

推論・知識ベース・サブシステム

ICOT 第4研究室

内 田 俊 一
近 山 隆
瀧 和 男
横 田 一 正

講演の概要

- | | |
|--------------------|----|
| 1. 研究開発の概要 | 内田 |
| 2. KL 1 処理系とPIMOS | 近山 |
| 3. 並列応用実験プログラムとPIM | 瀧 |
| 4. 知識ベース管理ソフトウェア | 横田 |

昭和63年度の研究開発の全体的特徴

中期計画から後期計画への橋渡し

1. ソフトウェアやハードウェアの要素技術の統合化
一まとまりのシステムとなりそうなものは、まとめる。
2. 後期に向けての研究開発の目標，方針の見直し
目標の再設定，他のテーマとの統合などを行う。

推論サブシステムの研究開発内容

1. KL1 処理系とPIMOS

マルチPSIを土台として，並列ハードウェアから並列ソフトウェアまでを，一まとまりのシステムに統合化。

- a) 並列応用実験プログラム(詰め碁，構文解析など)
- b) PIMOSとKL1プログラミング環境
- c) KL1分散処理系(ファームウェアによる実装)
- d) マルチPSIハードウェア(要素プロセッサ 64台)

* FGCS' 88で，デモ版として統合化。

* 後期の並列ソフトウェアの研究開発ツールとして，リリース予定。

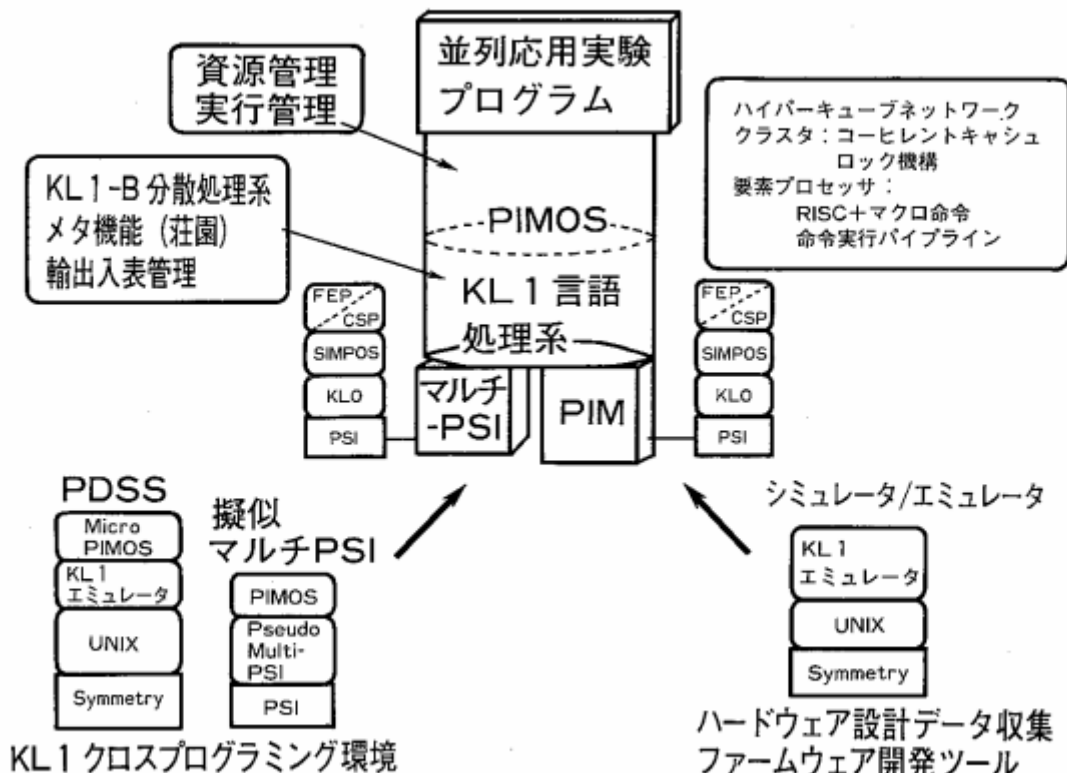
2. 並列推論マシン (PIM)

