転移率が高かった部分データを示している．この部分データはAの補集合になることから，ここではそれをA^cで参照する．A^cは4個の部分グループ(矩形)で構成されており，各部分グループの解釈の仕方はAの場合と同様である．A^cに属する患者数は21であり，リンパ節転移率は76.2%であった．A^cの1の付された矩形，すなわち{X線＝＋かつステイジ＝＋}の部分グループに属する患者は，この矩形と他の3個の矩形との重複の程度から，グレイド＝＋およびACP＝＋であることがわかる．CHAID法でも{X線＝＋かつステイジ＝＋}がリンパ節転移率の高かった部分データとして抽出されているが，その他のリスク因子に関する特徴まで把握されていないことに注意したい．



リンパ節転移率が低かった部分データ(A)　　リンパ節転移率が高かった部分データ(A^c)

図2．SPAN法の適用結果

3.4 MARS法

MARS法を実行するためのソフトウェアは提案者のFriedmanの所属するスタンフォード大学のサイトからFORTRANプログラムとしてダウンロードできる．MARS法の適用にあたり，ロジスティック回帰モデルの場合と同様に，応答をリンパ節転移の有無とし，リンパ節転移のなしに0，ありに1を与え，説明変数を5個の候補リスク因子として，－に0，＋に1を与えた．MARS法の適用結果は1個の基底関数で以下のように表現される．

$$y = 0.256 + 0.644\ \text{X線}\cdot\text{ステイジ}$$

ここに，X線とステイジには0か1の値が入力されることに注意したい．このモデル式からX線とステイジがリンパ節転移に影響を及ぼしており，この2個のリスク因子の間に交互作用の存在することが示唆される．MARS法のプログラムを実行すると，説明変数の相対的な重要度が併せて出力される．重要度は最も重要な説明変数を100として出力されるが，ここでの結果はX線とステイジがともに100であり，その他の3リスク因子の重要度はゼロであった．他の方法で示唆されたグレイドやACPがリンパ節転移に及ぼす影響はMARS法から示唆されなかった．

MARS法は主効果と交互作用効果を同時に検出できると期待されることから，私たちは，いまこの方法に最も熱い期待を寄せている．しかしながら，このリンパ節転移データについては，他の方法に比して示唆的な情報が得られなかった．この原因として，MARS法は最小2乗法を基礎にしているが，リンパ節転移の有無といった2値データに最小2乗法は適していないことが考えられた．最小2乗法の代わりに最大尤度を基礎にしたアルゴリズムを開発すればデータの特徴をより適切に把握できるようになるであろうと考えられる．

4．考察

ここでは統計学の分野から提案されたCHAID法，SPAN法，MARS法を統計的データマイニング法と呼び，これらの方法をリンパ節転移データに適用した結果を相互に対比した．階層的反復分割法のCHAID法はデータを階層的に互いに素な部分データに分割することから，説明変数間の交互作用，すなわちデータの局所的な特徴を検出するのに有効であると考えられる．実際に，リンパ節転移データに対する適用では，リンパ節転移率が高かった，あるいは低かった部分データの特徴だけでなく，リンパ節転移率が50%あるいは40%といった，いわゆる中程度であった部分データの特徴も描出していた．例えば，X線が＋であっても，ステイジが－であれば，X線とステイジがいずれも＋の場合に比べてどの程度リンパ節転移率が小さくなるかを把握できる．そのような意味で，CHAID法はデータの特徴を細部にわたって描出する，いいかえるとデータの「発掘」に適した方法であるといえる．しかしながら，データを階層的に互いに素な部分データに分割した結果に基づいてデータの全体像を眺めると解釈しにくい場合が起こる．例えば，リンパ節転移データにおいても，X線＝－を含む部分データにはリンパ節転移率が高かったものもあれば低かったものもあったため，X線のリンパ節転移に対する主効果を直接的に解釈できない．

そのような階層的反復分割法の解釈上の難点を動機付けとしてSPAN法が提案された．リンパ節転移データへの適用結果では，リンパ節転移率の高かった部分データが4個の部分グループで表現されており，この部分データの特徴が多面的に描出されている．また，4個の部分グループはすべて＋のリスク因子で表現され，リンパ節転移率が低かった部分データの部分グループはすべて－のリスク因子で表現されていたことに注意したい．このことは，リンパ節転移のリスクが各リスク因子の＋で高く，－で低くなるといった解釈を与える．SPAN法は，CHAID法の適用結果のように各リスク因子の＋でもリンパ節転移のリスクが高くなる場合があるという解釈を導かない．この意味でSPAN法はCHAID法と異なり，部分グループに登場した各リスク因子の影響を主効果として描出することができる．このことより，SPAN法はデータの特徴を全体的に把握する，いいかえるとデータの「帰納」を重視した方法であるといえる[10]．

MARS法はリンパ節転移データへの適用では満足のいく結果を与えなかったが，主効果と交互作用を同時に検出することができる，いいかえると帰納(全体像)と発掘(局所像)を統計的に調整できる方法論として私達は依然として注目している．現段階ではMARS法を実行するソフトウェアが整備されていず，その特徴を十分に評価できていないが，1999年8月にSalford Systems社から市販される予定のソフトウェアを用いて適用経験を積み重ねたいと考えている．

CHAID法とMARS法，とくにCHAID法には変数選択が関与している．変数選択は選ばれたものにのみ光をあてる．しかし，選ばれなかったものに注目することで別のストーリを構築できるかもしれない．本小文でとりあげたデータはいくつかのストーリで解釈できそうである．CHAID法に潜む変数選択の問題を含む他の問題はDoyle and Fenwick[11]やWatababe and Tasaki[12]などで論じられている．

上記の統計的データマイニング法と対比的に，古典的な多変量解析法として，統計学の分野で周知のロジスティック回帰分析法をとりあげた．モデルに基づく解析はモデルがデータの特徴を十分に反映するときにその有用性を発揮すると考えられるが，いわゆる最良モデルに到達するまでに試行錯誤を要することになる．最良モデルの探索には背景となる情報や経験などが総動員されることになるであろうが，そのような情報や経験が十分に整備されていなければ，統計的データマイニング法で示唆されたデータの特徴を最良モデルの探索に活かすことも考えられる．

5．結語

ここでとりあげた統計的データマイニング法を実行するにはソフトウェアの整備が不可欠であるが，このようなソフトウェアは誰しも整備できる状態にあるとは限らない．しかしながら，データの特徴はデータの単純なハンドリングなどからも明らかになる場合がある．田崎・松村[13]は多変量2値データをソーティングすることにより，データの特徴をどの程度まで把握できるかを検討している．そし

て，なんらかの方法による解析結果は行と列の並べ替えで表現できると仮定して，その効果を測定する方法を提案している．卑近なツールを用いてデータそのものを別な観点から眺めてみることもデータマイニングの出発点になるはずである．

参考文献

1) Brown,B.Wm.,Jr. Prediction analysis for binary data. In Biostatistics Casebook, ed.by Miller,R.G.,Jr.et al.,3-18, John Wiley & Sons,1980[医学統計研究会訳．２値データの予測解析．医学統計実践事例集．MPC, 1982].

2) Kass,G.V. An exploratory technique for investigating large quantities of categorical data. Appl. Statisit., 29,119-127, 980

3) Marshall,R.J. Partitioning methods for classification and decision making in medicine. Statistics in Medicine,5,517-526,1986

4) Marshall,R.J. A program to implement a search method for identification of clinical subgroups. Statistics in Medicine,14,2645-2659,1995

5) Friedman,J.H. Multivariate adaptive regression splines(with discussion). Ann. Statist.,19, 1-141,1991

6) Breiman,L.; Friedman,J.H.; Olshen,R.A.; Stone,C.J. Classification and Regression Trees. Wadsworth, 1984

7) 渡辺秀章；田崎武信．統計的データマイニング法 SPAN の評価．第 10 回日本計算機統計学会シンポジウム論文集，8-11,1996

8) Tasaki,T.; Goto,M. Statistical alternatives to neural networks. Kitagawa Symposium on Informatics and Statistics,1994.5.26-27,Kyushu University,1994

9) 田崎武信；後藤昌司．クラスター解析の個人学習，日本分類学会第 12 回シンポジウム予稿集，1-10,1994

10)渡辺秀章；田崎武信．データ科学における帰納と発掘．日本分類学会第 13 回研究報告会 研究報告予稿集，8-15, 1997

11)Doyle,P.; Fenwick,I. The pitfall of AID analysis. J. Marketing Res.,12,79-85,1975

12)Watanabe,H.; Tasaki,T. When recursive partitioning does not work well in classification/regression analysis. Proceeding of the ISM Symposium : Data Mining and Knowledge Discovery in Data Science,73-77,1999

13)田崎武信；松村智恵子．並べ替え効果の測定．日本分類学会第 15 回研究報告会 研究報告予稿集，29-34,1999

わたなべ ひであき WATANABE, Hideaki

たさき たけのぶ TASAKI, Takenobu（解析センター長）

連絡先 〒553-0002 大阪市福島区鷺洲 5-12-4 塩野義製薬（株）中央研究所

TEL：06(6455)2047，FAX：06(6458)9215

化学構造を対象としたデータマイニング

関西学院大学情報メディア教育センター　岡田孝
okada@kwansei.ac.jp

システムが直接構造式や３次元構造を取り扱う場合に絞って、データマイニングとの関係で研究動向を報告する。マイニング技法としては、機械学習とデータベースに関連した技法から、CASE システム、帰納論理プログラミング、相関ルール解析を取り上げる。また大容量データとの関係で、これからの方向性を論ずる。

1　はじめに

データマイニングあるいは知識発見の研究が、ここ数年非常に注目を集めている。そこで利用されている技法のもとになる分野としては、多変量解析、ニューラルネットワーク、パターン認識、人工知能、データベースなどがあげられるが、技法自体の発展とともに、マイニングの対象となる大容量データベースが利用可能になってきたことが注目されている主な理由であろう。化学の領域においても時を同じくして、コンビナトリアルケミストリとハイスループット・スクリーニング（HTS）が実用化され、検索に限らず解析においても、大規模データを取り扱う必要が生じているのは興味深い。

本特集の他の記事で紹介される多くの技法は、そのほとんどが属性：属性値の対で表現される表形式のデータ、あるいはマーケットバスケットデータのようにトランザクション形式のデータである。一群の化学構造を対象として解析を行う場合、表の各セル内の表現が数値または文字列であれば、技法の適用において特に化学固有の条件はない。適用対象や目的によって、適当な技法を選択すればよい。この場合、適切な記述子の獲得が決定的に重要になるのは当然であり、まさに化学者の知恵が発揮される所である。すでに決定木や rough set 理論を用いた研究が多数行われていることを指摘するに留める。

本稿では、構造式あるいは３次元構造をシステムが直接取り扱う研究、技法としても人工知能やデータベースの領域で発展してきたものに限って、その成果を紹介する。また、構造を取り扱う場合でもエキスパートシステムのように、専門家の知識として構造を与える場合は除外し、いわゆる帰納学習に限った。結果として、紹介する研究のほとんどが生理活性物質の構造活性相関に関するものとなっている。なお、遺伝子データバンクからのデータマイニング研究も現在非常に盛んに行われている。技法としては共通するものも多く、読者の興味も高いと予想されるが、筆者の能力の及ぶところではないため、一切触れていないことをご了承いただきたい。

2　　CASE, MultiCASE[1),2), 3)]

現在のようにデータマイニングが注目を集める以前から、グラフ表現された一群の化学構造から帰納的に知識を得ようとする研究がいくつか行われてきた。筆者も CLUSMOL, DNET 等を開発してきたが、簡潔な機構を用い現在も広く利用され続けている点で、Klopman による CASE 関係の研究が重要であろう。CASE は構造式から非水素原子数が 3-12 の部分構造をすべて切り出し、各部分構造毎にもとの化合物群で活性の有無が偏っているか否かを調べる。ただし、取り出す部分構造は枝分かれのない原子列とそれに付着する置換基に限られている。たとえば多環芳香族炭化水素の発ガン性に対しては、図１左の２種の構造（実線部）を活性有りに偏った構造として（Bay reigion に対応）、右端の構造を発ガン性を抑制する構造として出力する。

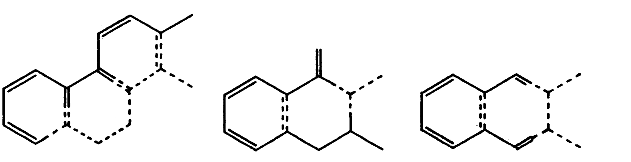


図１．発ガン性に関連する部分構造

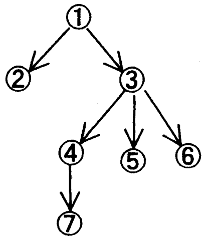
このソフトウェアは非常に単純な考え方ながらも、実用的に優れているためであろうか、現在は商用のソフトウェアとして米国の FDA などでも使用されているようであ

る。問題があるとすれば、部分構造として取り得る範囲が限られており、重要な構造を見逃すおそれがある点である。しかし、理論的には不完全でも、実質的には連結された重要な部分構造の大部分を記述していると考えられる。なお、CASE の出力は化学者の解釈を残した状態で出力されるため、これを整理し完全な知識の形で提供しようとする試みもある。[4)]

3　帰納論理プログラミング（ILP: Inductive Logic Programming）の方法

3.1　帰納論理プログラミングとは[5)]

Prologのような演繹型の論理プログラミングを帰納推論に拡張しようとする研究領域である。例えば、図２の左の家系図を考える（人名の代わりに数字を用いている）。この家系図での親子関係全６例を右側の箱内の parent 述語により前提知識として記述し、さらに grandparent 述語（述語の意味は一切与えない）の正例４例を入力として与えれば、結果の項に示したようにgrandparent述語の意味を学習させることができる。この例では、「XがZのparentであり、かつZがYのparentであるならば、XはYのgrandparentである。」という論理式を導き出している。結果の式自体が実行可能な論理プログラムであるから、この方法論はまさに例からの自動プログラム合成である。



前提知識
```
    parent(1, 2).
    parent(1, 3).
```
正例
```
    grandparent(1, 4).
    grandparent(1, 5).
```
学習結果
```
    grandparent(X, Y) :-
      parent(X, Z),parent(Z, Y).
```

図２．帰納論理プログラミング

属性：属性値の形式を前提とする命題論理では、たとえば分子 a が C=O 基を持つことを表すために、carbonyl(a) のように人間がcarbonylという述語（属性）を与える必要がある。他方、一階述語論理を基盤とする帰納論理プログラムでは、図３の２種の述語を定義し、これらを使用して前提知識を記述すればよい。例えば、前提知識の最初の述語は、分子 a 内の１番目の原子 a_1 は元素 C であると記述している。同様に３番目の述語は、分子 a 内の1番目と２番目の原子間に結合 double が存在することを表している。このようにして、対象とするすべての分子内の全原子および結合を記述し前提知識として与える。また、化合物分子 a がC=O基に特有のスペクトルを持つことを述語yellowで表し、化合物群を正例と負例に分けて与える。その結果、学習されたyellow定義の右辺には、carbonylという言葉はなくとも、C=O基構造の表現が自動的に出現することが期待できる。

このように ILP は、オブジェクト間の関係性を表現できること、また前提知識としては、個別の事実ばかりでなく、一般的な法則性を示すプログラムでも等しく取り込めることから判るように、非常に柔軟な表現力を持っている。

Srinivasan, Muggleton らは、上記のような入力から直接学習結果を出力する ILP システムを、精力的に開発している。特に Progol は逆伴意法により仮説空間の探索範囲を限定することに成功したため、あ

述語
```
  atom(分子, 原子序数, 元素).
  bond(分子, 原子序数 1, 原子序数 2, 結合型).
```
前提知識
```
  atom(a, a_1, C).
  atom(a, a_2, O).
  bond(a, a_1, a_2, double).
```
正例
```
  yellow(a).
```
学習結果
```
  yellow(X):-
        atom(a, a_1, C), atom(a, a_2, O),
        bond(a, a_1, a_2, double).
```

図３．構造式の記述と学習

る程度の規模の問題を解くことができる。[6,7] なお、帰納論理プログラミングといっても、推論結果を魔法のように出してくるわけではなく、実際には化学構造の重ね合わせによる共通部分構造の抽出や、提示された仮説がどれだけの事例をカバーするか、仮説の記述長は事実の列挙よりも短いか、等の多くの計算を行いながら、膨大な空間の探索を行っている。これらの処理内容は同じ ILP といっても、当然システムにより異なっている。ただし、化学構造に密着したレベルでアルゴリズムを記述しているのではなく、帰納推論問題を解くための一般的なシステムとして開発されている。従って、化学構造の問題に適用した場合、問題の与え方によりどのようにシステムの振る舞いが変わるか予測しにくい感は否めない。

3.2　変異原性の予測

このProgol システムによる化学構造取り扱い例を見てみよう。[8] 対象は Debnath らが SAR 研究の対象とした、芳香族ニトロ化合物 230 種に対する Ames test による変異原性データである。[9] なお、構造式以外に Quanta で計算した部分電荷も原子述語に取り入れている。解析は Debnath らに従い、188 種の回帰に適した化合物群と 42 種の多様な母核を持つ化合物群に分けて行われた。得られた知識は図4に示す活性部分構造である。図中の上方に示した5種の構造は、188 種の群から得られた結果である。これらを用いて未知化合物の活性を推定した結果は、記述子を用いて回帰やニューラルネット、決定木で解析した場合と同程度の精度であった。しかし、多様な構造を持つ 42 種からは、これまでに知られていない最下段の活性構造が得られ、その予測精度は他のいずれの方法をも上回っていた。

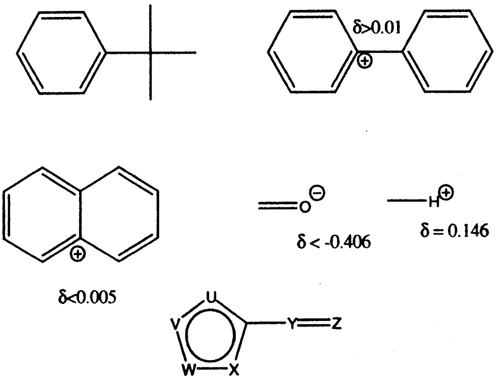


図4．変異原性に特徴的な部分構造

3.3　Pharmacophore の推定

3次元構造を対象とした帰納推論の例として、ACE（Angiotensin Converting Enzyme）活性阻害剤の pharmacophore 推定に Progol を適用した例を紹介する。[10] Mayer らは 28 化合物に対して、1種類の pharmacophore を提案している。[11] 前提知識としての3次元構造は Mayer らの調べた化合物群に対して、各 10 種の低エネルギーコンフォーメーションを計算して与えている。

構造表現は図5上端の述語 atom で記される。これは分子 m1 の原子 a1 が酸素であり、sp2 混成であること、またその 10 番目のコンフォーメーション c10 内での a1 の x, y, z 座標を最後の3項で表現している。さらに、hydrogen-donor, hydrogen-acceptor, zinc-site の定義、原子間距離の計算法、pharmacophore には3原子以上の関与が必要、等の記述を前提知識に書き込んで、学習を行うことになる。Progol により2種の pharmacophore が出力された。これらの内の1種は Mayer らにより得られているものと同一であった。図5の下部に出力結果の1例を示す。なお、計算に当たっては、距離の許容誤差を 1.0 Å に固定している。

```
atom(m1, c10, a1, O, 2, 3,43, -3.12, 0.05)
bond(m1, a2, a3, double)

Molecule A is an ACE-inhibitor if for some conformation B:
    A contains a hydrogen acceptor C,
    A contains a hydrogen acceptor D,
    the distance between C and D
        within conformation B is 3.236 +/- 1.0 Angstroms,
    A contains a hydrogen acceptor E,
    the distance between C and E
        within conformation B is 3.970 +/- 1.0 Angstroms,
    the distance between D and E
        within conformation B is 2.227 +/- 1.0 Angstroms,
    A can bind to zinc at a site F,
    the distance between C and F
        within conformation B is 5.472 +/- 1.0 Angstroms
    the distance between D and F
        within conformation B is 7.144 +/- 1.0 Angstroms,
    the distance between E and F
        within conformation B is 8.501 +/- 1.0 Angstroms.
```

図5．ACE pharmacophore の推定結果

3.4　他の応用例

Progol システムは、これまでに述べたように非常に強力であるが、化学の問題への適用は筆者の知る限りまだ少なく、^{13}C nmr スペクトルからの構造推定問題、[12] 後述する NTP の問題、[13] エナンチオマー分離問題 [14] 等に適用されている程度と思われる。

他方、方法論の発展は急であり、(1) ILP で得られた構造記述子を回帰変数に取り込む、[15] (2) 回帰法自体を前提知識として与え、回帰式をルールの構成要素として取り込む、[16] のように他の方法と組み合わせた定量的予測法への途を指向している。3.2 に述べた 42 化合物群の変異原性に対する、後者の結果を図6に示す。

```
If molecule A has:
    a LUMO value L1 of at most -1.35, and
    a carbon atom C in a 6-membered aromatic ring
                        with nitrogens around, and
    a LogP value L2
then its expected activity is -2.92* L1 +1.499*L2 - 12.72
Else if molecule A has:
    a carbonyl carbon C with partial charge at most
    0.608,
    C is connected via a single bond to some atom D
then its expected activity is 1.38
Else if molecule A has:
    a hydroxyl oxygen
  then its expected activity is very low.
```

図6．Pharmacophore の推定結果

4　グラフ構造を対象とした相関ルール探索

マーケットバスケット分析で用いられた相関ルールの方法は、大容量のデータに対して高速な計算を行えることが知られている（本特集での塩田氏の紹介を参照）。この方法の特徴は、ラティス内の各レベル毎に頻出 itemset を求め、これを次のレベルの節点生成時に、枝刈り情報として用いる点にある。

Dehaspe らは、[17] 枠組みとしては前節の ILP を用いながらも相関ルールの方法を採用し、各レベル毎に itemset に代わる部分グラフを生成して、化合物群に特徴的な頻出部分構造の発見を試みた。対象としたデータは後述の NTP による発ガン性データベース 337 化合物である。前提知識として、構造式情報以外には生物学的テスト結果のみを取り入れている。

彼らの実験結果によると、ラティスの第6レベルまでを展開したが、マーケットバスケット分析とは異なり、枝狩りの効果は小さいようである。また生成したラティス内では、発ガン性と相関する部分構造は見いだせていない。しかし、生物学的テストと組み合わせた場合には、右図に示す2種の発ガン性を抑制するルールが発見された。この方法は、前節と同じ ILP 表現を使っているが、頻出 itemset の制限、電荷を考慮しない点などが異なるため、Progol による実験結果 [13] とは一致しない。Progol による結果は電荷を含んだ構造がよく現れ、図5の結果よりも優れているように見える。しかし、筆者は、述語の選択、事例数が多い場合の計算効率などを考慮すれば、今後は相関ルール型のシステムの有効性が認識されてくると考えている。

```
cytogen_ca(C, n), sulfide(C, S) ->non-carcin(C)
        support: 0.06    confidence: 0.86
atomch(C, A, X), X≤-0.215, salmonella(C, n)
                              -> non-carcin(C)
        support: 0.27    confidence: 0.62
```

図5．Pharmacophore の推定結果

ごく最近、元田らは述語表現ではなくグラフ構造を記述する方法により、グラフ構造の相関ルール探索システムを作成した。[18] 化学構造への応用は未だなされていないが、HTS で出力される大量データの解析にも対応できる枠組みである。今後の発展が期待される。

5　まとめ

まず、これまでの節では触れられなかったいくつかの研究を紹介する。3次元構造の一部をハッシングする技法が開発されている。たとえば Wang らは、環構造の近似的な一致に基づくハッシュ技法を報告している。[19] 今後3次元構造検索を含むシステムの要素技術として重要になる可能性がある。吉田、元田らの GBI [20] や Cook らの SUBDUE [21] は、グラフ構造群からそれらの共通部分グラフを抽出し、それを使用してもとのグラフを圧縮して提示する機能を持つ。化学グラフにも適用可能で（SUBDUE で簡単な事例がある）、頻出する部分構造を検知できる。しかも、GBI では計算が非常に高速である。しかし、ラティスのように部分グラフのすべてを列挙するわけではないため、化学における有用性は未だ不明である。

米国の National Toxicology Program の一環としてネズミを対象とした発ガン性物質の評価が行われている(*http://dir.niehs.nih.gov/dirlecm/pte2.htm*)。このデータは、多様な化合物を含む構造活性相関において、各種の方法を比較検討するための標準データセットとなった感がある。筆者の狭い知識でも、北里大学や物質研がこのデータベースを対象とした研究を行っている。このデータセットに対し、Srinivasan が"The Predictive Toxicology Evaluation Challenge"を提唱している。正確でかつ理解の容易なモデルの構築を目指し、機械学習ばかりでなく広い分野の方法を対象として、その結果を比較できるようにしている。現在すでに 11 件の報告がなされており、その中には ILP によるものはもちろん、ILP で生成した記述子による決定木、回帰をはじめとする統計的方法の結果などが含まれている。WWW 上で公開されているため是非一度見ていただきたい。(*http://www.comlab.ox.ac.uk/oucl/groups/machlearn/PTE/*)。

日本国内において、構造活性相関や化学構造の取り扱いを含む研究は、質・量ともに誇るべきものを持っている。また、情報科学においても、機械学習やデータベースの研究は世界レベルの研究を推進している。しかし、本稿で取り上げた境界領域では、研究者数が決定的に少ないように思われる。すでに 15 年以上の歴史を持ち、一般の統計パッケージの中に組み込まれ始めた決定木でさえ、化学の領域へ適用しようとする研究者が非常に少ないことはなぜであろうか。さらに今、HTS の結果を理解が容易な形で提示し、ドラッグデザインを実りあるものとするには、大規模データマイニングの技術が不可欠であろう。この豊穣な木が国内で育つ契機として、本稿を役立てて頂ければ筆者の幸いである。

謝辞　本稿を書くにあたっては、本年の AAAI Spring Symposium の予稿集（文献[16]はその中の一報）を、下記の引用に現れる以上に大いに参考とした。予稿集を頂いた島根医大の津本先生に感謝いたします。

参考文献

1) Klopman, G Artificial Intelligence Approach to Structure-Activity Studies. *J. Amer. Chem. Soc.* **106**, 7315–7321, (1984).

2) Klopman, G MultiCASE 1. A Hierarchical Computer Automated Structure Evaluation Program, *QSAR*, **11**, 176–184 (1992).

3) Klopman, G; Rosenkranz, H. Approaches to SAR in carcinogenesis and mutagenesis. *Mutation Res.* **305**, 33–46 (1994).

4) Lee, V.; Buchanan, G; Aronis, J. Knowledge-Based Learning in Explanatory Science: Learning Rules to Predict Rodent Carcinogenicity. *Machine Learning*, **30**, 217-240 (1998).

5) 古川康一. 帰納論理プログラミング: チュートリアル. 人工知能学会誌, **12**, 655–664 (1997).

6) Muggleton, S.; Raedt, L.D. Inductive Logic Programming: Theory and Methods. *J. Logic Programming*, **19**, 629–679 (1994).

7) Muggleton, S. Inverse Entailment and Progol. *New Generation Comput.*, **13**, 245–286 (1995).

8) King, R.D.; Muggleton, S.H.; Srinivasan, A.; Sternberg, M.J. Structure-Acitivity Relationships Derived by Machine Learning. *Proc. Natl. Acad. Sci. USA*, **93**, 438–442 (1996).

9) Debnath, A.K. et al. Structure-Activity Relationship of Mutagenic Aromatic and Heteroaromatic Nitro compounds. *J. Med. Chem.* **34**, 786–797 (1991).

10) Finn, P.; Muggleton, S.; Page, D.; Srinivasan, A. Pharmacophore Discovery Using the Inductive Logic Programming. *Machine Learning*, **30**, 241–270 (1998).

11) Mayer, D. A Unique Geometry of Active Site of Angiotensin Converting Enzyme Consistent with Structure-Activity Studies. *J. Computer-Aided Molecular Design*, **1**, 13–16 (1987).

12) Dzeroski, S. et al. Applying ILP to Diterpene Structure Elucidation from ^{13}C NMR Spectra. *Proc. ILP-96*, 41–54, LNAI 1314, Springer-Verlag (1997).

13) Srinivasan, A.; King, R.D.; Muggleton, S.H., Sternberg M.J.E. Carcinogenesis Prediction Using ILP. *Proc. ILP-97*, 273–287, LNAI 1297, Springer-Verlag (1997).

14) Bryant, C.H. Data Mining via ILP: The Application of Progol to a Database of Enantioseparations. *Proc. ILP-97*, 85–92, LNAI 1297, Springer-Verlag (1997).

15) King, R.D.; Srinivasan, A. The Discovery of Indicator Variables for QSAR using Inductive Logic Programming. *J. Computer-Aided Mol. Design*, **11**, 571–580 (1997).

16) Srinivasan, A.; King, R.D. Using Inductive Logic Programming to Construct Structure-Activity Relationships. AAAI Spring Symposium, *Predictive Toxicology of Chemicals: Experiences and Impact of AI Tools*, 64–73 AAAI SS-99-01 (1999).

17) Dehaspe, L.; Toivonen, H.; King, R.D. Finding Frequent Substructures in Chemical Compounds. *Proc. KDD-98*, 30–36, AAAI (1998).

18) 猪口明博, 鷲尾隆, 元田浩. 膨大なグラフ構造データからの高速マイニング手法. 人工知能学会全国大会, **13**, 397–400 (1999).

19) Wang, X. et al. Automated Discovery of Active Motifs in Three Dimensional Molecules, *Proc. KDD-97*, 89–95 (1997).

20) Yoshida, K.; Motoda, H. CLIP: Concept Learning from Inference Pattern. *J. Artificial Intelligence*, **75**, 63–92 (1995).

21) Cook, D.J., Holder, L.H. Substructure Discovery Using Minimum Description Length and Background Knowledge, *J. A. I. Res.* **1**, 231-255 (1994).

