

TM-1296

ICOTで実施した技術開発の位置付けと
今後の動向について

内田 俊一

May, 1994

© Copyright 1994-5-11 ICOT, JAPAN ALL RIGHTS RESERVED

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03)3456-3191~5

Institute for New Generation Computer Technology

ICOT で実施した技術開発の位置付けと今後の方向について

内 田 俊 一
ICOT 研究所

1994 年 5 月 10 日

1 第 5 世代コンピュータ技術が開く新しい 情報処理の世界

1.1 第 5 世代コンピュータ・プロジェクト のはじまり

第五世代コンピュータプロジェクトの話は、昭和 53 年くらいに、電子政策課から電子技術総合研究所の測さんを中心とするグループに「第 5 世代コンピュータ」という名前をつけた革新的な研究開発プロジェクトをやりたいという話しが来たことに始まる。

中身を詰めるために元岡先生を委員長とする委員会が組織され、まず調査活動を行った。約 3 年間の活動の後に描かれた「第 5 世代コンピュータ」の中身を、左側にある丸 3 つの図で示した。(資料 1-p.2)つまり、知識ベースを用いた推論を中核とするコンピュータを作り、それを基にして知識情報処理を新たにやってみようということである。そのベースとなるアーキテクチャとしては高度並列処理を使うということだった。

プロジェクトが開始され、10 年余りに渡る関係者の協力と担当者の苦闘の末に右側に示したプロトタイプシステムが完成した。中身の説明は後に譲るが、一番下の並列ハードウェアから実験的な応用まで網羅するという、一つの新しいコンピュータの体系ができあがった。(資料 1-p.2)

このシステムは、プロジェクトの後期ではネットワーク経由で、Unix の世界とつながれ、国際共同研究により国際社会にデビューした。しかし、つながったとはいっても、Unix 系の汎用機の文化とは異質な、進歩した世界が離れ小島のようにできたと言える。1,000 台規模の並列マシンを駆使した並列知識処理の世界が忽然と出現したわけだから、世界の研究者は、目を見張った。しかし、当時のジャーナリストやメーカの技術者は、すぐには、その意義がわからなかった。今、

市場が分散、並列の世界へ急速に転換する状況となり、やっと、その意義が理解され評価が高まった。

1.2 第 5 世代コンピュータ・プロジェクト の計画(目標、予算、期間)

前期、中期、後期の各期の期間、目標、予算を一覧にした。後期は、当初の予定より約 2 ヶ月延びた 6 月に、国際会議を開き、お披露目をしたが、結局 1 年伸ばすことになり、最終的に 11 年間頑張った。これに、さらに後でお話する 2 年間の後継プロジェクトが続くこととなった。(資料 1-p.3)

1.3 第 5 世代コンピュータ・プロトタイプ システムの開発過程

最終目標である 1000 台規模の並列推論マシン(PIM)を、一足飛びに目指すことはできなかった。ほぼ、ゼロからのスタートであり、まず、人の育成と技術的蓄積の両面から取りかかった。(資料 1-p.4)

最初は、左側上に示されている逐次論理型言語を作り、逐次型推論マシン PSI というのを作った。これは大型の電気冷蔵庫くらいのサイズで大きかった。1 台で Prolog(プロログ)という論理型言語に関しては、当時の大型汎用機と同じ性能を発揮した。マルチウィンドウシステムなどの新しいマンマシンインタフェースを持ち、大型機をパーソナル化したという意味合いをもっていた。これを 100 台くらい作った。

さらに、その第 2 版として PSI-II を開発した。経験の蓄積がものをいい、カスタム LSI を作り、コンパイラの性能を上げたので、性能はほど 10 倍アップした。また、体積は 6 分の 1 に縮小された。PSI-II は 300 台くらい作り、研究者に使わせた。これが成功したので、図の右側に示すように、中期の中頃に、PSI-II の CPU を要素プロセッサとするマルチ PSI を作った。要素プロセッサは、CPU ボード 2 枚、メモリボード 2 枚

という構成だった。マルチ PSI はそれを 1 プロセッサとして 64 台つないだものとした。(資料 1-p.4)

マルチ PSI の上に、前期に設計完了していた並列論理型言語 GHC をベースとして、それをさらに発展させた実用的な言語 KL1 を実装した。その後、若手の研究者が育ち、彼らの情熱で、並列 OS である PIMOS と小規模な応用ソフトができた。マルチ PSI は PIM の実験機であり、この基本ソフトウェアを含むその完成は、技術的には 1 つの大きなエポックであった。また、これは知識処理の本格的な並列処理としては世界初のものであった。

これによって知識処理における並列処理の有効性を示し、世界の研究者から非常に評価された。また、この成功によって、我々担当者は、1000 台規模のプロトタイプシステムの開発に向けての技術的な諸問題は解決できるとの確信を深め、後期も走り続けることができた。

後期は、辛抱強い説得の後にメーカーの全面協力を得ることができた。しかし、PIM のためのチップ作りでは苦労した。このときのチップ作りで使った、0.8 ミクロンの ASIC の技術は各社においては、私の思っていたより未成熟で、結局 1 年くらい余計に時間がかかった。このために、全体計画が遅れてしまった。

しかし幸いなことに、マルチ PSI を 4 セットくらい作って、その上で進めていた並列応用ソフトウェアの開発は順調であった。また、KL1 言語が優秀だったので、PIM がぎりぎり、運び込まれてきたにも拘らず、256 台や 512 台のマシンができたときに、応用ソフトウェアは 64 台からすぐに規模を拡張し、256 台や 512 台の規模の PIM 上で順調に動いた。

並列ソフトウェア作り自身は、このように KL1 の優秀性のおかげで、まあ順調だったが、PIM のような大規模知識処理マシンの能力を発揮するような本格的なデモ用の応用問題をどのように捜し出すかについては苦労した。これは、結局、海外との共同研究によって、見つけ出し芽出しをした。

(写真 1:PIM/p)

この大きな箱は PIM モデル p の写真である。1 箱に 32 台入っている。コンピュータは、今はだんだん小さくなってきているが、並列マシンにするとかなり大きい。使用したデバイス技術は、4 年くらい前のものなので、現在の RISC 型のマイクロプロセッサより、クロックは少し遅い。しかし、知識処理に関しては最大規模で、今も相変わらず世界一速いマシンである。

(写真 2:PIM/m)

次のモデルは PIM モデル m である。4 筐体が並ん

で、128 台、これが 2 列あり、256 台のフル構成となる。

1.4 第 5 世代コンピュータ・プロトタイプシステムの構成

これが、第 5 世代コンピュータ・プロトタイプシステム構成図である。一番下には、5 つのモデルからなる、約 1000 台規模の並列ハードウェアがある。一番大きいモデルは、PIM/p と書いてあるクラスタ構成のものだ。8 台の要素プロセッサを共有メモリを接続して、一まとめにしたクラスタがある。クラスタ間は、ダブルハイパーキューブでつないでいる。(資料 1-p.7)

クラスタとは、会社のひとつの課のようなもので、1 つの課に 8 人の従業員がいて、共有メモリというのが会議用のテーブルになる。そのテーブルごとに、直接 8 人がアクセスするので、この 8 人の中では非常に速く通信ができる。しかし、隣の部屋にいくときには、郵便袋に手紙を入れて送らなければならない。通信が、このような二重構造になっている。この 8 人の中では、自動負荷分散される。仕事のあるうちは勝手に 8 人が手分けして仕事をやる。

クラスタ間は、パケット通信で、外の廊下を渡るようなものだ。これは、人間がプログラムで、コントロールする。このマシンの長所は、高速のコミュニケーションだ。欠点は、共有メモリが、いっぱいになるとガーベジコレクションといって机の上のゴミ片付けをするが、そのときに 8 人全員が休まなければならないところだ。

その右にある PIM モデル m は、16 × 16 の 2 次元のきれいな碁盤目構造になっていて、社員の一人一人が個室に入っているイメージだ。隣の部屋にいくときも手紙を郵便袋に入れて送るといったパケット交換をする。通信速度は相対的に遅くなるが、この方が構造がユニフォームなので、ソフトウェア開発者はこちらを好む人が多い。

これらのモデルは、どれも同じ言語 KL1 を実行し、OS も共通で PIMOS である。応用のモデルとアーキテクチャのモデルの適合については、当初は、ぴったり合う方が良く、合わないが遅くなるとも思った。

しかし、プログラムを書いてみると通信の頻度というのはそんなに多くないので、その影響は思ったより、小さかった。結局、知識処理の応用では、ネットワークはそんなに頑張らなくてもいいという結論になった。知識処理の応用は、数値計算の偏微分方程式を解くのととは大抵違う。

知識処理は、送る情報はそんなに多くない。だからデータパラレルの数値計算とは大分性質が違ふ。もちろんネットワークの通信速度は速いに越したことはないが。

並列ソフトは、並列OS、PIMOSが載っていて、その上にKL1等が載り、基本ソフトウェアを構成する。ハードウェアもそうだが、ソフトウェアも、この部分までは、かなり真面目に作っており、製品になるくらいのクオリティをもっている。そのために、KL1やPIMOSをUnixの載った超並列マシンに移植するだけで、かなり実用になる段階にある。

それに比べて、この上の「知識処理技術」の部分は、研究試作物的である。その中身は、定理証明、自然言語ツール、知識表現言語、また、その上のいろいろな応用システムであるが、みな研究試作物のレベルにあり、すぐに製品化できるレベルではない。(資料1-p.7)

しかし、研究という観点から見ると非常によい要素技術をとらえている。知識表現言語は、演繹的な言語とオブジェクト指向がうまくマッチングしている。言語処理系としては未成熟だが、「法的推論システム」といった法律のエキスパートシステムなどで用いている。新しい機能はあるが、実用的には、まだ不十分で、今後も研究が必要だ。「VLSI-CAD」はかなり実用的に使われてもいる。新しいところ、例えば「プログラム生成支援」は定理証明に基づいて、プログラム合成をやっている。研究としては、面白いソフトである。

将来の知識処理の研究材料としては、どれも非常に重要なので、これらをプロジェクト終了後、どのように育てていくかは、特に検討しなければならない。

1.5 知識処理の高速化

プロトタイプシステムの特徴としては、まず、高速性が挙げられる。世界に先駆けて、プロセッサ台数にほぼ比例した高速化が達成されている。応用問題に並列性が十分あれば、使用した台数の8割くらい、場合によっては9割程度の高速性が達成できる。(資料1-p.8)これは、今、世の中にある100MIPS級マシン上のC言語で書いたソフトと比べて、100倍くらい速い。

一方、プロセッサ一台あたりで比較してみると次のようなことがわかる。一台あたりでは、KL1言語で書いたプログラムをC言語のプログラムと比べると、5倍ほど遅くなっている。すなわち、PIMのプロセッサ5台分で、最新のワークステーション、たとえば、

SunのSS10の1台と同じになり、5台以上なら、それをしのぐ。

5倍遅いというのは、ハードの遅さが約2倍、KL1という言語のオーバーヘッドがまた、2倍と言うことだ。従って、後で述べるKLICという処理系で、KL1言語をUnixマシンに移植する2倍くらい遅くなるだけだ。

2倍という数字だが、現時点では、2倍なら非常に少ないという認識である。超並列マシンというのは、今売っているものでも60台、100台の規模があり、KL1で書けば台数比例で速度向上が得られるのに対して、Cで書いたらそんな何十台もあっても、うまく使えないので、圧倒的にKL1が有利となる。

1.6 汎用並列ソフトウェアの高い生産性 (KL1のメリット)

KL1に限らず、PrologやLISP系の言語の全てに言えることだが、プロセス間の同期は、データフローと違って、演算すべきデータが到着すれば直ちに人間の介入なしに計算が始まるという自動的な機構を使っている。バグが出ない。メモリは、ガーベジコレクションという機構が仕組みられているので、ダイナミックに割り当てられる。(資料1-p.9)

さらに、KL1という言語を使うとは、最初にプロセッサの数が無限にあるという仮定でプログラムを書ける。そこで、アルゴリズムを決めてプログラムしてしまう。それから、逐次型マシンの上でバグをとって、アルゴリズムの正しさを確認する。それから、並列マシンの上へもっていても、答えは正しく出ることが保証されている。

第1段階で、答えの正しさが保証されていれば、第2段階で、分割し、並列プロセッサに仕事を割り当てするときに、下手なやりかたをしても、速度は遅くなくても、答の正しさはやはり保証される。このように、プログラムの正当性が損われないということがミソだ。