後期目標である要素プロセッサ1000台規模のプロトタイプ、ハードウェアの試作を考慮して、研究開発目標を再設定。

- a) 後期PIMに用いる小型・高速の要素プロセッサ
- b) 8台の要素プロセッサを密結合したクラスタ
- c) クラスタ間の結合ネットワーク (ハイパーキューブなど)
- d) 小粒度の並列処理向きのKL1 クラスタ内処理系
(VPIM処理系)

- * 要素プロセッサやクラスタの設計完了, 専用LSIなどの製造に着手。
- * 1000台規模のハードウェア試作に向けての, 開発体制づくり。

推論サブシステムの研究開発項目の関連図



知識ベースサブシステム

1. 知識ベース管理ソフトウェア (KBMS)

多様な知識表現を，効率よく扱える管理ソフトウェアを目標。
(自然言語辞書，定理証明DB，設計用DB，遺伝子DBなど)

- a) 非正規関係モデルに基づくDB管理 (Kappa-DB層)
- b) 演繹・オブジェクト指向・DBに基づく知識ベース管理
(Kappa-KB層)
- c) 分散環境での知識ベース管理と問い合わせ最適化 (PHI)
- d) DNA知識ベース処理実験システム (KNOA)

*PSI-II上で開発されたKappa-IIは，ツールとして提供。

*KNOAは，CHI-II上で実験システム試作。

2. 知識ベース並列処理方式

論理型言語と整合性のよいデータや知識の表現方式を持ち，拡張した関係演算を効率的に並列実行する処理方式。

- a) 関係演算を単一化を用いて拡張した全解探索のPIM上での効率的な実現方式。
- b) 論理型言語と整合性のよい拡張関係演算の並列実行を，効率よく行うための並列演算ソフトウェア及び実験用並列ハードウェアシステム (Mu-X)。