おかだ　たかし　OKADA, Takashi
そろそろ化学構造も扱わなければ
連絡先　〒662-8501 西宮市上ケ原 1-1-155 関西学院大学情報メディア教育センター、Tel: 0798-54-6042

新任の挨拶

旭化成工業（株）コンピューターサイエンス(CS)室　　坂本 直紀

今年から日本化学会情報化学部会の編集委員に加えていただきました。現在はまだ、ベテラン編集委員の方々の足を引っ張っている状態なのですが、今後ともよろしくお願いいたします。

実は私、今でこそコンピューターサイエンス(CS)室に所属し、MD 等による高分子関連の計算を生業としておりますが、ついこの間までは、ねばねば、べとべと、悪臭の高分子実験屋でした。それで冒頭ように“足を引っ張る”状態となっているのですが、実験をしているときも、計算屋（理論物理屋）さんとは長く共同研究をしておりました。相手の方が非常に優秀であったこともあり、この共同研究は少なくとも私にとってはかなり有意義なものでした。現在、逆の立場で実験家といっしょに仕事をする機会が多く、私自身、この時のことが何かの参考になるかな、と思い、当時の状況を思い出してみました。

CS 室に来る前、私は高分子系における相分離現象のダイナミクスを散乱法や顕微鏡法により研究しておりましたが、当初は計算屋さんと共同で研究することはまったく考えていませんでした。ある時、透過型電子顕微鏡(TEM)観察で、これまでまったく観察されていない現象（構造）を捕らえることに成功しました。しかし、観察はできたものの、私自身も“これは真実を見ているのか？”といった疑問が常にあり、何か確認する手段はないかと、ノイローゼになるくらい考えていました。この時にはテーブルや天井の模様まで TEM 像に見えてきました。そんな時、理論高分子物理を専門としている方と共同研究をする機会に恵まれましたので、以下のようなわがままを言ってみました。

1、得られた実験結果(TEM 像)をシミュレーションで再現して欲しい。（裏付けが欲しい。）

2、実験で得られない部分を補強して欲しい。

その結果、実験をサポートする計算結果が得られ、自信をもって TEM 観察の結果を発表することが出来ました。

こうした訳で、計算（理論）科学を専門とされている方には今でも一目置いているのですが、現在、私もその一員になってしまいました。そして、逆に実験家から相談を受ける立場となりました。人間（私だけか？）は現金なもので、今では“実験でうまくいかない話ばかりが持ち込まれる“と愚痴をこぼしたりもしています。上の経験では、こうした”困っている“テーマのほうが双方協力し合い、良い結果がでるのでしょうが…。ただ、うまく行くのは計算屋が優秀な時だけでしょうから、今後、自分を磨いていかなければなりません。今後ともご鞭撻のほどよろしくお願いいたします。

部会役員会報告

第2回幹事会　兼　CICSJ Bulletin 編集委員会

日　時：4月1日　午後2時～

場　所：日本化学会会議室

出席者：吉田元二、時田澄男、三戸邦郎、伊藤眞人、船津公人、高木達也、相田美砂子、相馬融、高田章、工藤喜弘

議　題：CICSJ Bulletin vol. 17No. 2の編集作業を行った。

　　第22回情報化学討論会について

　　　工藤喜弘氏より経過報告があった。

　　　特別講演、基調講演、外国人講演などについて検討した。

　　講習会：

　　　「特許」をテーマとして行う案を前向きに検討することとなった。

　　論文誌：

　　　討論会のProceedingの形で刊行する方向で準備を始めることとなった。

　　国際交流の促進：

　　　部会設立20周年を見据えて、具体的な事業とその準備について検討した。

第3回幹事会　兼　CICSJ Bulletin 編集委員会

日　時：1999年5月25日（火）午後2時～

場　所：日本化学会会議室

出席者：吉田元二、時田澄男、三戸邦郎、伊藤眞人、高田章、工藤喜弘、船津公人、高木達也、相田美砂子、相馬融、坂本直紀、清水弘一

議題：CICSJ Bulletin 第17巻の進捗状況、18巻の準備

　　No. 3　：会員名簿編集作業

　　No. 4　：特集「データマイニング」執筆依頼中。

　　No. 5　：特集「第22回情報化学討論会」9月中に発行。

　　No. 6　：特集「材料を探る」の内容検討。12月刊行。

　　第18巻 No. 1　：特集テーマ案の検討

　情報化学討論会について

　　第22回進捗状況の報告

　　　プログラム概要の検討：特別講演、基調講演その他の決定

　　　プログラム編成会議の打ち合わせ

　　第23回の概要等を検討した。

　高分子学会・計算科学研究会と共催の講習会開催について（12月、名古屋）

4～8月部会員異動

［入会］

（個人正部会員　１０名）

白木　武夫　　　　　　　　　相馬　融（科学技術振興事業団）

岩浪　克之（日本グラクソ）　小塚　健志（室蘭工業大学）

神田　友司（丸石製薬）　　　伴　尚（琉球大学）

宮本　明雄（日本大学）　　　古明地　勇人（産業技術融合領域研究所）

相川　泰（東洋紡績）　　　　坂本　直紀（旭化成工業）

（学生部会員　　３名）

山本　敦巳（愛知教育大学）　　平田　晃義（高知大学）

門屋　元（岡山理科大学）

（順不同・敬称略）

◆部会行事

「第１２回情報化学講習会」参加のお誘い

日本化学会情報化学部会では，第 22 回情報化学討論会に併催して，講習会を開催いたします．今回は，下記のタイトルで平易な解説をお願いする予定です．これから，これらの分野について勉強したいと考えているお若い方々ばかりでなく，いままでの知識を整理して新しい発展を図ろうとしているベテランの研究者・技術者の方々にも，奮ってご参加いただきたいと思います．

日　時　1999 年 11 月 12 日(金)　13:30～16:00

場　所　グランドホクヨウ米沢　(山形県米沢市金池 2-3-7)

参加費　討論会参加者　無料
一般　5,000 円*
情報化学部会員　2,000 円
学生　無料
*情報化学部会にその場で入会されますと，4,000 円（年会費 2,000 円，参加費 2,000 円）となります．

講　演

1）「計算化学のすすめ － 電子スペクトル計算を例にして」
岡山理科大学客員教授、大阪市立大学名誉教授　西本吉助

2）「ケモメトリックス入門」
福島工業高等専門学校教授　石山純一

参加申込

当日会場でお申し込みいただきます。
（学生の方は身分証明書の提示をお願いいいたします。）

問合先　山形大学工学部電子情報工学科　工藤 喜弘
〒992－8510 山形県米沢市城南 4－3－16
Tel 0238－26－3360（注意:会誌会告の一部に誤り有り）
Fax 0238－26－3359
e-mail　cicsj@dip.yz.yamagata-u.ac.jp

第１１回ケモメトリックス・ワークショップ

コンピュータを利用した官能評価データの収集とケモメトリックス手法による解析

(Compusense Inc.,　カナダ) Dr.Chris Findlay

主　催　日本化学会情報化学部会
会　期　1999年11月19日（金）10～17時、懇親会 17～19時
会　場　日本化学会会館5階会議室　(定員80名)
（〒101-8307 東京都千代田区神田駿河台1-5、Tel: 03-3292-6161）
参加費　3,000円、当日徴収いたします。懇親会は無料です。

申込方法　参加ご希望の方は、氏名、所属、住所（電話、ファックス番号）を日本語と英語で、電子メールアドレスも明記の上、懇親会参加の有無と共にハガキ、ファックスまたは電子メールにて下記宛、11月10日までにお申し込み下さい。
申込先　〒278-0037 野田市野田399 キッコーマン（株）研究本部　相島鐵郎
Tel: 0471-23-5980、Fax: 0471-23-5550、電子メール:LDH00052@nifty.ne.jp

プログラム

１．官能評価へのコンピュータの導入
・実験計画の立案
・自動化、迅速化
・データの質向上、統合
２．多変量解析手法による解析
・複雑なデータの要約
・マッピングによる視覚化、官能特性の強度予測
・プロクラステス分析
３．時間強度（Time Intensity :T-I）データの収集と解析
・SATI：１特性のT-Iデータ
・DATI：２特性のT-Iデータ→相互作用の解析
・TAD：多特性のT-Iデータ→複雑な相互作用の解析
４．ソフトウェア・デモンストレーション
・日本語版 Compusense five によるデータ収集、解析
・プロクラステス分析手法による解析
５．パネリストの評価
・官能評価ISO化への対応
・卵殻図形の利用
・パネルチェック・プログラム
６．要約と結語
・客観的で再現性高い官能評価
・消費者嗜好の理解
・新製品開発への指針

本講演は英語で行われます。

CICSJ Bulletin Vol.17, No.5, October 1999

このページは「日本化学会・情報化学部会」の責任において運営されています。

目次

第22回情報化学討論会

（第 27 回構造活性相関シンポジウムと並催）

主催 日本化学会情報化学部会
共催 日本薬学会 日本農芸化学会 日本分析化学会 日本化学プログラム交換機構
日時 平成 11 年 11 月 10 日(水)～12 日(金)
会場 グランドホクヨウ米沢 (山形県米沢市金池 2-3-7)
発表形式 一般講演(討論時間を含めて 15 分)，ポスターセッション
参加登録予約申込締切 10 月 8 日(金)

第 1 日 (11 月 10 日)

12:30-13:15 **キーノートトレクチャー** 座長: 伊藤 眞人
J 01 化学教育への情報遠隔操作技術の活用 (仮題) (南ア) ○R. D. Sanderson

13:15-14:00 **一般講演 化学情報学** 1 座長: 相馬 融
J 02 tDNA の系統的検出 (山形大工) 工藤 喜弘, 金谷 重彦, ○木ノ内 誠
J 03 ヒトゲノムのエキソンとイントロンの識別 (山形大工) 工藤 喜弘, 金谷 重彦, 木ノ内 誠, ○坂井 孝光
J 04 情報化学システム CGS の開発(8)化学反応速度論的微分方程式系の符号化と速度論的シミュレータ (日本情報化学システムズ) ○内野 正弘

14:10-14:55 **一般講演 化学情報学** 2 座長: 船津 公人
J 05 JCICJ の分析を中心とした情報化学の計量文献学的研究 (図書館情報大) ○小野寺 夏生
J 06 JST における高分子データベースの開発 －IUPAC 構造基礎名の自動発生 Phase1－ (科学技術振興事業団) 飯室 茂, ○前田 知子, (三菱総研) 高枝 佳男, 長阪 匡介
J 07 マイクロソフト ACCESS 上の質量スペクトルデータベース (日本アステック) 六車 裕孝, (筑波研究学園専門学校) 仲久保 正人, (アイエヌジェイ) ○前田 浩五郎

15:05-15:50 **一般講演 化学情報学** 3 座長: 松岡 俊和
J 08 化学物質のリスクアセスメントシステムの開発 (日本化学工業協会) ○花井 荘輔
J 09 インターネット上の反応データベースシステム -3- (神奈川大理) ○中山 堯, 上野山 敦史, 兼子 崇, 柳沢 一博
J 10 インターネットにおける SDBS-NMR (物質研) ○早水 紀久子, 矢部 篤子, 柳沢 勝

16:30-17:30 **特別講演** 座長: 寺田 弘
KJ 生物活性物質の探索: 化学で探る生物の世界 (秋田県立大学学長) ○鈴木 昭憲

第 2 日 (11 月 11 日)

9:00- 9:40 **キーノートトレクチャー** 座長: 時田 澄男
J 11 第一原理計算による原子レベルデバイス設計の現状と将来展望 (東北大) ○川添 良幸

9:40-10:40 **一般講演 計算化学** 1 座長: 久保 百司

J 12 光合成反応中心における電子移動に関する理論的研究 (物質研) ○北尾 修, 青木 孝造, (電総研) 建部 修見, 関口 智嗣

J 13 TS-1/H2O2/PhOH 酸化反応活性種生成におよぼす溶媒構造の効果 (宇部興産) ○後口 隆, 八尾 滋,

J 14 ケミカル・ハードネスを用いた電荷移動を伴う分子シミュレーション手法の検討 (旭硝子) ○高田 章

J 15 QM/MM-vib 法に基づいた ab initio MD 法の開発と水溶液内反応への応用 (広島大院理) ○相田 美砂子, (大阪大産業科学研) 山高 博, (Pacific Northwest National Lab.) Michel DUPUIS

10:50-11:50 **一般講演 計算化学2** 座長: 高田 章

J 16 ナフタレン陽イオンを触媒とする水素分子生成反応 (静岡大理) ○徃住 貴顕, 萩原 紀子, 寺田 典弘, (静岡大工) 石田 俊正, (静岡大理) 相原 惇一

J 17 リチウム含有遷移金属酸化物の構造特性と電子状態に関する計算化学的検討 (東北大院工) ○鈴木 研, 小野津 崇之, 高見 誠一, 久保 百司, 宮本 明

J 18 ダイヤモンド膜の表面反応に関する計算化学的検討 (東北大院工) ○田村宏之, 周 慧, 高見 誠一, 久保 百司, 宮本 明, (無機材研) 蒲生西谷 美香, 安藤 寿浩

J 19 量子分子動力学プログラム'colors'の開殻系への拡張 (東北大院工) ○高見 誠一, 山田 有場, 田村 宏之, 久保 百司, 宮本 明, (広島国際学院大工) 今村 詮

13:30-16:00 **ポスターセッション**

<化学情報学>

JP 01 情報化学システム CGS の開発(9)正規4次グラフの生成、符号化、対称性認識およびデータベースの構築について (日本情報化学システムズ) ○内野 正弘

JP 02 二節部分グラフによる IUPAC 名の生成 (京都工芸繊維大院工芸) ○安代 容子, 藤田 眞作

JP 03 XyMML (XyM Markup Language)---構造式のコミュニケーションのためのマークアップ言語 (京都工芸繊維大工芸) ○藤田 眞作

JP 04 有機合成設計システム AIPHOS における戦略部位の優先順位付け手法の開発 (住友化学工業) ○田中 章夫, (豊橋技科大) 船津 公人

JP 05 有機合成設計システム AIPHOS-KOSP の開発(2) (第一製薬) ○佐藤 耕司, (豊橋技科大) 船津 公人

JP 06 化学物質生体影響データベース (科技団) ○志村 和樹

JP 07 化学物質生分解性予測システムの開発 (化学品検査協会, 大阪大) ○廣松 康一, (富士通九州システムエンジニアリング) 増本 茂和, (化学品検査協会) 矢可部 芳州, (大阪大) 西原 力

JP 08 生物情報統合データベースシステムの構築 －新しい発想に基づいた生物情報管理法－ (医薬分子設計研究所) ○福田 美紀, 瀧川 有紀, 豊田 哲郎, 板井 昭子

JP 09 錯体・有機金属化合物の双極子モーメントのデータベース作成(続報) (日赤看護大) ○山崎 昶

JP 10 立体化学の規範的コード化法の開発とコンホマーの識別 (理研) ○越野 広雪, (科技団さきがけ研究21) 佐藤 寛子, (豊橋技科大) 船津 公人, (理研) 鵜澤 洵, 中田 忠

JP 11 Circular permutation を取り扱うことのできる構造局所アラインメント法 (生物分子工学研究所) ○廣池 隆明, 藤 博幸

JP 12 二次元相関分光法による昇温脱離分析データの解析 (豊橋技科大) ○杉田 記男, (電子科学) 宮林 延良, (豊橋技科大) 阿部 英次

<計算化学>

JP 13 量子化学計算によるリチウムイオン電池材料の設計 (物質研) ○田辺 和俊, 松本 高利, (融合研) 長嶋 雲兵

JP 14 ライノウィルス外殻蛋白質の700ピコ秒の分子動力学計算 (北里大理) ○米田 茂隆, (北里大薬) 米田 照代, 梅山 秀明

JP 15 トライボロジーに関する計算化学的研究 (東北大院工) ○周 慧, 田村 宏之, 高見 誠一, 久保 百司, 宮本 明

JP 16 Dual Ensemble Molecular Dynamics (DEMD) 手法の開発とゼオライト膜による気体分離への応用 (東北大院工) ○小林 泰則, 森戸 英明, 水上 浩一, 高見 誠一, 久保 百司, 宮本 明

JP 17 タンパク質内電子移動グラフィックス (九工大情報工) ○柏木 浩, 佐藤 文俊, 吉廣 保, 岡崎 功, 上田 尚学, 保田 慎一郎, 恵良 信, 時枝 浩司, 佐藤 洋子

JP 18 タンパク質全電子計算プログラム ProteinDF におけるオブジェクト指向技術と分散並列処理 (九工大情報工) ○吉廣 保, 佐藤 文俊, 柏木 浩

JP 19 III-V 族窒化物系半導体の結晶成長過程に関する計算化学的検討 (東北大院工) ○稲葉 祐策, 小野津 崇之, 高見 誠一, 久保 百司, 宮本 明

JP 20 ヘプタフルバレンにおける擬 Jahn-Teller 効果の MCSCF 法による解析 (山形大教育) ○豊田 東雄, (三重大工) 小関 史朗

JP 21 第一原理計算による新規メタノール合成触媒の理論設計 (東北大院工) ○坂原 悟, 遠藤 明, 高見 誠一, Belosludov Rodion, Zhanpeisov Nourbosyn, 久保 百司, 宮本 明

JP 22 遺伝的アルゴリズムを用いる1本鎖コラゲンモデルの安定構造探索 (近畿大) ○高橋 紘一, 田中 克裕

JP 23 超高速分子軌道計算専用機 MOEngine の開発 (大正製薬創薬研) ○高島 一, 北村 一泰, (北海道教育大教育) 小原 繁, (島根大総合理工) 網崎 孝志, (九州大学院システム情報) 村上 一彰, (電総研) 早川 潔, (富士ゼロックス総合研究所) 山田 想, 稲畑 深二郎, 宮川 宣明, (物質研) 田辺 和俊, (融合研) 長嶋 雲兵

JP 24 分子の電子状態計算プログラム HAM/3 の修正 (信州大繊維) ○藤本 徳之, 成田 進, 渋谷 泰一

JP 25 Ab initio ペア近似法による巨大分子計算プログラムの開発 (国立衛研) ○中野 達也, 神沼 二眞, (生工研) 上林 正巳, (富士総研) 佐藤 智之, (RWCP) 秋山 泰, (阪府大) 北浦 和夫

JP 26 ウレタン化反応のアミン触媒機構の解析 (花王構造解析センター) ○山下 修, (花王生物科学研究所) 小澤 忠広

JP 27 C-13 化学シフトテンソルを利用したβ-ジケトンエノールの分子構造解析 (京大院理) ○今城 文雄

JP 28 フラーレン分子の局所的反応性と HOMO-LUMO エネルギー間隔の関係 (静岡大理) ○相原 惇一

JP 29 分子動力学法によるスチレン−エチレン共重合体のガラス転移温度計算 (電気化学工業中央研究所) ○戸木田 裕一, 鈴木 茂, 岡本 彰夫, (ナノシミュレーションアソシエーツ) 桑島 聖

<ケモメトリックス>

JP 30 ケモメトリックスによるプラスチック廃棄物の判別 (物質研) ○田辺 和俊, 松本 高利, (島津製作所) 鈴木 康志, (日産自動車) 岡田 行雄

JP 31 ニューラルネットワークを用いたフロン代替化合物の物性推算、および、新規化合物の逆設計 (旭硝子) ○山本 博志, (地球環境産業技術研究機構) 浦田 新吾

JP 32 ニューラルネットワークによる高分子の物性の予測 (物質研) ○田辺 和俊, 松本 高利

JP 33 ケモメトリックス手法を用いた材料開発支援プログラムの開発(3)−カウンタープロパゲ

ーションを用いた候補のクラスタリングー (豊橋技科大) ○棚田 東作, 荒川 正幹, 船津 公人, (三井化学) 西村 竜一

JP 34 Multiway PLS 法を用いた活性コンフォメーションと重ね合わせルール選択手法 (豊橋技科大) ○荒川 正幹, 船津公人

JP 35 遺伝アルゴリズムを用いた剛球モデルに基づくタンパク質立体構造比較 (図書館情報大) ○中山 伸一, (The University of Sheffield) Peter Willett, (図書館情報大) 吉田政幸

JP 36 低分解能質量スペクトルにおける部分構造識別関数構築のための変数選択手法の検討 (豊橋技科大) ○吉田 浩士, 船津 公人, (Technical University of Vienna) Kurz Varmuza

JP 37 カスケードモデルによる定性的構造活性相関 (関西学院大情報メディア教育センター) ○岡田 孝

JP 38 コンピュータ集約型検定法の一般化加法モデルへの適用と応用 (大阪大院薬) ○高木 達也, (大阪大薬) 渡里 智恵子, (大阪大院薬) 宮田 幸治, (大阪大薬) 黒川 絵美子, (大阪大医) 藤原英明, (大阪大遺伝情報実験施設) 黒川 顕, 長谷川 義高, 安永 照雄

JP 39 学習順序に依存しない自己組織化法の開発 (山形大工) ○阿部 貴志, 金谷 重彦, 工藤 喜弘, 木ノ内 誠, 大平 賢

<情報化学教育>

JP 40 SALS による数値計算曲線のカーブフィッティング -- 多段階プロトン解離平衡の精密解析 -- (山形大理) ○加藤 良清, (長岡技科大工) 山田 明文, (山形大工) 小野寺 準一, (山形大理) 鵜浦 啓

JP 41 インターネットを活用する「立体化学」独習システムの試作 (神奈川大理) 竹内 敬人, (広島大理) 吉田 弘, (創価大) 守谷 崇, ○伊藤 眞人

16:30-17:30 **特別講演** 座長: 吉田 元二

JK 高分子材料設計の最近の動向 (名古屋大学大学院工学研究科計算理工学専攻) ○土井 正男

第 3 日 (11 月 12 日)

9:00- 9:40 **キーノートレクチャー** 座長: 三戸 邦郎

J 20 ケモメトリックスが拓く設計・予測の世界 (豊橋技科大) ○船津 公人

9:40-10:25 **一般講演 ケモメトリックス** 座長: 高木 達也

J 21 3 次元的分子数値化法 FRAU の特性値を識別子とした試薬の自動分類と機能予測 (科技団さきがけ研究 21) ○佐藤 寛子, (豊橋技科大) 船津 公人, (お茶大) 鷹野 景子, (理研) 中田 忠

J 22 ニューラルネットワークによる化学物質の安全性の予測 (物質研) ○田辺 和俊, 松本 高利

J 23 X 線回折を用いた地下岩石情報の可視化 (一関高専) ○貝原 巳樹雄, 大岡 靖典, (帝国石油) 金田 英伯

10:35-12:20 **一般講演 計算化学** 3 座長: 相田 美砂子(4 件), 坂本 直紀(3 件)

J 24 有機電子論の数学的裏付けとその限界 (お茶大理) ○細矢 治夫

J 25 シュウィンガー・ダイソン形式による超重原子の理論 (政策研究大学院大) ○松浦 弘幸, (産業医科大) 中野 正博

J 26 分子力学法と分子動力学法による長鎖脂肪酸単分子膜の研究 (大阪電気通信大) ○橋口 聡, 藤田 岩男

J 27 ProteinDF によるタンパク質全電子計算と電子トンネリング過程 (九工大) ○佐藤 文俊, 吉廣 保, 岡崎 功, 上田 尚学, 保田 慎一郎, 恵良 信, 時枝 浩司, 佐藤 洋子, 柏木 浩

J 28 酸化物エレクトロニクス材料のハイオーダーエピタキシー機構の解明 (東北大院工) ○久保百司, 稲葉 祐策, 小野津 崇之, 高見 誠一, 宮本 明, (東工大院総合理工) 川崎 雅司, (東工大応セラ研) 吉本 護, (科技団戦略研、東工大応セラ研) 鯉沼 秀臣

J 29 シリコン酸化膜における応力特性の計算化学的検討 (東北大院工) ○山田 有場, 黒川 仁, 遠藤 明, 高見 誠一, 久保 百司, 宮本 明

J 30 サファイア基板上における GaN 薄膜初期成長の計算化学的検討 (東北大院工) ○小野津 崇之, 稲葉 祐策, 高見 誠一, 久保 百司, 宮本 明, (住友化学) 家近 泰, 前田 尚良

13:30-16:00 **第 12 回情報化学講習会**

計算化学のすすめ ― 電子スペクトル計算を例にして (岡山理科大学客員教授, 大阪市立大学名誉教授) 西本 吉助

ケモメトリックス入門 (福島工業高等専門学校教授) 石山純一

参加登録費 (一般) 当日 9000 円, 予約 7500 円, (学生) 当日 6000 円, 予約 4000 円 (第 27 回構造活性相関シンポジウムも聴講可, 要旨集合本 1 冊を含む), 要旨集前送を希望の方は送料として 1000 円を追加のこと

懇親会 11 月 11 日(木) グランドホクヨウ米沢, 会費: 当日 9000 円, 予約 7500 円

参加登録予約申込方法 申込 1 人毎に, 氏名(ふりがな)・所属・住所・電話・FAX・懇親会出欠・要旨集前送希望の有無・振込日・振込金額を記入の上, E-mail または郵送でお申込み下さい. 参加登録費等は郵便振替(第 22 回情報化学討論会 02290-5-82539, 通信欄に参加者氏名および内訳を記入)によりお振込み下さい. 入金確認をもって予約完了といたします。自前の領収書は発行いたしません.

申込先 〒992-8510 米沢市城南 4-3-16 山形大学工学部電子情報工学科 工藤喜弘 気付 情報化学討論会 Tel (0238)26-3360 Fax (0238)26-3359 E-mail: cicsj@dip.yz.yamagata-u.ac.jp ホームページ http://www.yz.yamagata-u.ac.jp/event/CICSJ.html

第 27 回構造活性相関シンポジウム

(第 22 回情報化学討論会と並催)

日時 平成 11 年 11 月 10 日 (水)～ 12 日 (金)
会場 グランドホクヨウ米沢
(山形県米沢市金池 2-3-7 Tel 0238-22-1238, Fax 0238-21-1067)
主催 日本薬学会医薬化学部会
共催 日本化学会, 日本農芸化学会, 日本分析化学会, 日本農薬学会, 構造活性相関研究会
講演時間 一般講演は口頭 25 分 (講演 20 分, 討論 5 分) または口頭 18 分* (講演 13 分, 討論 5 分)

プログラム
第 1 日
一般講演 14:00 ～ 15:51

J 28 酸化物エレクトロニクス材料のハイオーダーエピタキシー機構の解明 (東北大院工) ○久保 百司, 稲葉 祐策, 小野津 崇之, 高見 誠一, 宮本 明, (東工大院総合理工) 川崎 雅司, (東工大応セラ研) 吉本 護, (科技団戦略研、東工大応セラ研) 鯉沼 秀臣

J 29 シリコン酸化膜における応力特性の計算化学的検討 (東北大院工) ○山田 有場, 黒川 仁, 遠藤 明, 高見 誠一, 久保 百司, 宮本 明

J 30 サファイア基板上における GaN 薄膜初期成長の計算化学的検討 (東北大院工) ○小野津 崇之, 稲葉 祐策, 高見 誠一, 久保 百司, 宮本 明, (住友化学) 家近 泰, 前田 尚良

13:30-16:00 **第 12 回情報化学講習会**

計算化学のすすめ － 電子スペクトル計算を例にして (岡山理科大学客員教授, 大阪市立大学名誉教授) 西本 吉助

ケモメトリックス入門 (福島工業高等専門学校教授) 石山純一

参加登録費 (一般) 当日 9000 円, 予約 7500 円, (学生) 当日 6000 円, 予約 4000 円 (第 27 回構造活性相関シンポジウムも聴講可, 要旨集合本 1 冊を含む), 要旨集前送を希望の方は送料として 1000 円を追加のこと

懇親会 11 月 11 日(木) グランドホクヨウ米沢, 会費: 当日 9000 円, 予約 7500 円

参加登録予約申込方法 申込 1 人毎に, 氏名(ふりがな)・所属・住所・電話・FAX・懇親会出欠・要旨集前送希望の有無・振込日・振込金額を記入の上, E-mail または郵送でお申込み下さい. 参加登録費等は郵便振替(第 22 回情報化学討論会 02290-5-82539, 通信欄に参加者氏名および内訳を記入)によりお振込み下さい. 入金確認をもって予約完了といたします。自前の領収書は発行いたしません.

申込先 〒992-8510 米沢市城南 4-3-16 山形大学工学部電子情報工学科 工藤喜弘 気付 情報化学討論会 Tel (0238)26-3360 Fax (0238)26-3359 E-mail: cicsj@dip.yz.yamagata-u.ac.jp ホームページ http://www.yz.yamagata-u.ac.jp/event/CICSJ.html

第 27 回構造活性相関シンポジウム

(第 22 回情報化学討論会と並催)

日時 平成 11 年 11 月 10 日 (水)～ 12 日 (金)
会場 グランドホクヨウ米沢
(山形県米沢市金池 2-3-7 Tel 0238-22-1238, Fax 0238-21-1067)
主催 日本薬学会医薬化学部会
共催 日本化学会, 日本農芸化学会, 日本分析化学会, 日本農薬学会, 構造活性相関研究会
講演時間 一般講演は口頭 25 分 (講演 20 分, 討論 5 分) または口頭 18 分* (講演 13 分, 討論 5 分)

プログラム
第 1 日
一般講演 14:00 ～ 15:51

K 01 Thrombin と 4-aminopyridine および benzamidine とのリガンド－酵素結合モデルの ab initio 分子軌道計算法による水素結合の解析：Structure Based Design への応用の模索（分子研）藤田忠男

K 02* ニューラルネットワークを用いた薬物吸収性の予測（日本新薬）丹羽朋子

K 03* 二元的疎水性定数を用いた構造活性相関．X線造影剤関連物質への応用（阪大医）柳川康洋，野口武士，福田英司，木村敦臣，○藤原英明

K 04 メソトレキセート誘導体の有機アニオントランスポーターに対する基質特異性の構造活性相関解析（中外製薬化学研）○大田雅照，（東大薬）杉山雄一，韓　龍海，渡辺由香，鈴木洋史，加藤将夫，（中外製薬化学研）原村昌幸，松岡宏治

K 05 化学物質の水生生物毒性予測システム'AquaTox'の開発（豊橋技科大）○高橋由雅，伴　聡，Gary Admans，加藤博明，阿部英次，（日化協）花井荘輔

特別講演 16:30 ～ 17:30

生物活性物質の検索：化学で探る生物の世界（秋田県立大学 学長）鈴木昭憲

第2日

招待講演 9:45 ～ 10:30

K 06 ペプチド・蛋白質の構造とダイナミックス－ラマン分光法による研究（東北大院薬）竹内英夫

ポスタープレビュー 10:30 ～ 11:45

ポスターセッション 13:30 ～ 14:45（奇数番号），14:45 ～ 16:00（偶数番号）

（情報化学討論会と共通のセッション）

KP 07 ヒト血清アルブミン（サイト II ）結合薬物の結合配座解析および結合モデルの構築（北里大薬）○松下泰雄，広野修一，（杏林製薬中央研）岩瀬一彦

KP 08 GPIIb/IIIa 受容体アンタゴニストの結合コンフォメーションの推定と CoMFA 解析（北里大薬）○中込　泉，小早川寿一，山乙教之，広野修一

KP 09 光アフィニティラベル法によるヒト血清アルブミン上の薬物結合部位のミクロ環境解析（熊本大薬）○Chuang Tuan Giam，國安明彦，中山　仁，小田切優樹

KP 10 TFS による化学物質の疎水性プロフィールの記述と QSAR への応用（豊橋技科大）○金畑竜彦，高橋　由雅

KP 11 結合中心フラグメントを利用した薬物の構造類似性評価（豊橋技科大）○平井貴朗，高橋由雅

KP 12 化合物の部分構造と変異原性の相関－4224 件の変異原性試験結果より－（労働省産業医学総合研）○猿渡雄彦

KP 13 匂い分子の特徴解析（新潟大院自然科学）○吉井文子，（新潟大教育人間科学）荻野敏夫，（新潟大歯）山田好秋，（新潟大院自然科学）星　隆，萩原久大

KP 14 ニトロベンゼン類の植物毒性（東北薬大）○木村聖子，根岸布貴子，遠藤勝也

KP 15 The effect of the constant domains of Fab antibody on the antigen binding property: a normal mode analysis of HyHEL-10/HEL complex（北里大薬）○胡　国平，栗原庸次，竹田（志鷹）真由子，神谷健秀，梅山秀明

KP 16 ホモロジーモデリングによる A55T プラスミン変異体の活性低下機序の解析（北里大薬）○竹田（志鷹）真由子，梅山秀明

KP 17 計算化学的手法によるヒスチジノールデヒドロゲナーゼ (HDH) の酵素反応中間体物質 L-ヒスチジナール分子の酵素中での結合配座の予測（ノバルティスファーマ宝塚研）○合田圭吾，（大阪府大農）太田大策，（ノバルティスファーマ宝塚研）岩崎源司，（Novartis Crop Protection Basel AG）Peter Ertl，Olivier Jacob

KP 18　糖鎖付加カルシトニンの構造活性相関（旭化成基盤セ）○田頭瑞加，田中淳志，久谷邦夫，（旭化成骨体研）磯谷幸宏，堀　正幸，（福井医大）高松真二，藤林靖久，（野口研）吉野栄一，山本敬三，羽田勝二，稲津敏行，戸澗一孔

KP 19　立体疎水性パラメータの創製と分子設計への応用：ニトロフェノール類の皮膚透過性ならびに Brefeldin A 立体異性体のアポトーシス誘導活性（徳島大工）○堀　均，朱　継文，名倉ふみ，MOHAMAD, Saharuddin B，宇都義浩，永沢秀子，（徳島大工，名大院生命農学）大倉一人

KP 20　ヒト特異的細胞溶解毒素インターメディリシンの構造：ヒト細胞表面機能改変モジュールへの膜結合ドメインの応用を目指して（徳島大工，名大院生命農）○大倉一人，（徳島大工）後藤隆次，服部加奈子，大西大樹，（徳島文理大健康科学研）津下英明，勝沼信彦，（徳島大工）高麗寛紀，長宗秀明

KP 21　ニューラルネットワークを用いた化学物質の毒性予測（豊橋技科大）伴　聡，Gary Adamans，高橋由雅

特別講演 16:30 ～ 17:30

高分子材料設計の最近の動向（名大院工）土井正男

第 3 日

一般講演　9:30 ～ 12:27

K 22*　非ステロイド型昆虫脱皮ホルモン類縁体の殺虫活性の QSAR およびその昆虫種間での比較（京大院農）○中川好秋，（Department of Crop Protection, Gent University）Guy Smagghe，（京大院農）上野民夫，（京大学名誉教授）藤田稔夫

K 23*　新規ヒドロキシフェニルウレア誘導体の Acyl-CoA: Cholesterol Acyltransferase (ACAT) 阻害活性の定量的構造活性相関（田辺製薬創薬研）○清水　良，中尾和也，窪田　均，安原三紀子，斎藤敬子，鈴木利一，大水　博

K 24 ピレスロイド系化合物の構造類似性と QSAR（徳島大薬）○中馬　寛，後藤　了，（呉羽化学）唐沢真美，（セイラシステム）佐々木幹夫，（産業領域融合技術研）長嶋雲兵，（大阪府大）西村勁一郎，（エミルプロジェクト）藤田稔夫

K 25*　ForceFeedBack を用いた分子設計におけるバーチャルリアリティシステムの開発（国立がんセンター研）水島　洋，永田　宏，（三菱スペースソフトウェア）田中恵理子，若宮加奈子，初田正彰

K 26*　表面プラズモン共鳴センサーを用いた糖結合タンパク質の結合活性の解析（徳島大薬）○後藤　了，三浦美由紀，桝田恭子，中馬　寛，寺田　弘

K 27 ヒストン脱アセチル化酵素 (HDAC) 阻害剤 depudecin に関する研究と chemical probe を用いた HDAC の構造・機能解析（Harvard Univ.）○大和隆志，Ho Jeong Kwon，Christian A. Hassig，Jeffrey K. Tong，Junichi Shimada，Phyllis G. Grable，Stuart L. Schreiber

K 28 構造活性相関に基づくチミジンホスホリラーゼ阻害剤 (TPI) の分子設計（大鵬薬品工業化学研）○多田幸雄，矢野伸吾，数野秀樹，佐藤　務，（大鵬薬品工業第二がん研）福島正和，（大鵬薬品工業化学研）浅尾哲次

K 29 フィブリノーゲン受容体アンタゴニストの三次元構造と血小板凝集阻害活性との関係（京大院農）○赤松美紀，宮下正弘，（京都薬大）林　良雄，（京大院農）上野民夫

予約参加登録申込締切　10 月 8 日（金）

予約参加登録申込方法　申込 1 人ごとに Email 又は官製葉書に，氏名（ふりがな）・所属・住所・

Tel・Fax・懇親会出欠・要旨集前送希望の有無・振込金額・振込日を記入の上，送付して下さい．また，参加登録申込はホームページからも可能です．参加登録費等は郵便為替（口座名：第 27 回構造活性相関シンポジウム 02260-0-43478，通信欄に参加者氏名および内訳を記入）により振込下さい．入金確認をもって予約完了といたします．

予約参加登録費（同日開催の情報化学討論会と共通）一般 7,500 円（当日 9,000 円），学生 4,000 円（当日 6,000 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．要旨集の前送希望者は送料として 1,000 円を追加して下さい．

懇親会 11 月 11 日（木）場所：グランドホクヨウ米沢 平安，会費：予約 7,500 円，当日 9,000 円．

申込先

〒980－8578 仙台市青葉区荒巻字青葉　東北大学大学院薬学研究科反応制御化学分野　坂本尚夫　TEL 022-217-6865，Fax 022-217-6864　Email：qsar@mail.pharm.tohoku.ac.jp http://www.pharm.tohoku.ac.jp/depts/hannou/sympo/index.html

第 12 回情報化学講習会

（第 22 回情報化学討論会に引き続き同じ会場で開催）

主催　日本化学会情報化学部会 (第 22 回情報化学討論会に併催)

会期　11 月 12 日(金) 13:30～16:00

会場　グランドホクヨウ米沢 (山形県米沢市金池 2-3-7)

1. 計算化学のすすめ － 電子スペクトル計算を例にして (岡山理科大学客員教授，大阪市立大学名誉教授) 西本吉助
2. ケモメトリックス入門 (福島工業高等専門学校教授) 石山純一

参加費　第 22 回情報化学討論会参加者 無料，一般 5000 円*，情報化学部会員 2000 円，学生 無料．*情報化学部会にその場で入会されますと，4000 円(年会費 2000 円，参加費 2000 円)となります．

参加申込　当日会場でお申し込みいただきます．学生の方は身分証明書の提示をお願いいたします．

問合せ先　992-8510 山形県米沢市城南 4-3-16 山形大学工学部電子情報工学科 工藤喜弘 Tel (0238)26-3360 FAX (0238)26-3359 E-mail: cicsj@dip.yz.yamagata-u.ac.jp

Tel・Fax・懇親会出欠・要旨集前送希望の有無・振込金額・振込日を記入の上，送付して下さい．また，参加登録申込はホームページからも可能です．参加登録費等は郵便為替（口座名：第 27 回構造活性相関シンポジウム 02260-0-43478，通信欄に参加者氏名および内訳を記入）により振込下さい．入金確認をもって予約完了といたします．

予約参加登録費（同日開催の情報化学討論会と共通）一般 7,500 円（当日 9,000 円），学生 4,000 円（当日 6,000 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．要旨集の前送希望者は送料として 1,000 円を追加して下さい．

懇親会 11 月 11 日（木）場所：グランドホクヨウ米沢 平安，会費：予約 7,500 円，当日 9,000 円．

申込先

〒980－8578 仙台市青葉区荒巻字青葉 東北大学大学院薬学研究科反応制御化学分野 坂本尚夫 TEL 022-217-6865，Fax 022-217-6864 Email：qsar@mail.pharm.tohoku.ac.jp http://www.pharm.tohoku.ac.jp/depts/hannou/sympo/index.html

第 12 回情報化学講習会

（第 22 回情報化学討論会に引き続き同じ会場で開催）

主催 日本化学会情報化学部会 (第 22 回情報化学討論会に併催)

会期 11 月 12 日(金) 13:30～16:00

会場 グランドホクヨウ米沢 (山形県米沢市金池 2-3-7)

1. 計算化学のすすめ － 電子スペクトル計算を例にして (岡山理科大学客員教授，大阪市立大学名誉教授) 西本吉助
2. ケモメトリックス入門 (福島工業高等専門学校教授) 石山純一

参加費 第 22 回情報化学討論会参加者 無料，一般 5000 円*，情報化学部会員 2000 円，学生 無料．*情報化学部会にその場で入会されますと，4000 円(年会費 2000 円，参加費 2000 円)となります．

参加申込 当日会場でお申し込みいただきます．学生の方は身分証明書の提示をお願いいたします．

問合せ先 992-8510 山形県米沢市城南 4-3-16 山形大学工学部電子情報工学科 工藤喜弘 Tel (0238)26-3360 FAX (0238)26-3359 E-mail: cicsj@dip.yz.yamagata-u.ac.jp

第２２回情報化学討論会　発表の概要

J 01

Computer-based materials science education, is this a lesson for all of us?