今コネクションマシンなどの市販の超並列マシンが普及し始めていて、C++言語などでプログラムを書く人おり、いろいろコメントをいただく。C++で知識処理のプログラムを書くと、第2段階における変更・修正によって、第1段階のアルゴリズムの正当性がくずれ、答えが合わなくなるとのこと、これによる作業の手戻りが大変ということだ。

結局、これを避けるために、数値計算のような、きっちりしたアルゴリズムの応用しかやらなくなってしまう

うという。その意味でも、KLI や PIMOS のメリットは評価されている。Unix の上に移植すれば、また多くのユーザがつくだろう。

1.7 知識処理の新しい応用例

大規模並列処理のいろいろな応用を搜したが、知識処理の問題は、その前のノイマン型マシンが、あまり得意ではなかったために、大規模な応用ソフトウェアの研究開発が未発達だった。そのために、並列推論マシンをフルに使うような大規模な応用が準備されていなかった。結局、新しい応用をも自分達で探し、開拓することとした。(資料 1-p.10)

1.7.1 分子生物学データの配列解析への応用

遺伝子の情報処理をやることにした動機はアメリカでヒトゲノム・プロジェクトだ。これは、年間 200 億円くらい予算がつき、生物学者が、どんどんゲノムを解析して、その遺伝子のデータベースが莫大に蓄積した。その解説に、一種の略号解説のような処理をしなければならぬという需要があった。これを、日米共同研究でやろうとした。これにより、国際貢献もアピールできると思った。

ところが、共同研究の考え方や日米の研究風土の違いがあった。アメリカでは、研究者同士が、非常にコンペティティブだ。一方、こちらは、新しいことを教えてやるのだから、提供するシステムに、多少ムシがいてもいいという構えでいた。結果的に、彼らと組んでやると、このような考え方は、彼らの足を引っ張ることになった。初めの頃は、けんかになった。やっぱり向こうは、命懸けなのだから、こっちも命懸けにならないといけないという教訓を得た。

また、アルゴンヌ国立研究所のオーバヴィークさんという研究員が非常に乗り気で共同研究をやってくれて、話が進展し新聞で取り上げられた。アルゴンヌと ICOT が、遺伝子情報処理で共同研究をやっているという日本を評価する記事を書いてくれた。こちらは喜んだが、なんと、彼がクビになった。日米の技術摩擦のせいで、あちらの政府関係者で、面白くないと思った人がいたらしい。HPCC プロジェクトの成立前夜で、日本と共同研究していると、このプロジェクトの予算をもらいにくくなると思った人がおり、やめさせたかったらしい。研究予算獲得のコンペティションの激しさといい、このような政治的な背景を持つ叩き合いなど、米国の風土の激しさを実感した。

今、ICOT のやっている遺伝子情報処理は、まだブ

リミティブで、生物学者がためこんだ DNA や蛋白質の配列に対して、類似性の解析というのをやって分類しているレベルだ。もう 1 つ、一生懸命やっているのは、知識表現言語の応用としてやっているもので、主要な化学的、生物学的な反応や関連物質を知識表現言語で記述して、知識ベースを作っている。代謝反応や酵素反応などの知識ベースができると、遺伝子に書かれている蛋白質の合成などの指令を解説することが容易となる。(資料 1-p.15)

これは、まだ、知識表現言語の研究的な意味合いが濃くて、まだ生物学者に役立つレベルにはいっていないが、このようにしていくと、最後には、試験管の中でやっていた細胞の増殖などの実験をコンピュータの上で、シミュレーションできるのではないかという期待がある。今は、まだ、夢物語りだ。

1.7.2 並列定理証明への応用

この問題は、規模の小さいプログラムで膨大な計算を発生できる。前所長の測さんが、この専門家であることから、自らプログラムを書いた。これは、PIM のベンチマークとして非常に有効だったばかりでなく、スピード的にも世界一速かったから、これまで証明されていない定理が、いくつか証明できた。それで、高い評価を受けて、担当者が、人工知能国際会議で最優秀論文賞をもらったりした。(資料 1-p.11)

プロセッサ台数に比例して速度が向上しているグラフを挙げたが、これはトップデータだ。このような高い効率を得るのは、まだ大変であり、工夫がいる。200 台、500 台規模の並列処理となると各プロセッサのどれがどのくらい忙しく動いているかというのは、すぐには分からない。(資料 1-p.12)

したがって、この手のモニタリングツールが不可欠である。

(ここでカラーの OHP シートを提示)

この図は、PIM の動作を示しており、ブルーのところは暇で、赤いところが忙しいということだ。(資料 1-p.13) 縦軸は、0 番から、255 番までのプロセッサの番号である。横軸は、2 秒毎に時間軸がついている。この図では、最初は、0 番のプロセッサだけが赤くて、そこからメッセージが飛んでいくと、徐々に、忙しいプロセッサが増えていく。ここのリードタイムを、短くすればするほど速く立ち上がるわけだが、最初はこんなふういきなりには立ち上がっていかない。

プログラムの分割や割当をうまくやると、このように、最後は真っ赤になる。この真っ赤になっている段

階では全員が忙しいので、200倍というような高いパフォーマンスになる。(資料1-p.14)

これは遺伝子を解析したときのプログラムの実行例だが、これでは処理に周期性があるので、全プロセッサが一度の忙しくなり、答えが出ると、ブルーと黄色のところで、それらの解答をマスタープロセッサが引き集め、再評価して、新しい問題をバラ撒いて、また、いっせいに走る。そのため、このように赤い帯が規則的に出てくる。

こういうのは比較的アルゴリズムが規則的で、負荷発生パターンの予測がつくので、チューンナップが容易だ。一方、定理証明のようなものは、予測が難しいのでなかなか大変だ。チューニングには、最初のバージョンでは3ヶ月くらいかかった。

しかし、1度、こつが分かってしまうと、チューニング期間が、1週間になり、3日になるといった状況だ。今、定理証明の人達は、あまり赤くするよりも、探索の枝の数を少なくする方に一所懸命だ。無駄な枝まで捜せば、計算量はどんどん増えるのだから、赤くするためには、その方がよいが、枝を狭めれば、早く答が出る。

1.7.3 自動プログラム合成

これは、並列マシンのパワーで可能になったが、従来はできなかった。発電プランとの制御エキスパートシステムのKL1プログラムを、論理をベースとした仕様記述言語で書き、自動合成した。デバッグ期間が大幅に短縮されるという特徴がある。

1.7.4 法的推論

これは、並列定理証明システムをエンジンとして、過去の判例を表現して知識ベースに入れておく。そこに新たな事件を投入すると、その事件の被告が、どんな罪を得るかということを、類似性検索で検索をする。検索は、知的なもので、概念辞書等が入っている。自然言語処理の技術な要素が使われている。EDRの辞書なども参考にした。ICOTの成果の中では、これが一番、インテリジェントなシステムだ。(資料1-p.20)

1.8 第5世代コンピュータプロジェクトの成果に対する評価

第5世代プロジェクト11年の成果が、まとまることの分かった最終段階で、この成果をどうするかという

議論が起こった。そのために、評価作業が、通産省の諮問委員会で行われた。

明らかになったことは、1つは、学術的には、非常によい成果だということ、もう1つは、どのメーカーもすぐに製品化するのは、無理だということだ。これがはっきりしてきて、どうしたものかということで、いろいろな議論がおこった。

1.8.1 2年間の後継プロジェクトとUnix汎用機への移植

たまたま、その頃、市場では、メインフレーム技術が、分散処理や並列処理の技術に、シフトを始めていた。それが、思いがけなく急激だったため、評価の議論をしているうちに、並列の時代がくるという認識がどんどん高まっていった。

そこで、PIMの上でしか動かないKL1やPIMOSを、Unixベースの汎用機に移し、広く普及させることを目指す2年間の後継プロジェクトをやることになった。その中で、知識処理のソフトウェアもさらに磨きをかけ、Unixマシン上に移すことになった。実際には、マンマシンインタフェースをMotiefのようなX-Windowベースのミドルウェア使って使いやすくするようなことも実施している。(資料1-p.26)

このようにして新たに開発したソフトウェアは、第五世代プロジェクトのソフトウェアと同様にフリーソフト(ICOT無償公開ソフトウェア)として無償配布するということになった。

1.8.2 Unix化の対象とするハードウェア

KL1のプログラムの移植対象となるマシンは、27ページにあるようなものだ。KL1プログラムの規模が小さければ、C言語の動くパソコンやシングルプロセッサのUnixマシンでも可能で、規模が大きければ、たくさんプロセッサが接続されたUnixベースの並列マシンだ。(資料1-p.27)

最近、米国のオークリッジ国立研究所などでは、多くのプロセッサを含む分散システムを、1個の集合した並列マシンに見せかけるミドルウェア作られている。それは、パラレルバーチャルマシン(PVM)と呼ばれ、パブリックドメインソフトウェアとして公開されている。それを用いれば、ICOTのシステムは、分散システムの上にも載せることができる。遺伝子情報処理をやっている人の中には、分散したワークステーションをたくさん持っており、こういうので使いかたをサポートしようとするメーカーも出てきたりしてい

る。

1.8.3 第五世代技術と既存のコンピュータ技術との融合化

KL1 言語が C 言語にコンパイルされ、Unix マシン上で動くようになったことにより、KL1 は、分散システム上で、他の言語で作られたソフトウェア結び付ける役目もできることになる。

分散システムの個々のエレメントのソフトウェアモジュールは、C 言語だったり、フォートランだったりするが、それらの間を KL1 で書いてつなぐことができる。必要なソフトウェアを全て KL1 で書かなくてもよくなり、既存のシステムの拡張に便利に使える。

1.9 第5世代コンピュータプロジェクト以降の知識処理の展開

知識処理技術の将来を考えると、この図のようになることが予想される。まず、並列マシンが普及するようになると、知識表現言語を新たなプログラミング言語として使う時代が来ると思われる。(資料 1-p.29)

ユーザは、論理をベースとする知識表現言語で、自然科学系の応用のみならず、社会科学系、人文科学の応用も書くようになるだろう。そのときには並列記号処理システムは、下層のエンジンの役割になる。大規模並列処理のパワーがそのまま論理型の知識表現のパワーアップに使えるような技術が第5世代のおかげでできたことが、このような新しい展開を可能とする。

2 第五世代コンピュータの研究基盤化プロジェクトとその後の展開

ICOT の中で、今、議論しているのは、一番左側だ。第5世代プロジェクトの成果に、☆が印が3つある。上の2つが、先ほどからいっているハードウェアと KL1 を含む大規模な並列処理の技術、下の方が知識処理の技術だ。上グループと下グループに属する成果では、その後の発展の仕方が違う。(資料 2-p.2)

2.1 KL1 と PIMOS の今後の展開

研究基盤化プロジェクトによって、KL1 とか PIMOS は、1、2年後には、世の中にある Unix ベースの汎用機の上に移っていく。この汎用機上の KL1 処理系を KLIC(クリック)と呼んでいる。

普及は順調で評価も高い。KL1 という言語の素性の良さからすると、KL1 を元祖として、その子孫となる並列言語が生まれることも考えられるくらいだ。ともかく、KL1 を使った応用が、世の中にどんどん出てくるような方策をとることが必要である。

2.2 知識処理の成果の今後の展開

知識処理技術の方も筋はよいし、これからますます重要となる研究テーマだ。しかし、その成果を今後どのように普及し発展させていくかは、また、工夫を要する。

第五世代プロジェクトで開発したソフトウェアは、ICOT 無償公開ソフトウェアとして、オープンしており、多くの人が持っていていっている。覗きにきた人が、まず、1400 人位いて、その中で実際にソフトウェアを持っていた人は 1200 ~ 1300 人だ。

その国籍は、日、米、その他で、それぞれ3分の1くらいずつである。だから、各々、400 人、400 人、400 人位ということになる。日本は、その 400 人のうちほとんどが大学で、メーカは 10 社くらいしかない。

しかし、アメリカは、400 人のうち半分以上がメーカだ。当然のことながら、インテル、モトローラ、ヒューレットパッカード、シンキングマシン、IBM といった、並列マシンをやっているところが持っていていっている。

KLIC もまだリリースしてない期間の数字であるので、これらの持っていた人は、自分のマシンでは動かないソフトウェアを持っていたことになる。とりあえず、プログラムをテキストとして読んでいるのではないと思われるが、これから、汎用機上の処理系 KLIC をリリースすると、今度は本当に動くようになる。