*知識ベース処理を効率的に行うための要素技術として重要。

*実験用のソフトウェア，ハードウェアを試作して評価。

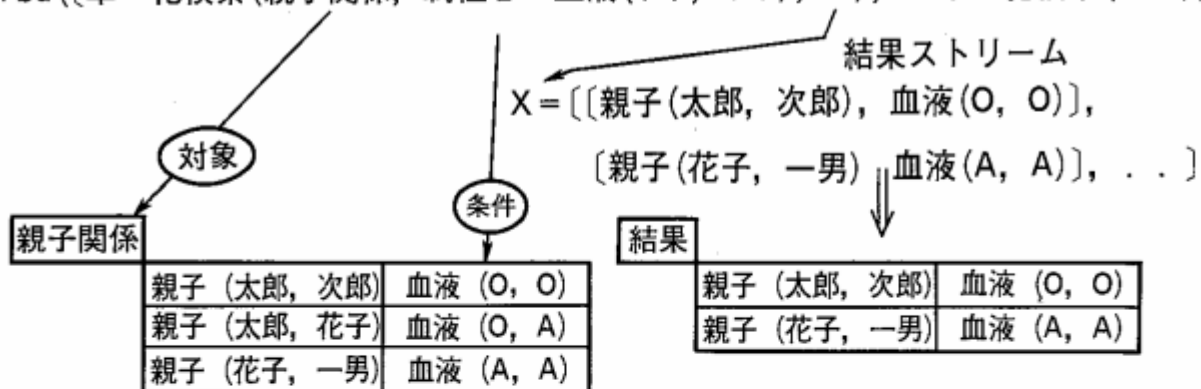
a) 関係演算のPIM上での実現方式

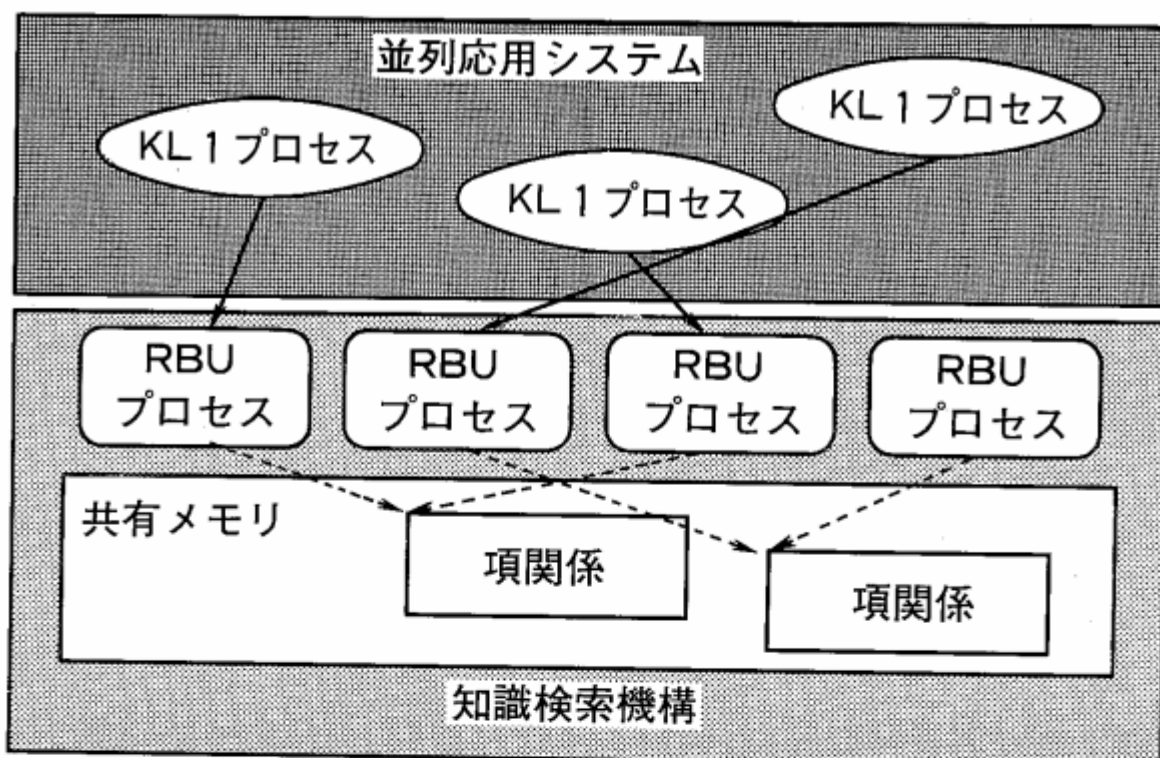
- 関係演算を単一化を用いて拡張
Retrieval By Unification (RBU)
- KL1の機能を強化 (集合演算: set-of, bag-of)
- 組込み述語として追加
- 検索コマンドごとの並列処理が可能
- KL1の処理系と統合したインタープリタの実装が問題

単一化による検索演算：検索の結果が、集合の形で得られる。

呼びだしの例

```
rbu([単一化検索(親子関係, 属性2 = 血液($1, $1), X), 単一化検索(. . .)])
```





b) 知識ベース並列処理実験システム (Mu-X)