○Ronald D. Sanderson　　University of Stellenbosch

J 02

tDNAの系統的検出

工藤 喜弘	くどう よしひろ	山形大工
金谷 重彦	かなや しげひこ	山形大工
○木ノ内 誠	きのうち まこと	山形大工

バクテリアを古細菌(A)、グラム陰性菌(N)、グラム陽性菌(P)、マイコプラズマ(M)という四つのクラスに分類したとき、それぞれのクラスに特有のtDNA並びがあるように見えることを利用して新しいtDNAを確立した。これにより、tDNAとしてのクローバ葉構造を形成する塩基配列を探すのにDNA塩基配列全体ではなく、ある見つかったtDNAを手掛かりにその上流、下流のみに絞ることに成功した。かくてたまたま見つかったtDNAか既存のtDNAを手掛かりに、まずそれがあるバクテリアの種類を調べ、そこでみつかって手掛かりtDNAを含んでいるものを引き出し、そのデータを参考に上流、下流に見つかる期待が大きいtDNA配列をさがす。

J 03

ヒトゲノムのエキソンとイントロンの識別

工藤 喜弘	くどう よしひろ	山形大工
金谷 重彦	かなや しげひこ	山形大工
木ノ内 誠	きのうち まこと	山形大工
○坂井 孝光	さかい たかみつ	山形大院工

mRNAのスプライシングはエキソンとイントロンの境界を中心にした数塩基配列に特徴があるとされており、イントロンがGTで始まりAGで終わるGT−AG則は最も有名であるが、DDBJのHomo sapiensのDNAデータ78425件に対して、実際に機能し構造が明確に記述されている451 5件のDNAデータの報告事例についてのイントロン5796件のうち、GT−AG則に従わないすなわち、GT−AG則以外の規則で説明すべきイントロンは162件見つかった。このうち109件はGT−AG則のい適用も可能であること、残りのうち22件はGC−AG則に従うと考えればよいということ、などがわかったが、最後の31件にはそれ以外の規則を見つける必要がありそうであるという結論を得た。

J 04

情報化学システム CGS の開発(8)
化学反応速度論的微分方程式系の符号化と速度論的シミュレータ

○内野 正弘　ウチノ マサヒロ　日本情報化学システムズ

学反応速度論的微分方程式系の符号化は、化学反応速度論的微分方程式系に関する知見の整理とデータベース化、並びに化学反応系のシミュレータの構築などにとって極めて重要である。微分方程式系、通常の化学者表現、グラフ表現、数値計算システム、化学的作業仮説としての簡単化、速度論的データ解析などを総合的に取り扱う理論とシステムについて報告する。

J 05

JCICS の分析を中心とした情報化学の計量文献学的研究

○小野寺 夏生　おのでら なつお　図書館情報大

文献データベースを用いて、情報化学分野の論文を計量文献学的に分析した。まず、Chemical Abstracts (CA)における情報化学に関する Subsection に収録された記事の掲載雑誌分布を、過去30年にわたり調査した。その結果、全期間を通して Journal of Chemical Information and Computer Science (JCICS)のシェアが他誌に比べ圧倒的に高いことが判ったので、JCICS の記事を中心に分析を行った。CA を用いて、JCICS の記事に多く付与された分類(Subsections)と統制キーワード(General Subject Index Terms)が、時代とともにどう変遷しているかを分析した。また、発表数の多い著者とその所属機関も調べた。次に、Science Citation Index (SCISEARCH)を用いて、JCICS の掲載記事がよく引用している雑誌、逆に JCICS の記事をよく引用している雑誌を調査した。この結果から、情報化学研究における相互依存関係を考察した。

J 06

JST における高分子データベースの開発
－IUPAC 構造基礎名の自動発生 Phase 1 －

飯室 茂　イイムロ シゲル　科学技術振興事業団研究基盤情報部
○前田 知子　マエダ チカコ　科学技術振興事業団研究基盤情報部
高枝 佳男　タカエダ ヨシオ　三菱総研
長阪匡介　ナガサカ キョウスケ　三菱総研

IUPAC 高分子命名法に準拠した構造基礎名を構成繰り返し単位の化学構造に基づいて自動発生させる機能を、一部の線状ホモポリマーを対象として開発したので報告する。(1)IUPAC 高分子命名法および有機化学命名法を高分子名称自動発生という観点から検討し、(2)名称自動発生用の命名規則を作成した上で、(3)名称発生用アルゴリズムを検討、これに基づいてプログラムを作成した。特に命名の副単位となる部分構造の認定方法、名称配列のための優先順位の適用や慣用的名称をどの程度許容するか、などが検討課題となった。当機能は JST 高分子データベースが持つ高分子用の化合物レジストリシステム(高分子辞書)の一部として開発中であるが、将来はインターネット上でデータベース利用者にも公開する予定。なお、ＪＳＴの高分子データベースは、プロトタイプシステムを無料提供中(http://kronos.tokyo.jst.go.jp:8085)。

J 07

マイクロソフト ACCESS 上の質量スペクトルデータベース II

六車 裕孝　ろくしゃ ひろたか　日本アステック
仲久保 正人　なかくぼ まさと　筑波研究学園専門学校
○前田 浩五郎　まえだ こうごろう　アイエヌジェイ

この処のＯＳのめまぐるしい変遷に対応する為と、既存のソフト資産の継用の為、同一ハードシステム内に複数ＯＳの共存可能なシステム実現に就いては既に報告した（熊本)。しかしその後 Windows の優位は動かず、その上での ACCESS がケミスト達にとってデータベース操作のユウティリティとして身近なものになる兆しが見えて来た。これに対処する為、先般来我々は NIST の約 4 万化合物の質量スペクトルデータベースをこのシステム内に移し、その操作可能性、実用性を吟味して来た。その結果、分子量、化学式、化合物名、ベースピーク等の単独のパラメータでの検索は全て 1～2 秒で、またそれらの複数の組合わせでも同様に 1～2 秒、更に化合物名では部分構造名の AND、OR 組み合わせが可能で、結果表示等も含めて充分実用性の有る事が判った。現在更に性能向上を含めて鋭意努力中。諸彦よりの助言を得たい。

J 08

化学物質のリスクアセスメントシステムの開発

○花井 荘輔　はない そうすけ　日本化学工業協会

事業者が化学物質のリスクを評価するためのツールとしての評価システムを開発した．システムは，フィジカルリスク(爆発燃焼等)，ヒト健康影響リスク，生態影響リスクを評価するそれぞれのサブシステムと，必要なら新しいリスク管理手法を検討・評価・選択するサブシステムを含む．リスクアセスメントにとって重要な暴露量評価には，作業者・消費者に対する直接暴露を評価する数理モデル式と周辺・一般住民あるいは環境への間接暴露を評価するモデル式を用意した．化学物質のリスクアセスメントは，化学物質の安全性・物理化学性・ヒトの行動パターン・気象・水文・物質管理等に関する多種多様の情報を必要とするものであり，データの検索・推算およびモデルによる動態シミュレーション等，化学情報システムの問題としても興味深い問題である．講演では，この点を強調し，今後当部会員がこの分野に積極的に関心を持ってくれることを訴えたい．

J 09

インターネット上の反応データベースシステム －3－

○中山 堯　なかやま たかし　神奈川大理
上野山 敦史　うえのやま あつし　神奈川大理
兼子 崇　かねこ たかし　神奈川大理
柳沢 一博　やなぎさわ かずひろ　神奈川大理

分子設計研究会の反応設計分科会において独自に構築して来た反応データベースの規模が，データ数にして６万件と一応の実用レベルに到達したので，その検索システムをインターネット上のクライアントサーバシステムとして開発し，一般に公開中である．公開法は，クライアントシステムを CD-R にコピーして希望者に配布する方式である．現在はユーザ認承は行っていない．ただし，バグレポートや反応データの提供など本システムの運用に対して何らかの貢献をする事を期

待している．1999 年 7 月から 2 年間はこの形態である。また，これまで構造検索のみであったが，テキスト検索の機能を実装中である．本データベースは，反応データ自体と反応データの表現が独自のものであり，他のデータベースに含まれていないデータや情報を有するという特徴を有している．さらに，本データベースをインターネットを介して更新(追加・修正)する機能を実装中である．更新についてはデータの信頼性を確保する上で考慮すべき事が多々あるので，実験を重ねながら実用システムとして実現するためのノウハウを蓄積する計画である．本システムは実用システムであって，多くのユーザに使用していただいて関心を持っていただき，反応データの提供ユーザが増加する事によって，本データベースが質的量的に充実して行く事を期待しているものである．

J 10

インターネットにおけるＳＤＢＳ－ＮＭＲ

○早水 紀久子　はやみず きくこ　物質研
矢部 篤子　やべ あつこ　物質研
柳沢 勝　やなぎさわ まさる　物質研

ＳＤＢＳは全面公開してから約２年の歳月がたったが、その間に合計３００万件以上のアクセスを得ている。アクセス数には季節変動があるが、全体として増加の傾向は顕著にある。地域別の統計から、我が国のアクセス増加率は低く”インターネット上で情報を取得する”思想に問題があることが推測される。　システムについては種々の改良を行っているがＮＭＲとＭＳではスペクトル検索ができるようにした。検索サービスについての方針と実際を報告する。またＮＭＲではスペクトル帰属をつけた化学構造式を公開しているが、ユーザから帰属についての種々のコメントがきており、それに対する我々の対応について報告したい。ホームページは www.aist.go.jp/RIODB/SDBS/ である。

J 11

第一原理計算による原子レベルデバイス設計の現状と将来展望

○川添 良幸　かわぞえ　よしゆき　東北大

J 12

光合成反応中心における電子移動に関する理論的研究

○北尾 修　きたお おさむ　物質研
青木 孝造　あおき こうぞう　物質研
建部 修見　たてべ おさむ　電総研
関口 智嗣　せきぐち さとし　電総所

植物は太陽光エネルギーを非常に効率よく化学エネルギーに変換しており、その初期過程は「光合成反応中心」において太陽光により生じた電子を決まった経路に沿って決まった方向に高速で運ぶ反応であり、この電子移動は光合成反応中心を構成する蛋白質が支配していると考えられている。我々は、理論化学・計算化学の手法で、その機構を解明する。

J 13

TS-1/H2O2/PhOH 酸化反応活性種生成におよぼす溶媒構造の効果

○後口 隆　　あとぐち たかし　　宇部興産
八尾 滋　　やお しげる　　宇部興産

Ti を格子内に含むゼオライト(TS-1)を触媒として用いると、フェノール(PhOH)酸化反応において生成するハイドロキノン(HQ)とカテコール(CL)の生成比および PhOH 転化率が溶媒により変化する。この溶媒の効果は、単に比誘電率によるものではないことが、実験的に確認されている。我々は各種溶媒を用いた実験とともに、想定される反応機構中間体に対する密度汎関数法による検討を行っている。先の計算化学討論会では水およびメタノールが、Ti 活性中心への過酸化水素(H2O2)分子の吸着を安定化することを報告した。今回は、さらにアルコールの構造を変化させ、エタノール、イソプロパノールとしたところ、H2O2 吸着はメタノール>エタノール>イソプロパノールの順に安定であることがわかった。この結果をもとに、実験による検討を行ったところ、HQ/CL 比はほぼ一定のまま、PhOH 転化率は吸着安定化の順と同じくメタノール>エタノール>イソプロパノールの順となった。このことは H2O2 の Ti 活性中心への吸着が全体の反応速度に影響を及ぼしており、その構造を安定化する溶媒がより効果的であることを示唆している。計算および実験の詳細に関しては当日報告する。

J 14

ケミカル・ハードネスを用いた電荷移動を伴う分子シミュレーション手法の検討

○高田 章　　たかだ　あきら　旭硝子

従来の分子動力学手法は固定電荷を用いるため構成原子の割合や構造が大きく変わったり、結合の解離や反応が起こるような、すなわち大きな電荷移動が伴う場合には適用限界がある。この問題に対する処方箋の一つとして、ケミカルポテンシャルが釣り合うように電荷の移動を許すモデル化が最近提案されている。しかしながら安定構造での議論がほとんどであり、モデルのパラメータは特定の状態について決められるため、大きく歪んだ構造や解離する場合には精度が悪いと考えられる。本研究では ab initio 法による電荷とエネルギーの解析結果をもとに安定構造から解離する場合まで応用できるような手法を検討する。解析対象としては水分子を選んだ。今回の検討結果により、原子種ごとに電気陰性度やケミカルハードネスを決定するだけでは不十分であり、結合状態等の環境の影響も取り込んだモデルが必要であることがわかった。

J 15

QM/MM-vib 法に基づいた ab initio MD 法の開発と水溶液内反応への応用

○相田 美砂子　　あいだ みさこ　　広島大院理
山高 博　　やまたか ひろし　　大阪大産業科学研
Michel DUPUIS　ミッシェル デュピュイ　Pacific Northwest National Lab.

化学反応をシミュレートするためには、経験的なポテンシャル関数を使用しないで分子動力学法計算の各時間ステップにおいて非経験的分子軌道法計算を実行することが必要である。私達は、絶対零度ではないある温度において運動エネルギーを持った状態において化学反応がどのように進行していくのかについて理解を深めることを目的として、この手法を発展させてきた。溶液内反応においては溶媒分子が重要な役割を果たしている場合が多い。しかし量子化学的取り扱いを

すべての溶媒分子に対して行うことはできない。そこで反応に直接関与している溶質および溶媒分子を QM 的に取り扱い、その周辺の多数の溶媒分子を MM 的に考慮に入れた分子動力学法計算を行う手法を発展させている。MM 部分には分子の構造および分極による影響を表現するポテンシャルを組み込んだ。

J 16

ナフタレン陽イオンを触媒とする水素分子生成反応

○佐住 貴顕	とこすみ たかあき	静岡大理
萩原 紀子	はぎわら のりこ	静岡大理
寺田 典弘	てらだ よしひろ	静岡大理
石田 俊正	いしだ としまさ	静岡大工
相原 惇一	あいはら じゅんいち	静岡大理

極低温・高真空の宇宙では、2つの水素原子が出会っても容易に水素分子が生成することはない。我々は数年来、宇宙に広く分布する多環式芳香族炭化水素の陽イオンが触媒となって、2個の水素原子から水素分子が生成する可能性を検討している。今回は、現実的な芳香族炭化水素分子としてナフタレンを選び、その陽イオンラジカルの1位または2位に水素原子が付加し、それをもう1つの水素原子が水素分子として引き抜く反応の経路を調べた。密度汎関数法を用いた分子軌道計算によると、水素分子生成の段階で、小さな活性化エネルギーを伴う遷移状態が認められたが、零点振動の補正をすると消滅した。以前に、ベンゼン陽イオンを触媒として活性化エネルギーなしに水素分子が生成することを見出しており、この種の反応は、宇宙における水素分子生成反応の有力な候補であると考えられる。中性の芳香族炭化水素を触媒とする反応では活性化エネルギーを必要とする。

J 17

リチウム含有遷移金属酸化物の構造特性と電子状態に関する計算化学的検討

○鈴木 研	すずき けん	東北大院工
小野津 崇之	おのづ たかゆき	東北大院工
高見 誠一	たかみ せいいち	東北大院工
久保 百司	くぼ ももじ	東北大院工
宮本 明	みやもと あきら	東北大院工

$LiCoO_2$ や $LiNiO_2$ をはじめとするリチウム含有遷移金属酸化物は、近年リチウム二次電池の正極活物質として注目を集めている。さらに最近では $LiNiO_2$ がメタンの酸化カップリング反応に対し、高い活性を持つことが確認されている。そこで本研究では MD 法や量子化学計算といった計算化学的手法を駆使し、層状リチウム金属複合酸化物（$LiCoO_2$、$LiNiO_2$）の構造特性や電子状態の検討を行った。$LiNiO_2$ の電子状態と構造の関係を密度汎関数法により検討を行ったところ、計算により得られた酸素とニッケルの八面体構造は長短二種の Ni-O 距離をもつ構造であった。長い部分で 2.12Å、短い部分では 1.93Åとなり、長短の Ni-O 距離で約 0.2Åほど差のある歪んだ 4+2 構造であった。これは $LiNiO_2$ 中の Ni の電子配置は低スピン状態の Ni^{3+}であり、t_{2g} 軌道よりエネルギー準位の高いσ型の反結合性軌道である e_g 軌道に一つ電子が入るため、電子格子相互作用に基づく構造不安定性（ヤーン・テーラー効果）をもつためと考えられる。

J 18

ダイヤモンド膜の表面反応に関する計算化学的検討

○田村宏之	たむら ひろゆき	東北大院工
周 慧	しゅう けい	東北大院工
高見 誠一	たかみ せいいち	東北大院工
久保 百司	くぼ ももじ	東北大院工
宮本 明	みやもと あきら	東北大院工
蒲生西谷 美香	がもうにしたに みか	無機材研
安藤 寿浩	あんどう よしひろ	無機材研

ダイヤモンドやc-BNは耐摩耗性の高い硬質膜として注目されている。ダイヤモンド薄膜はCVDによって形成されるが、表面の平滑化を高めるために、その構造および化学反応の機構に興味がもたれている。また、不純物をドープすることによる半導体としての特性、およびドーピングのメカニズムに関心が集まっている。本研究では、ダイヤモンド表面の構造と水素および酸素との反応機構、および不純物のドーピングのメカニズムを量子化学計算で検討した。

J 19

量子分子動力学プログラム'colors'の開殻系への拡張

○高見 誠一	たかみ せいいち	東北大院工
山田 有場	やまだ あるば	東北大院工
田村 宏之	たむら ひろゆき	東北大院工
久保 百司	くぼも もじ	東北大院工
宮本 明	みやもと あきら	東北大院工
今村 詮	いまむら あきら	広島国際学院大工

我々は従来の第一原理分子動力学法に比べより高速な、量子分子動力学プログラム'colors'を開発してきた。このプログラムは、拡張Huckel近似に基礎を置き、分子動力学の1ステップごとに電子状態を解いて原子間にはたらく力を算出するので、計算は高速でありかつ電子の授受を伴う反応も正確に記述できるという特徴を持つ。これまでのプログラムは閉殻系のみに適用できるものであったが、化学反応を記述する際には、反応性の高い化学種である、ラジカル種を考慮することが必須であるため、本講演ではプログラムの開殻系への拡張を試みた結果について報告する。開殻系を考慮するため、拡張Huckel近似におけるαとβ電子のクーロン積分の差を、開殻系でのFock演算子のαとβ電子の差に相当するエネルギーに等しいとおいた。発表では本プログラムをさまざまな反応系について適用し、その有効性を評価した結果について紹介する。

J 20

ケモメトリックスが拓く設計・予測の世界

○船津 公人	ふなつ きみと	豊橋技科大

ケモメトリックス手法は、実験や計算を通して得られる情報から、設計・予測のためのモデル構築する一般的な方法として定着し始めている。材料設計、分子設計のために、ケモメトリックス手法がどのような使われ方をされているのかを実例を通して解説する。

J 21

3次元的分子数値化法FRAUの特性値を識別子とした試薬の自動分類と機能予測

○佐藤 寛子　さとう ひろこ　科学技術振興事業団さきがけ研究21
船津 公人　ふなつ きみと　豊橋技科大
鷹野 景子　たかの けいこ　お茶大
中田 忠　なかた ただし　理研

コンピュータによる定量的反応予測のための知識・規則誘導の最初のステップとして自動反応分類研究を進めている．分類識別子となる特性値計算法として開発したFRAUは，仮想反応相手との静電的・立体的相互作用に基づいて3次元的に分子を特性値化する．今回は，FRAUによる特性値をもとに試薬の自動分類と機能予測を行なった．FRAUは各原子毎に特性値を与えるが，自動分類のためには種々の分子を同じ次元数で比較・分類する必要がある．そこでまず種々の試薬分子について計算したFRAU特性値のうち，試薬分子の構造と役割の類似と相違をよく識別する特性値の組み合わせで，その分子を代表させる方法を検討し，良好な結果を得た．つぎにこれら特性値の組合わせをデータセットとし，試薬分子をKohonenニューラルネットワークで分類した．本分類結果をもとに，さらに試薬－機能相関マップを自動作成した．本相関マップにより，マップの作成に用いた分子と用いなかった分子のいずれについても適切な分類・機能予測結果を得た．

J 22

ニューラルネットワークによる化学物質の安全性の予測

○田辺 和俊　たなべ かずとし　物質研
松本 高利　まつもと たかとし　物質研

化学物質の構造から安全性(毒性、生分解性)を迅速に予測する手法を開発するためにニューラルネットワークを用いる手法を検討した。3層構造のニューラルネットワークを用い、入力層には化学構造記述子として結合の数のみを入力し、出力層には毒性、生分解性の実測値を教師データとして入力して、back-propagation法でニューラルネットワークの学習を行った。leave-one-out test を行った結果、80%以上の的中率がえられ、既存の予測システムよりも高い的中率で毒性、生分解性を予測する手法を開発することができた。

J 23

X線回折を用いた地下岩石情報の可視化

○貝原 巳樹雄　かいはら みきお　一関高専
大岡 靖典　おおおか やすのり　一関高専
金田 英伯　かねた ひでのり　帝国石油

原油・天然ガスなどの地下資源を探索する方法中で最も一般的かつ普遍的なものは、地下岩石試料の孔隙率／浸透率などの物性値情報や元素組成・鉱物組成等の化学的情報を得ることである。しかし、岩石試料の分析値を求めるには、非常に労力がかかり、また、その費やした労力の割には得られる情報が少ないためか、これまでその手法について、またその結果の積極的な活用例は少ない。そこで、我々はX線回折データの活用を目的とした本研究を行うこととした。具体的には、南長岡ガス田のグリーンタフ貯留層を主にその対象として、ＸＲＤのピーク面積値から鉱物組成に関する化学的情報を得て、従来の地層区分との比較を行うこととした。本研究では、特に

化学情報の視覚的な表現方法について検討し、直感的にわかり易い、深さ方向の地層イメージを色相表示する方法（化学成分層、ケミカルレイヤー）を考案した。

J 24

有機電子論の数学的裏付けとその限界

○細矢 治夫　　ほそや はるお　　お茶大理

Robinson や Ingold が 1920 年代に提出したいわゆる有機電子論を使うと、不飽和共役系の中の π 電子の移動の範囲と度合いがかなり正しく予測できる。簡単な分子軌道法の計算もこれらの結果を多くの場合に再現することが知られている。従って現代の有機化学者も、反応の予想や反応機構の説明にこれを作業仮説として活用している。著者は最近、Coulson と Longuet-Higgins の摂動論にグラフ理論を組合せることによって、この有機電子論の数学的な裏付けに成功した（JMS,461-462,473(1999))。また、福井の superdelocalizability や Coulson 等の分極率等が、簡単な共役系で発散することも見出した（Theor.Chem.Acc.,102,293(1999))。更に、大きなベンゼン系炭化水素において、異常なトンネル電子移動の起こることも見出した。これらの最近の発見と現在解析を進めていることをもとに、表記の問題についての話題をいくつか提供し、討論を行いたい。

J 25

シュウィンガー・ダイソン形式による超重原子の理論

○松浦 弘幸　　まつうら ひろゆき　　政策研究大学院大
中野 正博　　なかの まさひろ　　産業医科大

本論では、相対論的場の理論に基づいて、有限電子密度系を記述する AtomicSchwinger-Dyson(ASD)法を展開する。相対論的平均場近似が、電子を主として解くのに対し、ASD の方法は、電子とフォトンを同等に取り扱う手法である。このことにより、相対論的平均場近似を超えた高次相関を取り上げることが出来る。この方法では、平均場は、古典場と見なされ、それを超えた量子論的振動をフォトンの Self-energy の形で取り上げる。フォトンの Self-energy は、また、物質中でのフォトン伝播とも関連している。ASD 法は、フォトンが物質中を走る時に、particle-hole 励起から受ける影響をフォトン伝播関数として表現している。多電子系に、相対論的場の理論が必要な理由は、他にもある。例えば、超重原子のような系では、陽電子生成が容易に起こり、真空からの電子ー陽電子対生成の効果を考慮しなくてはならない。この電子ー陽電子対生成も、ASD 法ではフォトン伝播関数の中に取り入れることが出来る。フォトンの物質中での伝播関数が free なものから変化すれば、当然それは電子の伝播関数にも影響を及ぼす。こうして、フォトンと電子の互いに結合した一連の非摂動的方程式が得られる。これが ASD 方程式である。以下に ASD 方程式の導出法の概略を示す。

J 26

分子力学法と分子動力学法による長鎖脂肪酸単分子膜の研究

○橋口 聡　　はしぐち あきら　　大阪電通大
藤田 岩男　　ふじた いわお　　大阪電通大

これまで飽和脂肪酸単分子膜については、アルキル鎖の傾きなど分子配向についての計算が多く

行われてきたが、不飽和脂肪酸についての計算はほとんど行われていない。不飽和脂肪酸の分子配置や傾き等の関係を明らかにするために我々は分子力学法(MM)と独立原子モデルを用いた分子動力学法(MD)の2種類のシミュレーションを用いた。力場は既存の力場のMM2力場を用い、比較検討のために飽和脂肪酸についても計算した。計算の結果は，飽和脂肪酸は占有面積を増加させたとき，分子の傾きが，ある占有面積のところで急激に増加した。この占有面積は，実験におけるπ-A曲線の折れ曲がり位置とほぼ対応する。また、不飽和脂肪酸は、trans-体では飽和脂肪酸と同様の特徴が得られたが、cis-体では、非常に秩序の低い結果となった。

J 27

ProteinDF によるタンパク質全電子計算と電子トンネリング過程

○佐藤 文俊	さとう ふみとし	九工大情報工
吉廣 保	よしひろ たもつ	九工大情報工
岡崎 功	おかざき いさお	九工大情報工
上田 尚学	うえだ たかあき	九工大情報工
保田 慎一郎	やすだ しんいちろう	九工大情報工
恵良 信	えら まこと	九工大情報工
時枝 浩司	ときえだ こうじ	九工大情報工
佐藤 洋子	さとう ようこ	九工大情報工
柏木 浩	かしわぎ ひろし	九工大情報工

本発表者らが独自に開発した Gauss 型基底関数による密度汎関数法プログラム（ProteinDF）を用いてタンパク質の全電子計算を実行した。現在、43残基のタンパク質の計算が終了した。これはトップクラスの大規模計算である。発表ではさらに大型の計算結果について報告するとともに、タンパク質内電子移動のシミュレーションについても発表する。

J 28

酸化物エレクトロニクス材料のハイオーダーエピタキシー機構の解明

○久保 百司	くぼ ももじ	東北大院工
稲葉 祐策	いなばゆうさく	東北大院工
小野津 崇之	おのづたかゆき	東北大院工
高見 誠一	たかみせいいち	東北大院工
宮本 明	みやもとあきら	東北大院工
川崎 雅司	かわさきまさし	東工大院総合理工
吉本 護	よしもとまもる	東工大応セラ研
鯉沼 秀臣	こいぬまひでおみ	科学技術振興事業団戦略研、東工大応セラ研

酸化物エレクトロニクス材料の合成過程をシミュレーションすることが可能な分子動力学計算プログラム MOMODY を開発し、酸化物薄膜のエピタキシー成長過程の検討を行った。MgO(001)表面上に MgO 薄膜を形成した場合には(001)配向の2次元エピタキシャル成長が確認されたのに対し、サファイア(0001)面上に同じく MgO 薄膜を形成した場合には3次元ナノドット構造の形成が予測された。また、このナノドットがアモルファス状ではなく結晶構造を維持しており(111)配向でエピタキシャルに成長することも明らかとなった。さらにサファイア(0001)面と MgO(111)面の格子定数比は約 8.4%もあることから MgO／サファイア(0001)界面でこの大きな格子定数比を緩和するハイオーダーエピタキシー現象が起こっているものと考えられる。そこで、この形成機構について検討したところ、MgO／サファイア(0001)界面の Al 原子が4～5配位のいフレキシブルな配位

数をとることで、ハイオーダーエピタキシーが実現されていることが明らかとなった。

J 29

シリコン酸化膜における応力特性の計算化学的検討

○山田 有場　やまだ あるば　東北大院工
黒川 仁　くろかわ ひとし　東北大院工
遠藤 明　えんどう あきら　東北大院工
高見 誠一　たかみ せいいち　東北大院工
久保 百司　くぼ ももじ　東北大院工
宮本 明　みやもと あきら　東北大院工

シリコンデバイスの微細化、高集積化に伴い、ゲート絶縁膜の薄膜化が進行し高品質な極薄絶縁膜の必要性が課せられている。近年、酸化層薄膜の微小表面応力の測定により、プラズマ酸化過程の極初期には引っ張り応力と圧縮応力が繰り返し観測された。そこで本研究では、酸化膜成長過程の応力変化を解明するために、分子力学(MM)計算を行なった。MM計算にはMSI社の開発したCerius2を用いた。Si(100)清浄表面に1～9層の酸化膜を形成し最安定構造を求めた。また、ステップを考慮したモデルについても検討を行なった。ポテンシャルはUniversalポテンシャルを用いた。周期境界条件の単位セルは、表面方向のみ緩和させた。スーパーセルのb軸(ダイマーに平行)とc軸(ダイマーに垂直)の変化から応力が異方性を持つ事が確認できる。またb軸とc軸の和の変化から始めに引っ張り応力、次に圧縮応力が確認された。

J 30

サファイア基板上におけるGaN薄膜初期成長の計算化学的検討

○小野津 崇之　おのづ たかゆき　東北大院工
稲葉 祐策　いなば ゆうさく　東北大院工
高見 誠一　たかみ せいいち　東北大院工
久保 百司　くぼも もじ　東北大院工
宮本 明　みやもと あきら　東北大院工
家近 泰　いえちか やすし　住友化学
前田 尚良　まえだ たかよし　住友化学

III-V族窒化物半導体であるGaNは、高出力の青色LEDが実現されており、室温での短波長レーザー発振の実現などから、次世代のDVD用光源としても期待されている。しかしながら、その製膜過程で結晶性に影響を与えている薄膜/基板界面での微細構造についてはまだ十分な理解は得られていない。そこで本研究では、分子動力学(MD)法と密度汎関数法を用いて、サファイア基板上のGaNの結晶成長過程に影響を与えている、ヘテロ界面の微細構造について計算化学的な観点から検討した。MD計算の結果より、N原子はAl原子と強く相互作用し、それほど大きく拡散しなかったが、Ga原子は基板表面を均一に拡散する傾向が見られた。このことより、界面ではAl-N結合が支配的であると考えられる。さらに、低温でGaN層を結晶成長させ、高温でアニーリングすることによって、GaN表面の結晶方位がどのような影響を受けるかについても検討した。

JP 01

情報化学システム CGS の開発(9)
正規 4 次グラフの生成、符号化、対称性認識およびデータベースの構築について

○内野 正弘　ウチノ マサヒロ　日本情報化学システムズ

2000 万個以上の正規 4 次グラフの生成、符号化、対称性の認識をおこなった結果について、報告する。節点等価クラス、全等価節点よりなる構造、節点照明関数の Resolution, 節点数とグラフ数の関係、データベース化する際に必要な媒体容量などについて報告する。

JP 02

二節部分グラフによる IUPAC 名の生成

○安代 容子　あじろ ようこ　京都工芸繊維大院工芸
藤田 眞作　ふじた　しんさく　京都工芸繊維大院工芸

本研究の最終目標は、ITS(虚遷移構造)に基づく反応検索システムの構築である。このシステムでは、ITS 結合表から抽出した二節部分グラフを用いることにより、反応物、生成物および反応過程に注目して検索を進めることができる。本研究は、この検索システムのサブシステムとして一般の化学反応に IUPAC 反応名を付与する方法の確立を目的とした。まず、あらかじめ反応名がわかっている化学反応を各基本反応グループに分類し、基本反応の ITS からグループ固有の二節部分グラフを決定するプログラムを作成した。次に、このようにして得た基本反応のデータと反応名未知の反応から得た二節部分グラフとを照合することにより、IUPAC 反応名を付与するプログラムを作成した。その結果、二節部分グラフを用いて一般反応を分類、検索する方法の妥当性と、その限界を明らかにした。

JP 03

XyMML (XyM Markup Language)---構造式のコミュニケーションのためのマークアップ言語

○藤田 眞作　ふじた しんさく　京都工芸繊維大工芸

インターネットや電子出版の普及にともなって、化学文書の流通があらためて問題になっている。とくに、構造式のコミュニケーションに関しては、まだ主流になるものは提案されていない。筆者は、すでに構造式のタイプセッティングシステムとして XyMTeX を公開しているが、この命令体系をさらに拡張して、XyM 記法(XyM Notation)を提案している。今回は、次期のネットワーク言語と目されている XML に適用することを念頭において、構造式のコミュニケーションのためのマークアップ言語として、XyM Markup Language (XyMML) を提案する。さらに、SGML 用の DTD および LaTeX トランスレーター用のフィルタープログラムを作成し、XyMML から出発して、版下用のプリントアウトを行うシステムを作成したのであわせて報告する。

JP 04

有機合成設計システムＡＩＰＨＯＳにおける戦略部位の優先順位付け手法の開発

○田中 章夫　たなか あきお　住友化学工業
船津 公人　ふなつ きみと　豊橋技科大

本発表では、LHASA 以降の多くのシステムの課題であるルートの拡散化を解消すべく、われわれがこれまで開発してきた知識指向型合成設計プログラム AIPHOS (Artificial Intelligence for Plannning and Handling OrganicSynthesis) についてターゲットや中間体から認識した戦略部位の優先順位付けについて検討した結果について述べる。

ルートの拡散化（１ステップで１０通りの提案を行うと、５ステップでのルート総数が原理的には１０の５乗となる状況）については、今日まで、解消するための有効な手法があまり開発されていなかった。そこで、今回我々は標的化合物や中間体から認識した戦略部位の優先順位付けについて検討した。客観的、定量的に取り扱うために、既存の反応データベースからの評価式の誘導を検討した。具体的には、反応データベースの生成系の構造全体と、優先順位付けの対象となる戦略部位をパラメータ化し、統計解析したものを評価式とした。ここから得られた各戦略部位の評価値で、有用なルートから提案する可能性を示すことが出来た。(戦略部位：合成前駆体構造を導く際の標的構造中の着目部位)

