

# カーネギー・メロン大学における分散処理 — 人工知能研究の立場から —

Distributed Processing at Carnegie-Mellon University

金 出 武 雄

Takeo Kanade

京都大学・工学部

Department of Information Science, Kyoto University

昭和52年8月末より昭和54年3月末まで19ヶ月滞在した、米国のカーネギー・メロン大学における分散処理の話題を紹介する。

## I. はじめに

筆者は、昭和52年8月末より昭和54年3月末まで、米国ペンシルバニア州ピッツバーグにあるカーネギー・メロン大学（以後CMUと略す）の計算機科学科に滞在する機会があった。CMUは、人工知能研究・ソフトウェア・C.mmpやCm\*の開発でよく知られており、米国ではMITやスタンフォード大学と並んで高く評価されている計算機学科をもつ。

筆者自身は人工知能（Artificial Intelligence: AI）研究の画像理解の研究を行ない、計算機システムの専門ではないが、人工知能研究の目から見た、CMUにおける分散処理の話題を紹介したい。

## II. C.mmp と Cm\*

CMUでは1970年以来マルチプロセッサの研究が精力的にすすめられており、C.mmp (multi-minicomputer) と Cm\* (modular multi-microprocessor) のシステムが開発されている。これらについては日本でもよく紹介されているので、詳細は他にゆずるとして、簡単にこれらの構成・特徴・現状をのべる。なお、これらの他に工業生産環境内での

fault tolerant なアーキテクチャ研究のため C.vmp (voted multi-processor) の設計製作が1975年から始まり、あでに完成している。

### II-1 C.mmp

#### (1) 構成

C.mmp はCMUで最初に作られたマルチプロセッサで、1970年に設計が始まり、1977年には完成・稼働している。そのPMS (Processor-Memory-Switch) 構造を図1に示す。16台のPDP11/20又はPDP11/40のプロセッサ(Pc)と16台の共有メモリ(Mp)が16×16のクロスポイントスイッチ(Smp)で結合されている。プロセッサ間の通信用のため各PcはKinterbusに接続されている。

#### (2) 特徴

C.mmp の設計を開始した1970年頃ミニコンの価格は\$10K程度であった。これを用いて高性能・低価格のマルチプロセッサを音声や画像理解の研究に提供するのが目標であった。

C.mmp はそれまでのマルチプロセッサとは①16台もの多数のプロセッサで②各々はミニコンであるという点で大きく異なる。

そして、① できるだけ対称型になる様に設計された。すなわち、メモリ、プロセッサ、I/O という資源のプールがタスクによって共有され、プロセッサ間にはマスター・スレーブの関係がなく、任意のユーザジョブはどのプロセッサでも実行される。これは OS についても言え、C.mmp の OS である Hydra のどの部分もどのプロセッサでも実行できる。

### (3) 現状

C.mmp は一応完成されており、ARPAnet にも接続され、日常的に使用できる。プログラミング言語として、BLISS, L\*, Algol 68 が使え、ユーティリティとしても SOS や TECO という PDP10 系のテキストエディタがインポートされている。man-machine 対話の方式で非常に高速でスクリーンの書き換え(4800 行)を要する ZOG システムのための端末もサポートある。

現在、研究の中心が次の Cm\* に移っており、C.mmp に対する新しいプロジェクトは特になく、また、マルチプロセッサを生かすプログラミングが、PDP11 のアドレス空間が小さいことも手伝って簡単ではないこともあり、活発に利用されているとはいえず、idle あるいは停止の状態であるのをよくみかけた。

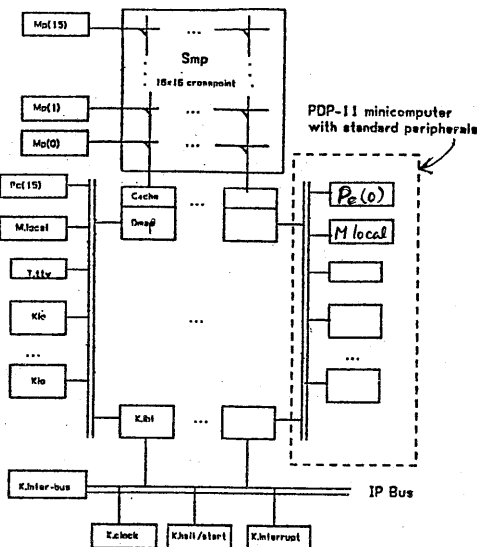


図1 C.mmp の PMS 構造

## II-2 Cm\*

Cm\* は 1975 年に設計が始った。近年の LSI 技術の進歩を利用し、LSI11 を主なプロセッサとして、極めて多数の安価なプロセッサの結合をねらったシステムである。かなり大きなプロジェクトで関係している faculty だけで 60 人以上、学生・職員を含め 300 人程になる。