日本のメーカは、現在、並列処理へ動きに乗り遅れている。しかし、日本でもそういうことが起こらないと問題だ。当初の第5世代プロジェクトの成果が世界に広まるという意味ではよいのだが、国内にも広めたので、何とかしたいものと思っている。

2.3 海外はインターネットを介して第5世代のリソースを利用

アメリカや海外で、第五世代のソフトウェアが迅速に広まる理由は、さっきのネットワーク、すなわち、インターネットだ。第5世代プロジェクトの国際共同研究でもそうだったが、ICOT のマシンとソフトがイ

インターネットにつながった瞬間に、彼らは ICOT のソフトとマシンが、自分のワークステーションの後につながったと考えた。

ある特定の計算処理だけは、ICOT のマシンを使っ
て行ない、ほかのマン・マシンインタフェースなどは、
自分のマシンの上におく。例えば、それまでアルゴン
ヌのスパコンを使っていた処理を、太平洋を越えて日
本の ICOT のマシンにすげかえるようなことが簡単に
できる。これは、オーストラリアの研究者にとっても
そう。定理証明のエンジンとしてだけ PIM を使っ
て、マン・マシンインタフェース、その他は、みんな
自分のマシンでやる。

しかし、ICOT の PIM は、512 台規模のもの、256
台規模のものの 2 システムしかない。そのため、外部
の多くの利用者に使わせることは無理だ。

それが、KLIC によって、KL1 言語や ICOT のソ
フトウェアが、市場にある Unix ベースのマシンに載っ
ていくと、この制約がなくなる。自分のもっているソ
フトを、全部 KL1 に作り変えなければならないのなら、
誰も KL1 を使わないだろうが、超並列マシンを使っ
た方が有利な部分のみを KL1 に書き直して使うの
なら、多くの人がやる。

超並列マシンは、これからいろいろな所で使える
ようになる。そのマシンが遠くにあっても、自分のマ
シンから、インターネットというネットワークに介し
て、特定部分の処理のみをその遠くのマシンに投げ、
答を戻すことができる。このような利用形態がこれか
ら増えていき、KL1 の出番も多くなるだろう。

2.4 KL1 が Unix マシンに載ると ICOT の 知識処理ソフトのプログラムの利用度 が高まる

KL1 が C 言語にコンパイルされ、Unix 環境で動作
可能となると、ICOT の知識処理のプログラムも、い
ろいろ人のプログラムに手軽に組み込めるようになる。
アメリカでは、多くの並列 Unix マシンが、ネット
ワーク経由で、研究者に使えるような状況となってお
り、ICOT の知識処理のソフトウェアの利用もおおい
にメリットがある。

それに比べてわが国は、メーカーの研究者や大学の
間で、相互にマシンやソフトウェアを利用するよう
になっていない。それに、インフラが貧しい。もっと
オープンにやれないかと思う。

2.5 アメリカでやっているよい意味での研 究のコンペティション

アメリカの研究は、オープンでかつコンペティティ
ブである。例えば、ICOT のソフトを使ってよい仕事
をすると、その人に多くの予算が付き、その人の研究
がさらに発展する。また、ICOT のソフトウェアも評
価が高まる。しかし、発展した人のかげには、必ず予
算を削られて泣いている人がいる。

この評価は、アニュアルコンファレンスのようなこ
とをやって、研究者同士がお互いに、たたきあいをや
る。予算を配分している責任者が、この叩き合いをき
いていて、成果の査定をやる。

我々の場合は、パートナーがアルゴンヌ研究所だっ
たので、エネルギー省の傘下で、遺伝子関係の研究を
やっている人が集まった。ワークステーションを並べ
て、できたソフトの品評会をやる。品評会の善し悪し
のほか、他の人に利用されたかどうかなども、評価の
大きなポイントとなる。

これにより、予算が削られたり、削られなかったり
するというようなコンペティションがある。彼らにと
って研究費の削減はクビ切りにつながるもので、すご
く真剣だ。そこまで真剣にやるのが、日本人によい
かどうかは分からないがそういうコンペティションの
結果できてくるソフトウェアは、なかなかのものとな
る。

なお、米国の場合、ナショナルプロで国の予算が入
るとできたソフトは、パブリックドメインとなる。ネ
ットワークがあるということと、このように制度的に
公開し競争させることで、先進的なソフトウェアを
生み出している。この競争に生き残ったソフトウェア
が、さらに磨かれて商品となる。これらは、競争力
がある。このようなメカニズムは日本にはない。

2.6 日本にもフリーソフトが開発される情 報インフラの整備が必要

米国ほど、激しくなくてもいいが、日本でも、少な
くとも大学の先生が書いたソフトやナショナルプロの研
究員の書いたソフトをもっとパブリックドメインとし
て流通するような制度を作る必要がある。

パブリックドメインのソフトウェアを組み合わせ
て、優れたシステムを作った人があつたら、どこの省
庁からでも、研究予算がつくとか、何らかの競争原理
の働く枠組が必要と思う。そうでないと米国の研究者
に勝てない。

また、先端ソフトウェアが米国にやられると後に続く、アーキテクチャやハードウェアもいずれ、産業レベルまで席卷されてしまう。このような枠組みも、新たな研究インフラの上部構造として、整備されるなら、その先に、先端的な情報処理産業が確立するだろう。

並列のKLICが定着して、それに合うようなハードが求められるなら、並列推論マシンのアーキテクチャの成果が役立つだろう。RISCのチップを作る研究は、そろそろ飽和している。これからのアーキテクチャを作ろうとすれば、新たな並列言語に準拠したプロセッサを作ろうということになるだろう。そのとき、KL1やその上のソフトウェアの研究成果が役立つ。

2.7 ICOT サテライト研究グループの試み

2枚目にICOTサテライト研究グループというのがある。ICOT OBを中心にして、KL1や第5世代技術を支援するグループを作っている。(資料2-p.6, p.7)

通産省電政課と相談して、ICOT OBで大学に出でいった方をまとめている。これがサテライト研究グループで、ICOTがセンターにいて、瀧先生を代表者とし、現在11の大学研究者グループを集めた。

そこに、少額の委託研究を出して、KL1の普及やソフトウェアの開発を行なっている。先生方には、いろいろな大学の事情はあるだろうが、ICOTフリーソフトウェアを使ってソフトを作ってくださいと依頼している。作ったソフトは、公開を原則としており、ICOTのフリーソフトが拡大再生産するようにしている。

海外のオレゴン大学やブリストール大学にも、同じような枠組みで研究を委託している。ICOTフリーソフトウェアと、同じ条件で公開できるよう交渉して相手を説得した。ICOT無き後も、こういうグループを何とか維持できる工夫を電子政策課にお願いしている。

ICOTのグループだけでなく、これをもっと広げた形で、センターのようなものを作るとか、そういう中にこういうのもはめてもらうというように、先導的な例に使ってもらうのもよい。ここにはメーカーの人も入っているのだから、そういった人を結ぶネットワークは絶対必要だと思う。

2.8 研究基盤化プロジェクトの終結

4番目は、研究基盤化プロジェクトの終結についてだ。4月から平成7年の3月末まで、あと1年はあ

るのだが、なにしろ、これまでが、大プロジェクトであつたので、作ったものが多く、処分だけを考えてもなかなか大変である。

今年の12月までで、研究開発活動は終了し、12月に最後の成果披露の国際シンポジウムをやることとしている。(資料2-p.8)

これまでのFGCS国際会議のように論文を公募して、学会的な研究発表会をやるのは、縮小した今のICOTではできない。そこは代りに、6つくらいのテーマのワークショップを開催しようとしている。最初の2日間は、後継プロジェクト成果のお披露目をやる。

研究所のクローズは諸設備等の撤収が結構大変なので、来年の1月に入ったらすぐに取り掛かり、3月末に人を返してから、さらに、数カ月頑張ることになるだろう。

海外共同研究は、覚書を交換している正式のものが5件ある。これらは6ヶ月前に予告をして終了するという契約になっているので、9月くらいに終結処理を始める。これも、研究員を派遣してワークショップなどをやりながら、お礼をしつつ、徐々に区切りをつけるつもりだ。

最後に残った問題はICOT無償公開ソフトウェアの配布サポートの継続だ。ICOTの研究所としてはクローズするが、フリーソフトは最後に残る。ファイルサーバーなどの機器を含めてしばらく維持する必要がある。そのために、サテライト研究グループとネットワークを含めたサーバの類、場合によっては一部の並列マシンをどこかで維持するような方策を検討している。

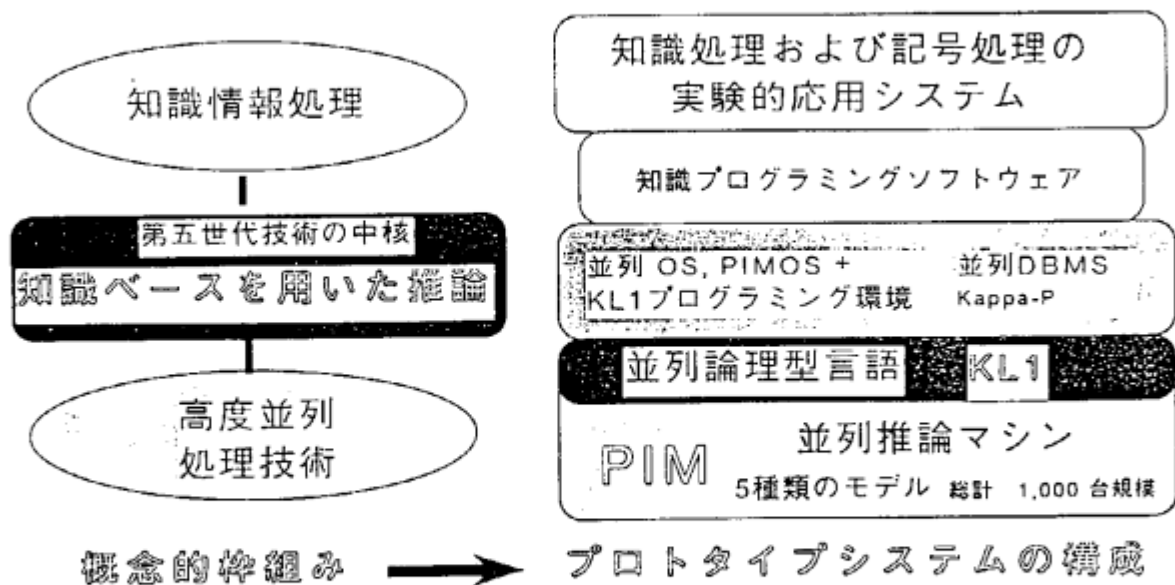
第五世代コンピュータ技術が開く 新しい情報処理の世界

内 田 俊 一

新世代コンピュータ技術開発機構 (ICOT)