JP 05

有機合成設計システム AIPHOS-KOSP の開発（２）

○佐藤 耕司　さとう こうじ　第一製薬
船津 公人　ふなつ きみと　豊橋技科大

合成設計システム KOSP(Knowledge base Oriented system for Synthesis lanning)は市販反応データベースより自動誘導された知識ベースを利用した新しいタイプの経験型システムである．昨年の情報化学討論会において，その開発コンセプトおよび動作概要を報告したが，一段階の逆合成ルート提案しか行なえなかった．今回，convergent synthesis の概念をシステムに導入し，多段階逆合成ルート提案を可能としたので報告する．また，GUI も向上させたので併せて報告を行なう．

JP 06

化学物質生体影響データベース

○志村 和樹　しむら かずき　科学技術振興事業団

科学技術振興事業団（ＪＳＴ）では、これまでＪＳＴが蓄積してきた化合物辞書、熱物性、質量スペクトル、結晶構造といった基礎的化合物データと、国立医薬品食品衛生研究所、国立環境研究所、工業技術院生命工学工業技術研究所、理化学研究所のデータベースとをインターネット上で結びつけたネットワーク分散型化合物データベースの実験システムを開発し公開を開始した。それぞれのデータベースは各機関で独自に開発されているものであるが、化学物質の生体影響に関する研究等の多面的なデータ検索のニーズに対応できるよう、化合物辞書番号をリンクキーとすることにより統合的に利用できるようにした。

利用者はインターネット上でＪＳＴのWWWサーバに接続することにより化合物名称、化学式、物性値、スペクトルピーク等の条件から化合物情報を検索することができる。そして、その検索

結果を、各機関のデータベースのデータに関連させることが可能となっている。

JP 07

化学物質生分解性予測システムの開発

○廣松 康一　ひろまつ こういち　化学品検査協会、大阪大
増本 茂和　ますもと しげかず　富士通九州システムエンジニアリング
矢可部 芳州　やかべ よしくに　化学品検査協会
西原 力　にしはら つとむ　大阪大

インターネット上で利用可能な化学物質の好気的微生物分解（生分解）性のデータベースとこれを活用した生分解性予測システムを開発した。本システムは、自動的にダウンロードされる JAVA Applet を用いて構造式を入力することにより、「化学物質の審査及び規制に関する法律(化審法)」の微生物による分解度試験データベース中の構造を検索（部分構造または完全一致）することができる。データベースに検索対象の化合物データが存在しない場合、入力された構造式から 106 種類の部分構造と 19 種類の原子数や環数などのトポロジカルな情報及び分子容や分子表面積・対称性等が自動的に計算され、経験則を体系化した予測フローチャートまたは置換基寄与法を用いた構造活性相関式から生分解性を予測してその結果を出力する。

JP 08

生物情報統合データベースシステムの構築
ー新しい発想に基づいた生物情報管理法ー

○福田 美紀　ふくだ みき　医薬分子設計研究所
瀧川 有紀　たきがわ ゆき　医薬分子設計研究所
豊田 哲郎　とよだ てつろう　医薬分子設計研究所
板井 昭子　いたい あきこ　医薬分子設計研究所

ゲノム解析が進む中、大量に排出される実験データの効率的な管理法・解析法の開発が急がれている。しかし現状では世界中のウェブサイトに各種データが無秩序に氾濫しており、どこにどのようなデータが公開されているのかを知ることさえ難しい状況にある。さらにそれらの生物情報同士の関連付けは各ページ間にはられたリンク情報に頼るしかなく、様々なデータを統合して扱うのは大変難しい。このようにデータの相互関係の把握が困難であるのは、生物情報の核であるアミノ酸や核酸の配列情報を一意に表すことのできるもの、人間でいえば「名前」にあたるものが存在していないためであると考えられる。そこで我々はこの点に着目し、配列情報に「名前」をつける方法を考案した。この方法を用いれば、散在している情報から容易に必要な情報のみを収集することができるようになる。今回は本法を利用したデータベース構築などについても紹介する。

JP 09

錯体・有機金属化合物の双極子モーメントのデータベース作成（続報）

○山崎 昶　やまざき あきら　日赤看護大

以前より継続して集積中の、金属錯体と有機金属化合物、および関連する配位子やイオン対などの双極子モーメントの実測値のデータを、関連の研究者の利用が容易となることを意図して、パ

ーソナルコンピュータ可読の形に整理してまとめた。現在までの所およそ三千レコードほどが集積された。多くの大権威の述べられるとおり、既報のデータ集や便覧の類にこの種の貴重・かつ有益な情報が含まれていることは確かであるが、あまりにも多くの不要なものの中から必要とされるものを手作業で探索するのは事実上不可能に近いから、権威ある成書や総説などを基礎としてデータを集め、利用者の便を考えて索引用フィールドの付加や化学式の誤りの修正などを行った。部分的なデータベースファイルを何人かの錯体化学の研究者に依頼して種々のコメントを頂戴し、以後のさらなる改善のための資料としている。

JP 10

立体化学の規範的コード化法の開発とコンホマーの識別

○越野 広雪　こしの ひろゆき　理研
佐藤 寛子　さとう ひろこ　科学技術振興事業団さきがけ研究 21
船津 公人　ふなつ きみと　豊橋技科大
鵜澤 洵　うざわ じゅん　理研
中田 忠　なかた ただし　理研

立体化学の情報は分子構造を取扱う上で重要な化学情報である．特に反応設計・予測や NMR 化学シフト予測等においては，特定の予測部位の周りの立体化学を含めた環境の類似と相違を識別することが必要である．これまでコンピュータによる分子の命名法などでは分子全体の立体化学を識別する方法は報告されているが，相対的な位置関係を一義的に表すことの難しさから，立体化学を反映させた NMR の化学シフト予測等に適う，特定部位からの立体構造環境情報の規範的コード化法はいまだ開発されていない．そこで，立体化学情報を規範的にコード化する新しい手法を開発した．今回はまず，コード化の方法と，本手法によりコード化した種々の有機化合物をデータベース化し，コンホメーションの類似と相違を識別した結果について述べる．

JP 11

Circular permutationを取り扱うことのできる構造局所アラインメント法

○廣池 隆明　ひろいけ たかあき　生物分子工学研
藤 博幸　とう ひろゆき　生物分子工学研

蛋白質の局所の立体構造を比較してアラインメントを行う、構造局所アラインメント法は、ダブル・ダイナミック・プログラミングを用いて、１９９３年にC.A.Orengo とW.R.Taylorにより開発されている。 しかし、この方法では、Circular permutationの関係にある蛋白質のアラインメントが良好には行われない問題があった。

そこで、我々は、蛋白質どうしの構造を比較をする為の structural nvironment を Circular permutation に影響されない様に再構築して、構造局所アラインメントを行う方法を開発した。我々の方法により、通常のCircular permutationの関係にない蛋白質はもちろんのこと、Circular permutationの関係にある蛋白質のアラインメントも良好に行うことが出来る様になった。

JP 12

二次元相関分光法による昇温脱離分析データの解析

○杉田 記男　すぎた のりお　豊橋技科大
宮林 延良　みやばやし のぶよし　電子科学
阿部 英次　あべ ひでつぐ　豊橋技科大

昇温脱離分析法(Thermal Desorption Spectroscopy)は検体を昇温加熱し、放出される脱離ガスを質量分析する分析法である。この手法は検体に含まれる極微量の不純物の検出に威力を発揮する。しかし、複数の不純物が同じ温度(あるいは近い温度)で脱離ピークを生ずる場合、これらを別々の物質として認識することは困難である。
本研究ではこの問題の解決に二次元相関分光法を用いることを試み、成果を得たので報告する。

JP 13

量子化学計算によるリチウムイオン電池材料の設計

○田辺 和俊　たなべ かずとし　物質研
松本 高利　まつもと たかとし　物質研
長嶋 雲兵　ながしま うんぺい　融合研

リチウムイオン二次電池は携帯電話やノートパソコンなどに使用されているが、小型化、大容量化などの機能向上が求められている。そこでリチウムイオン二次電池の機能向上を目指して、負極材の炭素材料へのリチウムイオンの吸蔵の構造や機構を解明するために量子化学計算による解析を行った。リチウムイオンーグラファイトのモデル系について、プログラムQ-Chemを用いて非経験的分子軌道計算を行い、モデルサイズの影響などを考察した。その結果、モデルサイズに関してきわめて興味深い結果がえられ、炭素電極の設計に有用な指針を得ることができた。

JP 14

ライノウィルス外殻蛋白質の７００ピコ秒の分子動力学計算

○米田 茂隆　よねだ しげたか　北里大理
米田 照代　よねだ てるよ　北里大薬
梅山 秀明　うめやま ひであき　北里大薬

風邪の主要な病原であるライノウィルスの外殻蛋白質は外径３００Åという巨大な球殻状の分子集合体であり、単に内部に遺伝子を保護、運搬するだけでなく、ライノウィルスの感染、増殖に関して精巧な機能をもつ。我々は、特に宿主細胞への吸着と脱殻のメカニズムの解明を目的として、我々が開発した回転対称性境界条件という独自の方法により、ライノウィルスの外殻蛋白質の７００ピコ秒の分子動力学計算を行った。このような巨大分子集合体のこのような長時間計算は少ない。その分子動力学計算の原子運動の軌跡をもとに、従来のような個々の蛋白質構造の変化よりも、むしろ、外殻蛋白質中で各蛋白質がどのように並進、回転の運動をしているかに焦点をあて、その集団運動の解析を行った。その結果より、ライノウィルス外殻蛋白質が脱殻を開始する部分を議論した。

JP 15

トライボロジーに関する計算化学的研究

○周 慧	しゅう けい	東北大院工
田村 宏之	たむら ひろゆき	東北大院工
高見 誠一	たかみ せいいち	東北大院工
久保 百司	くぼ ももじ	東北大院工
宮本 明	みやもと あきら	東北大院工

機械や装置の小型化が急速に進むにつれ、原子、分子レベルの摩擦特性の解明が求められている。しかし、トライボロジー材料の摩擦特性は、分子構造、分子量、温度、滑り速度、負荷などに大きく依存し、一般的な法則性は導かれていない。そこで、本研究ではトライボロジー材料の中でも、最も重要な潤滑剤分子の設計を目指して、分子動力学法を活用し様々な分子量、分子構造を有する潤滑剤分子のせん断場におけるダイナミックスを検討した。また、油性向上剤及びさび止め添加剤などの金属表面への吸着特性を向上させるためには、金属新生面と潤滑剤の相互作用についても検討が必要である。ここでは、密度汎関数理論（DFT）に基づく第一原理量子化学計算を活用し、金属新生面上での数多くの有機化合物の吸着エネルギーを検討した。本研究により、実験的には困難であった潤滑油分子の分子構造とそのダイナミックスさらには吸着特性などの関係を明らかにすることができた。

JP 16

Dual Ensemble Molecular Dynamics (DEMD) 手法の開発とゼオライト膜による気体分離への応用

○小林 泰則	こばやし やすのり	東北大院工
森戸 英明	もりと ひであき	東北大院工
水上 浩一	みずかみ こういち	東北大院工
高見 誠一	たかみ せいいち	東北大院工
久保 百司	くぼ ももじ	東北大院工
宮本 明	みやもと あきら	東北大院工

無機多孔質膜は有機膜に比べて熱的安定性や化学的な安定性が高いため、発電施設などから排出される高温の排ガスからの直接二酸化炭素分離への応用が期待されている。分子シミュレーション手法は膜分離プロセスでの気体の吸着・拡散機構に関する知見が得られるため、非常に有用である。近年では、グランドカノニカル分子動力学法(GCMD)と呼ばれる新しい分子動力学法(MD)に関する研究が行われている。GCMD 法は定常的な圧力勾配下での気体透過などの従来の NVT アンサンブル MD では取り扱うことができない非平衡過程を取り扱うことができる手法であり、透過実験から得られる膜透過係数を直接算出できるという利点がある。我々は圧力勾配を駆動力とした気体の透過シミュレーションなどの非平衡 MD が行える Dual Ensemble MD 手法を開発し、ゼオライト膜による二酸化炭素分離シミュレーションを行った。

JP 17

タンパク質内電子移動グラフィックス

○柏木 浩　かしわぎ ひろし　九工大情報工
佐藤 文俊　さとう ふみとし　九工大情報工
吉廣 保　よしひろ たもつ　九工大情報工
岡崎 功　おかざき いさお　九工大情報工
上田 尚学　うえだ たかあき　九工大情報工
保田 慎一郎　やすだ しんいちろう　九工大情報工
恵良 信　えら まこと　九工大情報工
時枝 浩司　ときえだ こうじ　九工大情報工
佐藤 洋子　さとう ようこ　九工大情報工

本発表者らは密度汎関数法による大規模分子軌道計算プログラム（ProteinDF）を開発し、これを高効率に並列化した。本プログラムを光合成反応中心であるクロロフィルダイマーのスピン密度分布の計算やヘムタンパク質の電子状態計算に応用した。

JP 18

タンパク質全電子計算プログラム ProteinDF におけるオブジェクト指向技術と分散並列処理

○吉廣 保　よしひろ たもつ　九工大情報工
佐藤 文俊　さとう ふみとし　九工大情報工
柏木 浩　かしわぎ ひろし　九工大情報工

発表者らはタンパク質の性質や反応性を電子レベルで理解することを目的に、ガウス型基底関数を用いた密度汎関数法に基づく分子軌道法プログラム ProteinDF を開発・発展させてきた。これまでの計算で、我々の興味の対象であるタンパク質（数千軌道以上）の全電子計算へ適用するならば、10～100G FLOPS 程度の計算能力が必要となることがわかっている。この計算能力を得るために、最近のワークステーション（WS）の高性能化、ネットワークの高速化を理由に、WS クラスタによる分散並列処理システムへの適用を考えて ProteinDF の並列化を行ってきた。本研究では、オブジェクト指向言語 C++を使用し並列化を行ったプログラムを、ワークステーションクラスタ上で実行させ数十残基のタンパク質の全電子計算を行った。講演では並列化のために用いたオブジェクト指向的な技法とタンパク質の全電子計算の結果について報告する予定である。

JP 19

III-V 族窒化物系半導体の結晶成長過程に関する計算化学的検討

○稲葉 祐策　いなば ゆうさく　東北大院工
小野津 崇之　おのづ たかゆき　東北大院工
高見 誠一　たかみ せいいち　東北大院工
久保 百司　くぼも もじ　東北大院工
宮本 明　みやもと あきら　東北大院工

窒化物系半導体による青色発光ダイオードやＬＥＤでは、活性層として InGaN が用いられ、可視光から紫外光までの波長をカバーする発光素子材料として注目されている。しかし，サファイア基板上に成長した InGaN 量子井戸構造では In 組成の不均一化や結晶欠陥生成など、その成長過程

に関して不明な点が多い。そこで本研究ではIII-V族GaN系結晶の成長過程に関して計算化学的手法を用いて検討を行った。周期的DFT量子化学計算にはMSI社のCASTEPプログラムを用いた。擬ポテンシャルおよび平面波基底関数を用い、交換相関相互作用についてはPerdewとWangが開発したものを用いた。InGaN混晶のモデルを作成し、その最安定構造とトータルエネルギーを計算する事によってInGaNの結晶性について検討を行った。

JP 20

ヘプタフルバレンにおける擬Jahn-Teller効果のMCSCF法による解析

○豊田 東雄　　とよた あずまお　　山形大教育
小関 史朗　　こせき しろう　　三重大工

非交互炭化水素に付随する特異な化学的性質は、擬Jahn-Teller効果の観点から興味が持たれ、理論・実験の両面から研究されてきた。 最近、私たちは共役分子を対象にして、非対称平面構造の安定化の起源についてab initio RHFとMCSCF MO法を用いて検討した。 既知の非交互炭化水素の中で、ヘプタレンやヘプタフルバレンの分子構造は、非平面でしかも周辺炭素骨格に二重結合が局在化していることがNMRスペクトルやX線結晶構造解析によって明らかにされている。本研究では、擬Jahn-Teller効果の物理的描像を明らかにするため、つまり、非平面構造の安定化の本質を考察するため、ヘプタフルバレンを取り上げ、平面構造から非平面構造への分子対称性の低下に寄与する因子を解明する。 具体的には、平面な対称構造(D2h)および非平面なD2, C2v, そしてC2h構造の幾何学的なパラメーターをab initio MCSCF MO (6-31G(d)) 法を用いて最適化し、それぞれの核配置における全エネルギーを、エネルギー分割法に基づいて各成分に分割して解析する。

JP 21

第一原理計算による新規メタノール合成触媒の理論設計

○坂原 悟　　さかはら さとし　　東北大院工
遠藤 明　　えんどう あきら　　東北大院工
高見 誠一　　たかみ せいいち　　東北大院工
Belosludov Rodion　　ぺろするどふ ろでぃおん　　東北大院工
Zhanpeisov Nourbosyn　　ざんぺいそふ のるぼしん　　東北大院工
久保 百司　　くぼ ももじ　　東北大院工
宮本 明　　みやもと あきら　　東北大院工

現在、エネルギー利用の高効率化と低環境負荷の観点から高品位燃料の合成が期待されている。高品位輸送燃料の一つであるジメチルエーテルは、低温でのメタノール合成が早急な課題である。そこで本研究では著者らが提唱しているコンビナトリアル計算化学を活用し、新しい触媒材料の設計を試みた。その結果、メタノール合成用触媒として実験的に工業化されているCu^+を活性中心とする触媒が計算化学を用いても有効であることが確認され、さらにAg^+、Au^+がCu^+より高い活性を持つ可能性があることが示唆された。さらに本研究では別の反応経路、担体効果に関する知見を明らかにした。

JP 22

遺伝的アルゴリズムを用いる１本鎖コラゲンモデルの安定構造探索

○髙橋 紘一	たかはし こういち	近畿大理工
田中 克裕	たなか かつひろ	近畿大理工

ペプチドの安定な立体構造をアミノ酸残基の配列情報から物理的に決定することは，蛋白質の立体構造予測のための伝統的なアプローチである．昨年 Kramer らにより human type III コラゲンに含まれる合成モデルペプチドの X 線結晶構造解析が行われ，コラゲンの 3 本鎖ヘリックス構造が確定された．各鎖は Pro-Hyp-Gly の繰り返しからなるゆるいヘリックス構造をとり，鎖間に水素結合がある．

そこで，大域的安定解探索法である遺伝的アルゴリズムを用いて，1 本鎖コラゲンモデルの安定なコンフォメーションを求めた．アルゴリズムは Mathematica で記述し，適応度関数として半経験的 MO 計算(WinMopac)を使い，最終的には ab initio MO 法で安定構造を求めた．求めた安定構造は X 線構造の 3 本鎖中の構造と同じであり，コラゲンの分子構造が本質的に，繰り返し単位 Pro-Hyp-Gly に依存するものであることがわかった．

JP 23

超高速分子軌道計算専用機 MOEngine の開発

○髙島 一	たかしま はじめ	大正製薬創薬研
北村 一泰	きたむら くにひろ	大正製薬創薬研
小原 繁	おばら しげる	北海道教育大教育
網崎 孝志	あみさき たかし	島根大総合理工
村上 一彰	むらかみ かずあき	九州大学院システム情報
早川 潔	はやかわ きよし	電総研
山田 想	やまだ そう	富士ゼロックス総合研究所
稲畑 深二郎	いなばた しんじろう	富士ゼロックス総合研究所
宮川 宣明	みやかわ のぶあき	富士ゼロックス総合研究所
田辺 和俊	たなべ かずとし	物質研
長嶋 雲兵	ながしま うんぺい	融合研

Hartree-Fock 計算において計算時間の殆どを占めるのが二電子積分の計算である。この計算は基底関数の４乗に比例する演算量を必要とするため、現実に即した巨大分子系の非経験的分子軌道計算を手軽に実行できない。そこで、現在の数 10 倍～100 倍程度のコストパフォーマンスをもちワークステーションや PC に簡単に接続できる形態の専用並列計算機 MOEngine の開発を試みたので報告する。MOEngine は、専用 LSI を搭載した複数のボードを含む筐体で構成され、二電子積分及びフォック行列の計算を高速に実行する。積分計算にはデータ並列性の高い小原の方法を採用し、漸化展開で用いられる積和演算をパイプライン処理し高速計算を行なう。また、我々は、(専用計算機のような) 並列各プロセッサのメモリ量が小さいの場合にも大次元のフォック行列を計算できるような新規並列フォック行列作成アルゴリズムを考案したのでこれについても報告する。

JP 24

分子の電子状態計算プログラムHAM/3の修正

○藤本 徳之　ふじもと のりゆき　信州大繊維
成田 進　なりた すすむ　信州大繊維
渋谷 泰一　しぶや たいいち　信州大繊維

分子の電子状態計算プログラムHAM/3はLindoholmらによって開発された半経験的計算プログラムの一つである。それは、分子の軌道エネルギーや生成熱、イオン化エネルギー、励起エネルギーなどの色々な物理量を分子の原子座標と原子の種類を元にして計算するプログラムである。しかし、対称性を持つ分子の計算結果に不具合を生じるという欠点を持つ。このことから、長年このプログラムは対称性を持たない分子にしか使われてこなかった。当然の事ながら、対称性の優れた分子であるフラーレンC60に対しての計算結果は、その対称性が若干崩れて出る。本研究では、対称性に関係するトラブルの原因を突き止め、そのプログラム修正を試みた。

JP 25

Ab initioペア近似法による巨大分子計算プログラムの開発

○中野 達也　なかの たつや　国立衛研
神沼 二眞　かみぬま つぐちか　国立衛研
上林 正巳　うえばやし まさみ　生工研
佐藤 智之　さとう としゆき　富士総研
秋山 泰　あきやま ゆたか　RWCP
北浦 和夫　きたうら かずお　阪府大

生体高分子のような巨大分子への適用を目的として、北浦らが提唱したab initioペア近似法に基づいた巨大分子計算プログラム ABINIT-MP の開発を行った。

JP 26

ウレタン化反応のアミン触媒機構の解析

○山下 修　やました おさむ　花王
小澤 忠広　おざわ ただひろ　花王

ウレタン樹脂製造触媒反応の量子化学による解析の報告です。イソシアネートとOH基とのウレタン化反応を無触媒とアミン触媒の両機構を非経験的分子軌道法（Gaussian）で解析した。無触媒系では構造的にタイトな4中心の遷移状態を経由する。一方、アミン触媒が存在すると複数の遷移状態をもつ逐次反応で進行することにより活性化エネルギーが低下している事が分かった。これらの遷移状態の形態は、アミンにより活性化されたイソシアネートがOH基からプロトン移動、生成したアルコキサイドのカルボニル炭素の攻撃、触媒アミンの脱離、と特徴付けられた。この機構からアミンの塩基性度の最適値の存在が理論的予測され、その後の実験データの整理により確認できた。

JP 27

C-13 化学シフトテンソルを利用したβ-ジケトンエノールの分子構造解析

○今城 文雄　　いましろ ふみお　　京大院理

分子内水素結合型のβ－ジケトンエノールでは速い水素移動の存在が考えられている。溶液状態では平均化されて対称構造となっているものでも、結晶状態では非対称構造となっている場合が多い。これらの状態でも水素移動は存在するはずであるから構造の平均化が存在すると考えられる。1,1,2,2-テトラアセチルエタンも結晶状態では非対称型分子内水素結合をもつエノールとして存在している。そこでモデル化合物を用いて ab initio 計算により分子構造と C-13 化学シフトテンソルを計算した結果、35% 存在する反転型エノールとの間での速いエノール－エノール異性化を仮定すれば実測値を説明することができることがわかった。また、中性子線回折による分子構造もこの結果と矛盾しない。

JP 28

フラーレン分子の局所的反応性と HOMO-LUMO エネルギー間隔の関係

○相原 惇一　　あいはら じゅんいち　　静岡大理

一般に、HOMO-LUMO エネルギー間隔が小さい分子ほど反応性は大きい。HOMO-LUMO エネルギー間隔が小さいということは、HOMO のエネルギーが大きいか、LUMO のエネルギーが小さいか、またはその両方であることを意味し、フロンティア軌道理論からも大きな反応性が予想される。我々は最近、炭素数 90 までのフラーレンのすべての i solated pentagon 異性体について結合共鳴エネルギーの最小値（min BRE）と T 値を計算し、両者の間によい相関があることを見出した。min BRE は、トポロジー的共鳴エネルギーに対する個々のπ結合の寄与の最小値であり、そのπ結合付近での反応性と関係している。T 値は HOMO-LUMO エネルギー間隔に共役原子の数を掛けたものである。このことから、フラーレンでは、反応性の大きな部分構造があると HOMO-LU MO エネルギー間隔が小さくなる傾向があるといえる。

JP 29

分子動力学法によるスチレン－エチレン共重合体のガラス転移温度計算

○戸木田 裕一　　ときた ゆういち　　電気化学工業中央研究所
鈴木 茂　　すずき しげる　　電気化学工業中央研究所
岡本 彰夫　　おかもと あきお　　電気化学工業中央研究所
桑島 聖　　くわじま さとる　　ナノシミュレーション

スチレン－エチレンランダム共重合体（St－mol%:0,10,30,50,100）のガラス転移温度（Tg）を分子動力学法により計算を行った。スチレンの含量が増加するに従って、Tg が増加するという実験の傾向を再現することができた。絶対値でも、最大で約６０K のずれにとどまった。また、シミュレーション前の初期構造依存性についても検討を行った。その結果、同じ１次構造でも、発生させる初期構造により無視できない程度の計算値のばらつきが生じた。この結果から、同じポリマーでもいくつか初期構造を変えた結果を平均する必要があることが示唆された。

JP 30

ケモメトリックスによるプラスチック廃棄物の判別

○田辺 和俊　たなべ かずとし　物質研
松本 高利　まつもと たかとし　物質研
鈴木 康志　すずき やすし　島津製作所
岡田 行雄　おかだ ゆきお　日産自動車

プラスチックのリサイクルのためのプラスチック廃棄物の判別では近赤外反射測定が現在有力な方法であるが、この方法では家電製品や自動車用品に使用されている黒色のプラスチックの判別が困難である。また、家電製品や自動車用品のプラスチックリサイクルでは、プラスチック廃棄物のグレードの判別も必要である。そこで黒色プラスチックのグレードの判別手法を開発するために、赤外分光測定とケモメトリックス解析を組み合わせた方法を検討した。実際の自動車用の黒色プラスチック試料について試験した結果、この方法で黒色プラスチックのグレードの判別手法を開発することができた。

JP 31

ニューラルネットワークを用いたフロン代替化合物の物性推算、および、新規化合物の逆設計

○山本 博志　やまもと ひろし　旭硝子
浦田 新吾　うらた しんご　地球環境産業技術研究機構

オゾン層を破壊しない、かつ、地球温暖化に対する影響の小さいフロン代替化合物の開発は非常に重要な課題である。しかしながら、こうした含ハロゲン化合物の物性は多くは知られていない。また、実際に候補となりうる化合物を合成するにしてもこうした化合物は非常に高価であることが知られている。われわれはこうした問題に対し、既知のハロゲン化合物のデータを収集しそのデータを解析することによって物性推算式を構築した。物性推算式としてはニューラルネットワークを用いた。こうした推算式を用いることによって、実際に化合物を合成する前にその化合物の物性を予測することが可能となった。次に、このように構築された物性推算式を用いて逆に欲しい物性を持つ化合物をリストアップするシステムを開発した。当日はデモを交えてこうしたシステムを紹介したい。

JP 32

ニューラルネットワークによる高分子の物性の予測

○田辺 和俊　たなべ かずとし　物質研
松本 高利　まつもと たかとし　物質研

新規高分子の開発では高分子の物性の予測が重要である。高分子物性の予測手法としては分子動力学法などの計算化学的方法も試みられているが、計算時間が猛烈にかかる上に系統的誤差を与える欠点がある。一方、原子団寄与法は迅速、簡便な方法として実用性は高いが、精度を上げようとすると補正項が際限なく増加するという問題点がある。これは高分子の物性が本来、原子団に割り当てた定数の線形結合では表せないためであり、このような非線形な関係の解析にはニューラルネットワークが有効である。そこでニューラルネットワークを用いて高分子物性の予測を検討した。高分子の物性の中で最も重要なガラス転移温度をモノマーユニット構造から予測する

ために、モノマーユニット構造における C-H などの結合の数を入力して学習を行った。テストを行った結果、温度差 600K 以上の Tg の実測値をよく再現する手法を開発することができた。

JP 33

ケモメトリックス手法を用いた材料開発支援プログラムの開発(3)
ーカウンタープロパゲーションを用いた候補のクラスタリングー

○棚田 東作　　たなだ とうさく　　豊橋技科大
荒川 正幹　　あらかわ まさもと　　豊橋技科大
船津 公人　　ふなつ きみと　　豊橋技科大
西村 竜一　　にしむら りゅういち　　三井化学

計量化学の分野では、PLS(Partial Least Squares)法や主成分分析法などの手法を用いて、化学データをモデル化する研究が行われている。そして、そのモデル式による物性予測は、材料開発などさまざまな分野で応用されている。しかし、このような通常の物性予測法では複数の目的物性値を同時に満たす説明変数の組を見つけることは困難である。そこで、これを実現する逆解析プログラムを開発した。

前回（日本化学会第 77 秋季年会）は線形重回帰分析、PLS 法、QPLS(Quadratic PLS)法、ニューラルネットワークなどのモデリング手法を用いて作成した予測モデルにより逆解析を行い、目的物性値を満たす説明変数の候補（つまり材料候補）を提案するプログラムを作成した。今回は、逆解析により提案された候補をカウンタープロパゲーション(CP)を用いて、その特性に応じてクラスタリングを行うプログラムを開発した。CP は高次元空間での応答の変化の様子が Kohonen マップ上での応答の変化として三次元的に視覚化できるため、候補の解析が容易にできる。

JP 34

Multiway PLS 法を用いた活性コンフォメーションと重ね合わせルール選択手法

○荒川 正幹　　あらかわ まさもと　　豊橋技科大
船津公人　　ふなつ きみと　　豊橋技科大

3 次元構造活性相関において活性コンフォメーションと重ね合わせルールを決定することはモデルの適用能力を左右する重要な因子である。これら因子はお互いに密接な関連があり、X 線結晶解析などの構造情報がない場合には、容易には推定できない。本研究では、Multiway PLS 法による合理的な活性コンフォメーションと重ね合わせルール推定法を提案する。Multiway PLS 法は PLS 法を多次元解析に拡張した手法であり、PLS 法と比較してノイズに強い安定なモデルを与えることができる。サンプル、フィールド変数、コンフォメーションセット、重ね合わせルールからなる４次元配列を説明変数とし Multiway PLS 法で阻害活性との関係を検討した。モデル係数値から阻害活性に大きく寄与する活性コンフォメーションと重ね合わせルールを推定することができた。

JP 35

遺伝アルゴリズムを用いた剛球モデルに基づくタンパク質立体構造比較

○中山 伸一　なかやま しんいち　図書館情報大
Willett, Peter　ウィレット，ペーター　The University of Sheffield
吉田政幸　よしだ まさゆき　図書館情報大

タンパク質の立体的な類似度は，その機能や進化を考える上で重要である。我々は，タンパク質の立体的な類似度として，アミノ酸を剛球としたときの重なり容積を考えた。しかし，この類似度は二つのタンパク質の最適な重なりを求めるのに膨大な計算を必要とする。遺伝アルゴリズムは多数の状態の中から最適状態に近い状態を求める高速な方法の一つである。そこで遺伝アルゴリズムを用いてタンパク質間の類似度を求める方法を検討した。

遺伝アルゴリズムは，様々な変型が提案されており，問題ごとに妥当な方法論を検討する必要がある。本研究においても係数を変えた幾つかの遺伝アルゴリズムを比較検討した。その結果，剛球モデルによるタンパク質の構造比較に最適なアルゴリズムとその係数を明らかにした。さらに，この方法論のデータベース検索への利用の可能性についても考察する。

JP 36

低分解能質量スペクトルにおける部分構造識別関数構築のための変数選択手法の検討

○吉田 浩士　よしだ ひろし　豊橋技科大
船津 公人　ふなつ きみと　豊橋技科大
Kurz Varmuza　クルツ・ファルムーツァ　Technical University of Vienna

低分解質量スペクトルデータは離散的でコンピュータで扱いやすく、少量の試料でも測定できるという長所を持つため、今日構造解析に対し重要な方法の一つとなっている。しかし部分構造-部分スペクトル間の相関が一般的に成り立たず、測定条件により測定結果が変わるなどの欠点がある。本研究では上述の長所を積極的に利用し、有機化合物の構造推定に役立つ知識の獲得を目指すものである。特に本研究では、スペクトルデータから部分構造の「ある」「なし」といった2クラス分類のための識別関数を構築する。識別関数を構築するために、まずスペクトルデータの特徴を400個の記述子で表した。しかし400個という記述子は実用上好ましくなく、変数選択を行う必要がある。変数選択には1)Fisher比、2)遺伝的アルゴリズムが考えられる。今回は統計手法としてLDAおよびPLSを用いることとし、変数選択を行い識別関数を構築した。この結果30個程度の変数を用いて識別関数を構築することができた。

JP 37

カスケードモデルによる定性的構造活性相関

○岡田 孝　おかだ たかし　関西学院大情報メディア教育センター

データマイニング特にマーケットバスケット分析で注目されている相関ルール探索の技法は、非常に柔軟な入力が可能であり、多様な分野への適用が期待されている。筆者はこれを一般的なデータ解析技法へと拡張するカスケードモデルを提案した。このモデルでは、カテゴリー変数の平方和分解を規範として、説明力の強いルールを導出することができる。得られたルール群は全体の平方和をできる限り多く説明する簡潔な表現を取るため、重回帰分析や主成分分析が行おうとする解析をルール表現でわかりやすく実行できる。他方、化学グラフの表現として、視点となる

原子を与え、グラフ内の各原子をそこからの距離と原子の対で表現することできる。この表現は ambiguous であるが、活性情報と併せて、分子の情報をアイテム群で表現できるため、カスケードモデルによ取り扱いが可能となる。本報告では、簡単な生理活性物質のデータセットを対象として、上記手法を適用した結果を発表する。

JP 38

コンピュータ集約型検定法の一般化加法モデルへの適用と応用

○髙木 達也	たかぎ たつや	大阪大院薬
渡里 智恵子	わたり ちえこ	大阪大薬
宮田 幸治	みやた こうじ	大阪大院薬
黒川 絵美子	くろかわ えみこ	大阪大薬
藤原 英明	ふじわら ひであき	大阪大医
黒川 顕	くろかわ けん	大阪大遺伝情報実験施設
長谷川 義髙	はせがわ よしたか	大阪大遺伝情報実験施設
安永 照雄	やすなが てるお	大阪大遺伝情報実験施設

近年、一般化加法モデル（GAM)、交替条件付期待値（ACE)、Multiple Additive Regression Trees（MART）などの、平滑化法、ノンパラメトリック回帰法が盛んに応用され始めている。しかし、国内の薬学、化学分野への応用はまだ数えるほどしかない。また、属性の選択、検定に関しても、多くの手法では客観的な基準を設定することが容易でない。すでに当研究グループでは、ブートストラップ法に代表されるコンピュータ集約型検定・推定手法の薬学・化学分野への応用を積極的に試みてきただけでなく、独自に拡張シフト検定法のアルゴリズムを開発し、ANN などに適用してきた。今回、他分野ではすでに多くの応用例のある GAM を取り上げ、拡張シフト検定法を含むコンピュータ集約型手法を導入、S-Plus 付属の標準的医薬学データの解析に適用、オリジナルの結果と比較してみたところ、比較的良好な結果を得ることができた。