### (1) 構成

図2は Cm\* の構成を示す。基本構成要素は Computer Module と呼ばれ(図中右側の [ ]) LSI11 をプロセッサとし、ローカルなメモリ、I/O 装置が LSI11 のバスにつながっている。これが Slocal というアドレスマッピング機構を通じて Computer Module 間のバスにつながれる。いくつかの Computer Module (14 個まで) が 1 つの Mapbus につながり上位のマッピングプロセッサである Kmap を共有したものを Cm\* Cluster という(図中左側の [ ])。そして、図2に示る様に Kmap は 2 つまでの Cluster 間バスに結びことができ、Cluster のネットワークを構成したものが Cm\* である。Kmap はマイクロプログラムされた特殊な処理装置で Mapbus の制御、Inter-clusterbus とのインタフェース、Cluster 内及び Cluster 間のアドレスマッピングを司る。

### (2) 特徴

a. 拡張性 Cm\* はモジュール構成のため、プロセッサ、メモリ、結合を追加することで、処理能力・メモリ容量・通信速度を増加できる。システムのトポロジーも、個々の応用に合う様にあればよい。1977 年には 10 プロセッサシステム、1979 年には 50 プロセッサシステムが稼働している。10000 プロセッサのシステムも十分可能という。

b. アドレスマッピング Cm\* は  $2^{28}$  バイトの仮想アドレス空間を持つ。メモリ参照はプロセッサのローカルメモリ(即ち Cm 内)、同一クラスタ中の他のプロセッサのローカルメモリ、別のクラスタ中のプロセッサのローカルメモリ という 3 レベルがあり、それぞれ、3  $\mu$ sec, 9  $\mu$ sec, 26  $\mu$ sec 程度の参照サイクル時間である。

c. OS のファームウェアによるプリミティブ Kmap のマイクロプログラムがプロセッサ間

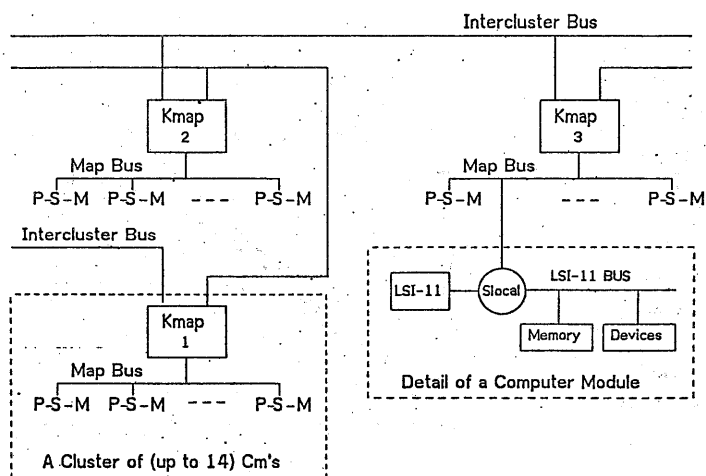


図2 Cm\* の PMS 構造

のメッセージ通信をサポートする。

### (3) 現状

1977年で 10 Cm, 3 Kmap からなる Cm\* が稼動状態で、いろいろなベンチマーク的な並列プログラム・音声理解システム H ARPY がインプリメントされ、評価データも出ている。1978年には 50 Cm, 5 Kmap にある予算もつき、1979年夏には稼動している。OS は基礎的なものがあり、Cm\* 用の StarOS 及び分散型の OS が設計・インプリメント中とのことである。プログラミング言語のプロジェクトには Algo/ 68 がある。また画像ファイルが分散されている場合を想定しての画像理解システムも応用例題として考えている。