2

新しい概念的な枠組みから 現実のシステムへ

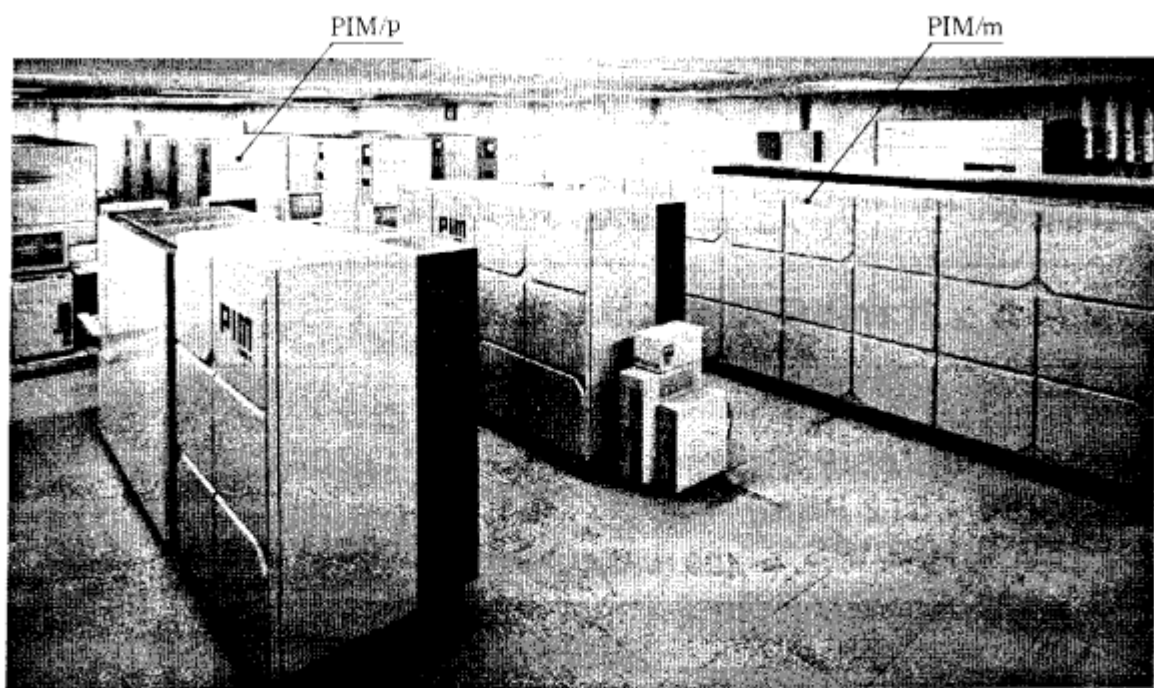
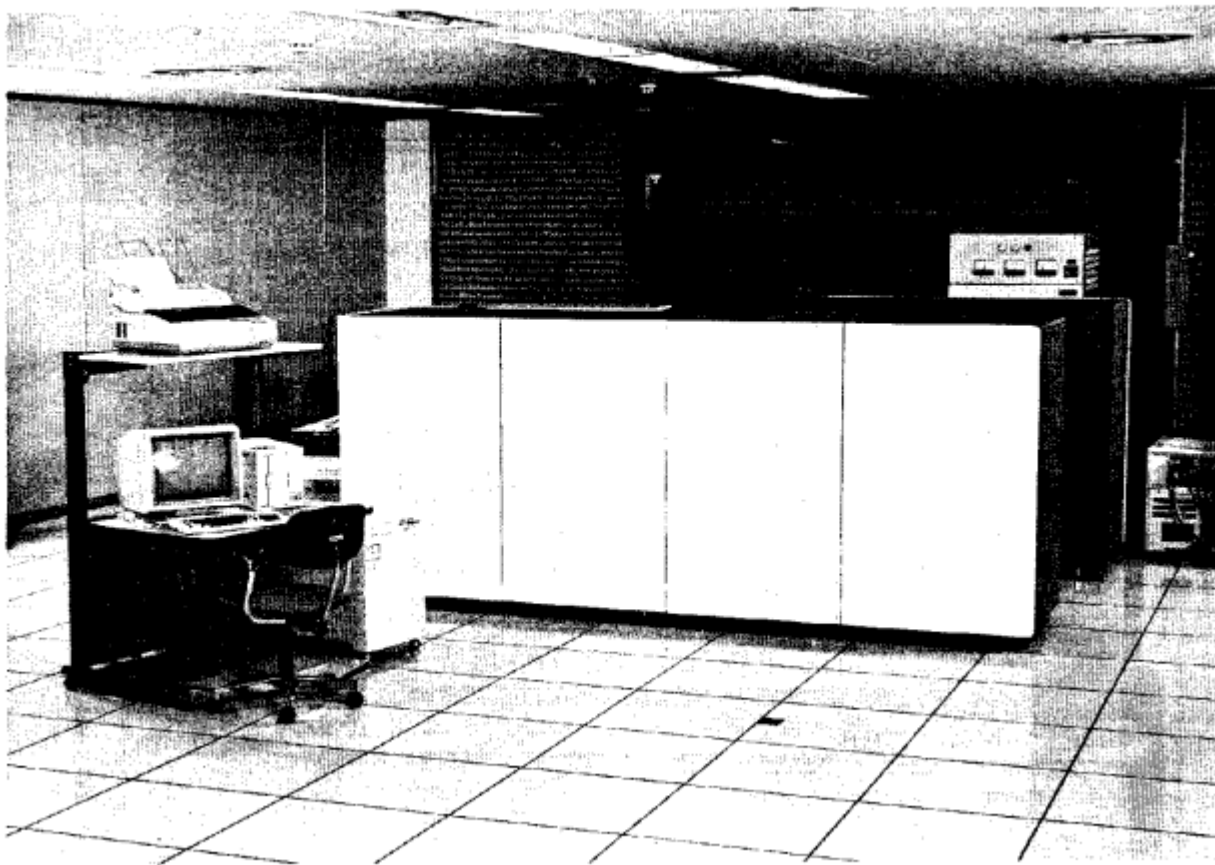


第五世代コンピュータプロジェクトの計画

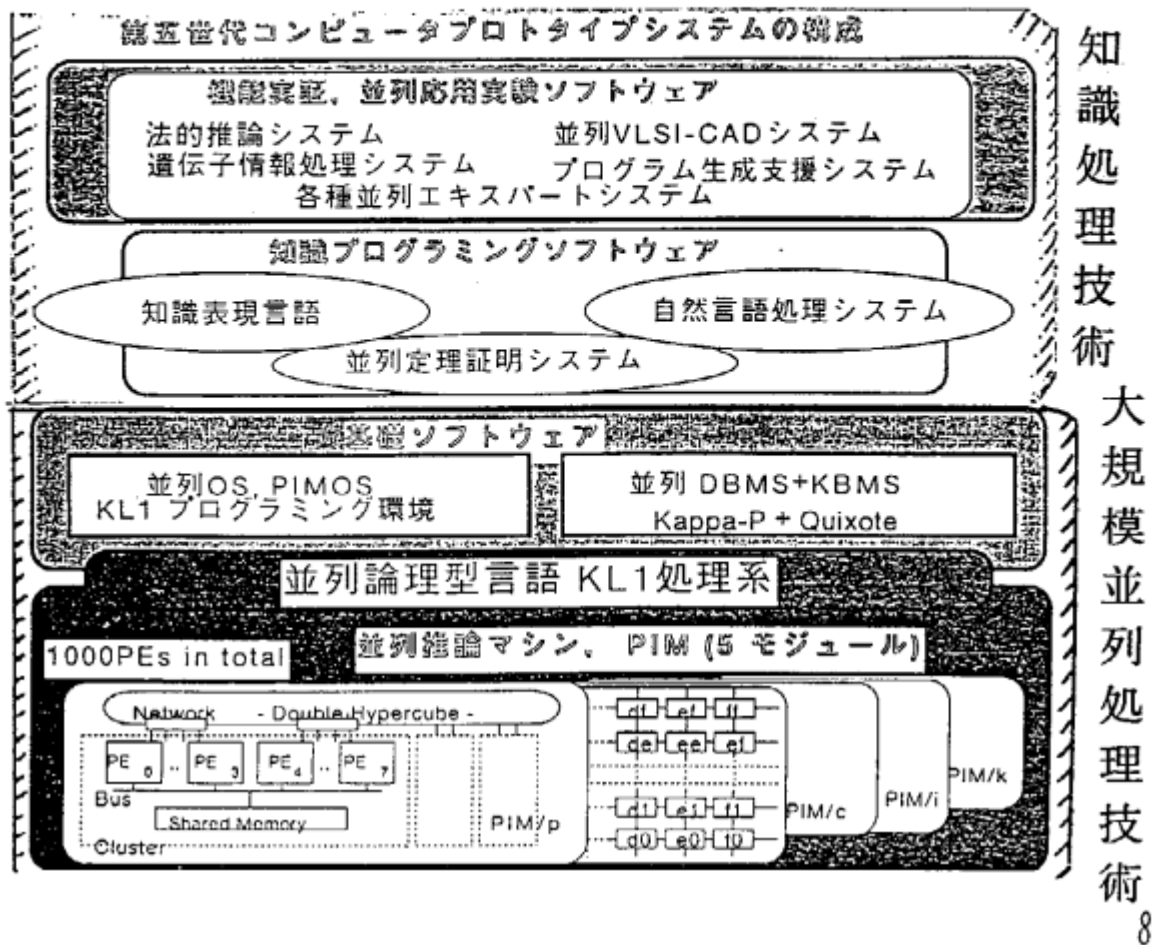
年度	各期の目標	予算額	
'82 ~ '84 昭和57 - 59	前期 要素技術とソフト開発ツールの開発	83	億円 542
'85 ~ '88 昭和60 - 63	中期 実験的中規模サブシステムの開発	216	
'89 ~ '92 平成 1 - 4	後期 総合的プロトタイプシステムの開発	243	

第五世代コンピュータプロトタイプの開発過程

	逐次推論の技術	並列推論の技術
'82-'84 前期	逐次論理型言語、 KL0とESPの設計 逐次型推論マシン、 PSI-Iの開発 40KLIPS for KL0	並列論理型言語、 GHCとKL1の設計
'85-'88 中期	第2版、 PSI-IIの開発 400KLIPS for KL0	PIMの実験機、 マルチPSI, 64台版 600KLIPS / 64PEs for KL1 並列OS, PIMOSと 小規模応用ソフト
'89-'93 後期	第3版、PSI-IIIの開発 1,400KLIPS for KL0	並列推論マシンPIM -1000台規模 1000KLIPS 超 / 512PEs for KL1 並列基本ソフト PIMOS+Kappa-P 知識処理ソフト+並列応用ソフト



ICOT Machine Room



5Gプロトタイプシステムの特徴

知識処理 (非定型的な処理) の高速化

プロセッサ台数にほぼ比例した高速化を達成

並列推論マシン(PIM)

100-200 MLIPS

(256 - 512 PEs)

5G - 10G IPS $\approx 100 \text{ MIPS/5台}$

汎用逐次型マシン

1 - 2 MLIPS

(1 PE)