JP 39

学習順序に依存しない自己組織化法の開発

○阿部 貴志	あべ たかし	山形大院工
金谷 重彦	かなや しげひこ	山形大工
工藤 喜弘	くどう よしひろ	山形大工
木ノ内 誠	きのうち まこと	山形大工
大平 賢	おおだいら まさる	山形大院工

生物には種固有のコドン利用特性があることが知られている。本研究では、遺伝子をコドン利用特性により分類することを検討した。初めに主成分分析を用いて、線形解析を行った。次に、コホネンの自己組織化法(Self Organizing Map)を用いて、非線形解析を行なった。この時の、自己組織化法の学習法の改善を行なった。1999 年現在までに、21 種のバクテリアの全塩基配列が明らかとなっている。これらの生物種を用いて検討した。

JP 40

SALS による数値計算曲線のカーブフィッティング
-- 多段階プロトン解離平衡の精密解析 --

○加藤 良清　かとう よしきよ　山形大理
山田 明文　やまだ あきふみ　長岡技科大工
小野寺 準一　おのでら じゅんいち　山形大工
鵜浦 啓　うのうら けい　山形大理

最小二乗法標準プログラムシステム SALS は、古くから化学分野で利用されて来た。SALS には、解析解がない場合も、数値計算が可能であれば、計算曲線を実験値に最小二乗フィッティングする機能が早くから組み込まれている。本研究では吸光度測定および希薄溶液の伝導度測定から水の自己解離を考慮し、多段階プロトン解離平衡を精密計算するプログラムを例に、SALS を用いて数値計算結果の実験値へのあてはめを実施するプログラムの開発とその応用例を報告したい。このようなプログラムは既に多くが公表されているが、SALS のライブラリーパッケージ を利用する方法は、読みやすい FORTRAN で記述できること、など見直されるべき特徴を有している。

JP 41

インターネットを活用する「立体化学」独習システムの試作

竹内 敬人　たけうち よしと　神奈川大理
吉田 弘　よしだ ひろし　広島大理
守谷 崇　もりや たかし　創価大工
○伊藤 眞人　いとう まさと　創価大工

インターネット、特に WWW は、その双方向性を活用したインターラクティブな遠隔教育や独習システムの新しい媒体として注目されている。化学教育で必要不可欠な構造と運動や反応に関する内容を、効果的に理解させるのに不可欠な図表や動画像を活用するには、java や VRML を用いて開発された素材を併用するのが有効である。われわれは、このような化学教育教材の一例として「立体化学」を取りあげ、インターネット上で動作する独習システムを試作した。このシステムの特徴は、(1)cgi によるインターラクティブなプログラム学習：質問に対する学習者の回答に応じて、次の質問に進むか解説を見せるかを判断する、(2)java や VRML による動画や三次元グラフィクス画像：分子の構造とダイナミクスを視覚的に体験させる、および(3)インターネット上に分散した素材の有機的活用：遠く離れたサーバーにある素材にリンクして活用する点である。

第２２回情報化学討論会　特別企画

ミキサー 「化学に於ける計算機利用の明日を語る」

主　催：　日本化学会 情報化学部会
日　時：　平成１１年１１月１０日（水）
　　　　　１７：４５〜１８：４５
会　場：　グランドホクヨウ米沢（山形県米沢市金池 2-3-7）・飛天
　　　　　（第 22 回情報化学討論会　第一日目「鈴木昭憲先生特別講演」終了後同会場にて）
参加費：　無料（多少飲み物も用意してあります。）

開催趣旨：

化学・農学・薬学等の分野で、研究・開発を進める際に抱えている、「膨大な情報を如何に効率的･効果的に活用するか」、「電子化、情報化の弊害を如何に克服するか」、「研究を進める上で如何に計算機を活用するか」など様々な課題や、学会・討論会に期待すること等々を、自由に語り合いたいと考えております。お時間の許す方はお気軽にご参加下さい。

話題提供：

1）学会・討論会の役割
　情報化学部会活動に関するアンケート結果（９８年実施）
2）「情報化学部会による論文集の電子化刊行」企画の紹介
3）実験化学者と計算機利用専門家との関わり
4）他

幹事：三戸（三井化学）、相田（広島大学）

情報化学部会　総会

期　日：　平成 11 年 11 月 11 日（木）17 時 30 分より
場　所：　グランドホクヨウ米沢（山形県米沢市金池 2-3-7）
　　　　　（第 22 回情報化学討論会　特別講演終了後　同会場で）
議　題：　3 月よりこれまでの活動報告
　　　　　これから年度末までの活動予定紹介

◆部会行事

第２２回情報化学討論会　特別企画

ミキサー 「化学に於ける計算機利用の明日を語る」

主　催：　日本化学会 情報化学部会
日　時：　平成１１年１１月１０日（水）
　　１７：４５～１８：４５
会　場：　グランドホクヨウ米沢（山形県米沢市金池 2-3-7）・飛天
　（第 22 回情報化学討論会　第一日目「鈴木昭憲先生特別講演」終了後同会場にて）
参加費：　無料（多少飲み物も用意してあります。）

開催趣旨：

化学・農学・薬学等の分野で、研究・開発を進める際に抱えている、「膨大な情報を如何に効率的･効果的に活用するか」、「電子化､情報化の弊害を如何に克服するか」、「研究を進める上で如何に計算機を活用するか」など様々な課題や、学会・討論会に期待すること等々を、自由に語り合いたいと考えております。お時間の許す方はお気軽にご参加下さい。

話題提供：

1）学会・討論会の役割
　情報化学部会活動に関するアンケート結果（９８年実施）
2）「情報化学部会による論文集の電子化刊行」企画の紹介
3）実験化学者と計算機利用専門家との関わり
4）他

幹事：三戸（三井化学)、相田（広島大学)

情報化学部会　総会

期　日：　平成 11 年 11 月 11 日（木）17 時 30 分より
場　所：　グランドホクヨウ米沢（山形県米沢市金池 2-3-7）
　（第 22 回情報化学討論会　特別講演終了後　同会場で）
議　題：　3 月よりこれまでの活動報告
　これから年度末までの活動予定紹介

◆部会記事

部会役員会報告

第4回幹事会　兼　拡大編集委員会議事録

日　時：平成11年7月5日（月）午後4時～
場　所：日本化学会会議室
出席者：吉田元二、時田澄男、三戸邦郎、伊藤眞人、髙田章、工藤喜弘、船津公人、高木達也、相田美砂子、相馬融、坂本直紀、清水弘一

議　題：アンケート集計結果の取り扱いと公表方法について検討
CICSJ Bulletin
・17巻4号：特集「データマイニング」
・17巻5号：特集「第22回情報化学討論会」
・17巻6号：特集「材料を探る」
第22回情報化学討論会の準備状況についての報告
・11月12日（金）午後の講習会の検討
・11月10日（水）の夜の企画
・プログラム編成会議：8月9日（月）10:30～14:00東京で。
情報化学講習会のテーマ検討
論文集の電子出版について検討
ロボット合成研究会シンポジウム（11月26日、化学会館7階）への協力について

◆関連記事

5th International Conference on Chemical Structure

時実　象一 (tokizane@pc.highway.ne.jp)
科学技術振興事業団

International Conference on Chemical Structure <http://chemweb.com/conference/5iccs/structure.html> は 3 年毎に開かれる化学情報の国際会議である。米国化学会年会の化学情報部会に対応するものであるが、欧州、特に英国中心である。1999 年は 6/6-10 に定例会場であるオランダのハーグの近郊の Leeuwenhorst Congress Center,でおこなわれた。今回の参加者数は 171 名であった。写真は最終日に会場前で撮ったものである。参加者には Steve Heller (NIH), Wendy Warr (Wendy Warr & Associates), John Barnard (Barnard Chemistry Information, LTD), Janet Ash, Peter Willet (Sheffield University), William Town (ChemWeb), Guenter Grethe (MDL) などがいた。また展示社は CAS, MDL, ISI, FIZ Karlsruhe, Oxford Molecular, MSI, Tripos, ACD, Cambridge Soft, SymBioSys, ID Business などであった。

今回の特徴はコンビナトリアル・ケミストリに関連する発表がほとんどをしめたことである。ソフトウェア業者の発表・製品説明もこの点に重点が置かれていた。詳しい参加報告は「薬学図書館」に発表の予定であるので、ここでは主要な発表のスライドが掲載されている Web サイト <http://chemweb.vei.co.uk/> を紹介する。ここにアクセスするには ChemWeb の会員となる必要があるが、登録は無料である。RealPlayer をインストールすればスライドを見ながら実際の講演を聞くことができる。ここに掲載されている講演は次のとおりである。

A. Trends in Combinatorial Chemistry (Wendy Warr)
B. Enhancing Hit-To-Lead Properties of Diverse Libraries (Stephen Pickett)
C. The Internet and Electronic Publishing-a Disruptive Technology (Stephen Heller)
D. Computer Manipulation of Large Virtual Combinatorial Libraries for Diversity Analysis and Subset Selection (John Barnard)
E. Experiences With Web-Based Tools In Teaching Chemical Information (Englebert Zass)
F. The Chemical Workbench - A Reaction-Centered Synthesis Design Tool (Hartmut Braun)
G. Field-Based Similarity Forcing: A Conformationally-Flexible Approach to Molecular Matching (Gerald Maggiora)
H. Example Of Successful Prediction Of Biological Activities (Eugene Babaev)
I. Closing remarks and discussion

フォトニクス材料の分子設計に関する国際ワークショップ

International Workshop on Molecular Design of Photonic Materials (MDPM99)

主　催　通商産業省　工業技術院　大阪工業技術研究所
　　　　科学技術庁（(社) 科学技術国際交流センター受託)
協　賛　日本化学会、高分子学会、近畿化学協会（予定）
会　期　１２月９日（木）～１０日（金）
会　場　宝塚ホテル（兵庫県宝塚市、阪急電車今津線宝塚南口駅前）

事前参加登録及び一般講演申込締切　１０月２７日（水）

トピックスの範囲　非線形光学、電界発光、フォトクロミック材料、機能性色素など種々の光機能材料の理論計算による分子設計法の現状と将来方向について討論します。

招待講演（予定）　Hans Agren (Linkoeping University, Sweden), Jean-Luc Bredas (University of Arizona, USA), Benoit Champagne (Facultes Universitaires Notre-Dame de la Paix, Belgium), Kimihiko Hirao (University of Tokyo,Japan), Shashi Karna (Air Force Research Laboratory, USA), Masayoshi Nakano (Osaka University, Japan), Hideo Sekino (Analatom Inc., USA), David Yaron (Carnegy-Mellon University, USA), 他

参加費　無料

参加及び講演申込方法　参加及び講演希望者は下記ホームページからサーキュラー及び申込用紙を入手（pdf 形式）するか下記申込先宛にご請求下さい。

http://www.onri.go.jp/labo/photchem/mdpm99.html

申込先

〒563-8577　池田市緑丘 1-8-31
大阪工業技術研究所
光機能材料部　光電子化学研究室
太田浩二　又は　鎌田賢司
Phone:　0727-51-9523
Fax:　0727-51-9628
E-mail: mdpm99@onri.go.jp

先導研究シンポジウム

化学反応シミュレータ

主　催　　化学技術戦略推進機構　新エネルギー・産業技術総合開発機構
後　援　　通商産業省　日本化学会
会　期　　１２月１７日（金）9時50分～17時15分
会　場　　中央大学駿河台記念館６７０号室（千代田区神田駿河台3-11-5）
［交通］JR「御茶の水駅」徒歩３分

講　演　　永瀬 茂（都立大）、寺倉 清之（融合研）、土井 正男（名大）、諸熊　奎治（エモリー大）、銅金 巌（住友化学）、船津 公人（豊橋技科大）、山田 知純（旭化成）

参加申込締切　定員（約120名）になり次第締切

参加費　無料

参加申込方法　（1）氏名、（2）勤務先、（3）連絡先を明記し、下記宛てFAXまたはE-mailでお申し込みください。

申込先　〒101-0051
千代田区神田神保町 1-3-5
財団法人 化学技術戦略推進機構
戦略推進部　安田　宏
FAX　(03)5282-0250
E-mail : yasuda@jcii.or.jp
TEL　(03)5282-7276

◆編集後記

今年はまだ4分の1を残していると言うのに、アルメニア（コロンビア）、イズミット（トルコ）、台湾と、相次いで大きな地震が起こり、残念なことに多数の方々が亡くなられました。そしてつい先日、久々に「高潮」などと言う嫌な名詞を見ることになってしまいました。

災害が起こる度、私たちは、どのようにしたら少しでも被害を小さくできるかと言うことを、考え、あるいは話し合います。こうした努力は、大局的には報われているものだと信じています。ですが、地震にしても高潮にしても、被害が起こってから、しばしば、「危険であることは指摘されていた」と言う声を耳にします。今年の地震でもそれに近い話を聞いたことがあります。それは恐らく事実なのでしょう。では、その「情報」はどのように利用されたのでしょうか？「本当に的確に利用されていれば、今少し被害を小さくできたのではないか？」と言う問いかけには、答えはYesの場合もあればNoの場合もあるのでしょうが、どうも、何度も問いかけずにはいられなくなってしまいます。

「情報」は、存在するだけでは役に立ちません。収集、解析／分析、要約、伝達などの過程を経て、はじめて何を行えばいいのかがわかり、実際に行動を起こせます。情報は適正に伝達される必要があるのだと言うこと、情報を取り扱う私たちの役割が重大であると言うことを、悲しい災害の後、改めて認識することができました。

（T. T.）

CICSJ Bulletin Vol.17, No.6, November 1999

このページは「日本化学会・情報化学部会」の責任において運営されています。

目次

◆特集：材料を探る

「材料を探る」を特集するにあたって

豊橋技術科学大学・知識情報工学系　船津　公人
funatsu@tutkie.tut.ac.jp

私たちの日常生活の中で使われている化学物質には様々なものがあるが、その中で生活に直接関わっている材料として最も多く用いられているものの一つが高分子材料である。物性や機能を比較的自在に引き出せることから生活のいろいろな場面に入りこんでいる便利なものである。したがって、最近はこの高分子材料については環境問題への対応も念頭に置いた新しい機能の実現や、これまでの物性バランスなどは維持したまま、使用する素材の内容変更が求められたりもしている。このように高分子材料が持つ材料としての意義に注目し、特集「材料を探る」は特に高分子（次号では触媒、電子材料）を選び、新たな高分子材料設計のためにどのようなことが問題とされ、そのためにどのようなコンピュータ利用の手法が必要とされているか、一方その手法の利用者の立場から見た場合、どのような問題が指摘されているかなど、その現状が垣間見えるように各記事の中で触れて頂いている。

高分子物性予測については、要因の複雑さのために計算化学的なアプローチよりむしろ経験的な手法（たとえば原子団寄与法）が実用性の観点から広く定着している。今回の特集でも随所にこのことが述べられている。たとえば膨大な実験データや計算結果に対するケモメトリックス手法を用いた解析や予測モデルの作成は当然のこととして、さらにケモメトリックスの手法を土台に候補材料の提案および候補の評価を可能とする逆解析手法の開発が進められていることは実用の観点からの新しい動きである。ただ、ここで常に気をつけなければならないことは、このアプローチの有用性と信頼性を支える大きな要素の一つが、体系的に整理された信頼のできるデータベースの構築にあるという点である。科学技術振興事業団(JST)は、かねてより高機能基盤データベースPoLyInfoの開発を進めているが、データベースそのものと併せて、物性予測などの機能の充実も予定している。この全体構想が実現すると高分子材料設計を指向した世界でも類を見ない本格的な高分子データベースシステムとなる。データベースを軸に、日本でもやっとこうした根っこからの動きが見え始めたことは喜ばしく、今後の開発の動向に注目して行きたい。

また、高分子材料設計をシームレスズーミングのキーワードの下に計算化学の立場から推し進めている土井プロジェクト(JCII)における最新の研究成果の一部と今後の見通しなどについても執筆願った。今後、公開に向けた研究の進展に期待したいところである。この他、高分子膜についての物性・機能予測の取り組みについても寄稿して頂いた。設計・予測のための基礎となる理論的研究成果を経験的な手法と組合せることで、構造—物性—機能との関係の定式化図る道筋を示したものとして示唆に富んだ内容となっている。

最後に、今回の特集では扱わなかったが、たとえば生分解性プラスチックなど、ゼロエミッションを視点に据えた高分子材料設計の動きも当然のことながら指向されている。いずれにしてもコンピュータ支援による高分子材料設計の研究は日ごとに進展し、その成果の果たす役割はますます重要となってきている。この特集でも見られるように、その特徴の一つが、「計算化学や情報化学の領域を問わないさまざまな手法の高分子材料設計への複合利用」の動きであろう。この企画を担当した者として、そうした動きの一端をお伝えできたとしたら幸いである。最後に、今回の特集にあたりご執筆頂いた方々にこの場を借りて謝意を表す次第である。

ふなつ　きみと　FUNATSU, Kimito
連絡先　〒441-8580　豊橋市天伯町雲雀ヶ丘1-1　豊橋技術科学大学・知識情報工学系
Tel: 0532-44-6879　Fax: 0532-47-9315

機能材料分野における高分子と計算科学

三菱化学（株） 四日市事業所　岡崎 慶二
1404726@cc. m-kagaku. co. jp

機能材料分野において、高分子等が創り出すメゾ構造のかかわる問題に対し、計算科学を、複合的な視点やアプローチでもって生かそうと模索している姿を紹介する。

１．はじめに

高分子材料と一口に言っても、純粋な高分子物質、可塑剤や安定剤などを含む実用的な樹脂、繊維強化プラスチックでのマトリクスや有機感光体のバインダーのように付加価値の高い製品において重要な役割を担っているものまであり、主役としても脇役としても重要な素材である。関係する構造性は、重合度・分岐度・架橋度とそれらの分布やポリマーの幾何学的形状といった分子構造、非晶相・結晶相・相分離構造といった高次構造、ガラス繊維による強化といった巨視的構造と非常に多様で、それだけに材料として多くの可能性を秘めている。また、そこでの特性の発現には、極めて幅広い空間・時間スケールの現象が絡んでいる。こういった、極めて多様性のある（特に構造性の面で）広がりのある（合成高分子だけでなく無機高分子や生体高分子へのという意味で）高分子にこだわり、その可能性を最大限引き出していくことは、新たな機能材料を生み出す引き金になっていくであろう。

この高性能化、高機能化には、いくつかの（往々にして相反する）性能や機能のバランスや新たなものの付与が求められている。それらを、化学構造、分子構造、高次構造、巨視的構造といったいろんなレベルの構造性の工夫で実現させる。この構造性は、何らかの反応なり製造プロセスで、コスト、安全性、品質、リサイクル・廃棄物処理の側面を考えながら作り出さなければならない。したがって、材料開発では、キーとなる構造性をおさえ、それを通じて特性の発現機構を理解し、さらに、構造形成の機構を捉えることが極めて重要である。このためのひとつの、しかも有効なアプローチとして計算科学的なものがある。1)

２．計算科学の主な要素技術

一般に計算科学の要素技術としては次のようなものがある。ミクロなレベルから分子や超分子の構造・性質を扱う分子軌道法。原子数が多くなって分子集合体としての性質・構造が問題になってくると分子シミュレーション（古典的分子動力学法、第１原理分子動力学法など）が威力を発揮する。2)これらの手法は、理論的、ミクロな見方、計算負荷が大きいという特徴を持つ。さらに、時間・空間スケールの大きな現象や状況を扱うのに、粒あるいは場という描像でもって粗視化する様々な手法（粒子シミュレーション、動的平均場理論など）がある。3) 一方、これらと全く性格の異なるものが原子団寄与法である。4) 化学構造をいくつかの原子団に分割し、それにあらかじめ割り当てられている値を加算することによって望みの特性値を得る。通常は、既存の実測データに基づいて、特性と化学構造情報とを、直接、統計的手法やニューラルネットワークで結びつけている。したがって、経験的で、マクロな見方、計算負荷が極めて小さい。薬物の構造活性相関のように分子軌道法や分子力場法から得られる情報を指標として用いて、適用性を広げたものもある。

これらの要素技術はそれぞれ、当然、扱える対象、精度、適用性に限界があるが、高分子のかかわる材料分野でも、たまたまうまく適合した問題に対しては有効な武器となる。中でも、分子軌道法は、ここ１０年間、企業の中で、たとえば、樹脂重合触媒開発といったところにおいて、確実に認知されてきている。可能な作用機構を絞り込み、触媒のプレスクリーニングにまで使われている。このような計算結果を生かした特許も、最近、精力的に模索されている。しかし、ここでも、要求されることが、さらに巨大系、ミクロ不均一系（界面での現象）、励起状態に関係してくると、まだ、大きな壁があり、分子軌道法を中心とした今後の発展が望まれる。

３．材料開発の場での問題のさらなる複雑性

現在の計算科学で扱えるかどうかには関係なく、機能材料開発では、もっと多くの複雑だが重要な問題を抱えており、そこでの貢献が強く望まれている。その状況を、電子写真プロセスの有機感光体やトナーの例で見てみる（図１）。そこには、多様な実用特性（基本特性ではなく、感度、帯電性、定着性等）や、その性能、機能を具体化させる物作りで多彩な現象（分散、乾燥、熱定着等）が関係

する混合系（物と物理的状態、界面が重要となる不均一系）の姿がある。無論、ここでも高分子は重要な役割を担った構成要素のひとつである。感光体にしてもトナーにしても、所望のものを得るための物作りをしながら、そこでの支配因子が何かを知るため、できるだけ系統的な、しかし往々にして絨毯爆撃的な実験をする。その中で、キーとなる構造性を見出し、構造制御や特性発現での支配因子を捉えようとする。ただ、そこで見ているものは、いろんな因子が複雑に絡み合ったものである。問題の所在さえ明らかでない場合もある。このため、実用特性に関しては、そのものを直接考えるのではなく、実際に使われる場面での状況（プロセス的な側面）をある程度解析したり、実用特性をさらに基本的な特性にブレークダウンして捉えたりする。

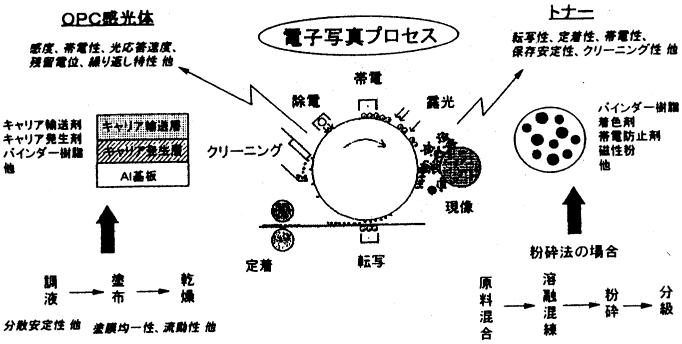


図１ 材料開発の場での問題の例
-多様な現象や実用特性が関係する混合系（物と物理的状態）-

４．計算科学的なアプローチへの期待と現実的な対応

では、計算科学的なアプローチにどういうことが期待されているか？ 本質を捉えて、独創的な物を早く発見し、広くて強い特許を取得する。実験と相補的で、定量的な結果が望ましいが、定性的にでも明確で有効な道筋を示す。さらに計算科学というものの特徴を踏まえると、具体的には次のようなことなどが上げられる。

①実験では容易に作れない条件、状態、物での情報。理想状態としての、たとえば、欠陥のない、１００％立体規則性、単分散で。極限条件としての、たとえば、高温・高圧、超高せん断場で。
②複雑な因子や条件が絡み合っているとき、それぞれが示す単独の効果。
③ありすぎる可能性の絞り込み。すなわち、スクリーニングや方向性の提示。
④想像できる単一の状況でなく、複数の因子や条件が組み合わさったときの状況。たとえば、分子量・組成・シークエンスの分布、相互作用のバランス、混合物、複数の現象。

ただ、現状では、まだまだ、求められていることとできている（できる）こととのギャップは大きい。

我々は、こういった期待に対し、計算科学という立場から、実験グループとうまく連携しながら、その問題解決に挑み、現象の理解、新たな研究指針の提示、独創的で新規の材料の提案とその特許化、時には、新たな問題を提起していこうとしている。目指す“新たな材料開発”には、新しい材料を創出したいという想いと、従来の材料開発を新たなものに変えていきたいという想いがある。

広い意味での計算科学を生かした現実的な迫り方としては次のようなものがありうる。

①物なり作り方なりを特徴づける量と、目的とする量とを、直接ニューラルネットや統計解析で結びつけて、支配因子を抽出したり、最適な処方箋や原料を見出す糸口を得る。
②いろんな形で実用特性をブレークダウンし、問題とすべき特性が物の種類や使用条件でどう変わるかを、キーとなる構造性を想定しながら、分子軌道法や分子シミュレーションなどによって、定性的に、できれば定量的におさえ、材料開発の方向性を決める判断材料にしていく。
③キーとなる構造性などの問題の所在やより本質的な因子を、各種分析や物性評価技術とともに、現実を単純化したモデルでのシミュレーションにより、基本的な描像を提示しながら見出していく。それをベースに、物や物作りへの提案を試み、必要があればさらに、もっと化学的個性が反映したアプローチを模索していく。
④流体解析のようなマクロ的な製造プロセスシミュレーションとうまくリンクできれば、そこで必要とされる化工物性を求めるという極めて明確なターゲットを設定できる。

５．現状と共通した課題

このような対応で、実際にどういうレベルの貢献ができているかを表１にまとめてみた。限定された問題の、しかもあるひとつの側面を見ているにすぎないが、いずれのところでも、四苦八苦しているところは共通している。それは、メゾ構造、すなわち、数十nmからミクロンオーダーの空間的不均一性や界面の問題がかかわっているところである。ここでも、ミクロやマクロな視点でのアプローチ

が必要とされるが、それらについては、比較的実績のある汎用的な手立て（たとえば、ＭＳＩ社の汎用ソフトウエア）がある。ところが、メゾ構造そのものに真っ向から立ち向かおうとすると、途端に実験的に捉えるのが難しく、まして、計算は、その道筋なり、拠り所になる汎用的な手法が全く不十分なのが実状である。必然的に、実験も含めた複合的なアプローチをとらざるを得ないし、また、本来そういう分野とも言える。それは、ミクロ、メゾ、マクロを行き来する視点なりアプローチでもある。このようなメゾ構造が絡む問題に対してはその取り組みはまだ緒についたばかりである。

表１ 機能材料分野での計算科学の現状

－有機／無機／高分子材料が複合化したものが多い－

樹脂可塑剤	触媒開発の目標提示、パラメータ特許（QSPR）
架橋樹脂	不均一性の指標と強度との関係への糸口（ゲル化前までの架橋反応シミュレーション）
合成ガラス	ゲル化過程の描像とゲル構造の特徴づけ（分子動力学シミュレーション）
OPC材料	プレスクリーニング、パラメータ特許、特許強化（分子軌道法、エネルギー極小化法）
乳化剤	特徴づけの指標提案、相図予測への見通し（動的平均場理論）
表面処理剤	表面偏析と性質、プレスクリーニングへの糸口（動的平均場理論）

メゾ構造が本質的に支配している問題では、まだ入り口に来たばかりの状況

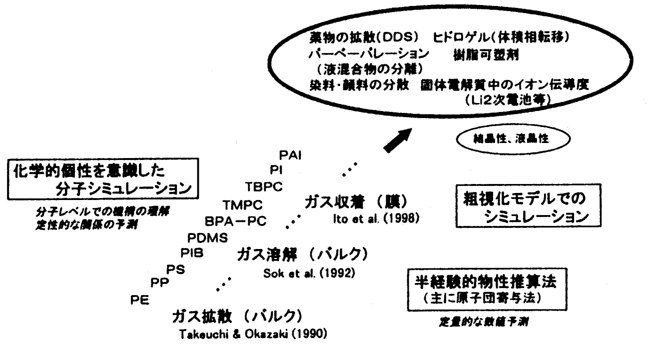


図２ ガス透過性の問題からの展開

６．複合化された基盤技術の姿

複合化と言っても、一般論として捉えていては、現実に役立つ技術としての実感がなかなか持てない。そこで、私なりに、複合的に迫る基盤技術の姿をひとつ、これまでに経験してきた範囲で描いてみる。

ガス透過性のような広がりのある特性に着目する。それは、機能と言われるものの代表例であるとともに、この現象は物質移動として考えれば、広く様々な場面で遭遇する、共通性の高いもので、高機能化や高付加価値化につながる現象の根っこにあるものである(図２)。すなわち、未反応モノマーの除去、液体混合物の分離、吸湿性、可塑剤・染料・酸化防止剤等の添加剤の分散性、膨潤性、ヒドロゲル、ＤＤＳでの薬物の放出、高分子電解質電池でのイオン伝導性、塗布・乾燥、酸化分解反応等での劣化、発泡成形などの問題と関係している。無論、それぞれの個別的な問題の詳細に入れば入るほど、共通の問題としては捉えにくくなる。しかし、基盤技術的な見方で捉え、そこでの技術の体系化を試みることは意味があり、現状である程度それができるように思う。ただ、根っこにあるとした高分子中のガス透過性の問題自身、実はそう単純でなく、ポリマーアロイで高分子膜の場合、大きく、各成分自身の性質、界面での状況、モルフォロジー、膜としての形状等が関係してくる。さらに、個々の要素である単品の性質も、結晶性や液晶性に影響されるし、バルクで均質な非晶状態での性質が大きく効いてくる。また、膜中の低分子成分の種類や量の影響も。これらに対し、既に、計算科学的なアプローチ（分子シミュレーション自身が広がりをもって、他の手法と相補的に）で、比較的系統的で継続的な研究がなされてきている。その結果、定性的な面、定量的な面で、計算ならではの重要な知見が得られつつある。5)

一方、下記のような問題をまだ抱えていることも事実である。最近、Liら(1997)は、ＤＤＳへの応用を意識して、４種類の非晶ポリマーに対し、分子量が１６から４５２までの１２種類の低分子化合物の拡散係数を、市販のソフト（分子動力学法）で計算し、実測値との比較を行っている。その結果は、依然、相対値を得るのに有効なレベルにすぎない。結晶部のモルフォロジーについては、粗視化モデルを用いたモンテカルロ法での研究があるが、そのサイズ、形状、配置、濃度に制限があり、現実のポリマーで見られる多様な半結晶性を考慮できるまでには至っていない。非晶部と結晶部の界面層の扱いも問題である。さらに、膜というものの扱いについては、まだ、膜全体での透過現象そのものではなく、これを、バルク領域と界面領域に分けてアプローチしている段階である。Fritzらは、原子レベルでの分子動力学法で、PDMS中での水－エタノール混合物の拡散挙動を研究している。吸着、脱着の過程をおさえるには、こういった膜表面のミクロな描像も重要である。以上のことは、やっとここまでできるようになってきたということでもある。今後、ますます多様な技術を駆使して、相互に関連した広がりのある分野、技術になっていくのではなかろうか。

無論、実際的な場面では、基盤技術全体が組み立てられていなければ何もできないということはな

い。問題を選べば、出来上がっている技術レベルに応じた貢献が期待できる。ただ、その場その場での断片的なもの（利用技術や実績）にとどめず、全体を見通せるようになることで、さらなる可能性を刈り取っていくことができる。

７．おわりに

社会全体がリストラの嵐で少し明かりは見えるものの依然冬の時代の中、計算科学に大いに夢を持ち、積極的な提案や特許の模索によって、その可能性を着実に実りあるものにし、新たな材料開発へのひとつの、しかし大きな起爆剤にしていきたい。材料自身が、単品の素材、混合物、各種空間スケールでの複合材料、さらには、いろんな機能を材料の中に閉じ込めて融合させたインテリジェント材料という方向に向いていっているように、計算科学にまつわる技術も複合化、融合したものが求められてきている。無論、いろんな観点の複合化がありうる。既に示した複合化の方向の他に、ややこじつけがましいかも知れないが、さらに、いくつかの複合化の姿を上げておく。

ひとつは、計算科学を生かすことを徹底的に意識して分析・物性評価技術とをうまく組み合わせる試み。次に、材料データベースと有機的につながった統合化システム構築の方向。今、先導研究″化学反応シミュレータ″の中で、主に有機合成経路探索に関し、産学官で、このような方向の可能性と姿が議論されている。また、科学技術振興事業団では、既に高機能基盤物質データベースの開発に取り組み、プロトから実用システムへ、さらにシミュレーション機能までも取り込もうとしている。6)

最後に、産と学でのプロジェクト。メゾ構造が重要な役割を担う問題に対し、昨年、土井プロジェクトが始まった。7) ４年後に、ミクロ、メゾ、マクロがシームレスにつながったひとつの具体例の提示、複雑な現象や系に迫れる独創的の理論の開発、発展性のあるユニークなズーミング機能を有するプロトタイプシステムの完成を目指している。ここから、計算科学の要素技術のブラッシュアップ、これらを組み合わせて実体に迫る利用技術と基本的な事例が生み出されてくることを期待している。企業の中では、こういった要素技術、複合的な基盤技術、それでの実績を踏まえて、より複雑な実際の問題に取り組んでいくことになる。

さらに、産学のプロジェクト間での複合化。精密重合は、それ自身、高分子の可能性を大きく膨らませる原動力になりうる。と同時に、計算結果を分子設計に跳ね返したり、計算結果の検証を行いやすいという面で、さらに、精密重合で得た高分子の可能性やそこでの目標を提起するという面で、本来高分子計算科学の最も近くにある分野である。精密重合関連プロジェクトでなされている最先端の研究とのうまいタイアップで、計算科学が実用面や基礎的な面で大きく飛躍していくことを期待している。

参考文献

1) 高分子, 43(11), 775 (1994) 特集ー高分子科学と計算機. ; 高分子, 48(5), 311 (1999) 特集ーコンピュータが高分子を科学する.
2) L.Monnerie, U.W.Suter : Atomistic Modeling of Physical Properties of Polymers, Springer, New York (1994).
3) G.J.Fleer, M.A.Cohen Stuart, J.M.H.M.Scheutjens, T.Cosgrove, B.Vincent : Polymers at Interfaces, Chapman & Hall, London (1993).
4) A.A.Askadskii : Physical Properties of Polymers : Prediction and Control, Gordon and Breach Publishers, Amsterdam (1998).
5) F.Mueller-Plathe, Permeation of Polymers - A Computational Approach, Acta Polymer., 45,259-293 (1994).
6) 公開データベース ホームページ (http://kronos.tokyo.jst.go.jp:8085/).
7) 吉田元二,『高機能材料設計プラットフォーム』プロジェクト,JCPE NEWSLETTER, 10(3),17(1998).