なお、C.mmp, Cm\* を通じハードウェアの中心人物であった S. H. Fuller はすでに DEC 社に移った。

## III. 並列アルゴリズムのインプリメンテーションとマルチプロセッサの評価。

いくつかのベンチマーク的な問題の並列アルゴリズムが C.mmp と Cm\* にインプリメントされ、評価データが出ている。又、CMU で成功した HARPY 音声理解システムの C.mmp 版、Cm\* 版がある。

### III-1 C.mmp

簡単な並列アルゴリズムをインプリメントすると、プロセス数に対し、能力向上が理論値とくらべて意外と早く悪くなることがある。単調増加関数の根を見つけるのに、複数のプロセッサを用い、ある区間内のいくつかの適当な点での関数値を各プロセッサで並列的に計算し、その結果から根の存在する区間の範囲をせばめていく、というのと線形探索アルゴリズムを使い、性能低下の原因を詳細に検討

した結果、次の6つの原因が（この順の重要性で）あることがわかった。

- (1) 繰り返し計算をする際、計算時間が点ごとに違う。
- (2) メモリの有限バンド幅のためのメモリコンテンション
- (3) OS のスケジューリングプロセス
- (4) 各プロセッサのスピードの違い。
- (5) I/O 装置サービスルーチンに関係する割込み。
- (6) 各メモリバンクのスピードの違い

また、各プロセスの同期をとる方法の違いによって、アルゴリズムの性能が大きく影響されることがわかった。

### III-2 Cm\*

- (1) ラプラスの偏微分方程式を解く問題 (PDE)
- (2) クイックソートの問題 (QUICKSORT)
- (3)  $A$  を  $M \times N$  の2値マトリクス、 $e$  を  $(c_1, c_2, \dots, c_N)$  なる  $N$  次元ベクトル、 $e$  を  $(1, 1, \dots, 1)$  という  $M$  次元ベクトルとあるとき  $x = (x_1, \dots, x_N)$   $x_i = 0$  or  $1$  を  $Ax = e$  で  $\sum c_i x_i$  が最小となるように  $x$  を決める整数プログラミングの問題 (INTEGERPROG)
- (4) HARPY 音声理解プログラム (HARPY) に対する並列アルゴリズムを Cm\* にインプリ

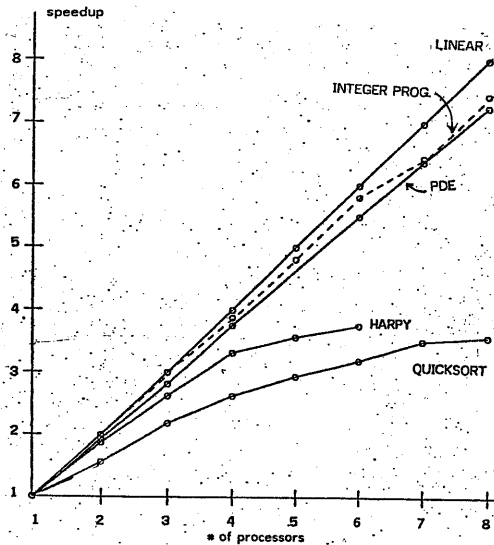


図3 Cm\*におけるマルチプロセッサによるスピードアップ

メントして、プロセッサ数によるスピードアップを調べた結果が図3である。

プロセッサ数に対しかなり線形な能力向上に近いものが得られている。ただ、このグラフはGlobalな変数のみかmappingされ、他はローカルなCm中のメモリ参照で済む様にした場合であり、変数やスタックの置き場所を変え、メモリの参照パターンを変えるとPDEの場合では図4に示すような性能向上となる。

### Ⅲ-3 音声理解システムHARPYの各種システムでのインプリメント

HARPYは1971年～1976年のARPAによる音声理解プロジェクトにおいて、初めに設定した仕様を満たす結果を得たシステムとして、CMUの誇るプログラムである。これを、CMUのいろいろな計算機にインプリメントして、処理時間を比較したのが図5で、DESCALと呼ぶ算術文の認識タスクの例である。タスクが簡単であるためPDP10/KL-10(DEC10の1番速いシステム)かC.mmpより速いかもっと複雑な1000語を含むタスクの場合では図6に示す様に4台のプロセッサによるC.mmpでKL10を上回る。