50M - 100M IPS

言語処理や知識処理では、約10-100倍高速

1 MIPS = 1000 IPS Inference Per Second 1 MIPS = 50 - 100 IPS

応用における量的な変化 質的な変化をもたらす。



5Gプロトタイプシステムの特徴

9

汎用並列ソフトウェアの高い生産性

高レベルの並列論理型言語KL1

- データフロー方式によるプロセス間の同期機構
- メモリーの自動管理機構

効率的な2段階の並列プログラミング手法

- 第1段階：並列アルゴリズムの記述
- 第2段階：ジョブの分割とプロセッサ割当ての指定

使いやすいデバッグとモニタリングツール

↓ **C言語等に比べ1.0倍以上の生産性** ↓

記号処理，知識処理の新しい応用例

10

遺伝子／ゲノム情報処理

生物学データの配列解析

生物学データの総合データベースシステム

高次推論

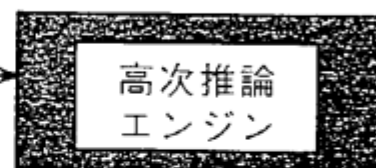
並列定理証明システム，MGTB

自動プログラム合成

並列プログラム生成システム
MENDELS ZONE

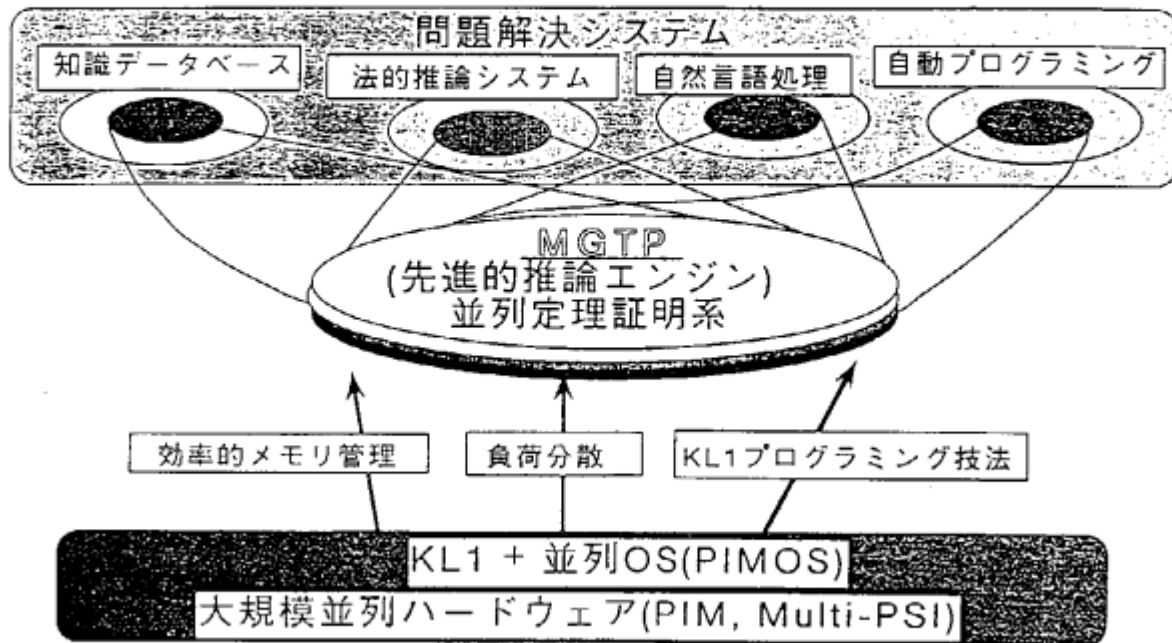
法的推論

法的推論システム，HELICS-II

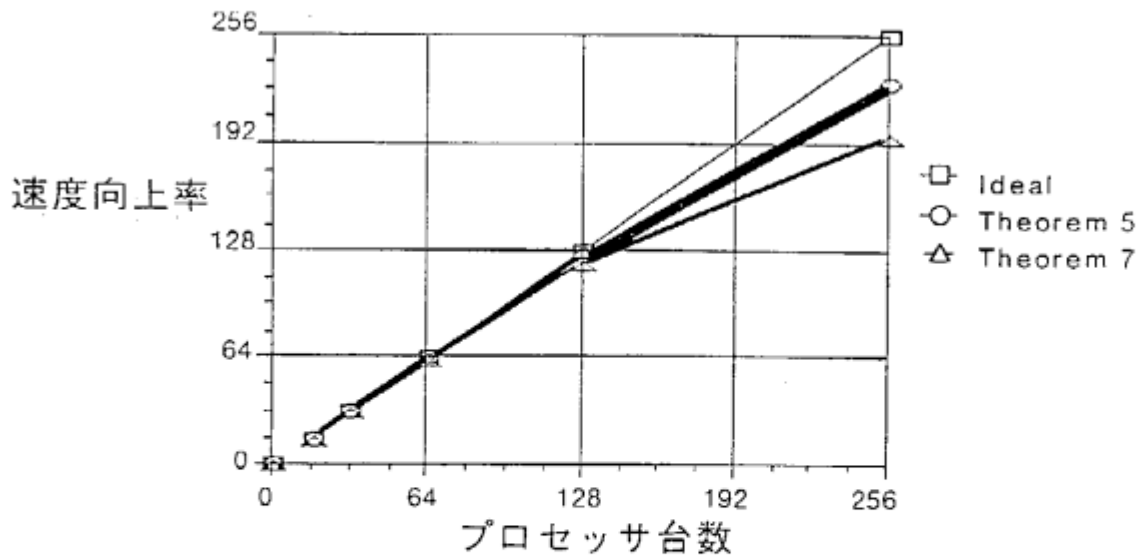


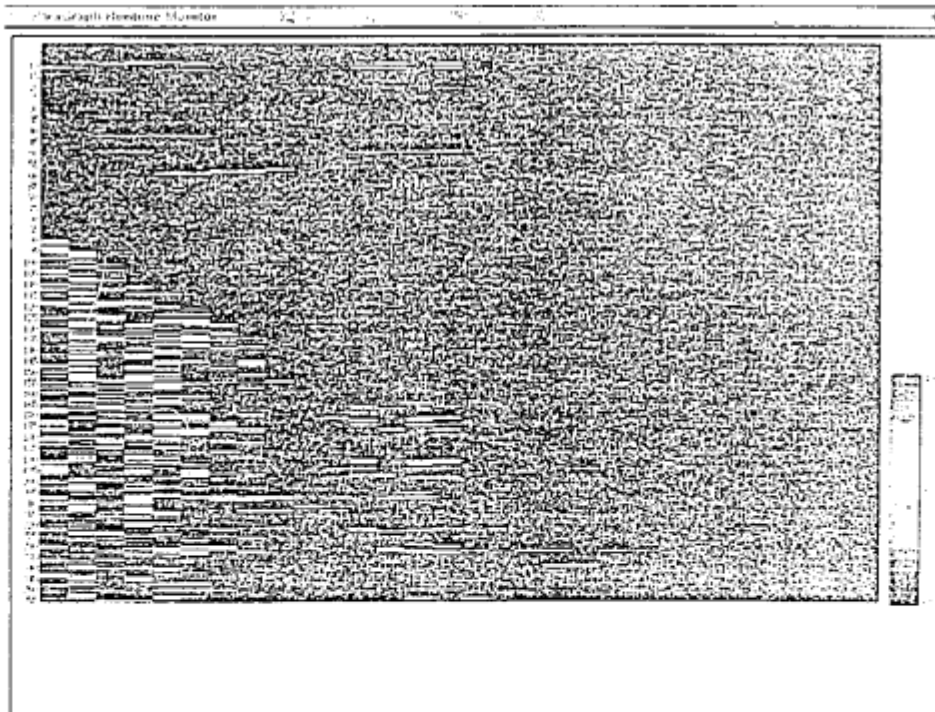
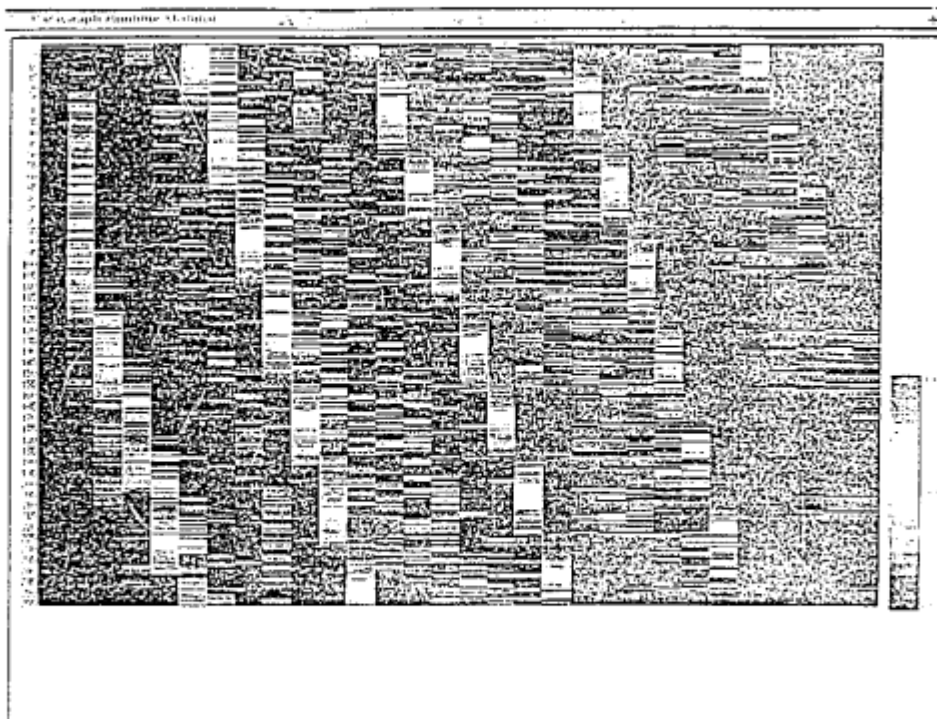
並列定理証明システム MGTP

FGCSにおけるMGTPの位置づけ

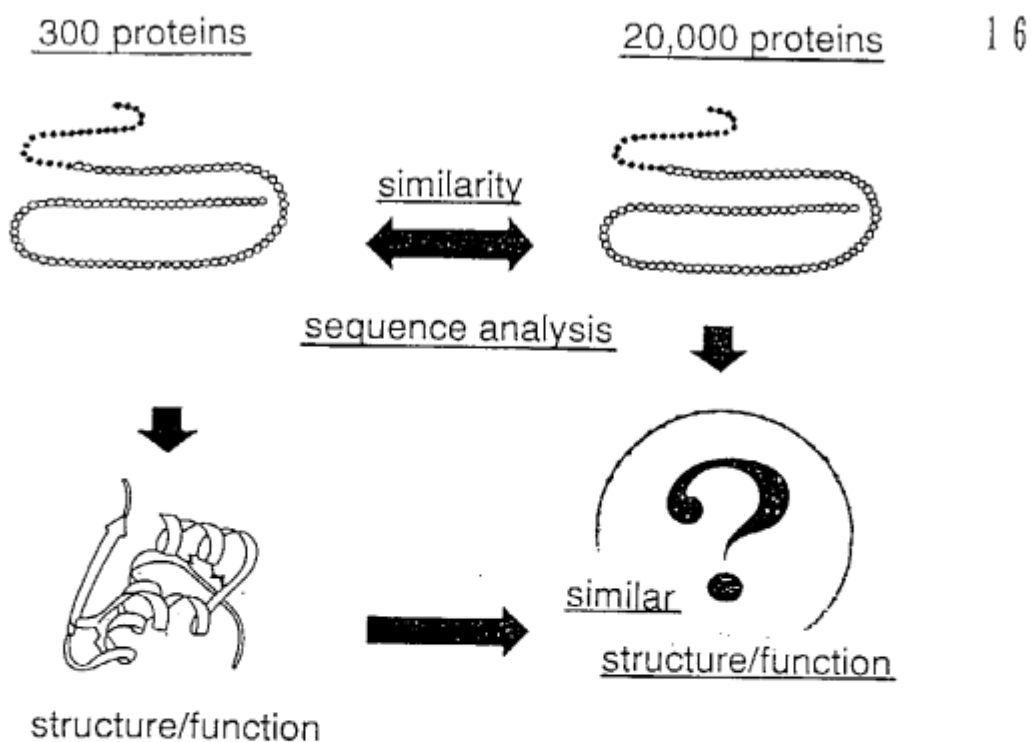
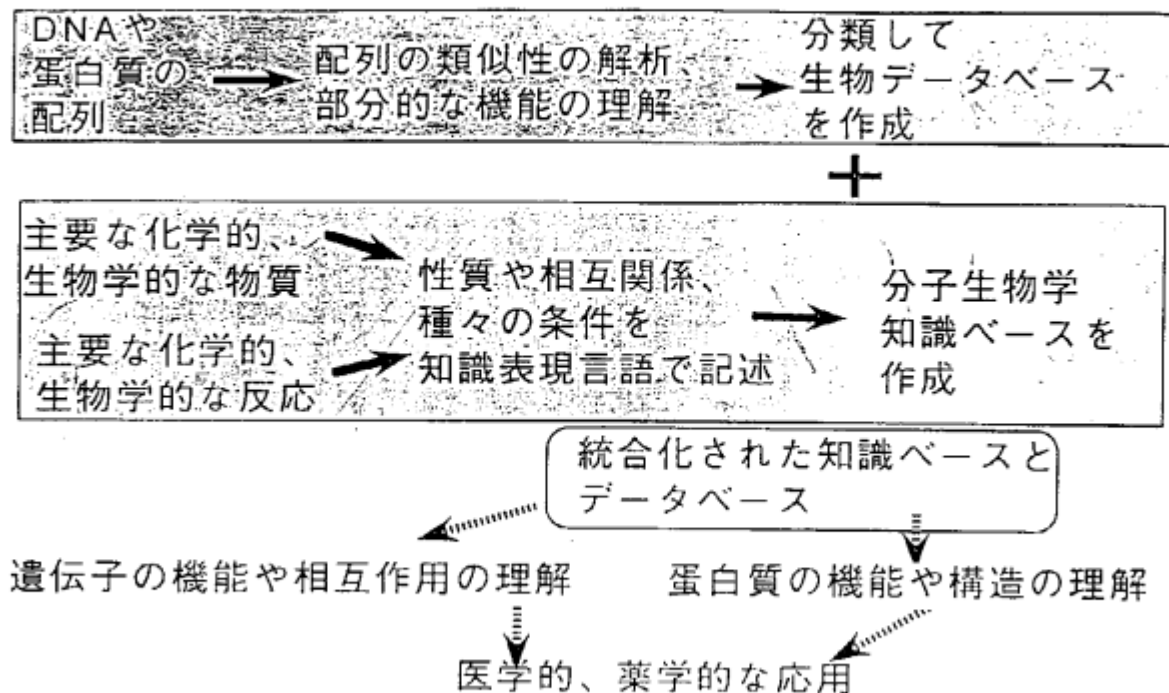