おかざき けいじ OKAZAKI, Keiji

プロセス解析、化工物性推算、高分子材料設計、分子設計、反応設計、医薬設計、メソ構造制御といったこれまでの雑多な経験を生かして、機能材料分野での新たな計算科学の姿を、開発現場に近いところで模索しています。

連絡先 〒510-8530 三重県四日市市東邦町１番地 三菱化学株式会社 四日市事業所 技術開発センター 基盤技術研究所 電話代表0593-45-7096

物性推算法を利用した分離膜のシステム設計

東京大学大学院工学系研究科　　山口猛央、王　保国、中尾真一
yamag@chemsys.t.u-tokyo.ac.jp

計算により機能材料を設計する場合、フィッティングパラメーターを一切含まない計算モデルが必要であり、計算によるスクリーニング試験から素材および材料構造を決定することになる。現実的には大まかな計算精度で構わないから、多くの素材、素材同士の組み合わせ、様々な構造から最適な機能材料を探し出す必要がある。候補がある程度絞れれば、実験により最適解を検討することも困難ではなくなる。ここでは溶媒分離膜の設計を目的として、フィッティングパラメーターを含まずに膜透過性能を予測する手法を紹介する。膜性能は膜溶解性、拡散性、膨潤抑制効果により決まるが、それぞれの物性値を計算より予測し、その結果を組み合わせることにより機能材料としての性能を予測する。これより膜性能の計算からのスクリーニング試験が可能となりつつあり、計算だけから分離系に最適な膜材料および構造を選択することも夢ではなくなってきた。

1.　はじめに

有機溶媒を分離する膜、VOC（揮発性有機物）など環境汚染物質を水や大気中から除去する膜など、溶媒分離膜技術は省エネルギー技術として注目されている。この技術の他法に対する優位性は、高い性能を示す分離膜が開発できるかどうかにかかっている。分離法としては膜供給側を液とし透過側を真空にする浸透気化(PV)法および供給側を溶媒蒸気とし透過側を真空にする蒸気透過(VP)法の 2 法がある。両法とも透過機構としては同一であり、膜への溶媒の溶解、溶解した溶媒の膜内での拡散により、供給側から透過側へ溶媒が透過する。

溶媒分離と言ってもプロセスにより様々な分離系を考えなければならない。それぞれの溶媒系に対して最適な膜を実験から探すことは困難であり、膜の設計法が望まれる。

高分子分離膜では溶解性および拡散性の差が選択性を決めるが、一般に溶媒の膜への溶解性の差が主に選択性を左右する。選択性を高めるために特定溶媒に対し溶解性の高い素材を用いると、高い親和性のため膜が膨潤してしまい高分子鎖の過疎化が起こり、選択性が大幅に低下する。本来、膜には溶媒間に対して大きな溶解性の差を示す素材を用いるべきであるが、膜の膨潤を抑えることも重要となる。通常の膜では架橋により分子内で膨潤を抑制するが、架橋により膜膨潤を抑制すると高分子鎖自身の運動性が低下し、溶媒の拡散性が極端に低下する。つまり高い選択性と透過性を兼ね備えた膜の開発が困難である。我々は膜の膨潤抑制と溶解拡散性による選択性発現を別々に分担するフィリング重合膜を提案した。常温では全ての有機溶媒に膨潤しない素材でできた多孔膜の細孔を、特定成分にだけ溶解性を有するポリマーで充填する。すなわ

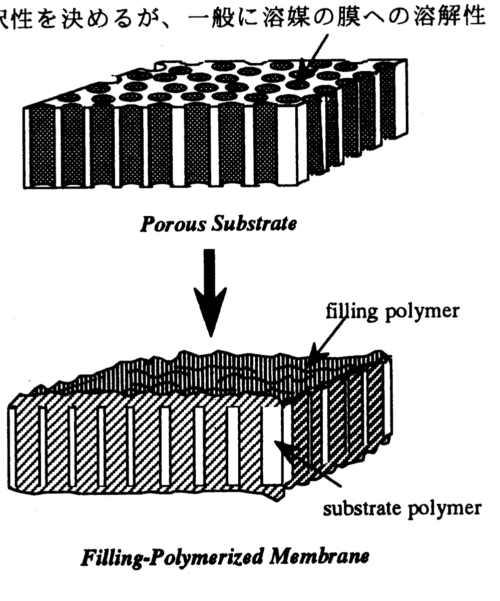


図 1.　フィリング重合膜の概念図

ち、多孔膜のマトリックスが充填されたポリマーのたがの役割をして膨潤を防ぐわけである。細孔を充填するポリマーは、細孔表面に化学結合させた直鎖状ポリマーを用いることにより安定性と高い拡散性が得られる。フィリング重合膜のコンセプトを**図 1** に示す。この重合膜は、プラズマグラフト重合法を用いて多孔性基材の細孔を細孔表面から成長する直鎖状のグラフト鎖で埋めることにより作製される [1]。この膜では、役割分担しているために高い選択性と透過性の両方が得られるとともに、これから述べる計算による膜設計へと展開できる特徴を持つ。

2. 推算モデルの考え方

我々が提案している推算法は単一素材の物性値を推算し、複数の素材性能を組み合わせることにより透過性能など機能材料の性能を予測する方法である。材料設計には計算によるスクリーニング試験が必要だと考えているが、このためにはフィッティングパラメーターを全く含まないモデルが必要であり、また多くの素材を一度にテストするために計算時間があまりかからないモデルが望ましい。現状のコンピューター性能を考え、原子団寄与法を対象とした。通常のガス透過膜性能の推算モデルでも原子団寄与法を用いたものもあるが、この原子団は膜透過実験値に直接あうように決められているため物理化学的な意味は希薄である [2]。ここでは膜溶解性、拡散性、膨潤抑制効果という物性値をそれぞれ推算し、その結果を組み合わせて膜透過性能を予測する。この単一素材の物性値の積み上げ法であれば他の材料設計へと展開することも可能であろう。細孔中のグラフトポリマーの溶解性、溶媒の拡散性は、単純な直鎖状アモルファスポリマーの理論で表現できると仮定した。違いは基材による膨潤抑制効果だけである。基材の膨潤抑制効果と組み合わせて、実験からのパラメーター無しに膜透過性を予測する [3]。

3. ポリマー・溶媒間における溶解性の推算

GCLF-EOS（原子団寄与-状態方程式）モデル [4]および UNIFAC-FV モデル [5]は、どちらも原子団寄与法によりポリマー溶解性を推算するモデルであり、モノマー単位および溶媒の構造式だけから計算できる。2 つのモデルの違いは、UNIFAC-FV 法が活量係数予測モデルであり、混合に伴う Excess Volume を理想型として扱っているのに対し、GCLF-EOS モデルでは状態方程式を基礎としているため、系の体積挙動を正しく表せる点である。本研究ではポリマーおよび溶媒の構造式だけから溶解性を予測できる両法を比較し、膜設計法へ適応した。

ポリマーとしては各種アクリレートおよびメタクリレート類を、溶媒としては各種芳香族を

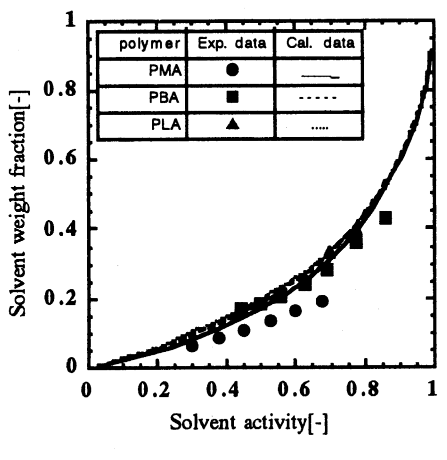


図 2(a). 25℃におけるアクリレートポリマーに対するベンゼンの溶解性および UNIFAC-FV model による推算結果の比較

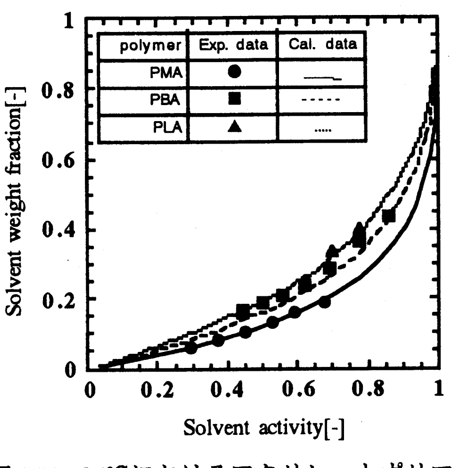


図 2(b). 25℃におけるアクリレートポリマーに対するベンゼンの溶解性および GCLF-EOS model による推算結果の比較

用い、純ポリマーに対する溶解性を石英スプリング法により測定した。ゴム状のアクリレート類に関してベンゼンの溶解性試験結果と Unifac-FV および GCLF-EOS 法による予測値との比較を**図 2(a)**、**(b)**にそれぞれ示す。予想通り Excess Volume を考えない Unifac-FV 法と比較して GCLF-EOS 法による予測は実験値をより正確に再現した。GCLF-EOS 法を用いればゴム状ポリマーの溶解性を原子団より予測することが可能である。

図 3 にガラス状のポリメチルメタクリレート(PMMA)およびポリエチルメタクリレート(PEMA)に対するベンゼンの溶解性を示す。ガラス状ポリマーも溶媒により膨潤すると、過疎化効果により定温でゴム状態へと転移する。この現象は定温ガラス転移と呼ばれるが [6]、グラフ中にその組成 w_{gm} を書き入れた。PMMA 系ではベンゼンの重量分率が 0.19 でガラス転移するが、この点以下では高分子・溶媒混合系もガラス状態である。また、ガラス状ポリマーは本来非平衡状態であり、ポリマー鎖の濃度が希薄な部分が存在する。乾燥したガラス状ポリマーの場合、希薄な部分に Langmuir 的な収着が起こり、通常の溶解である Henry 的収着と併せ溶媒活量の低い部分で溶解性が急激に増加する。これを二重収着現象と呼ぶ [7]。しかしながら、図 3 に示すように膨潤した系では、ガラス状態であるにも係わらず溶媒の低活量領域で急激な溶解性の増加は観察されず、溶解性はゴム状のポリマーに適用される GCLF-EOS モデルの予測値と一致している。ガラス状ポリマーであっても溶媒によりある程度膨潤すれば、ゴム状と同様に均一な溶液として扱え Langmuir 的な吸着は無視でき、GCLF-EOS 法により溶解性が推算できる。この考え方は高分子がある程度膨潤したときには適用できるが、ガス分離膜のように乾燥したガラス状ポリマーの場合には適用できず、現在のところガス収着においてフィッティングパラメーター無しに推算を行う方法は報告されていない。本研究のように溶媒分離膜を考える場合、膜ポリマーは膨潤するためこの考え方を用いることができる。

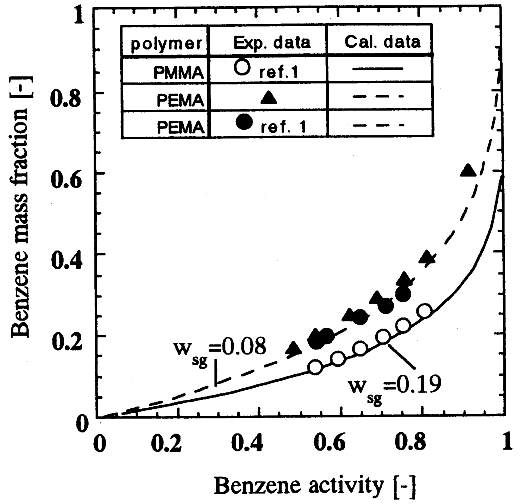


図 3. 25℃におけるメタクリレートポリマーに対するベンゼンの溶解性および GCLF-EOS model による推算結果の比較 (ref.1: Tait, P.J.T.; Abushihada, A. M., *Polymer*, **1977**, 18, 810.)

4. ポリマー中での溶媒拡散性の推算

充填ポリマーを直鎖状アモルファスポリマーと仮定しているので、溶媒拡散性は自由体積理論 [8]より計算できる。この理論ではポリマー体積を 0[K] における体積である Core Volume、主鎖および側鎖の熱振動に起因する Interstitial Free Volume および主鎖の Dynamic な運動に起因する Hole Free Volume の 3 つに分ける。溶媒は系の Hole Free Volume 中を拡散すると考えるため、溶媒・ポリマー混合系の Hole Free Volume が求まれば拡散性が推算できる。溶媒およびポリマーそれぞれが決まった Hole Free Volume を持ち、その和が溶媒・高分子混合系での Hole Free Volume となる。したがって混合系に対するパラメーターは無く、それぞれ純成分の Hole Free Volume より溶媒の高分子中での拡散性が求められる。

これらのパラメーターの中で原子団寄与法により推算できないのはポリマーの自由体積パラメータ、K_{12}/γ、K_{22} [cm^3/gK] だけであるが、それぞれ純ポリマーの動的粘弾性測定より得られる。すでに 30 種類以上のポリマーについて求められており、これらのポリマーについては

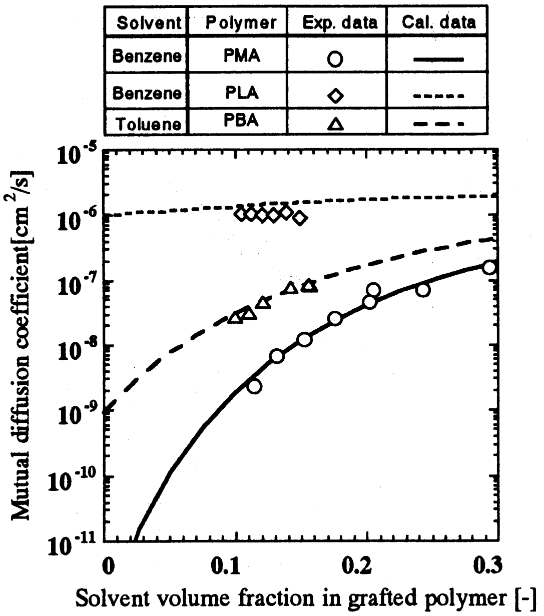


図 4. 25℃における芳香族溶媒のアクリレートポリマー中での拡散係数および自由体積理論による推算結果

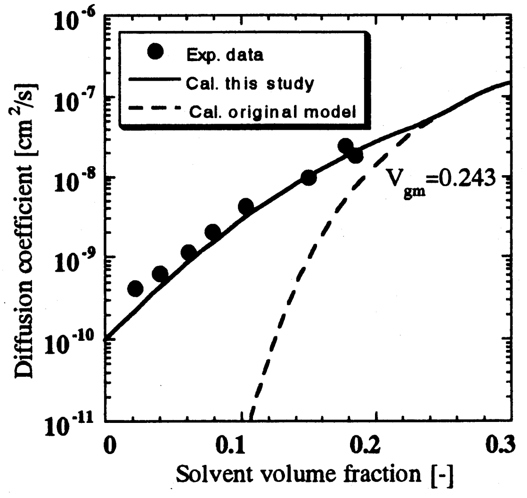


図 5. 25℃におけるベンゼンのガラス状 PMMA 中での拡散係数および修正自由体積理論による推算結果

多くの溶媒の拡散性を推算できる。

図 4 にゴム状の各種アクリレートポリマー中におけるベンゼンの相互拡散係数を示す。PMA, PBA, PLA はそれぞれポリメチルアクリレート、ポリブチルアクリレート、ポリラウリルアクリレートであり、線は自由体積モデルによる推算結果である。いずれのポリマー中でのベンゼンの拡散係数も、自由体積モデルにより良好に再現できる。

本来、自由体積理論はゴム状ポリマーにしか適用できなかったが、ガラス状ポリマーへ拡張する試みもなされている [9]。しかしどのモデルも実験値を再現することができていないため、新たなモデルを検討した。上記で説明したように、ある程度膨潤しているガラス状ポリマーでは均一系として扱えるが、この状態でもゴム状ポリマーと比較してポリマー鎖の運動性は極端に抑制されている。つまり自由体積の温度依存性は殆ど熱振動による Interstitial Free Volume の変化であり、ポリマー鎖の Dynamic な運動に起因する Hole Free Volume の温度変化は無視できるほど小さいと考えた。この仮定に基づき、自由体積モデルをガラス状ポリマーに適用した。また、ガラス転移点としては純ポリマーの値の代わりにポリマーと溶媒の混合物の転移点, T_{gm},を用い、T_{gm} 以上ではゴム状ポリマーと同一のモデルを適用した。T_{gm} の値は、純ポリマーおよび純溶媒それぞれの物性値より推算が可能である [10]。その他の考え方およびパラメーターは自由体積理論と同一であり、新たなパラメーターは導入していない。ガラス状ポリマーに関してもフィッティングパラメーターを用いずに拡散性の推算を試みた [11]。

図 5 にガラス状ポリマーである PMMA 中でのベンゼンの拡散係数の濃度依存性をそれぞれ示す。オリジナルの自由体積モデルではガラス転移点以下におけるガラス状態での拡散性を予測できないが、我々が提案している Modified model によるガラス状態での予測値は実験値と良好に一致している。ポリスチレン、PEMA およびトルエン、メチルアセテートなどの系についても予測値と実験値が一致することを確認しており、提案しているモデルによりガラス状態でも拡散係数の推算が可能になりつつある。ただしこのモデルは、ガスの拡散など乾燥状態にあるガラス状ポリマー中で

の推算には適さず、ガス分離については
さらに検討が必要である。

5. 膜の膨潤抑制モデル

本研究で対象とするフィリング重合膜では基材のたがにより充填ポリマーの膨潤を抑える。今回用いている基材素材は結晶性ポリマーである高密度ポリエチレンであり、結晶間を結ぶタイセグメントの弾性力により細孔に充填したポリマーの膨潤を抑制する。この考え方に基づいて膨潤抑制モデルを構築した[12]。パラメーターはタイセグメントの濃度だけであるが、同一の基材を用いれば常に同じ値を用いることができるはずである。細孔充填膜を用いた収着実験より今回用いる多孔性基材の f 値を求めたが、0.15 であった。この値は充填するポリマーまたは収着する溶媒に係わらず、常に同じ値を用いることができることを確認した。

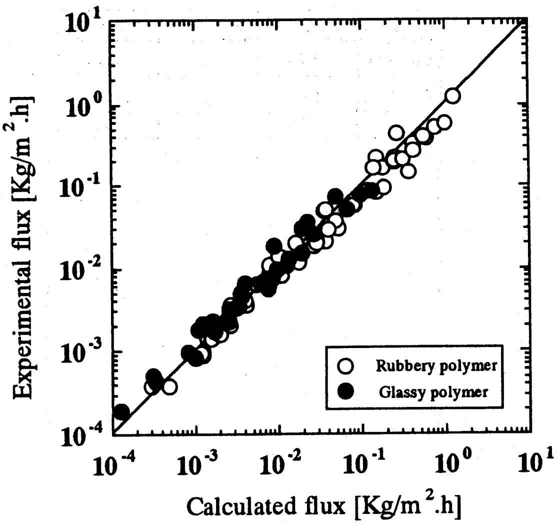


図 6. モデルによる膜透過性の計算結果および実験結果の比較

6. 膜透過性予測

これらの推算結果を基に、Fick の式に従って膜透過性を計算した[3]。ここで膜としては先に挙げた細孔充填膜を用いたが、上記のアモルファスポリマーの理論がそのまま充填したポリマーに適用できるとして、計算を行っている。基材として多孔性高密度ポリエチレン膜（東燃化学社製、細孔径 0.02μm、膜厚 5μm）を、グラフトモノマーとして様々なアクリレートおよびメタクリレートを用いた。細孔充填膜はプラズマグラフトフィリング重合法により作成した。作成法の詳細は既報に示してある[1]。

3 種類の芳香族系溶媒を用いた場合のアクリレート、メタクリレートポリマーを重合した膜の透過性に関して、実験値と計算値の比較を**図 6** に示す。計算にはフィッティングパラメーターが入っていないにも関わらず、モデルはゴム状、ガラス状ポリマーの両方について、実験値をほぼ正確に予測できることが分かる。つまりこれらの系に関して、ポリマーおよび溶媒の化学式さえ分かれば、それぞれのポリマーを用いた細孔充填膜の透過性が計算だけから予測できることになる。

7. おわりに

単成分系ではあるが、計算だけから溶媒分離膜の性能を予測することに成功している。これら計算によるスクリーニング試験より、分離系に対して最適な膜材料の選択が計算からできるようになってきた。今後は多成分系への拡張、他の多くのポリマー・溶媒系での推算精度の確認が必要だと考えている。今回のように単一材料の物性を予測すれば、それを組合わせた機能材料の性能を予測することも可能である。この考え方は分離膜に限らず、他の機能材料の設計へも発展させられると考えており、材料のシステム設計と呼んでいる。

引用文献

1. T. Yamaguchi, S. Nakao, S. Kimura, *Macromolecules*, 24 (1991) 5522; *J. Polym. Sci., Polym. Chem. Ed.*, 34, (1996) 1203.
2. J.Y. Park, D.R. Paul, *J. Membrane Sci.*, 125 (1997) 23.
3. T. Yamaguchi, Y. Miyazaki, S. Nakao, T. Tsuru, S. Kimura, *Ind. Eng. Chem. Res.*, 37 (1998) 177; B.-G. Wang, Y. Miyazaki, T. Yamaguchi, S. Nakao, ***J. Membrane Sci.***, in press
4. B.C. Lee, R.P. Danner, *AIChE J.*, 42 (1996) 837.
5. T. Oishi, J. M. Prausnitz, *Ind. Eng. Chem. Process Des. Dev.*, 17 (1978) 333.
6. P.R. Couchman, *Macromolecules*, 11 (1978) 1156.
7. W.R. Vieth, J.M. Howell, J.H. Hsieh, *J. Membrane Sci.*, 1 (1976) 177.
8. J.S. Vrentas, J.L. Duda, *J. Polym. Sci., Polym. Phys. Ed.*, 15 (1977) 403; J.M. Zielinski and J.L. Duda, *AIChE J.*, 38 (1992) 405
9. J.S. Vrentas, C.M. Vrentas, *Macromolecules*, 27 (1994) 5570.
10. P.R. Couchman, F.E. Karasz, *Macromolecules*, 11 (1978) 117.
11. B.-G. Wang, T. Yamaguchi, S. Nakao, ***Ind. Eng. Chem. Res.***, in press.
12. T. Yamaguchi, S. Nakao, S. Kimura, *J. Polym. Sci., Polym. Phys. Ed.*, 35 (1997) 469.

やまぐち　たけお　YAMAGUCHI, Takeo

研究テーマ：材料のシステム設計。様々な機能を有する素材を適切に組み合わせ（システム化）、新しい材料を構築しようとしています。分離膜に限らず、環境応答ポリマーとホストゲストの組み合わせや有機・無機を組み合わせた燃料電池用電解質膜の開発など。今回のように各素材の物性を予め予測できれば、材料システムの最適化も可能となります。

連絡先　〒113-8656　東京都文京区本郷 7-3-1　東京大学大学院化学システム工学専攻
Tel: 03-5841-7345

高分子データベースと物性予測

(株) 三菱総合研究所　長阪 匡介

nagasaka@mri.co.jp

新規高分子材料の設計・開発では、研究開発の効率化に有効な手段として、既存高分子に関するデータの分析やこれに基づいた半経験的な物性予測がしばしば行なわれている。これまで、高分子材料向けの設計支援システムとしてデータベースと物性予測機能を併せ持つ統合型システムが幾つか開発されている。本稿では、これらのシステムの構築事例を紹介し、それぞれのシステム機能とデータベースの設計コンセプトについて述べる。

1. はじめに

高分子材料の物性は多くの因子が複雑に影響しあって発現するため、新規高分子の開発や材料設計が困難であり、研究開発の効率化のためにコンピュータの利用が望まれている。物質・材料分野でのコンピュータ利用法としては、量子化学や分子動力学等のような計算化学的な手法が良く知られているが、高分子材料の場合には、問題のサイズが大きく、また、内容も複雑であるため、現在のコンピュータの能力では計算化学的な手法をそのまま適用することはできない。そのため、系の粗視化などを用いたシミュレーション手法などの研究も行なわれているが、材料設計に利用できるようになるためには、まだしばらく時間がかかるものと考えられる。

このような状況下にあるため、高分子材料の設計・開発ではデータベースや半経験的な物性予測、ケモメトリックスなどのような情報化学的な手法が有用であり、これまでにこのような手法を実装した設計支援システムが開発されている。(現在開発が進められているシステムもある。) しかし、高分子材料の研究分野は広範であり、一口に設計支援システムと言っても、取り扱うデータや要求される内容によって、システムの構成や機能はかなり異なったものとなる。したがって、このようなシステムの開発を行なう際には、対象とする高分子材料の種別や情報、システムの利用方法等を考慮した設計が重要になってくる。本稿では幾つかのシステムについてそのシステム機能や設計のコンセプトについて紹介するとともに、今後の展望について述べる。

2. システムの構築事例

高分子材料に関しては幾つかのデータベースや物性予測システムが知られているが、データベース機能と物性予測機能をあわせ持つような設計支援システムの構築事例はそれほど多くはない。ここでは、このようなシステムとして筆者が開発に関わってきた以下のシステムについて紹介する。

1. 高分子材料設計支援システム EXPOD
2. ポリマーアロイ設計支援システム PALSAR
3. 高機能高分子データベース PoLyInfo

2.1 高分子材料設計支援システム EXPOD

EXPOD は昭和電工 (株)、住友化学工業 (株)、東レ (株)、三菱化成 (株) の協力を得て、(株) 三菱総合研究所で開発されたシステムであり、線状ホモポリマー (構成繰返し単位が確定しているポリマー) の基本的な物性に対するデータベースと物性予測機能を装備した高分子材料設計支援システムである [1)]。このシステムは密度やガラス転移温度などの基本物性を対象としており、既知ポリマーに対する測定データの整理や半経験的な物性予測を行なう環境を提供することによって、新規ポリマー開発における研究や設計作業を支援することを目的としている。また、ユーザ自身による

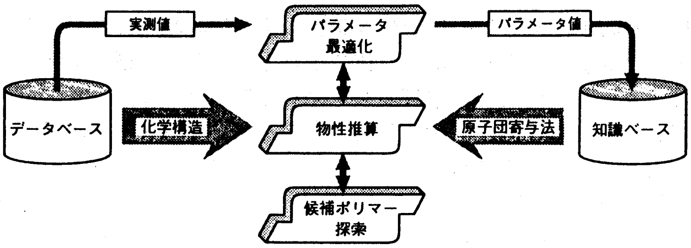


図 1: EXPOD のシステム構成

予測手法の開発や与えられた物性条件を満足するポリマー (候補ポリマー) の探索等の機能も同じ環境で利用できることを目指しており、図 1に示すような構成となっている。

図からわかるように、システムには線状ホモポリマーの物性データを格納するデータベースと物性予測に関する情報を格納する知識ベースが装備されており、これらを利用した物性予測関連機能としては以下のようなものがある。

- 物性推算機能
  データベースに格納されている構成単位の化学構造と知識ベースに格納されている原子団寄与法の計算手法や原子団パラメータを利用して、高分子の基本物性に関する物性推算を行なう。
- パラメータ最適化機能 [2)]
  データベースに格納されている物性データの測定値に基づいて、原子団寄与法で用いられる原子団パラメータの最適化を行ない、その結果を知識ベースに格納する。
- 候補ポリマー探索機能 [2)]
  原子団寄与法と遺伝的アルゴリズムを用いて、要求される物性条件を満たすと推測される高分子の構成単位化学構造を自動的に生成する。

EXPOD のデータベースは線状ホモポリマーに対する物性予測機能と連携することを想定しており、線状ホモポリマーに関するデータの特徴を考慮して設計された。その結果、このデータベースは図 2に示すように、「繰返し単位」レベル、「分子構造」レベル、「物性」レベルの 3 階層からなる構成になっている。各レベルには以下のようなデータが格納されている。

- 繰返し単位レベル … 構成繰返し単位の化学構造情報、名称情報など
- 分子構造レベル … 平均分子量、分子量分布、立体規則性情報、配向度、結晶化度など
- 物性レベル … 測定値、試料作成条件、測定条件など

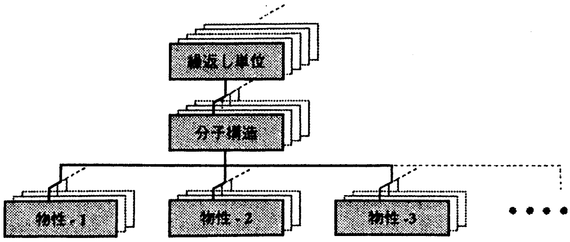


図 2: EXPOD データベースの構成

データベースをこのような構成にすることにより、同一種のポリマーに対する同一物性の測定であっても、分子構造や高次構造による試料の違いや測定方法や測定条件による測定値の違いを整理して格納することが可能となっている。

EXPODで物性予測に用いられている原子団寄与法は、構成繰返し単位の化学構造から推算を行なうため、分子の一次構造 (平均分子量、立体規則性など) や高次構造 (配向度、絡み合いなど) を考慮することができない。そのため、推算の対象となる物性は、原則として、高分子量領域で分子量依存性が無視できるような物性に限定される。ただし、分子量や立体規則性に関して系統的に測定されたデータを用意し、これを解析することにより、原子団寄与法を拡張した物性予測手法も開発することが可能になるものと考えられる。EXPODのデータベースは知識ベースはユーザ自身によるデータや知識の変更・追加が可能であり、ユーザが保有する系統的なデータを整理・格納し、様々な物性推算法を開発することを可能にしている。

2.2 ポリマーアロイ設計支援システムPALSAR

PALSARは昭和電工 (株)、住友化学工業 (株)、東レ (株)、三菱化成 (株) の協力を得て、(株) 三菱総合研究所で開発されたシステムであり、ポリマーアロイ材料の設計・開発を行なう上で必要となる、データの整理/解析作業の支援を目的としている[3)]。

ポリマーアロイ材料の物性は、ベースとなるホモポリマーの物性に加え、成分組成や混練条件、相溶化手法等様々な因子の影響を受けるため、データの解析や物性予測を行なうための手法として確立されたものがあまりなく、現在も種々の手法が開発・提案されつつある。PALSARのシステム開発の際には、このような状況を考慮し、特定のデータや手法を対象とせずに多様なデータや解析手法取り扱えることを優先して設計が行なわれた。この結果、システム構成としては図3に示すような構成が採用された。

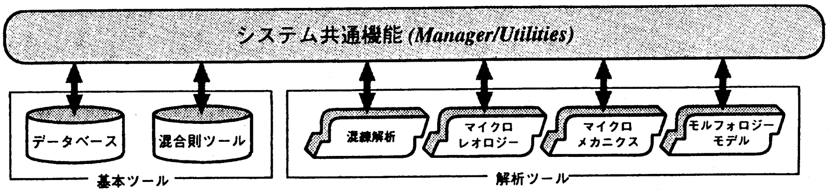


図3: PALSARのシステム構成

この図からわかるように、このシステムはポリマーアロイ向けデータベースおよび混合則ツールからなる基本ツールと、簡単な流動解析や構造解析等を行なうための種々の解析ツールを装備し、これらが共通の環境で利用可能となる統合型設計支援システムの形をとっている。

PALSARでは、ポリマーアロイ物性の解析/予測に関する基本ツールとして混合則に注目し、これらを格納・利用するための混合則ツールを装備している。混合則は、元来、物性の成分組成依存性を記述するために、粗視化されたモデルに基づいて提案された半経験式であるが、ベースとして考えたモデルに応じて多様な関数形が存在しており、式中の変数を組成以外の量と解釈することによって、相溶性や混練条件、モルフォロジー等のような様々な依存性を取り扱うことも可能である。PALSARの混合則ツールには混合則の内容を記述した知識ベースと計算ツール、測定値を利用して式中にあらわれる任意パラメータの最適化を行なうツールなどが含まれており、各ユーザがポリマーアロイ材料の物性解析を行なうために各種半経験式を利用/考案する環境を提供している。

PALSARのデータベースはポリマーアロイ材料に関する多様なデータの格納を目的としており、特に成分ポリマーに関する情報や成型加工条件、依存性データ等を扱いやすいように設計されている。また、物性値や測定条件のような数値・文字列情報だけでなく、モルフォロジーなどのような画像情報を取り扱うことも考慮されている。このデータベース構成のイメージを図4に示す。

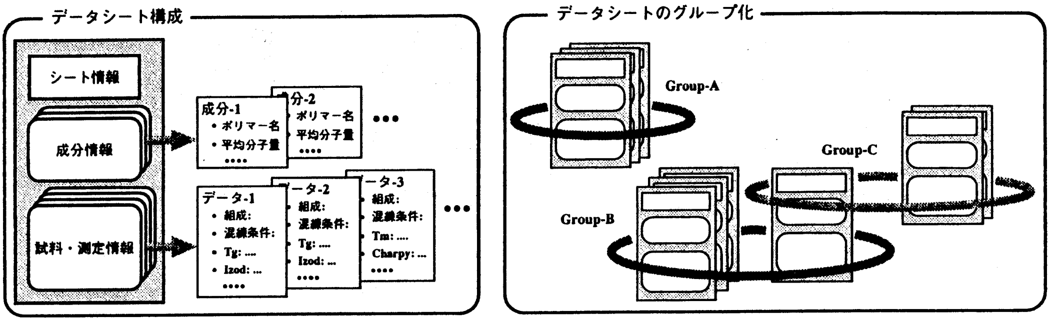


図 4: PALSAR の DB 構成のイメージ

図に示すように、各種物性の測定値はデータベース中では"データシート"という形で管理され、各データシートは「シート情報」部、「成分情報」部、「試料・測定情報」部という3つの部分から構成されている。これらの各部には以下のような情報が格納される。

- シート情報部
  測定試料名、文献情報等のように測定値(群)全体に関する情報を格納する。
- 成分情報部
  成分ポリマーの名称や分子量、物性等のように試料中に含まれる成分に関する情報を格納する。一般にポリマーアロイは複数の成分からなるため、各成分についての情報が個別に格納できるようになっている。
- 試料・測定情報部
  試料の成分組成、相溶化手法、混練条件、測定方法、測定値等のような試料の作成や測定の方法/条件、測定値に関する情報を格納する。組成や混練方法、相溶化剤の添加量、モルフォロジーなど様々な条件の違いを含む複数の測定値が格納できるようになっている。

このデータベースでは、データシートに複数の測定値を格納可能とする設計に加えて、複数のシートをグループ化して関連づけるための機能も考案されている。この機能は、各ユーザが物性発現のメカニズムを考察する場合や混合則ツールでデータを利用する場合に、類似系列や測定精度等によって測定値群を定義すること可能にしている。データベースに対してこのような構成や機能を持たせる設計を行なったのは、測定データに付随する多様な情報をできるだけ整理した形で保持させたいというニーズと、システムにある程度の柔軟性を持たせ、ユーザの経験的な知識や発想を盛り込みたいというニーズを考慮した結果といえる。このようなデータベースは、ポリマーアロイのような"複雑系"に対して、詳細な情報を整理するとともに、人間による"曖昧な整理"も受け付けるデータベースとしての1つの形であろうと考えている。