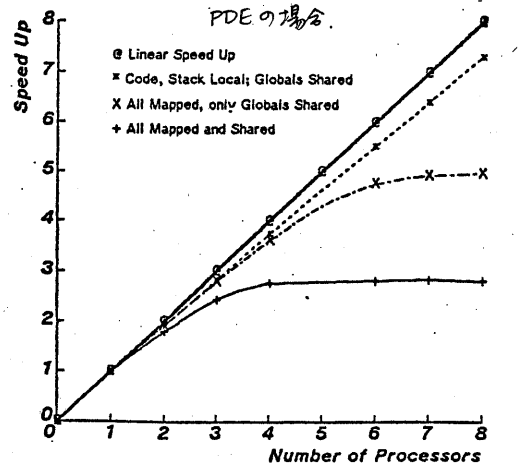


図4 メモリ参照パターンの違いによるスピードアップ効果の違い。このグラフのx---x---xが図3のPDEのグラフと同じ。

## Ⅳ. 人工知能研究と分散処理

### Ⅳ-1 音声理解

すでに述べたように、HARPY音声理解システムは(もともとPDP10上にインプリメントされたが)マルチプロセッサの有効性も示す応用例題として使われている。

CMUではこれ以前からHEARSAYと呼ばれる音声理解システムが研究されており、この音声理解システムに十分な処理能力を与えることがマルチプロセッサ開発の応用面からの動機とも言われていた。

さて、HEARSAY-IIはBlackboardモデルと名づけられた構成をしており、このモデルは人工知能システムのモデルとして評価されている。図7に示す様にKnowledge Source(KS)という、ひとまとまりの知識に対応するプログラムモジュールが多数あり、Blackboardという共通のデータベースを通じインタラクションし、仮説の生成・検証という作業をしながら問題を解決していくというものである。このHEARSAY-IIのアーキテクチャを余り変えずに、ひとつのKSをひとつのプロセッサに対応させ、分散処理向きに分解し、マルチプロセッサ上にインプリメントすると、プロセス間の

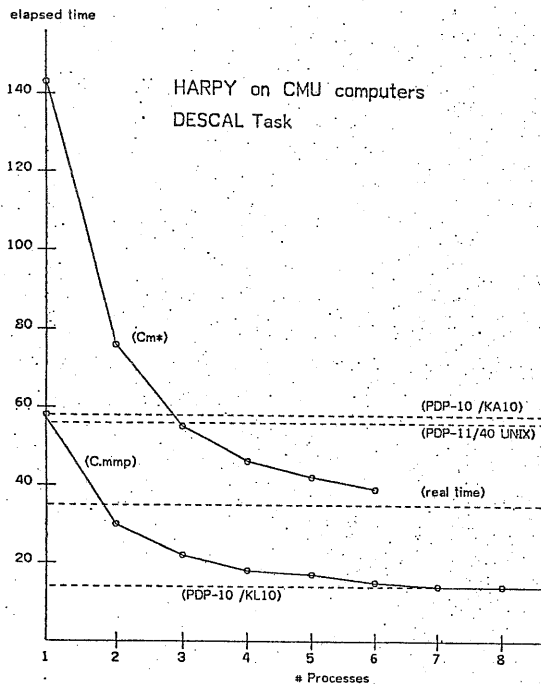


図5 HARPYのCMUの計算機による処理時間 (DESCALタスクの場合)

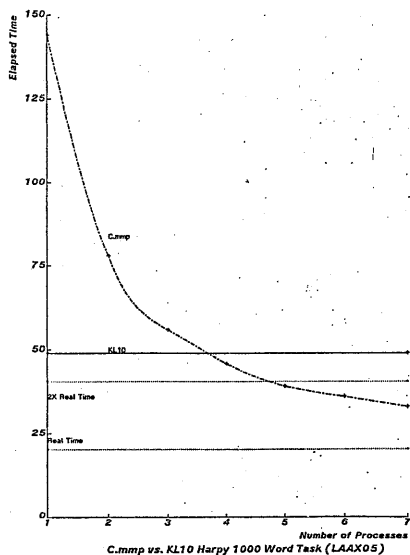


図6 1000語を含むタスクに対するHARPYのC.mmpとKL10でのインプリメントの比較

explicitな同期をとらなくても、認識の能力が余り劣らないことがわかった。従って、このアーキテクチャは、分散処理の性能低下とプログラムの難しさの大きな要因である同期の問題の負担を軽くするので、後のDistributed Interpretation Systemの良いモデルであるという。