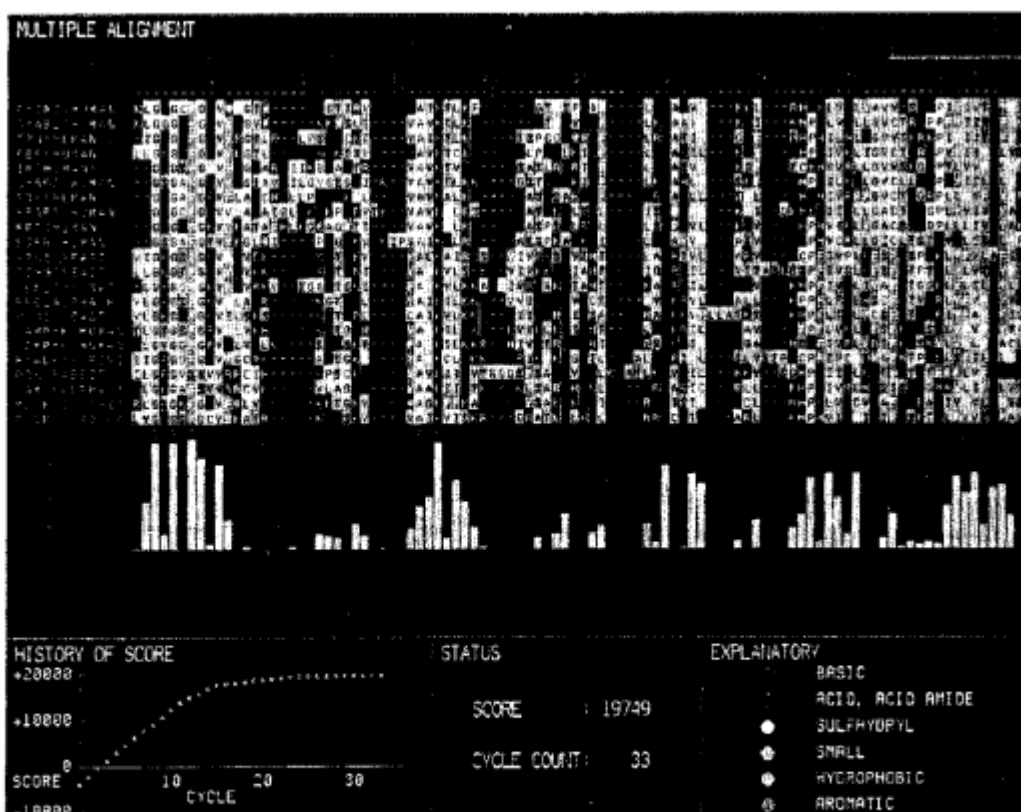
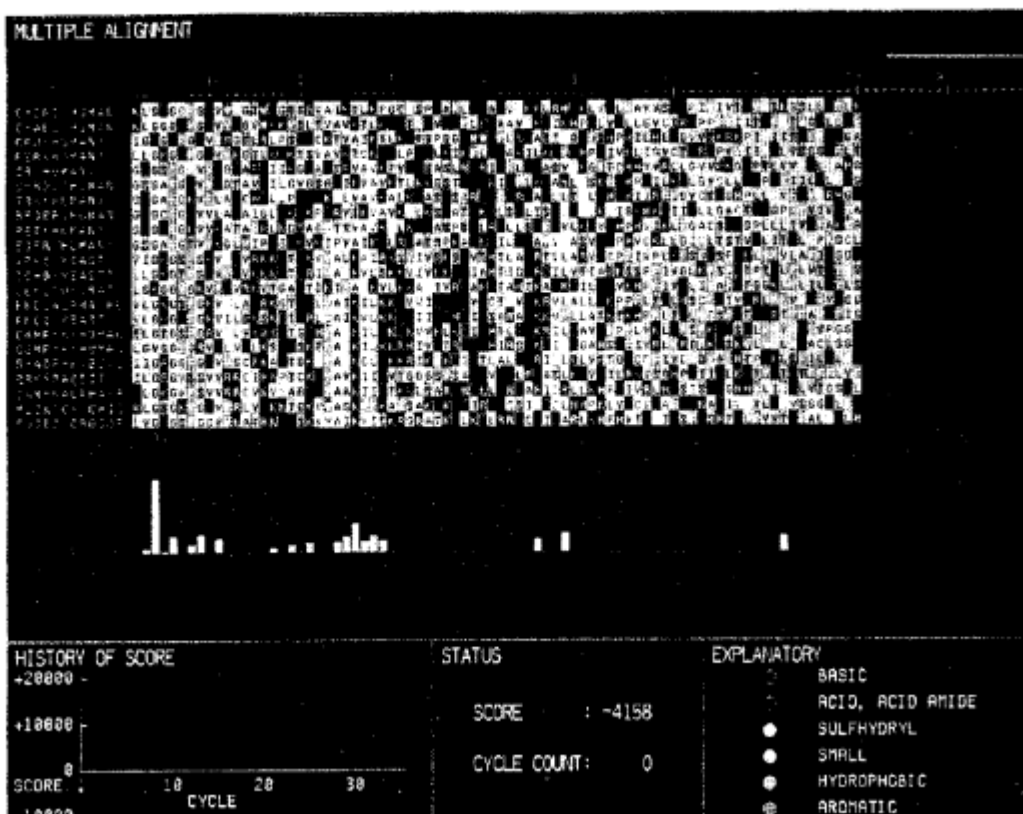
並列定理証明システムの高速度化 (MGTP/N on PIM/m)

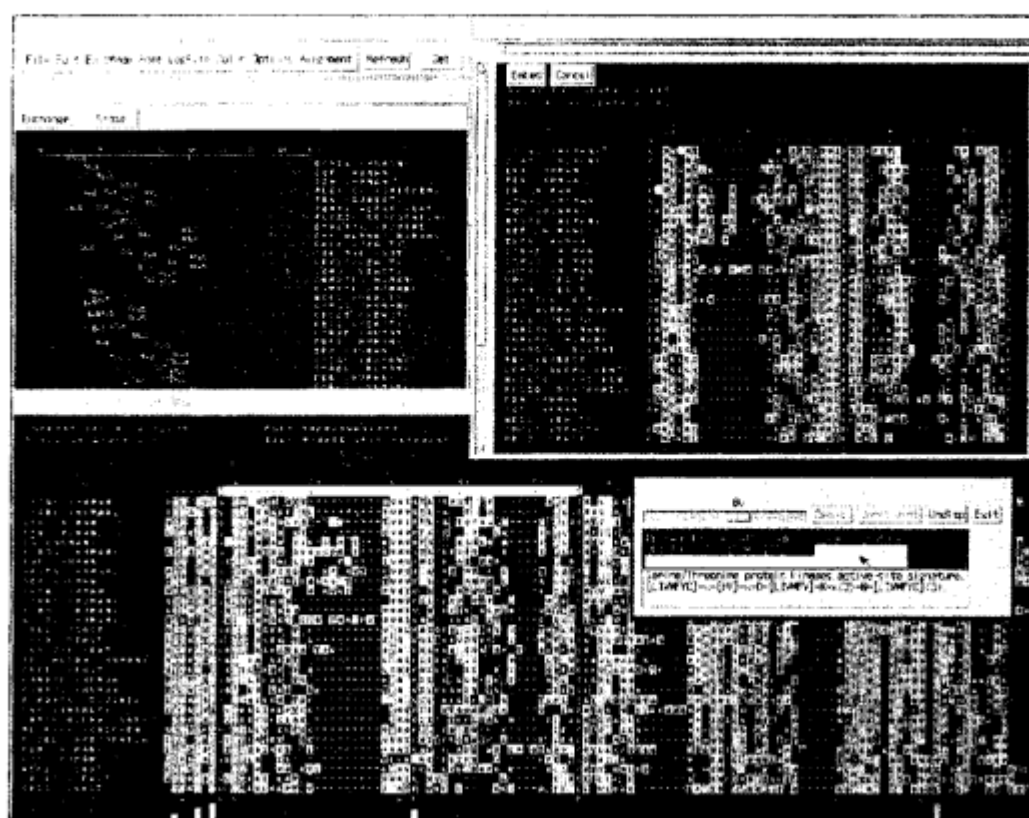




遺伝子情報処理への知識処理技術の応用¹⁵

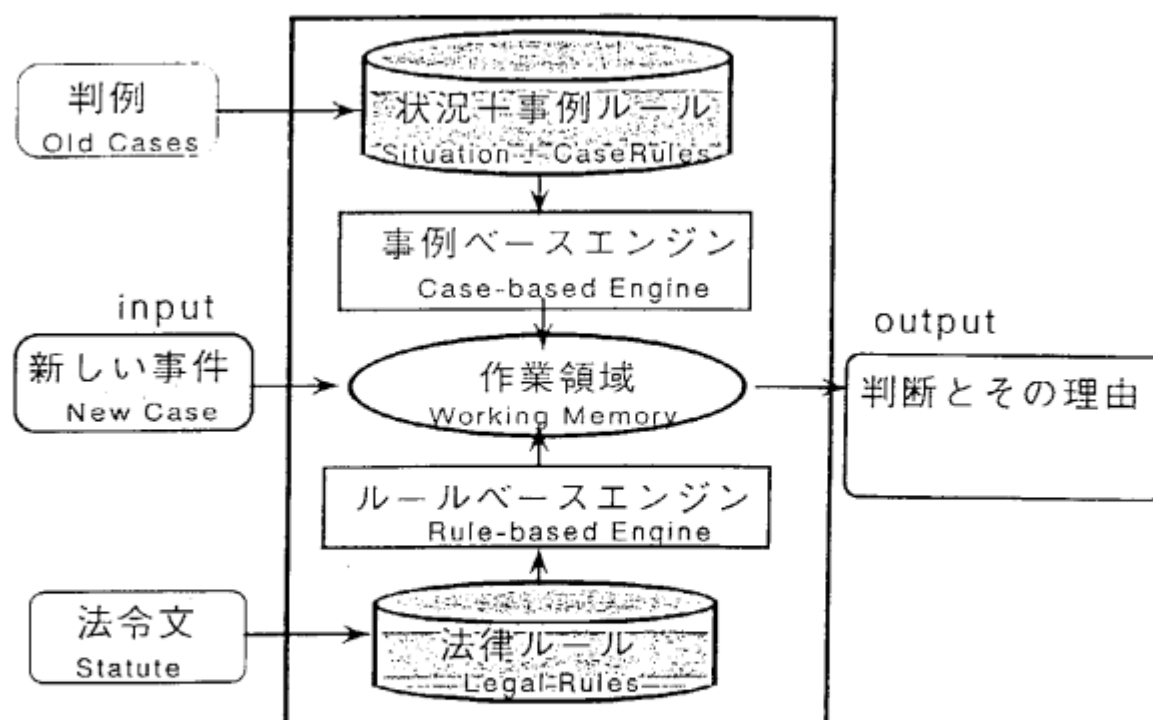






法的推論システムHELIC-IIの構成

20

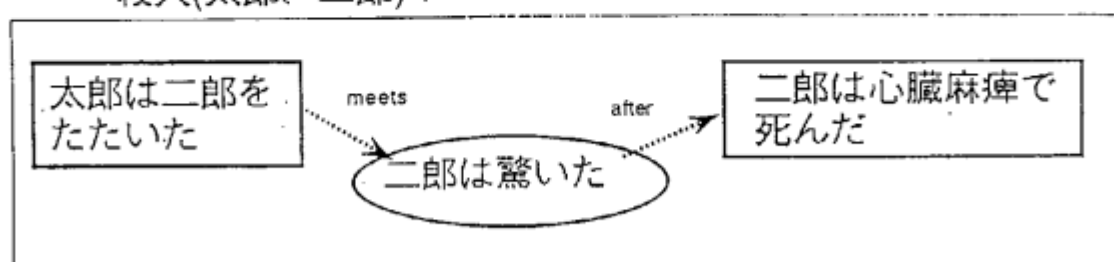


条文ルールとあいまいな概念

IF 殺人(A, B) THEN 殺人罪(A)

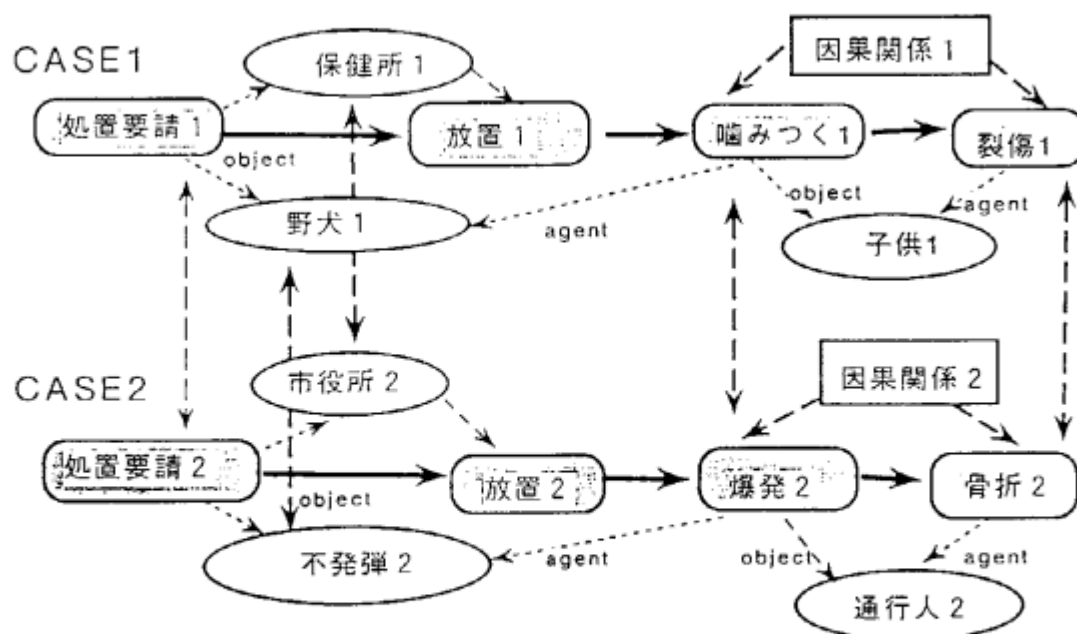
↑
因果関係?
↓
殺人(太郎、二郎)?