2. 3　高機能高分子データベース PoLyInfo

PoLyInfoは科学技術庁で予算化されている高機能基盤データベース開発事業の一貫として、科学技術振興事業団(JST)で開発が進められているシステムである。この事業では「高分子材料設計・開発分野における研究者の発想を支援するデータベースシステムの構築」を目的としている。具体的には、計算機に習熟していないユーザでも容易にアクセスでき、簡単にデータを検索することができ、検索結果に対して多様な取り扱いができ、物性予測などのシミュレーション機能も有するシステムということになる。現在、このようなシステムの構築を目標として、図5に示すような構想で開発が進められている[4),5)]。

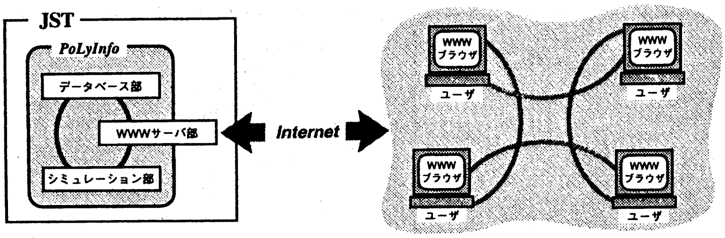


図 5: PoLyInfo システムの構想

図からわかるように、PoLyInfo システムの実体は JST 内のサーバ計算機上にあり、データベースやシミュレーションの機能をインターネットを介して提供する形態となる。ユーザは特別なコンピュータを用意する必要はなく、インターネットに接続されたパソコンなどから WWW ブラウザを用いてアクセスすることができる。現在、PoLyInfo システムはデータベース部のプロトタイプシステムが完成しており、この部分がテスト公開されている。(ユーザ登録することにより、無償で利用することができる[7]。)

このデータベースでは、対象とする高分子材料の種別や物性を限定しておらず、また、格納する情報についても、高分子の物性情報以外にも試料作成に用いられた重合反応に関する情報や原料モノマーに関する情報など広範囲にわたる情報を扱う予定にしている。このように多様で複雑な情報を格納するデータベースを開発するためには、データベースの設計時に対象データの分析を行ない、適切なデータモデルを構築しておく必要がある。PoLyInfo データベースの開発では、オブジェクト指向分析法の一種である OMT 法 (OMT: Object Modeling Technique)[6] を用いて、高分子材料関連データの分析を行なった。この手法では対象となるデータ (事象) の中で共通性の高いものを「クラス」としてまとめることによって整理し、クラス間の関係を明確にすることによってデータモデルを構築する。この分析の結果として得られたデータモデルのクラス図を図 6に示す。

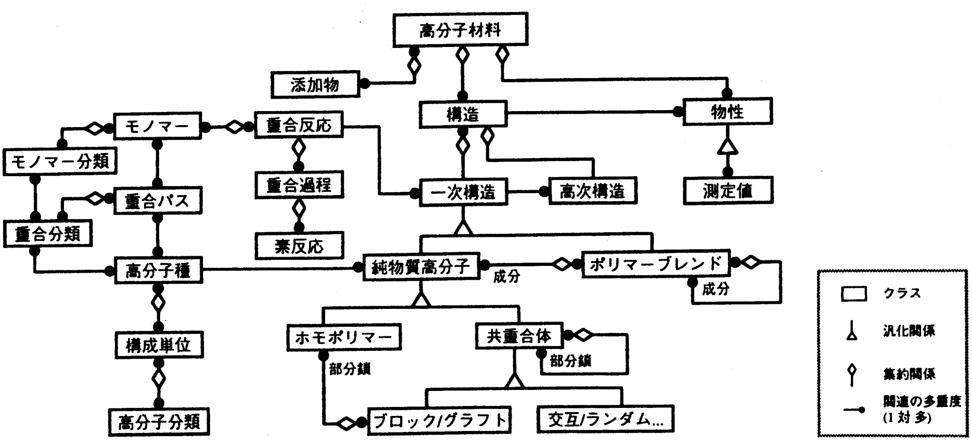


図 6: PoLyInfo で構築されたデータモデル (オブジェクト指向分析図)

紙面の都合上、モデルの詳細についての説明は省略するが、図 6 に示されたモデルでは、高分子材料に関する多様な情報の関係性がかなり詳細に記述されていることがわかる。このようなモデルに基づいてデータベースを開発し、データを整理・格納しておくことにより、様々な形でのデー

タの参照や利用が可能となる。特にPoLyInfoでは物性予測などの機能を持つシミュレーション部を開発する予定になっており、今後開発される種々のシミュレーション手法で利用できるような形でデータを保持しておくことは重要なポイントと考えられる。

現在、PoLyInfoシステムは実用システムの設計・開発が進められているが、プロトシステムのテスト公開を通じて得られるユーザの声をシステムに反映させたいと考えており、是非、多くの方々からのアクセス/ご意見をいただきたい[7]。

3. 今後の展望

ここ数年の間にコンピュータの高性能化・低価格化とネットワークの普及が急速に進み、「コンピュータは専門家だけが使うもの」というイメージは一掃されてしまった。今後もこの動きは加速されるものと考えられ、それとともにコンピュータの利用状況もますます変化していくものと考えられる。高分子関連の分野でも、すでにWWWを介した電子ジャーナルの配布・閲覧サービスや自社樹脂製品カタログのネットワーク提供等も行なわれている。

情報の電子化が進むにつれ、データマイニングや知識獲得などのようなソフトウェア技術も進展し、コンピュータによる知識処理の適用分野も広がりつつある。高分子材料分野においてもこのような手法による新たな展開の可能性もあるが、高分子材料の挙動について解明されていない部分も多く、現状ではまだ人間の経験/発想に基づく判断の方が有効な場合が多い。したがって、コンピュータによる研究支援についても、既存のデータや知識に基づいた手法の有用性も当分の間は変わらないものと考えられる。

このような状況で重要となるのが情報の流通・共有化の問題であろう。ゲノム・タンパク等の情報を扱うバイオインフォマティクスや気候変動予測等の研究を行なう地球科学などの分野では、公的な機関や公的資金を利用した研究を中心に、測定結果や解析ツール等を無償または安価で相互に利用する制度が発達している。このような情報インフラストラクチャーの整備は、分野全体が発展する上で非常に重要な要素になっている。高分子材料分野においても、データ/知識の流通・共有化の枠組みを用意し、共通基盤の上に立って独自研究を進められるような環境の整備がますます重要になってくるものと考える。

参考文献

1) 長阪 匡介. コンピュータによる高分子の物性予測. 高分子, **40**, 744–747 (1991).
2) Nagasaka, K.; Yada, T., *Reverse Design of Polymers: Application of Genetic Algorithm*, PMSE, **69**, 45–46 (1993).
3) 長阪 匡介. ポリマーアロイ設計支援システム. 日本ゴム協会誌, **68**, 891–898 (1995).
4) Maeda, C.; Tsuchiya, E.; Nagasaka, K., *Development of New Polymer Database System at JST*, 5th International SAMPE Symposium, 95–100 (1997).
5) 前田 知子. 高分子データベースの構築と利用. 高分子, **48**, 320–323 (1999).
6) Rumbaugh, J. et al., "*Object-Oriented Modeling and Design*", Prentice Hall, Englewood Clifs, 1991.
7) PoLyInfoプロトタイプシステムは 以下のURLでユーザ登録と利用が可能になっている。
http://kronos.tokyo.jst.go.jp

ながさか きょうすけ NAGASAKA, Kyosuke

もともとは分子論的高分子モデルを用いた理論的研究をやっていましたが、会社に入ってからのここ10年間は高分子関連のシステム開発に従事しています。

連絡先 〒100-8141 東京都千代田区大手町二丁目3番6号
(株) 三菱総合研究所 フロンティア科学部 (Tel: 03–3277–0556 Fax: 03–3277–0568)

材料設計・開発とケモメトリックス

三井化学（株）　西村竜一
Ryuichi.Nishimura@mitsui-chem.co.jp

ケモメトリックス（計量化学）は、「統計あるいは数学的手法を化学に適用した学問分野」である。その適用範囲は、分析化学を中心に様々の分野にわたっており、またその手法も多変量解析、ニューラルネットワーク、ＡＩなど多岐にわたっている。本稿では、三井化学における材料設計・開発分野へのケモメトリックス活用状況について報告する。

1　はじめに

パソコンの高性能化・低価格化にともない、実験化学者自らが各種の化学計算を実行する環境が整ってきた。CAE 計算はもとより、MO,MD 計算も可能である。また、実験データを整理、解析する上でも Excel などの表計算ソフトを使用することが当たり前である。実験化学者にとって、ケモメトリックス手法はデータ解析の方法として比較的なじみやすく（「ケモメトリックス」と言う呼び方は認知度が低いが）、程度の差はあるが利用している。しかしながら、ケモメトリックスの利用の仕方は、あくまでも解析が中心であり、そこから新材料などを考え出すのは研究者のアイデア・ひらめきに依存している。

本稿では、ケモメトリックスは解析型だけではなく、逆解析を行うことにより提案型の手法としても利用できることを紹介する。

2　なぜケモメトリックスを利用するのか

2.1　三井化学における計算手法の位置づけ

材料開発分野における計算手法は、高分子材料開発を例に取ると、図１に示すように計算対象物の大きさにより異なる。ミクローメソーマクロ領域で、MO,MM,MD,CAE 手法を使い分けしている。一方、DB やケモメトリックス手法は計算対象物の大きさにかかわらず、全領域で利用可能である。材料開発における膨大な実験データや各種計算結果は、DB 化そしてケモメトリックス解析をして、効率的にかつ有効的に使われるべきである。

新機能材料設計ではミクロ領域から計算を始める場合が多いが、製造（事業化）する段階で不具合が生じマクロ領域からメソ、ミクロ領域にスケールをダウンして計算する場合もある。MO,MM などを用いた材料設計技術も重要であるが、企業においては CAE を用いた製造技術を確立することも大切である。

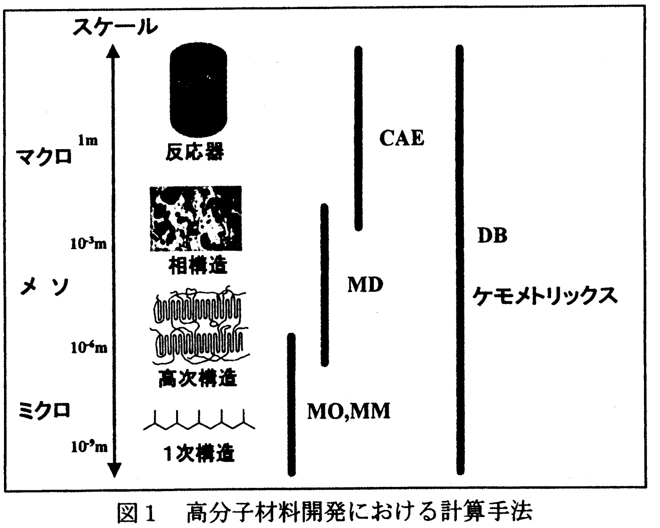


図１　高分子材料開発における計算手法

2.2　理論化学とのシナジー効果

コンピュータの高速化とともに材料設計分野においても理論化学計算が実用的になりつつある。しかしながら、企業における材料（製品）開発期間は、顧客のニーズに合わせて短期化する傾向にあり（通常３～５年）、研究開発のスピードアップが要求されている。また、高分子材料のように高次構造に起因する諸物性を理論的に実用レベルで計算することは、コンピュータ能力や時間的制限から達成されていない。実験現場では、ある程度の理論計算を参考にしながらも試行錯誤を繰り返し、材料開発を行っているのが現状である。

このことから、少しでも早く研究開発を進めるためにケモメトリックス手法の利用が期待されている。最近、注目を浴びているコンビナトリアルケミストリは、合成ロボットなどハード面は別として、ソフト的な手法はケモメトリックスと密接な関係にあるといえる。また、理論化学においても3D-QSARを筆頭に、ケモメトリックス手法とリンクさせたソフト技術が開発されている。理論化学手法とケモメトリックス手法の両方を利用することで、シナジー効果が生み出されるわけである。

3　材料設計のためのケモメトリックスソフトウェア

市販されている通常のケモメトリックスソフトウェアは、多変量解析やニューラルネットワークを計算できる機能があり、価格も１０万円から高いもので１，０００万円近くするものまである。目的を特化した形のニューラルネットワークソフトウェアでは、予測機能や目標値を満足する説明変数値を創出する機能などが盛り込まれている。材料設計を行う上では、このような提案型の機能を持つソフトウェアを利用することがベストである。しかし、大半は解析型のソフトウェアであり、逆解析などを用いた提案型のものは少ない。また、企業において実験化学者が日常的にケモメトリックスを利用するため、みんなが容易に使える環境にしておくことが望ましい。

三井化学では、豊橋技術科学大学：船津研究室およびＣＡＣフォーラム [1]で開発されたケモメトリックスソフトウェア「Chemish」[2,3]（図２）を実験化学者に配布している。市販のソフトウェアに比べると操作性は若干劣るが、機能的には同等以上かつ、他のソフトウェアでは実現されていない変数選択などの機能があり、導入コスト的にも非常に大きなメリットがある。また、このChemishに逆解析機能を追加した材料設計支援プログラム「Chemish Pro」[4,5,6]の開発も進めており、まもなく提案型のソフトウェアが完成する。

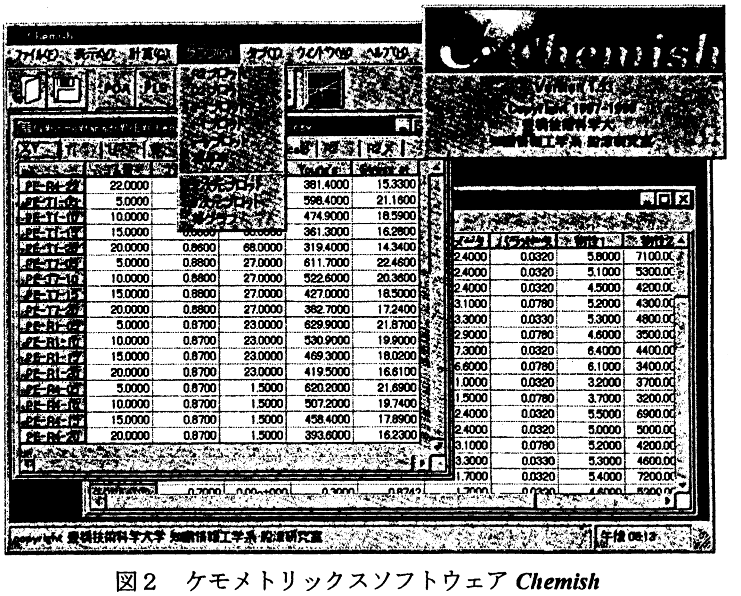


図２　ケモメトリックスソフトウェア *Chemish*
http://www.cac-forum.gr.jp/chemish/

4　ケモメトリックスを用いた高分子材料設計

4.1　高分子物性推算

高分子物性は、ポリマーユニット構造や分子鎖構造（分子量、分子量分布、立体規則性）などの一次構造、およびポリマー鎖集合体の構造（結晶構造など）に起因している。また、ポリマーアロイ・ブレンドの物性では、さらに因果関係が複雑になっている。このため、理論だけでは物性を推算することは難しい。三井化学では、高分子物性を推算する手法として、原子団寄与法、半理論式、経験式を用いている。これらの方法および物性推算する上で基盤になるデータベースについて以下に説明する。

【データベース】

高分子のデータ項目は低分子のものとは比較にならないくらい複雑で種類が多いのが難点である。ＪＳＴの高分子データベースシステム PoLyInfo[7]は、この点をよく考慮されて設計されており、本格稼働が待たれる。三井化学においても高分子データベースを構築しているが、統一的に整備されていない部分も多い。データベース管理者としては、統一されたフォーマットで専用データベースシステムに登録していく方法を期待するが、それに固執すればデータ整理に時間がかかりすぎてデータ蓄積ができなくなる。このため、現在は Excel で作成した実験データシートを Web システム上に登録する方法に変更している。この方法であれば、実験化学者が定期的に作成した実験データをそのまま利用することができる。

【原子団寄与法】

原子団寄与法では、原子団パラメータを作成した元データにより推算値が左右される。市販ソフトウェアでは幅広い種類の高分子データを用いているため、企業で取り扱うような特定高分子領域では推算値と実験値の相対的な逆転が見られる場合がある。このため、高分子の種類毎に原子団パラメータを作成して、より現実的な推算ができるようにしている。材料設計の初期段階で、スクリーニングに用いる場合や、置換基を変化させて置換基効果を調べることに利用したりする場合が多い。また、意外にこの手法は手軽に指標が得られるという理由で、実験化学者の間では利用頻度が高い。

【半理論式】

高分子の溶融物性は、ある程度理論式を用いて推算することが可能である。三井化学では、この理論式を基に、実験データから各高分子固有のパラメータを作成して、半理論式を作成している。図３は PE,PP の溶融粘度を計算した例である。この手法の利点は、理論式がベースであるため各高分子固有パラメータの適用範囲であれば、どのような条件でも計算できる点である。例えば、高分子の分子量や分子量分布が物性にどのように影響するかを検討することができる。

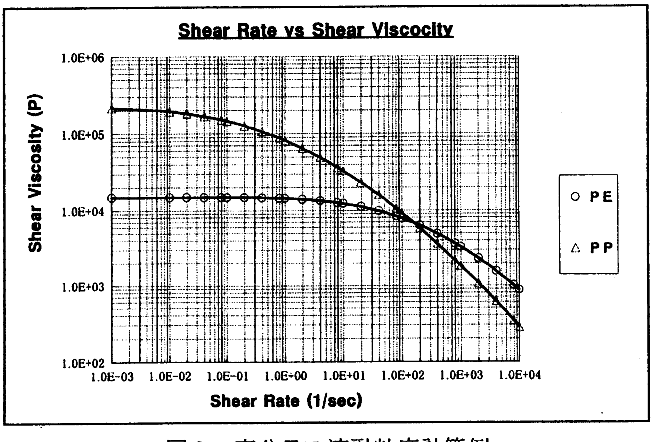


図３　高分子の溶融粘度計算例

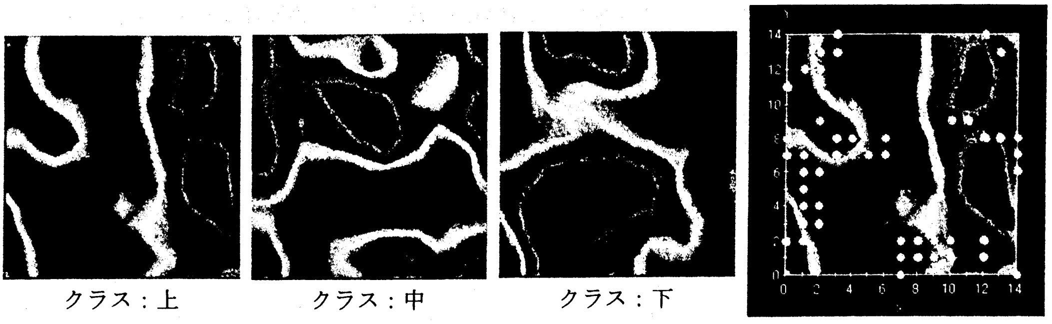


図５　ＰＰアロイ材料設計における逆解析候補のクラスタリング例

5　おわりに

ケモメトリックスはデータ解析を行うためだけの手法ではなく、うまく利用すれば材料設計・開発においても候補材料を提案する手法として使える。また、理論化学手法などとリンクさせることでその材料設計手法の可能性は大きく広がり、シナジー効果が期待される。

ケモメトリックスの利点のひとつは、実験化学者がパソコン上で容易に利用できることである。データの性格をよく理解している人が解析するわけであるから、その効果は大きい。しかしながら、企業における材料設計へのケモメトリックス利用は、意外と少ないないように感じられる。この原因のひとつには手軽なソフトウェアがないことであろう。今後、安価で良いソフトウェアが入手可能になることを期待する。

参考文献

1) CAC フォーラムホームページ http://www.cac-forum.gr.jp/
2) 荒川正幹；木村敏郎；船津公人；神村基和；西村竜一．ケモメトリックス・ソフトウェア Chemish の開発およびその Web チュートリアルの作成．第２１回情報化学討論会講演要旨集，150-153，1998.
3) CAC フォーラム Chemish ホームページ http://www.cac-forum.gr.jp/chemish/
4) 荒川正幹；棚田東作；西村竜一；船津公人．ケモメトリックス手法を用いた材料開発支援プログラムの開発．日本化学会第７６春季年会講演予稿集Ｉ，542，1999.
5) 棚田東作；荒川正幹；船津公人；西村竜一．ケモメトリックス手法を用いた材料開発支援プログラムの開発（２）．日本化学会第７７秋季年会講演予稿集，275，1999.
6) 棚田東作；荒川正幹；船津公人；西村竜一．ケモメトリックス手法を用いた材料開発支援プログラムの開発（３）．第２２回情報化学討論会講演要旨集，1999.
7) JST PoLyInfo ホームページ http://kronos.tokyo.jst.go.jp:8085/

にしむら　りゅういち　NISHIMURA, Ryuichi

数年前から高分子材料設計の研究にケモメトリックスを利用していましたが、最近はケモメトリックスの比重が大きくなりつつあります。

連絡先　〒740-0061 山口県玖珂郡和木町和木 6-1-2
三井化学（株）研究開発本部　マテリアルサイエンス研究所　計算科学室
Tel:08275-3-9215　Fax:08275-3-8870

【経験式】

実験データからケモメトリックス手法を用いて経験的（統計的）に物性推算をする方法はもっとも一般的な方法である。一つの物性についての重回帰分析ぐらいであれば、Excel の分析ツールでも十分実用可能である。しかしながら、現実の材料設計では、満足しなければならない目的物性は一つではなく複数物性あるのが通常である。この場合、ＰＬＳやニューラルネットワークを用いれば、複数の目的変数（物性）を同時に計算するモデル式（経験式）を作成することができ、物性推算するときに有効である。

４.２　逆解析による候補材料の設計

複数の目的物性を同時に満足する材料を探し出すことは容易ではない。しかしながら、ケモメトリックス技術を駆使した逆解析を行えば候補材料を設計（提案）することも可能になってくる。Chemish Pro での逆解析の流れは次のようになる（図４）。

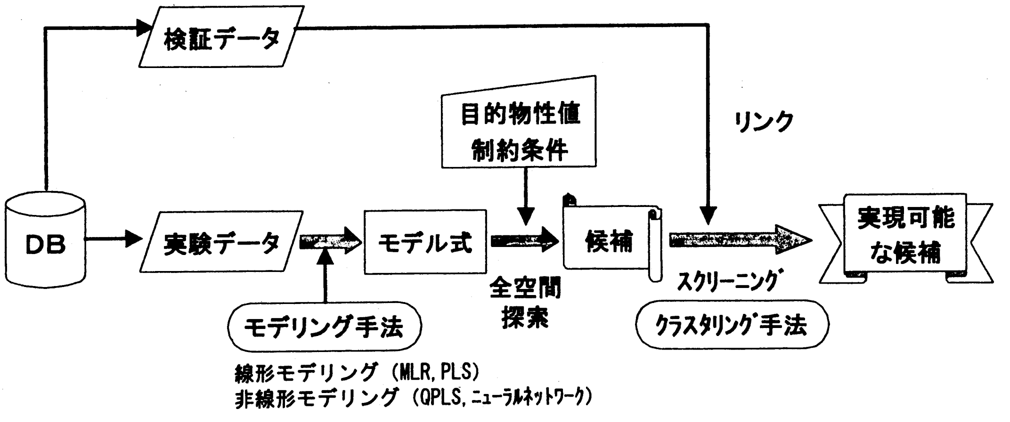


図４　逆解析のフローチャート

①モデリング手法を用いて複数の目的物性を予測するモデル式を作成
②目的物性値および各種の制約条件を入力
③全空間探索により条件を満たす候補の創出
④ＤＢを用いた検証やクラスタリング手法で候補をスクリーニング

全空間探索（③）では、複数の目的物性値を満足する候補を網羅的に創出するため、制約が甘ければ相当数の候補が発生する。この候補をスクリーニング（④）するために、モデル式作成に用いた実験データをカウンタープロパゲーションネットワークで学習させ、その応答曲面をコホーネンマップ上に投影し、実験データを性能別にクラスタ分けする。そして、このクラスタ分けされたコホーネンマップ上に候補を発火（マッピング）させることにより、その候補の優位性を判断する。つまり、このクラスタリング手法を用いると、元の実験データ特性を考慮した性能が高い候補を絞ることが可能になるわけである[6]。

図５にＰＰアロイ材料設計における逆解析候補のクラスタリング例を示す。性能クラス別（上、中、下）の応答曲面投影図から、このデータがうまくクラス分けされていることがわかる。性能クラス：上のマップ上に候補を発火させた図では、全空間探索により創出された候補が広く分布しており、更に絞り込める可能性を示している。

粗視化分子動力学法の概要と応用

（財）化学技術戦略推進機構 名古屋大学・土井プロジェクト　青柳　岳司
aoyagi@zoom.cse.nagoya-u.ac.jp

材料設計に求められるメソスケールシミュレーションの一つのアプローチとして、分子構造を粗視化した単位で取り扱い、分子鎖のダイナミクスをシミュレーションする、粗視化分子動力学法があげられる。本稿では粗視化分子動力学の考え方、モデルの解説、および粗視化分子動力学シミュレーションの高分子材料への適応例を幾つか紹介する。

1. はじめに

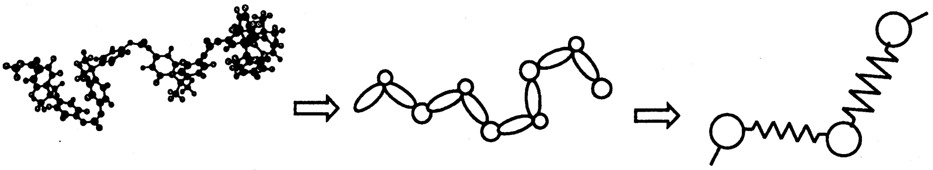
高分子の分子動力学シミュレーションは 1990 年頃より広く用いられるようになり、弾性率、ガラス転移点温度、低分子の拡散挙動等様々な機械的、熱的物性の予測が行われ、高分子材料設計に寄与してきている。しかしこのような原子レベルのシミュレーションは、たとえ古典的でシンプルな運動方程式を解く分子動力学においても、計算量は膨大なものになり、取り扱える高分子の分子量、分子数は限られたものになる。現実的には分子鎖間のからみ合いをあらわせるような分子量の分子を複数存在させて計算することは不可能に近い。

そこで数原子、あるいは数－数十モノマーユニットを一つの構成単位として、その単位がつながった様な簡略化した鎖を考え、その運動をシミュレーションするという“粗視化分子動力学”という手法が、高分子の構造とダイナミクスを調べるために用いられる。以下に粗視化分子動力学に用いられるモデルの解説と汎用粗視化分子動力学プログラムの開発、および応用例に関して述べる。

2. 粗視化分子動力学のモデル

粗視化分子動力学とは先に述べたように、数－数十モノマーユニットを一つの単位（例えば一つの球状のビーズ）としたモデルを用いて計算を行うが、どれくらいの原子数を粗視化の一単位とするかにより様々な粗視化のレベルが考えられる。最も原子レベルに近い粗視化としては、メチレンユニット（CH_2）を一つの球と置き換えるようなユナイティッドアトムモデルがある。さらに粗視化を進めていく際、図 1(b)に示すようにミクロレベル（原子レベル）の形状、排除体積、フレキシビリティーを出来る限り忠実に再現しようとするレベルでの粗視化や、(c)のように、ビーズースプリングモデルと言われる、質点が単純なバネでつながれたガウス鎖に近づく粗視化が考えられる。

粗視化分子動力学とひとくくりで言っても、これらの粗視化のレベルを、シミュレーションの目的に応じて使い分けることが重要である。材料の弾性率やガラス転移点温度を半定量的に求めようとす



(a) ミクロレベルモデル　(b)分子の特徴を残した粗視化　(c)ガウス鎖に近づく粗視化

図 1 粗視化分子動力学モデルの例

れば、原子レベルにおけるシミュレーション、あるいは分子構造の特徴を残した、粗視化のレベルの低いモデルを用いる必要がある。一方、溶融高分子のダイナミクスのように、分子鎖のからみ合いなど、多数の分子の相互作用を考慮することが必要であるが、反面どの様なモノマー構造であっても、ある程度共通の挙動を示すような場合は、ビーズースプリングモデルのような粗視化の程度の高いモデルを選択することが出来る。

3. 汎用粗視化分子動力学エンジン COGNAC の開発

このように粗視化分子動力学にも様々なモデルがあり、実際にシミュレーションを行おうとする場合、現状では各研究者が各自の研究に応じて特化した用途のプログラムを作成していて、一般に普及している粗視化分子動力学ソフトウエアというものは存在しない。

この粗視化分子動力学を含めた、いわゆるメソスケールのシミュレーションエンジンの開発を一つの目標として、1998 年度より、通産省の出資・補助により、新エネルギー・産業技術総合開発機構（NEDO）により「高機能材料設計プラットフォーム」というプロジェクトが立ち上げられ、（財）化学技術戦略推進機構が委託を受けて大学連携型として名古屋大学内で実際の研究·開発が始動した[1]。

このプロジェクト、通称「土井プロ」において開発が進められている、汎用粗視化分子動力学エンジン COGNAC(<u>CO</u>arse <u>G</u>rained molecular dynamics program by <u>NA</u>goya <u>C</u>ooperation)は種々の粗視化モデルに対応したポテンシャル関数を標準で持っている。またこのプログラムは C++言語で開発されており、若干のプログラミングの知識が有れば、ユーザーによるポテンシャル関数の拡張も容易に行えるように設計されている。分子動力学に用いられる種々のアンサンブルも各種装備されている上、拡張することも可能である。それに加えてニュートン運動方程式のみでなく、ブラウニアンダイナミクス等の各種運動方程式へも柔軟に対応できる構造になっている。

現段階ではユーザーインターフェースなど未完成な部分が多く、プロジェクト期間中は一般に公開するレベルには至らないと考えられるが、2002 年 3 月のプロジェクト終了時点での公開を目指して開発が進められている。

また、これら粗視化分子動力学の適応例として高分子結晶化、液晶相転移、高分子電解質、熱可塑性エラストマー等、種々の研究が現在プロジェクトにおいて進められている。

4. 粗視化分子動力学の応用

以下に、これまで検討を行ってきた粗視化分子動力学の応用例、特に粗視化レベルの高いビーズースプリングモデルを用いた検討例を紹介する。

4.1 せん断流動下における高分子鎖の挙動

溶融高分子のせん断粘度、および流動下における分子配向を知ることは高分子材料の加工成形性を検討する場合に非常に重要である。特に分子配向を定量的に実測、あるいは理論的に予測することは困難である。そこで我々は粗視化分子動力学を用いて、せん断流動下における溶融高分子の挙動に関して、シミュレーションを行った。

ここで用いたビーズースプリングモデルは Kremer ら[2]が提案したモデルで、式(1)(2)に示すようにすべてのビーズ間に斥力 U^{LJ} を考慮する。それに加えて結合しているビーズ間には FENE 型のポテンシャル U^{ch} を与えて結合長が安定な長さを持つようにしている。

このビーズースプリングモデルからなる溶融状態の高分子の集合体に一定速度 $\dot{\gamma}$ でせん断変形を加え、応力 σ をシミュレーションより求めると、せん断粘度 η は $\eta=\sigma/\dot{\gamma}$ の関係より求められる。

$$U_{ij}^{LJ} = \begin{cases} 4\varepsilon\left[\left(\dfrac{\sigma}{r_{ij}}\right)^{12} - \left(\dfrac{\sigma}{r_{ij}}\right)^{6} + \dfrac{1}{4}\right], r_{ij} \le 2^{1/6}\sigma \\ 0 \qquad\qquad\qquad\qquad\qquad , r_{ij} \ge 2^{1/6}\sigma \end{cases} \quad (1)$$

$$U_{ij}^{ch} = \begin{cases} -0.5kR_0^2 \ln\left[1-\left(r_{ij}/R_0\right)^2\right], r_{ij} \le R_0 \\ \infty \qquad\qquad\qquad\qquad , r_{ij} \ge R_0 \end{cases} \quad (2)$$

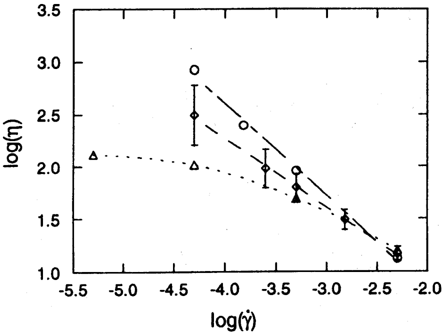


図 2 粗視化分子動力学によるせん断粘度のせん断速度依存性
重合度 N=100(△), 200(◇), 400(○)

図 2 に重合度 N=100,200,400 単分散高分子におけるせん断粘度のせん断速度依存性を示す[3]。計算結果はすべて無次元量で示しているが、実際の高分子の慣性半径、拡散定数などでスケールすることにより、実際の高分子の定量的な粘度を予測することが出来る。なお、ここでもちいたモデルにおいて、からみ点間重合度は 35 程度となることが知られており、ここで計算している重合度はからみ合いの存在する系である。今回計算した範囲ではほぼ $\log(\eta)\sim\log(\dot{\gamma})$ の Power-Law が成り立つ領域であることがわかる。さらにせん断速度を下げていくとニュートン領域が観察されると予想されるが、低せん断速度において定常状態に達するまでには、非常に計算時間がかかるため、現実的にはこのような粗視化モデルでも困難である。

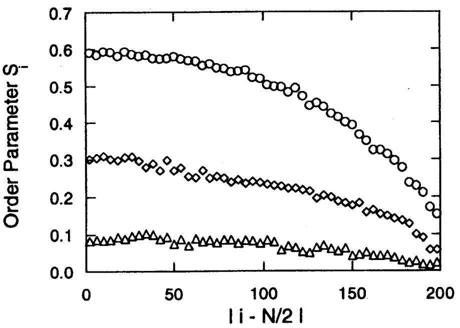


図 3 せん断流動下における結合ベクトルの配向度パラメータ S_i。鎖の端から i 番目の結合ベクトルの配向度パラメータを分子鎖中央からの位置 |i-N/2| を横軸に取りプロットしている。N=400, $\dot{\gamma}=5.0\times10^{-2}$ (○), $\dot{\gamma}=5.0\times10^{-3}$ (◇), $\dot{\gamma}=5.0\times10^{-4}$ (△)。

また、このような分子動力学シミュレーションを行った場合、粘度のようなマクロな諸量と同時に分子鎖の構造変化、配向状態を解析することが出来る。例えば個々の結合（スプリング）ベクトルのオーダーパラメータを調べると、分子鎖のからみ合いの効果が強く現れるため、図 3 に示すように、分子鎖の中央部分の配向が大きく、比較的からみ合いの影響が少なく自由に動ける末端部分の配向は小さくなると言う結果が得られる。

4.2 壁間にはさまれた溶融高分子の挙動

上記の例はバルク状態における溶融高分子の挙動に関してシミュレーションした例であるが、狭い壁間に拘束された場合、同じ溶融高分子においても挙動の変化が現れることが知られている[4]。このような狭い空間内に拘束された分子の挙動はナノスケールの微細加工、トライボロジーなどの観点から興味が持たれる。そこで我々は上記と同様のビーズースプリングモデルの粗視化分子動力学シミュレーションにより壁による拘束の効果を検討した[5]。

計算条件は重合度 N=40 という比較的短く、からみ合い影響の無い系を用いて、高分子の慣性半径 Rg($\sim3.2\sigma$: σ はビーズのファンデルワールス直径)に比較して 1 桁程度大きい 20-50σ の間隔に壁を配置している。また壁と高分子鎖の相互作用も、引力がある場合、反発力のみの場合など数通りシミュレーションを行っている。

シミュレーション結果の例として図 4 に式(1)で定義される Normal modes X_p の自己相関関数 $\langle X_p(t)\cdot X_p(0)\rangle$ の p=1 のモードを $\exp(-(t/\tau_R)^{\beta})$ でフィッティングすることにより求めた緩和時間 τ_R を

示す。

$$\mathbf{X}_p = \frac{1}{N}\sum_{i=1}^{N}\mathbf{r}_i \cos\frac{p\pi(i-1)}{N-1} - \frac{1}{2N}\left[\mathbf{r}_1 + \mathbf{r}_N\right] \qquad (3)$$

ここで N は一分子あたりのビーズ数、$\mathbf{r}_i$ は i 番目のビーズの座標である。

この結果より壁との相互作用が大きい場合(ε=5.0)、壁との引力により分子鎖の運動が拘束を受けτ_R が長くことがわかる。しかし同時に壁との相互作用の違いに依存せず Z=30~40σ あたりの距離からτ_R が大きく変化するという結果が得られ、高分子のサイズに比較して、かなり広い間隔から壁の効果が現れることが、シミュレーションにより明らかになる。

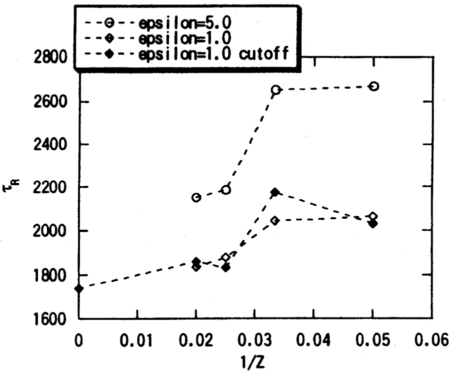


図 4 Normal modes Xp(p=1)の自己相関関数より求めた緩和時間 τ_R の壁間隔依存性

5. おわりに

今回紹介したような粗視化分子動力学法は、高分子というサイズの大きい分子の構造、およびダイナミクスを現実的な計算時間でシミュレーションすることが出来る有力な方法である。しかし当然、モノマーレベルにおける分子構造など、粗視化により失われる情報も存在する。同時に効率的とは言っても、現実的に扱うことの出来る粒子数は限られているため、実際の材料を扱う時間やサイズのスケールをシミュレーションすることは不可能である。そのため、よりミクロなレベルにおけるシミュレーション、およびマクロなレベルにおけるシミュレーションを組み合わせて使いこなすことにより材料開発に真に貢献できる計算機シミュレーションが可能になるであろう。

謝辞

本研究の一部は、通産省工業技術院の出資・補助を受け、NEDO が（財）化学技術戦略推進機構に委託した、大学連携型産業科学技術研究開発プロジェクト「高機能材料設計プラットフォーム」（通称「土井プロ」）の下で行われたものである。

参考文献

[1]青柳岳司,土井正男, 高分子, **48**(5), 316 (1999).
[2] K. Kremer and G. S.Grest, *J.Chem.Phys.*, **92**(8), 5057 (1990).
[3]青柳岳司,豊島日出夫, 高分子学会予稿集, **44**, 3251(1995).
[4] Y.Cho, H.Watanabe and S.Granick, *J.Chem.Phys.*,**110**(19),9688 (1999).
[5]青柳岳司,庄司達也,澤史雄,福永宏雄,滝本淳一,土井正男, 高分子学会予稿集, **48**, 3562 (1999).