## IV-2 画像理解

ARPA援助による画像理解プロジェクトは音声理解プロジェクトの後を受け、MIT、スタンフォード大、ロチェスター大、USC、SRI CDC、Honeywellなど10程の大学・研究機関・企業などを契約者に進められているもので、CMUも参加している。

CMUの研究テーマの1つには、画像理解システムのための計算機アーキテクチャがあり、その関係で SPARC (Symbolic Processing Algorithm Research Computer) という画像処理用高性能並列計算機を CDC と共同で開発している。

SPARC processor の構成は図8に示す様に、加算・乗算・シフト・データ記憶・I/O など多数の functional unit がスイッチで相互に結合される。スイッチは一般的で

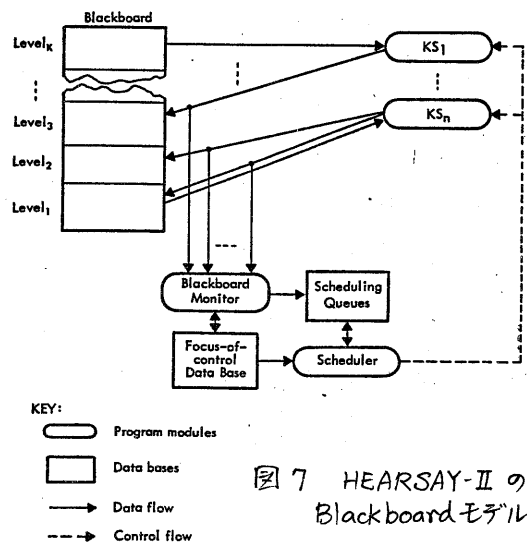


図7 HEARSAY-IIのBlackboardモデル

あり、マシンサイクルの速度で切換え可能で、これを通して  $10^{10}$  bit/sec のデータ転送速度をもつ。Memory Port は最大  $2^{35}$  アドレス空間、800 M bit/sec の転送速度を与える様に設計され、データ limited になりがちな画像処理の問題に対処しようとしている。Ring Port は複数の SPARC をつないでマルチプロセッサにあるため、及びホスト (PDP11 の Unibus) へのインターフェースのためである。SPARC は ECL を用い、 $10^8 \sim 10^{10}$  operation/sec を目標とし、1979 年秋に CMU へ納入の予定となっている。

画像理解研究のグループが専用に用いている計算機システムは PDP11/40E (C.mmp プロジェクトの過程で改造され、動的マイクロプログラム可となったもの) を UNIX で動かして、2 MB の IC メモリ、200 MB ディスク × 2 台、カラーディスプレイなどを備えたものであったが、これを現在、次の V で述べる学科全体の facility 改良とも関連して VAX11 に変更する予定である。

#### IV-3 広域 Distributed Interpretation System とそのモデル

Distributed Interpretation System は環境からデータを集めるセンサが広域に分布しており、それらのデータをすべて中央に集めることが困難な望ましくなく、比較的ローカルないくつかのセンサ site からのデータを速く解釈して全体として良い結果を得ようとするシステムである。レーダ、音響・光検出器などからなるセンサネットワーク、交通ネットワーク、動くロボットのタスク制御などの応用がある。

このような問題と討議するため、ARPA の援助で Distributed Sensor Nets のプロジェクトがあり、1978 年 12 月 CMU で Workshop があった。人工知能分野からのアプローチもいくつかあるが、その中で CMU で HEARSAY-Ⅱ を研究していた研究者達 (すでに他機関に移っている) は HEARSAY-Ⅱ のアーキテクチャが Distributed Interpretation を構成する際のモデルとして適当であるという主張をしている。

#### V. 研究 facility

CMU の計算機科学科における研究・教育用 (大学院しかないから、ほとんど研究用と言えるが) の計算機システムは

PDP10 3 台 (KA10 が 2 台, KL10 が 1 台ですべて TOPS10 time sharing)

C.mmp

C.m\*

音声研究用 PDP11/40E (UNIX time sharing)

画像研究用 PDP11/40E (UNIX time sharing)