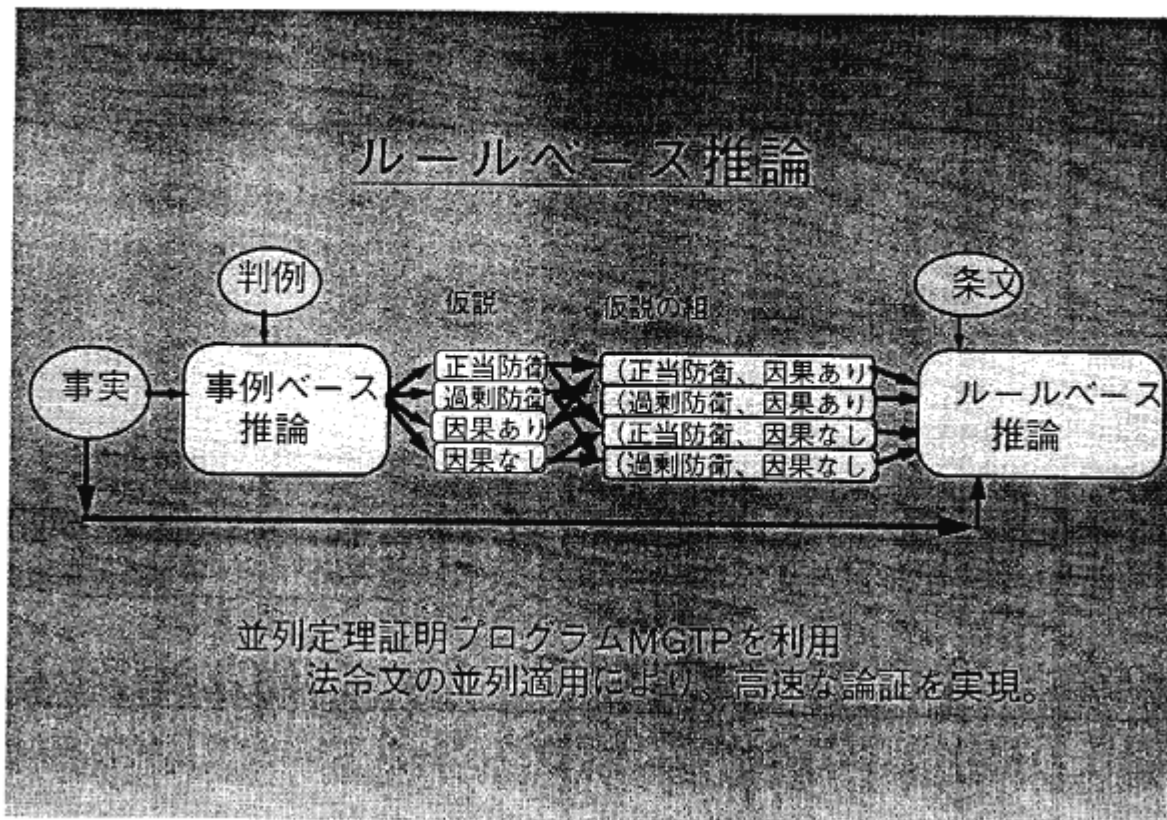
殺意?



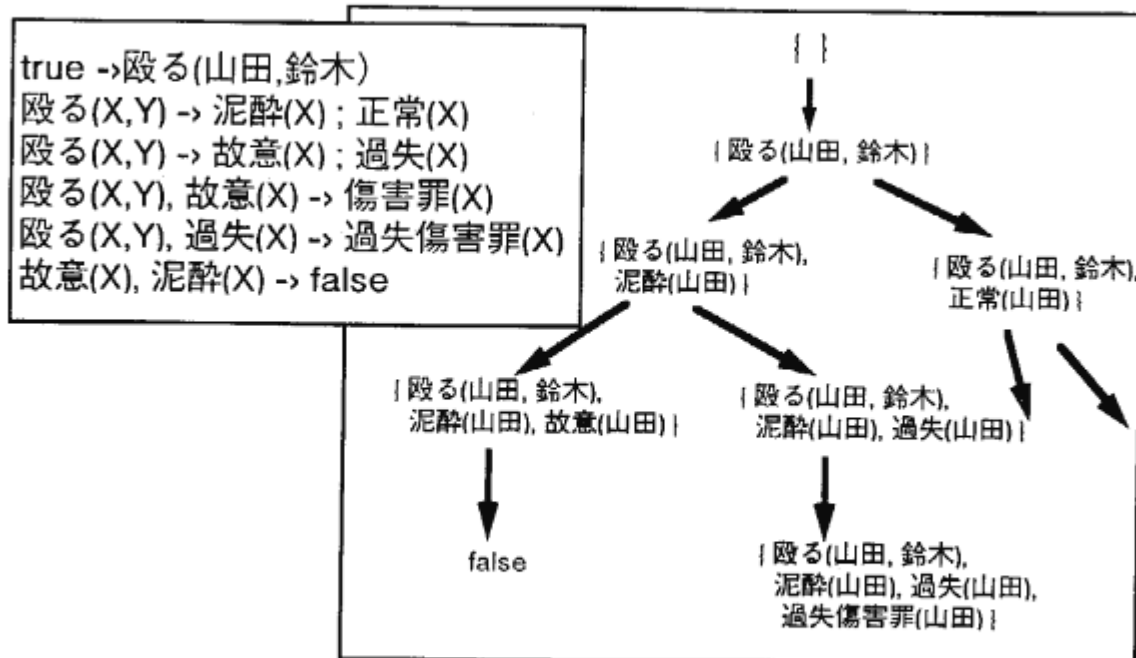
類似検索による判例の検索

事例ベース推論エンジンにより並列実行





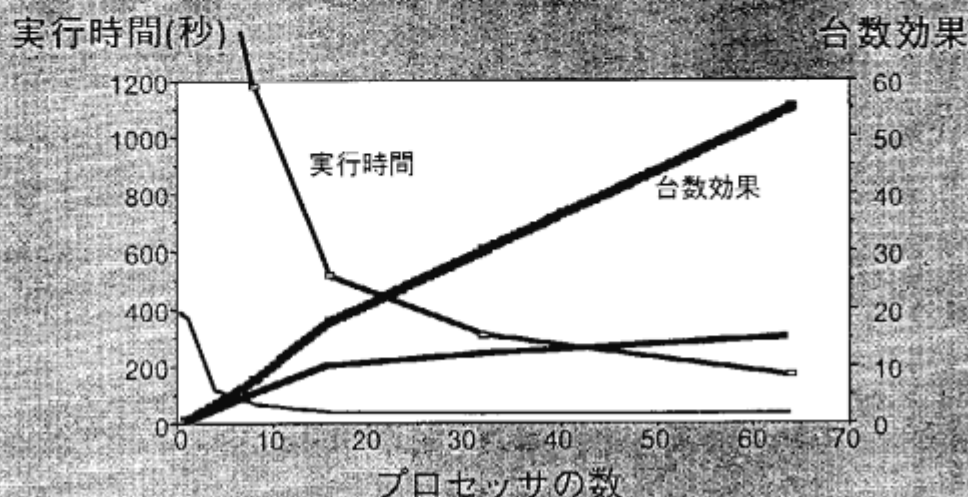
ルールベース推論



並列推論の効果

判例の並列検索

仮説の組み合わせごとに法令文を並列に適用



第五世代技術の基盤化プロジェクト

26

(平成5年4月 より 2年間)

第五世代技術とUnixベースの汎用マシン技術の融合

大規模並列処理

KL1や並列OS, PIMOSをUnixベースの汎用チップを使った並列マシンに移植する

記号処理や知識処理のツールや応用ソフトウェアを汎用マシン上で稼働させる



新たなコンピュータ技術開発の基盤とする

KLIC
システム

↓
・Unixマシン
・パソコン

知識処理

知識処理のソフトウェア技術のさらなる研究開発

知識プログラミングや知識ベース構築技術

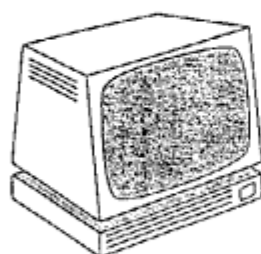
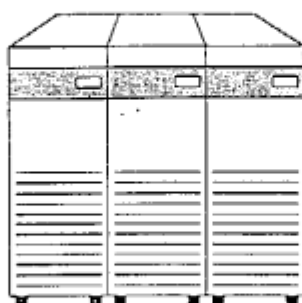
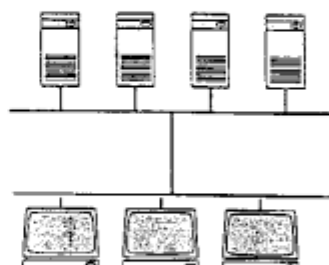
知識処理の新しい応用分野の開拓（社会科学の分野など）