あおやぎ　たけし　AOYAGI, Takeshi
1998 年 8 月、土井プロジェクト始動時より旭化成工業（株）から出向中。
連絡先　〒464-8601 名古屋市千種区不老町名古屋大学共同教育研究施設 1 号館 4 階
（財）化学技術戦略推進機構　名古屋大学土井プロジェクト
Tel 052-789-5828 FAX 052-789-5831

シミュレーションによる高分子の粘弾性予測

(財)化学技術戦略推進機構(土井プロジェクト)　田﨑弘恭
tasaki@zoom.cse.nagoya-u.ac.jp

高分子材料の開発を進める上で、耐衝撃強度や耐熱性といった材料自体の性能と同時に、その材料の成形加工性、すなわちいかに少ないエネルギーで成形できるか、人手をかけずに連続成形できるか、そもそも寸法精度良く製品になるのかといった点が重要なポイントとなる。この時、高分子の構造から直接的に性能を予測することができれば、要求性能を発現させるために的確な材料設計をおこなうことができるはずである。現在コンピュータシミュレーションにより、分子鎖のアーキテクチャや分子量分布から高分子のレオロジーを予測する試みを実施しているので、その現状を紹介させていただく。

1. はじめに

近年、重合技術の進歩、例えばメタロセン系触媒の実用化等により、分子鎖のアーキテクチャ、すなわち分岐の有無、数、長さといったパラメータや分子量分布のパターンに特徴のある高分子材料が工業的に得られるようになってきている[1]。長鎖分岐や分子量分布を制御した、機械的性質と成形加工性の両面に優れた高分子材料を設計する方向性の一つとして期待されている。

従来より高圧法による低密度ポリエチレンやCr系触媒で重合された高密度ポリエチレンのように長鎖分岐を有する高分子や、多段リアクターでの重合やブレンドによって分子量分布を制御した高分子材料が生産されてきた。これまでに成形加工性と高分子のレオロジーとの関係には多くの知見が蓄積されている。しかし分子鎖のアーキテクチャや分子量分布から直接的にレオロジーを予測することは非常に難しく、高分子の構造と成形加工性との関係は必ずしも明確になっていない。長鎖分岐や超高分子量成分のキャラクタリゼーション自体も単純な問題ではない。また、高分子の構造は製品の機械的性質とも密接に関係しているため、材料を設計する上で、更に問題を複雑にしている。

このような状況において、多種多様なアーキテクチャを持つ任意の高分子を自由にシミュレーションする技術を開発し、高分子の構造からそのレオロジーを予測することができれば、材料開発をおこなう上で、有用なツールとなりうるものと考えられる。

2. モデル

高分子溶融体中の分子鎖の運動を記述できるモデルとして、土井-Edwardsの管模型(チューブモデル)[2]がある。このモデルでは注目する分子は、絡み合ったまわりの分子によって形成される管の中をレプテーション運動し、徐々に古い管を抜けて新しい管に移動する。レプテーションとは分子鎖が管に沿って１次元的に拡散する運動である。

この管模型の物理的描像の有効性はすでに広く確認されていて、応力緩和のダンピング関数は定量的に正しく予言される。しかし、ずり粘度の分子量依存性など、定量的に実験と異なる結果を与える場合も知られている。また、ずり速度が速くなると応力が減少する領域があるなどの定性的な困難もある。これらの困難を解決するため、いくつかの改良が提案されている。その中でも重要なのは、Contour length fluctuation(管にそっての分子鎖長のゆらぎ)とConstraint release/creation(絡み合いの相手の分子が運動することにより絡み合い点が消える/生じる)の２つの効果である。

高分子溶融体のレオロジーにおいては、絡み合いに起因する非常に長い緩和時間の運動が重要であり、従って計算時間上の制約から、通常の分子動力学(MD)シミュレーションは用いることができない。そこで上述の２つの効果を共に取り入れた管模型を用い、さらに任意の分子量分布が扱えるシミュレーション手法が滝本(山形大工)、土井 (名大院計算理工)によって考案された。

fig.1 に示すような一本の分子鎖を考え、他の分子鎖との絡み合い点間を結ぶベクトルを $\mathbf{r}_j$(j=1,2,...,n)とする。この分子を拘束する管は $\mathbf{r}_j$ を結んだものと考えられる。管の長さを L_t とし、分子鎖がこの管の両端から出ている長さを s_1,s_2 とすると、分子鎖の全長 L は $L = L_t + s_1 + s_2$ で表される。

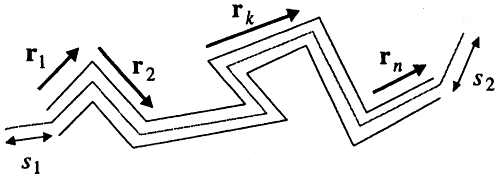


fig.1　シミュレーションモデルにおける分子鎖(primitive path)と管(tube)の模式図

高分子溶融体をこのような分子鎖の多数本の集合と考える。そして各々の分子鎖上の絡み合い点を互いにランダムに対応させる(パートナーを形成する)。分子鎖間の相互作用は、この絡み合い点の対応関係としてだけ取り入れられ、各分子鎖はそれぞれを拘束する管の中で独立に運動する。

具体的には、各タイムステップにおいて各分子鎖に対し以下の操作を適用する。

①流動による管の変形

試料にマクロな流動が加えられている場合、それに従って絡み合い点の位置を移動させる(管のアフィン変形)。

②Contour length fluctuation

分子鎖の全長 L を平衡長 $L_{eq} = Za$ のまわりでランダムに増減する。a は絡み合い点間の平均距離、Z は M/M_e、M は分子量、M_e は絡み合い点間分子量である。

$$\frac{dL}{dt} = -\frac{1}{\tau_R}\left(L(t) - L_{eq}\right) + \dot{L}_{affine} + g(t) \quad (1)$$

ここで $\tau_R = \tau_e Z^2$ は分子鎖のラウス緩和時間、τ_e は分子鎖が管による拘束を感じ始める時間。また $\dot{L}_{affine}$ は流動による変化速度で、$g(t)$ はランダム力である。

③レプテーション

分子鎖(の重心)を管に沿ってランダムに動かす。拡散係数 D_c は(2)式で与えられる。

$$D_c = \frac{1}{3\pi^2}\frac{a^2}{\tau_e}\frac{1}{Z} \quad (2)$$

④Constraint release/creation

②③の結果、分子鎖の末端が管の末端の絡み合い点の内側に入ったら($s < 0$)、その絡み合い点とそのパートナーとを取り去る。あるいは逆に分子鎖の末端が管の末端から長さ a 以上出たら($s > a$)、新しい絡み合い点とランダムに選んだ分子鎖上にそのパートナーとなる絡み合い点を追加する。

3. 計算例

これまでに分岐の無い直鎖状高分子に対し、ずり粘度の分子量依存性やずり速度依存性、瞬間歪を与えた場合の応力緩和等を正しく予測することが可能になっている。単分散の各成分からなるブレンド系として、多分散系を取扱うことも可能である。以下に計算例を紹介する。

3-1 定常ずり流動

分子量(Z)、ずり速度($\dot{\gamma}$)を変化させて定常ずり流動のシミュレーションをおこなった。

分子量が十分に大きい時、実験ではゼロずり粘度 η_0 は分子量の 3.4～3.5 乗に比例することが知られているが、シミュレーションでは Z の 3.5 乗に比例する結果が得られ、実験の結果を再現している(fig.2)。

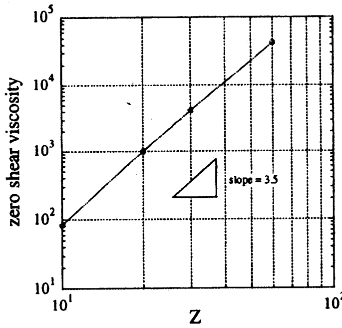


fig.2 ゼロずり粘度の分子量依存性

3-2 応力緩和

系に瞬間的に歪γを加え、

$$G(t,\gamma)=\sigma_{xy}(t,\gamma)/\gamma \quad (3)$$

で定義される非線形緩和弾性率 $G(t,\gamma)$ の時間変化を調べる(fig.3)。ここで$\sigma_{xy}(t,\gamma)$ はずり応力である。

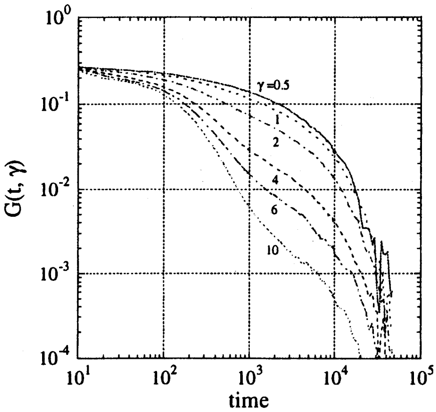


fig.3 歪を変化させた時の非線形緩和弾性率

長時間部分では、$G(t,\gamma)$が時間に依存する項と歪に依存する項とに分離できることが実験、理論により良く知られている。すなわち

$$G(t,\gamma)=h(\gamma)\cdot G(t) \quad (4)$$

ここで、$h(\gamma)$はダンピング関数、$G(t)$は線形緩和弾性率である。シミュレーションから求めたダンピング関数を土井-Edwards の管模型による理論値と合せて示した(fig.4)。シミュレーションの結果は理論値に近い値を示している。

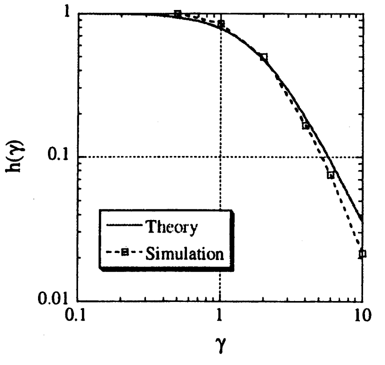


fig.4 ダンピング関数(理論値との比較)

緩和弾性率のフーリエ変換から得られる複素粘度$\eta^*(\omega)$とずり粘度とを比較した結果、シミュレーションにおいても実験的に知られる Cox-Merz 則が成立していることが確認できた(fig.5)。

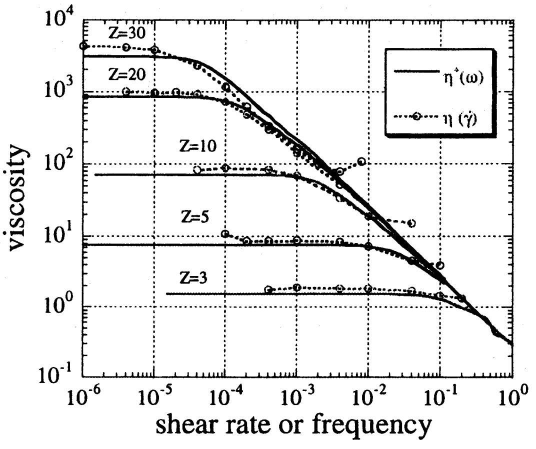


fig.5 Cox-Merz 則

また、分子量の異なる直鎖高分子の混合系に対する計算もいくつかおこない、ほぼ良好な結果を得ている。緩和弾性率のフーリエ変換から求めた貯蔵弾性率 G'、損失弾性率 G''について、分子量分布の狭いポリスチレン(PS)のブレンド実験[3]と比較した結果を fig.6, 7 に示す。各成分の分子量及びその設定値を table 1 に示した。実験値と比較するために H-PS の G'が実験値と一致するように縦、横のシフトファクターを決めた。

table 1 ブレンドの成分

	Experiment		Simulation
H-PS*	Mw	581,000	Z 32
L-PS*	Mw	58,700	Z 3 (=Mw/Me**)

*H(or L)-PS :High(or Low) molecular weight PS

** Me :18,000

4. まとめ

最終的な目標は、高分子材料のレオロジーをシミュレーションで予測し、分子量分布や長鎖分岐の制御による成形性に優れた材料の設計を支援することにある。これまで述べてきたよう

に、分岐の無い直鎖状高分子に対しては予測が可能になってきた。伸長粘度などの成形性にとって重要な物理量の計算も可能であり、これらを含め現在実験との広範囲の比較を進めている。次のステップとしては、分岐高分子の中で最も単純な構造で比較的実験データも豊富な、星形高分子への拡張を予定している。更に、商業的に生産されているような広い分子量分布や長鎖分岐構造を有する高分子にも対応できるように検討を進めているところである。

我々とは別のグループでも、レプテーションモデルに基づく確率論的な方法による高分子のレオロジー研究が進められている [4]。また解析的な取扱いが可能な新しいレプテーション理論が提案されてきている [5]。

材料設計において、シミュレーションによる高分子のレオロジー予測のように、計算機の果たす役割は今後益々重要になってくる。

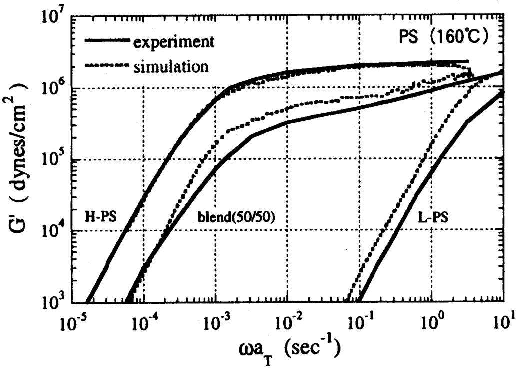


fig.6 貯蔵弾性率(G')の周波数依存性

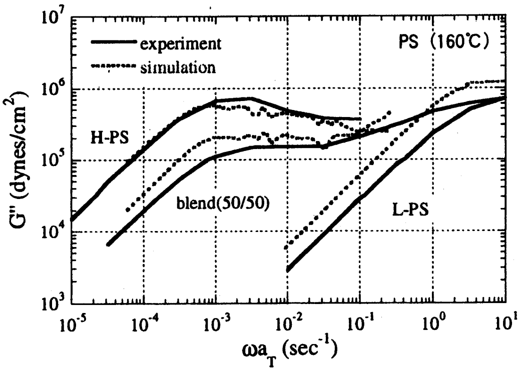


fig.7 損失弾性率(G'')の周波数依存性

参考文献

1)小松公栄, 小野進, 今泉文武, メタロセン触媒でつくる新ポリマー　工業調査会

2) M.Doi, S.F.Edwards, The Theory of Polymer Dynamics(Oxford University Press, New York, 1986)

3) T.Masuda, K.Kitagawa, T.Inoue, S.Onogi, Macromolecules vol.3, No.2, 116, 1970

4) C.C.Hua, J.D.Schieber, J. Chem. Phys., vol.109, No.22, 10018, 1998

5) D.W.Mead, R.G.Larson, M.Doi, Macromolecules vol.31, No.22, 7895, 1998

※本研究は、通産省工業技術院の出資・補助を受け、NEDO が(財)化学技術戦略推進機構に委託した大学連携型産業科学技術研究開発プロジェクト「高機能材料設計プラットフォーム(通称「土井プロ」)」の下でおこなわれたものある。

たさき　ひろやす　TASAKI, Hiroyasu

日本ポリオレフィン(株)から現所属に出向しています。これまでは、半絶縁性 GaAs(LEC 法)単結晶育成技術開発やポリオレフィン系材料(ポリエチレン、ポリプロピレン)開発等、主として材料開発に携ってきました。今後ともよろしくお願いいたします。

連絡先　〒464-8601　名古屋市千種区不老町　名古屋大学　共同教育研究施設 1 号館 4 階

TEL 052-789-5828　FAX 052-789-5831

◆部会記事

第22回情報化学討論会の報告

工藤喜弘　ykudo@eiemx.yz.yamagata-u.ac.jp

標記討論会が平成11年11月10日から三日間、山形県米沢市のグランドホクヨウ米沢の大ホールを仕切りで二分して日本薬学会主催の第27回構造活性相関シンポジウムとの併催の形で開かれた。構造の世話人（東北大学坂本尚夫薬学部長）と連携しつつ労作業の大半は情報側が担当した。金谷重彦助教授監修のもと、木ノ内誠博士が発表環境の整備、佐藤典子さんが事務を含む残り一切を仕切り、研究室の学生全員がそれを助けた。参加者（会費納入）数と研究発表数は下表の通りである。

参加者数	合同	情報化学			構造活性相関			発表数	総数	情報	構造
	合計	予約	当日	計	予約	当日	計	計	103	72	31
総数	264	84	28	112	115	37	152	特別	2	1	1
一般	236	70	22	92	107	37	144	招待	4	3	1
学生	28	15	5	20	8	0	8	口頭	41	27	14
懇親会	115	38	9	47	48	20	68	ﾎﾟｽﾀｰ	56	41	15

熊本（第２０回）で米沢開催案が再び打診されたときは実は一抹の不安があった。というのは討論会は初期のころは方法の真の創始者や自称創始者の間の学問的生命を賭けた真剣勝負の場であったが、いまは「どれを選ぶかの時代」（藤原厳氏）に変っているからである。座長として議論をけしかけると穏やかに纏めるのが役目だと非難されるようになった。最先端技術の導入はきわめて重要であるが、討論会でそういうものの紹介ばかりが並べば、単純な応用を研究と言おうものなら袋叩きにあうという空気の中で物心がついた世代としてはつらい。二番煎じは単純に無知によるものはまだ許せるが、五番列車の機関士に線路をちょっと補修させて路線開発者と煽てる方とそれにのる機関士の方、他人の作図を著書や講演に無断で使う方、（さらには）他人の業績を自分のもののように錯覚させるためにマネーならぬ論文ロンダリング手口を使う方、など、自分の業績が汚される痛み、悲しみを知らない方の噂を聞くたびに、先達の業績を尊重すべきという他学協会がしたような啓蒙も手遅れかも知れないなどと落ち込んだが、そういう方の応募を受けるのもつらい。また、最先端の強力ではあるが自分の発明でない方法―――主成分分析、ニューラルネットワーク、ケモメトリックス、遺伝アルゴリズム、など―――を意味もなく適用しているだけだから、 内容がまだ薄いから、討論会は遠慮せよとは前科があるから言えない。前科とは応

募した短い講演の部の講演数が少ないから討論会に吸収したいという飯塚幹事の急の要請にうかつにのったことである。事情を知らない参加者から図々しいと顰蹙を買ったが言い訳はできない。このように心配事がたくさんあったので昨年１２月に米沢開催が本決りになると、討論会が変身して「欧米産の教義を宣教するためだけの会」になってしまう悪夢を見たのである。

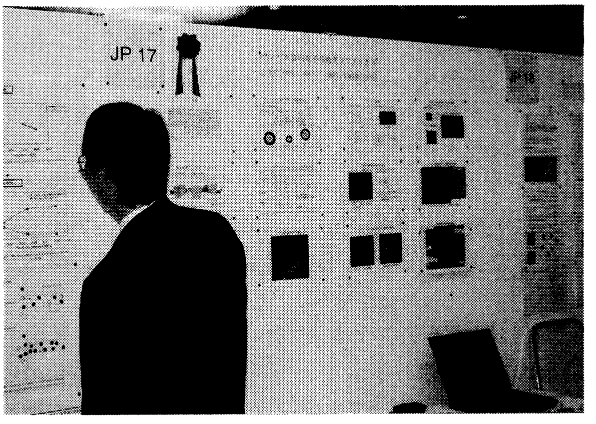
ところが奇跡が起きた。吉田部会長以下、幹事諸氏が奮い立って操縦桿をはっきり上昇側へ切った。それが「要旨を A4 版２ページにすること」、「ミキサー開催」などの決定となって現れた。要旨２ページはProceedings 発刊を前提としていたから構造活性側だけが直前に４ページも可と変更しても動揺しなかった。ミキサーでは「討論会や電子ジャーナルを仲良しクラブに堕落させるな」、「指導者はレビューを纏めよ」、など参加者約五十名の意見が尽きず、予定の１時間では三戸副部会長は終了を宣言できなかった。レビュー云々は他人に頼るなと反論されたが、真意は「過去の研究事例を簡潔に系統的に紹介することによる二番煎じの防止」であった。電子ジャーナル発刊は、多くの方々、中でも米田幸夫部会長(当時)、が強く推進した夢の少し姿を変えた形での達成である。来年の京都ではProceedings や電子ジャーナルの論文の独創性、図の色彩の豊かさなどは言うまでもないことでもっと高度のことが話題になるのだろうか？ 楽しみである。

幹事の皆様の活躍の分だけ暇になった世話人に残されたことは、討論会の成功と部会の将来性を祝福する方法を謡曲金剛流の師範に考えていただくことぐらいであった。その回答が懇親会での３人の師範による厳粛で華やかな三謡一舞であった。最後に、このような意義あるものとなった討論会を担当する機会、小さな学会にさえ出向くのが難しい学生諸君が最先端の討論の雰囲気にどっぷり浸る機会、を与えてくださった関係者の皆様に感謝する。三件のポスター賞の内の一つに我が研究室のものを選んでいただけたのは望外の喜びである。小出しの小うるさい質問に丁寧に応対して下さった第 20 回の松岡俊和、第 21 回の米田茂隆の両先生にも篤く御礼申し上げる。なお好評だったインターネット接続は米沢では一般市民が気軽に買える市販技術である。

第２２回情報化学討論会　ポスター賞発表

（講演番号順）

JP 17　タンパク質内電子移動グラフィックス（九工大情報工）佐藤 洋子・上田 尚学・佐藤 文俊○柏木 浩

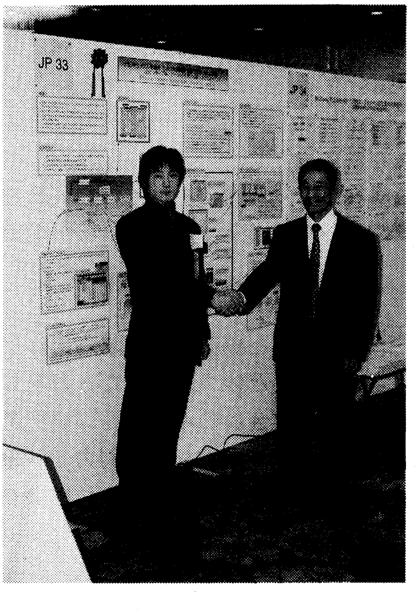


（左下）

JP 33　ケモメトリックス手法を用いた材料開発支援プログラムの開発(3)－カウンタープロパゲーションを用いた候補のクラスタリングー（豊橋技科大）○棚田 東作・荒川 正幹・船津 公人（三井化学）西村 竜一

（右下）

JP 39　学習順序に依存しない自己組織化法の開発（山形大・工）○阿部 貴志・金谷 重彦・工藤 喜弘・木ノ内 誠・大平 賢



第２２回情報化学討論会特別企画
ミキサー「化学における計算機利用の明日を考える」の報告

広島大学大学院理学研究科化学専攻　相田美砂子
maida@ipc.hiroshima-u.ac.jp

平成１１年１１月１０日から１２日まで開催された第２２回情報化学討論会の初日の特別講演終了後を、本討論会の初の試みであるミキサーにあてました。あまり行き届いた宣伝ができなかったので、どのくらいの方々が特別講演終了後残ってくださるか心配していましたが、約５０人の方々が参加してくださり、大変有意義な会にすることができました。参加してくださったみなさまに、ここであらためてお礼を申し上げます。

ミキサーは午後５時４５分から一時間の予定でしたが、議論が白熱したため予定を大幅に超過し、終了したのは午後７時すぎでした。次のような話題が提供されました。討論会担当の工藤先生の御理解により飲み物とおつまみも準備されて、討論全体も盛り上がりました。

1. 情報化学部会活動に関するアンケート（1998 年実施）の結果説明。
   企画担当幹事の高田氏よりアンケートの結果および解析結果の説明があった。部会員の関心ある分野が広がっていること、ホームページの活用拡大の必要性、また討論会のレベルアップへの要望があったことも報告された。
2. 情報化学部会による電子ジャーナル刊行企画の紹介。
   担当幹事の高木氏よりまずおおまかな説明があり、その後活発な意見の交換があった。情報化学討論会の Proceedings を発行すること、そしてそれを電子ジャーナルとすること、に関してはみなさんが賛成してくださったようだ。電子ジャーナルはまだ世界的にも歴史が浅いため、いろいろな問題も生ずるであろうが、とにかく前向きに取り組んでいく、という点では合意していたと思う。
3. 実験化学者と計算機利用専門家との関わり。
   今年から情報化学部会の幹事に加わった坂本氏からまず御自身の経験をふまえた、コンピューターシミュレーションのメリットについての意見が提示され、その後、情報化学部会の役割や情報化学討論会の質についての議論が活発に交換された。情報化学部会あるいは情報化学討論会を基盤として進展した研究のレビューを CICSJ Bulletin でとりあげていくことも提案された。
4. グリーンケミストリーの紹介。
   物質研の北尾氏から、グリーンケミストリーの紹介と部会としての参入の勧めがあった。CICSJ Bulletin でとりあげて広く部会員へ紹介していくことが提案された。

ミキサーは、情報化学部会として、また、情報化学討論会としての初の試みでしたが、今後もこのような企画（飲み物付きで！）を続けてほしい、という意見もだされました。今回出された多くの意見をふまえたうえで、情報化学部会が会員のみなさんの研究あるいは業務の推進のうえで大きなメリットがあるような組織にしていくための努力を、みんなでしていきたい、と思っております。今回のミキサーは、情報化学部会の三戸副会長と編集担当幹事の相田が担当いたしました。

第12回情報化学講習会を終えて

平成11年11月12日（金）の午後、情報化学討論会の余韻が残る中、第12回情報化学講習会を盛会の内に終えることができました（参加者６５名）。

さて前半は「計算化学のすすめ－電子スペクトルのMO計算を例にして」というタイトルで西本先生からご講演がありました。「若い研究者に夢を持ってもらえるようなお話」をという幹事の勝手な要望にお応えいただき、先生の名著「量子化学のすすめ」同様に、数式をあまり使わずに量子化学のエッセンスをわかりやすくご説明いただきました。「アウトプットを理解できなきゃ ただのゴミ（若い研究者向け）」及び「Computational Chemsistry 馬鹿にしてたら こけまっせ（企業トップ向け）」というユーモアのあるお言葉にはこの分野の本質が感じられました。また、パソコンをお持ちになられて、種々の分子の光吸収の特性をその場で計算してお示しになられました。電子スペクトルのMO計算がとても身近なものに感じられる時代になったことを痛感いたしました。

後半は石山先生から「ケモメトリックス入門」というタイトルでご講演がありました。ケモメトリックスの手法である多変量解析については一つ一つ丁寧にご説明していただきました。参加者の皆さんもデータを解析・料理する機会が多いと思いますが、今回のご講演で頭の中の整理ができたのではないかと思います。「仮説の明確な構築」、「手法の基礎的な理解」、「データの信憑性」、「結果の慎重な検討」等、ケモメトリックスを使う上での本質についても聞くことができました。

今回始めて情報化学討論会との併催という形を取りましたが、討論会に参加していただいた方々には、「参加しやすかった」とのご意見をいただきました。また時間の制約のため、「時間をもっと長くして欲しかった」とのご意見もいただいております。計算手法の一般的な紹介等は他所で開催している講習会等もたくさんありますので、部会員の皆さんの関心の高いテーマを中心に今後の企画を考えていきたいと考えております。企画等についてのご意見を随時皆様からいただければ幸いに存じます。最後に、お忙しいところ米沢まで来てご講演いただいた西本先生、石山先生にこの場をお借りして厚くお礼申し上げます。

（文責・髙田 章）

新任の挨拶

科学技術振興事業団　文献情報部　　　　相馬　　融

今年の３月から編集委員として加わることになりました。私はデータベース作成という、アカデミックとはかけ離れたかなり泥臭い世界で長年仕事をしており、文献データを加工処理する作業は行ってきましたが、Bulletin を編集する立場というのは初めての経験であり、どこまでお役に立てるか心許なく感じているところです。既に半年以上も過ぎ、今更新任の挨拶というのも間が抜けてはいると思いますが、半年過ぎた視点で感じたこと書いてみたいと思います。

現在年６回 CICSJ Bulletin を刊行しているわけですが、これにあわせて二ヶ月に一度の割合で編集委員会が開催されます。この編集会議での話題が、次号の内容をどうするかよりも「情報化学部会をいかに活性化させるか」について多くの時間が費やされていることは、この部会に対する歴代の役員の方々の思い入れが、現在の役員に引き継がれてきているからではないかと云う思いが強く致しました。毎年部会の活性化についての方策を考えて来られていると思いますが、やはり内側から見ないと分からない部分があります。今回の情報化学討論会において特別企画のミキサー「化学に於ける計算機利用の明日を語る」を実現させたり、講習会を同時開催したことは、４月から重ねて来た議論の成果と云えるでしょう。

もう一つ以前より検討課題となっていた活性化の方策として、討論会の Proceedings あるいは論文誌の発行がありました。部会全体及び討論会の活性化ばかりでなく、論文発表の機会を提供できれば特に若手研究者の求心力なると思われます。ただし従来のような冊子体での刊行はそれなりの手間と経費が必要であり、実行に移すためにはかなりの準備が求められます。欧米ではだいぶ以前より普及してきている「電子ジャーナル化」という動きが、最近我が国においてもにわかに風雲急を告げて来ております。冊子体を作成しないで電子ジャーナルのみの発行というのは国内ではあまり例がないようですが、少なくとも経費の面はかなり軽く実現できる状況になって来ています。このあたりで何かお役に立てることがあればと思っているところです。

こうして書いてくると、全体的に現役員会の自我自賛というか手前味噌的内容ばかりのようですが、これはまだ私が新参者で半ば部外者として眺めているからであり、諸先輩の方々に対する敬意の表れとしてご容赦ください。これからは部会員からの厳しい評価を受ける側の一員として、微力ながら協力していきたいと思いますので、よろしくお願い致します。

部会役員会報告

平成11年度第2回役員会議事録

日　時：平成11年11月10日（水）11時～12時30分

場　所：グランドホクヨウ米沢（山形県米沢市）

出席者：吉田元二、時田澄男、三戸邦郎、伊藤眞人、髙田　章、有國　尚、松岡俊和、久保百司、工藤喜弘、船津公人、相田美砂子、相馬　融、坂本直紀、中馬　寛、清水弘一

議事内容

（1）第22回情報化学討論会について

工藤委員より準備状況について報告があった。

三戸副部会長よりミキサーの準備状況について報告があった。

ポスター賞について検討し、従来どおりとすることになった。

（2）第12回情報化学講習会について

髙田委員より準備状況について報告があった。

（3）論文集の電子化刊行準備について

準備WG主査の時田副部会長より進捗状況について説明があった。JSTのJ-Stageシステムを活用する方向で、前向きに進めることが承認され、ミキサーで検討し、総会で報告することとなった。

（4）第23回情報化学討論会について

藤田眞作氏（京都工織大）に世話人をお願いし、2000年10月26～27日に構造活性相関シンポジウム（世話人は京大の赤松美紀氏）と共に京都リサーチパークで開催することになった旨、吉田部会長より報告があった。

（5）総会について。

吉田部会長より総会の議事次第が紹介され、承認された。

総会を年1回、討論会会期中に行うよう部会規定第17項を改定することが提案され、承認された。

（6）その他

・有国委員より春年会プログラム委員会での春年会のプログラム編成について報告があり、部会員に申し込みを勧めるよう要望があった。

・船津委員よりCICSJ Bulletinの原稿の集まり具合について報告があった。

・平成12年度役員人事の準備について報告があった。