などがフロントエンドプロセッサ (PDP11/40) につながり、ほとんどのターミナルはこのフロントエンドにつながっているため、どのシステムも自由にアクセスできる形になっている。

便利な Editor, XGP (ゼロックスプリンタ) を使った良い文書作成システム、Computer による Mail システムが整備されているので、事務や日常的活動にも便利である。faculty のメンバーは筆者の様な visiting faculty も含め、オフィスと自宅にターミナルをもち、計算機は年中無休 1 日 24 時間運転だから、いつでも使えるという恵まれた環境であった。

しかし、それでも彼らには不満が多く、主計算機である PDP10 のアドレス空間が 256 K しかないことから大きな AI のプログラムを作りにくい、ほとんどの端末がキャラクタ端末である、多数の人が logon している昼ごろはレスポンスが遅い、PDP10 から DEC の生産ラインからはあれていく、などから学科の facility を改善する計画が 1978 年ごろから教官・学生を含む faculty 委員会を中心に検討された。改善の方向は、高性能パーソナル計算機端末、TSS 機、ファイルマシン、高性能プリンタをネットワークでつないだ分散処理ネットワークが目標であった考えである。

このような研究用 facility の分散システムとしては、米国では Xerox 社 Palo Alto Research Center におけるものが有名である。Alto という高性能端末 (端末といっても、高解像度ビットマスク方式のディスプレイ、マイクロプログラム可能なミニコン、磁気ディスクが 1 セットとなった計算機) が各オフィスの机

A - ADDER  
F - FILE  
K - CONTROL  
I/O - INPUT OUTPUT  
M - MEMORY  
P - MULTIPLIER  
R - RING PORT  
S/B - SHIFT/BOOLEAN

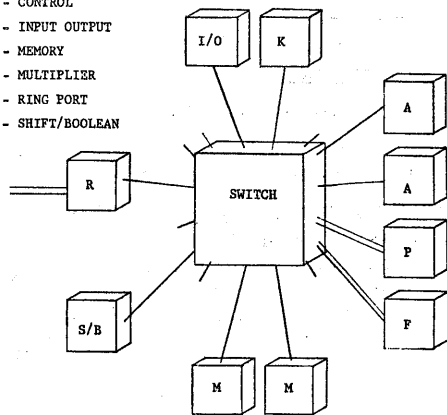


図8 SPARC の構成

の所にあり、EtherNet と呼ばれるネットワーク方式でつないだものである。このような高性能個人用計算機のネットワークというのが主要研究所のfacilityの方向で、MITではLisp Machineを多数つないだChaos ネットというのが作られてある。

CMUでも、VAX11 を主計算機とする分散システムにすることになった様で、最近聞いた所では、すでに7台のVAXが発注され、そのうち2台をプロジェクトマシン(画像とVLSIのグループ)用に使い、5台を学科共通の計算機とすることである。端末もカラーを含む高性能端末、前述のLisp Machine、更にAltoの流れをくみPascalを端末で翻訳実行するPascalto(仮称?)を導入することのことであった。

この改善を検討している時にXerox社よりMIT, Stanford, CMUに対しそれぞれ、Alto 15台、Dover(高精度レーザプリンタ)、Ethernet一式と寄付するというニュースが入り、寄付の大きさとタイミングに驚かされた。

## Ⅵ. おわりに

問題やシステムを分割し、並列アルゴリズムを考え、並列プロセッサにインプリメントして大きな効果を得るには、その並列プロセッサの

構造に適した分割や解法を考える必要があることがよくわかっている。事実、C.mmpやCmpのようなマルチプロセッサの開発に対し、「アルゴリズムの分解に関する一般的・自動的な解法がない現在では、先にマルチプロセッサの構造を与えても、ユーザに負担を与えるだけでメリットが少ない」という批判をする人もある。

いずれにしても、並列アルゴリズムに関する理論的研究は重要で、CMUでも若手の優秀な研究者が居て、良い成果をどんどん出している様である。VLSI化に適した解法というのも研究対象になっている。

約1年半CMUに滞在してみて、研究内容や設備の良さは勿論、先導的な計算機学科としての意気と誇り、改善に対する教官・学生の学科全体としての熱心な取り組み方にも学ぶ所が多かったと思う。

謝辞 日頃、御指導いただき、またCMUに滞在することにも御力添えいただいた本学、坂井利之教授に感謝します。