⇒ Unix上へ

ICOTフリーソフトウェア(IFS)として無償公開

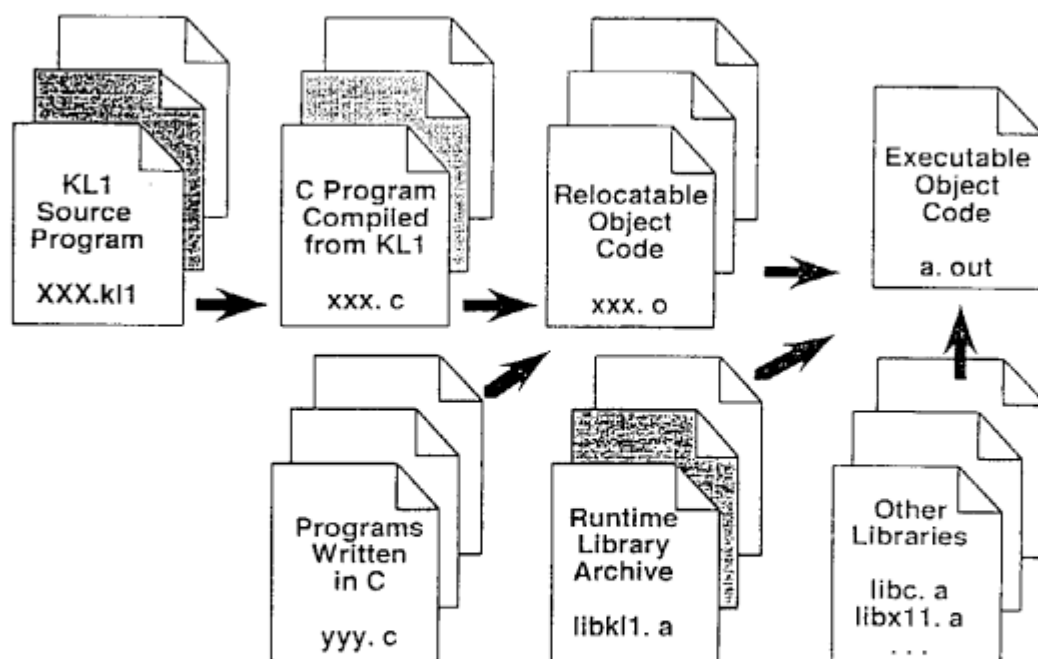
インターネット経由で世界中に配布中（現在84本）

対象とするハードウェア

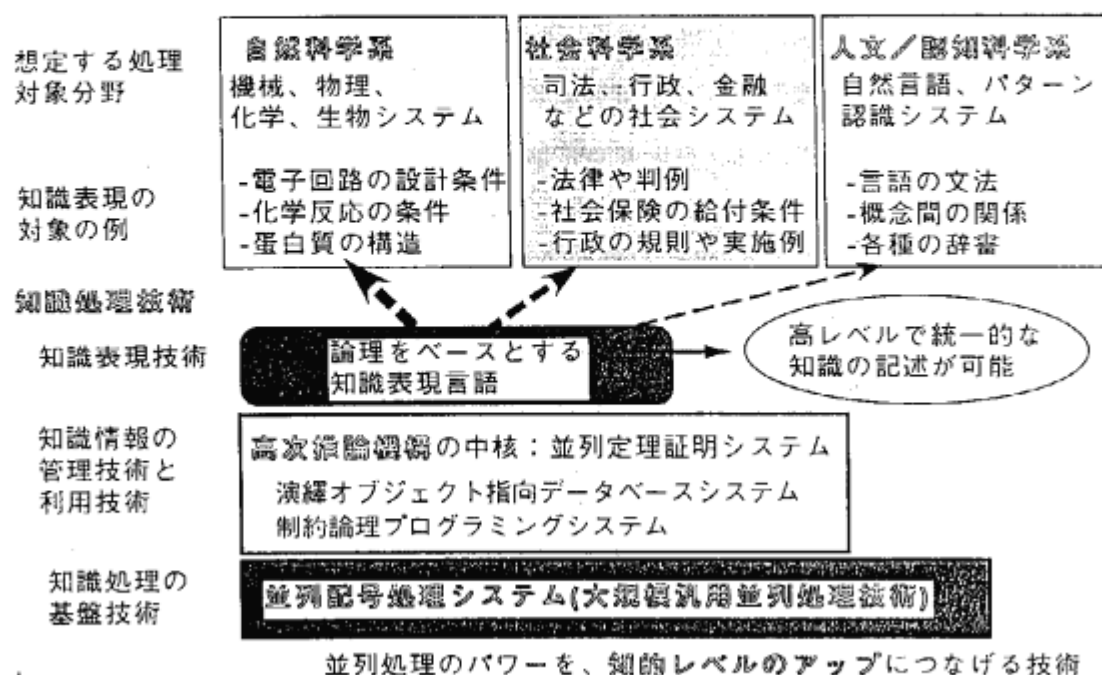
Single-Processor
UNIX SystemMulti-Processor
UNIX SystemsNetwork-Connected
UNIX Systems

KLIC システム

KL1とPIMOSの環境をUNIXベースの
逐次型 および 並列型マシンに移植する



第五世代コンピュータ技術を基盤とする知識処理の展開



新年拡大所議 平成6年1月

第五世代コンピュータの研究基盤化プロジェクトを
どのようにまとめるか？

I C O T 研究所 内 田 俊 一

1. 研究基盤化プロジェクトの目的とその後の展開

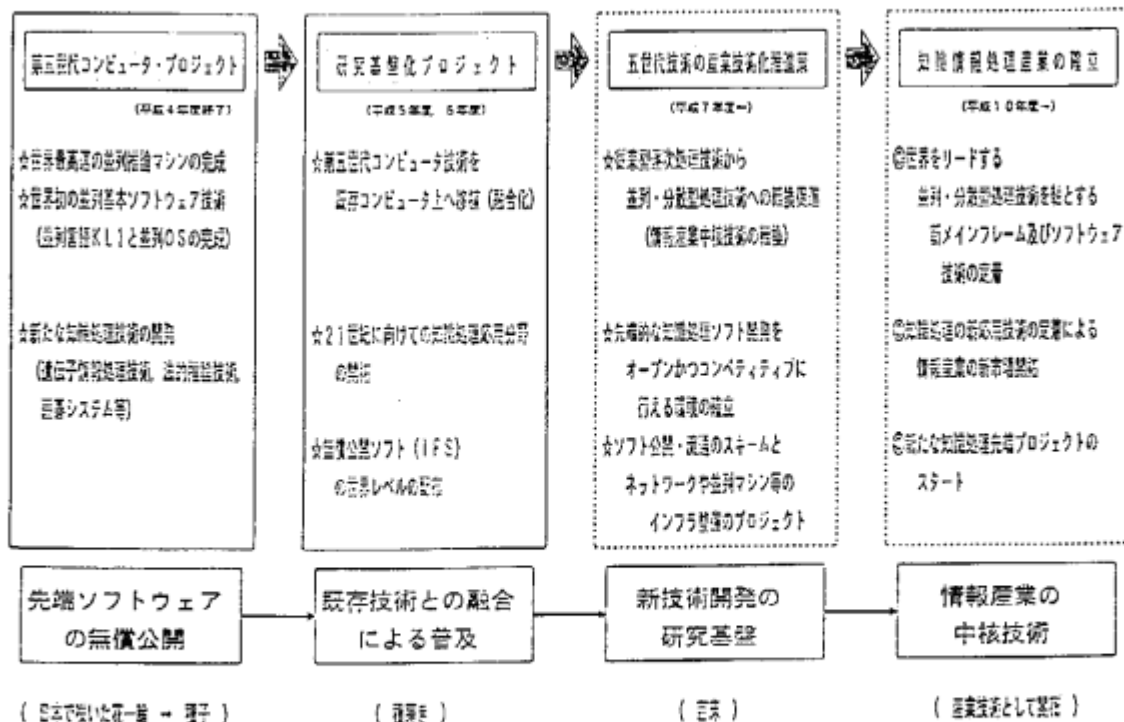
第五世代コンピュータプロジェクトの研究開発成果 ⇒ 産業技術としての利用はまだ先
(市場の技術が追い付いていない)

⇒ 成果を普及しやすい形として 研究コミュニティーに技術移転 ⇒ 開花に向けての育成
(2-5年)

⇒ 情報産業の中核技術として定着
- 第五世代の大規模並列処理技術 (3-5年)
- 第五世代の知識処理技術 (5年-10年)

2

第五世代コンピュータの研究基盤化プロジェクトとその後の展開



2. 第五世代コンピュータ技術の未来展開の予測

1) 大規模並列処理技術

市場技術の動向：急速に分散・並列処理に向っている。

- ハードウェア技術 => MIMD型の大規模並列処理の時代到来
研究・製造部門は技術革新のまっただ中
=> 逐次型マイクロプロセッサをそのまま利用する形態から
並列言語を意識したアーキテクチャへの移行

=> * 並列高級言語指向アーキテクチャ => PIMの技術の生きる時代へ

- ソフトウェア技術 => 従来の言語やOSの追加拡張でしのいでいる。
=> すぐに本格的な技術革新が要請される時代へ

=> * 第五世代の言語やソフトウェア技術が本流となる時代へ

2) 第五世代の知識処理技術

その現在の姿は:

- 述語論理を基礎とする知識プログラミング言語
演繹+制約+オブジェクト指向
- 並列定理証明系は、言語処理系のエンジン
- 論理を基礎とする知識表現言語と知識ベースを活用した応用システム
法律、遺伝子などの情報処理

その将来の姿は:

- 将来のナショプロの一大研究テーマ
- 大規模並列処理による強力な計算パワーを備えた
記号処理システムが実現したときの有望な応用分野
=> 5年から10年先の大きなマーケット

3. 開花に向けての準備と技術の移転・普及

1) 既存技術との融合化

- a) 大規模並列技術: KLICによるUnix環境との融合
=> 研究開発は順調に進展、
講習会、IFSとしての普及も軌道に乗る。
- b) 知識処理技術: さらなる発展と移植に向けてのコンパクト化
=> 研究進展も良好。学会発表、デモ、
海外WSなどで、もっと宣伝する。

2) 技術の移転と普及

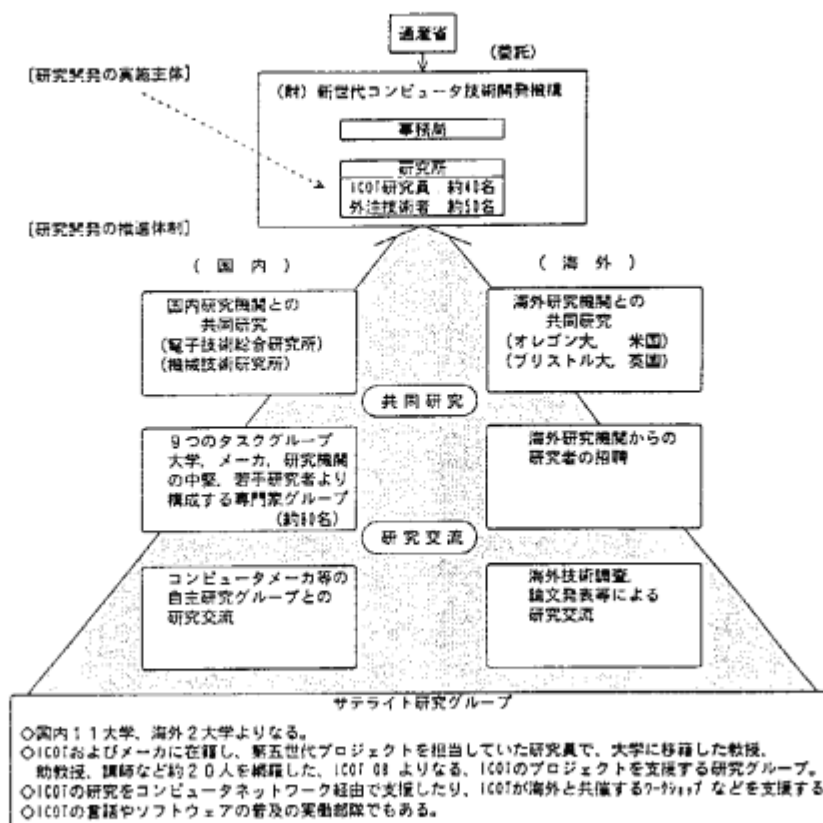
- a) ICOTフリーソフトウェアの配布
予想より順調。米国の企業よりのアクセスが多い。
- b) 学会発表やデモ
国内、海外ともにもっとやろう。
- c) 普及・移転のための広報活動
ICOT TODAY, IFSニュースレターなど。
国内は軌道に乗る。海外向けがこれからの課題。
- d) KLIC講習会、大学での講演や講義など
今年も大変だが、積極的に実施する。

3) ICOTサテライト研究グループの組織作り

- a) KL1や第五世代技術を育成する支援グループの設立。
まず、ICOT OBや研究協力者のいる大学を中心。
(11大学+メーカ自主研究グループ+TG)

=> 形はできたので、中身を充実させる。
参加者の意識の昂揚とまとまりを強める努力をしよう。
(近い将来の日本のコンピュータ技術を背負うグループを目指す)
- b) 共同研究のセットアップと海外ワークショップ
(Oregon, Bristol, ANU, SICS)
海外大学や研究グループでの第五世代技術の育成を狙う。

=> これも形ができたので、中身を充実を。
国内の大学のグループとの連係を強める努力。
ICOTサテライト研究グループの海外展開。



4. 研究基盤化プロジェクトの終結

1) 成果のまとめ

成果披露の国際シンポジウム ⇒ 平成6年12月半ばを予定
 研究発表 (デモ付き) 2日間 + 主要テーマのWS 2日間

⇒ 研究開発活動は、このシンポジウムまでで終結。
 その後は、ICOTフリーソフトウェアとしてのとりまとめやドキュメント、報告書の作成にかける。

2) 研究所のクローズのスケジュール

a) 諸設備、スペースの撤収予定

平成7年の1月より、順次開始予定。研究所は3月末にてクローズ。

b) 海外共同研究の終結処理

平成6年9月頃より、最後のWSなどを開催。
 区切りをつけつつ、平成7年3月に終結。

3) ICOTフリーソフトウェアの配布とメンテ体制の確保

IFSの配布用機器、ネットワークへの接続、メンテ要員の確保。
 IFSの普及・発展は、ICOTサテライト研究グループで実施。