- 関係演算に推論処理向き機能を追加・拡張
- 並列処理による高速化

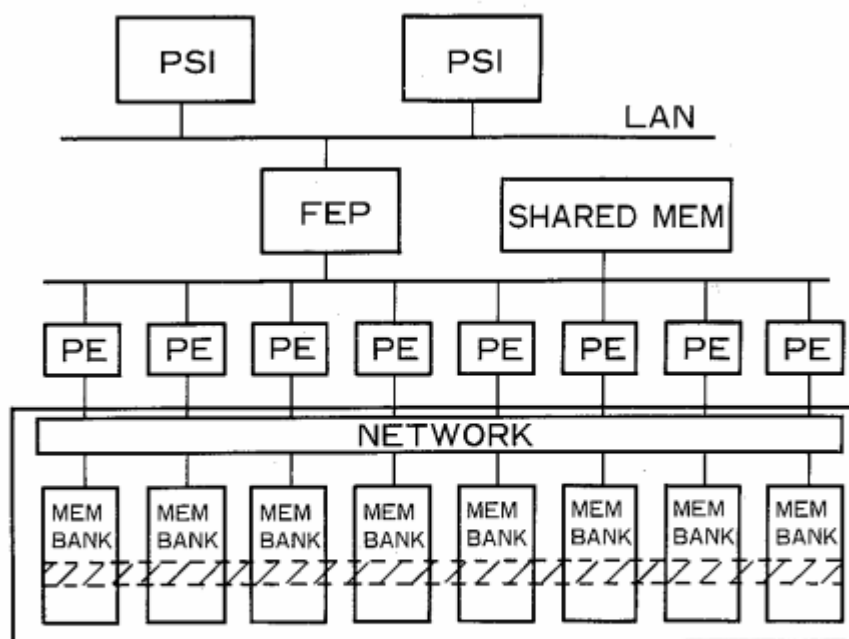
ソフトウェアの機能

- 項データ (構造データ) が扱える。
- マルチトランザクションのサポート
- 関係演算の並列処理

実験システムの特徴

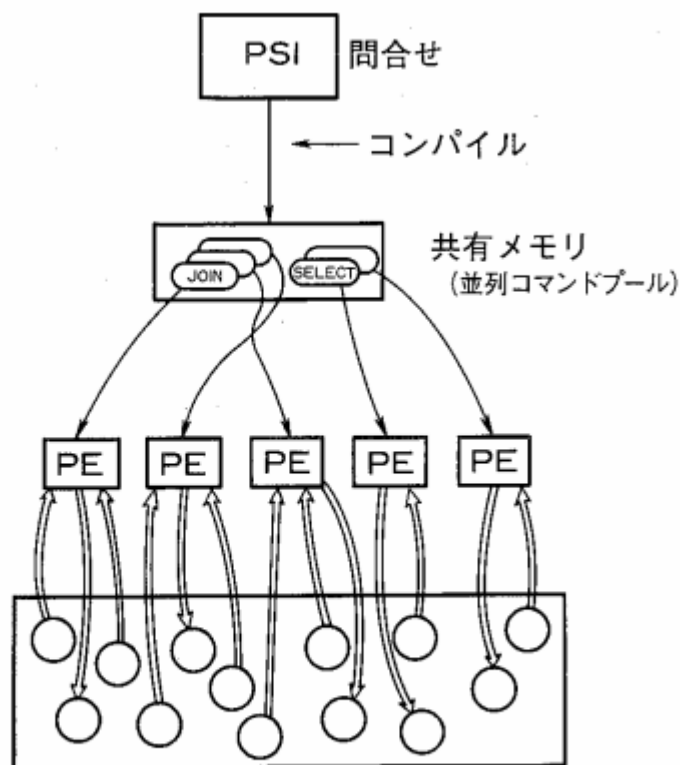
- 汎用マルチプロセッサによる並列処理
- マルチポートメモリの利用

実験システムのハードウェア構成



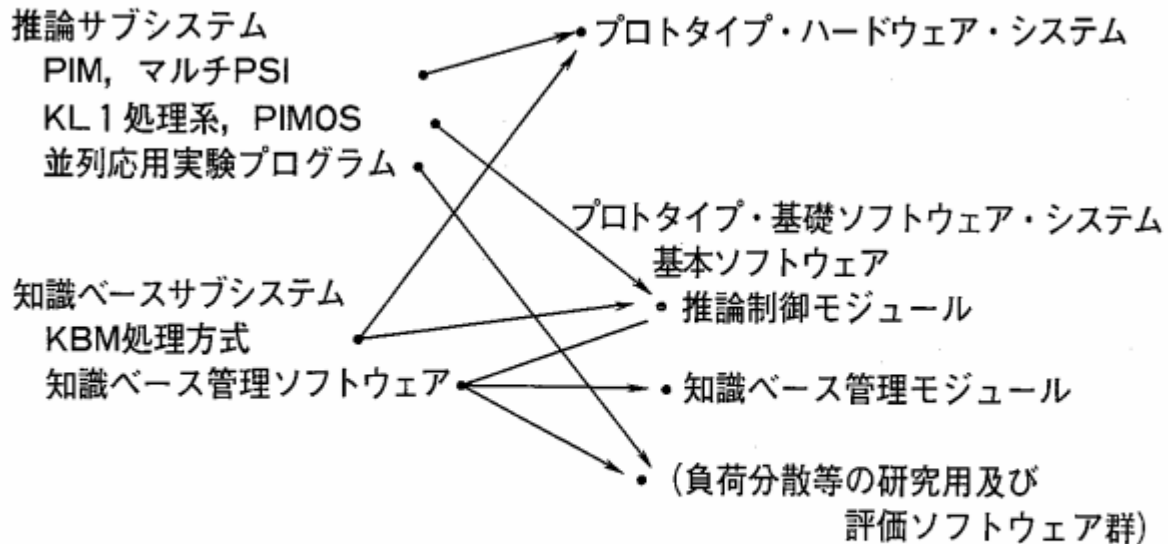
マルチポートページメモリ

関係演算の並列処理方式

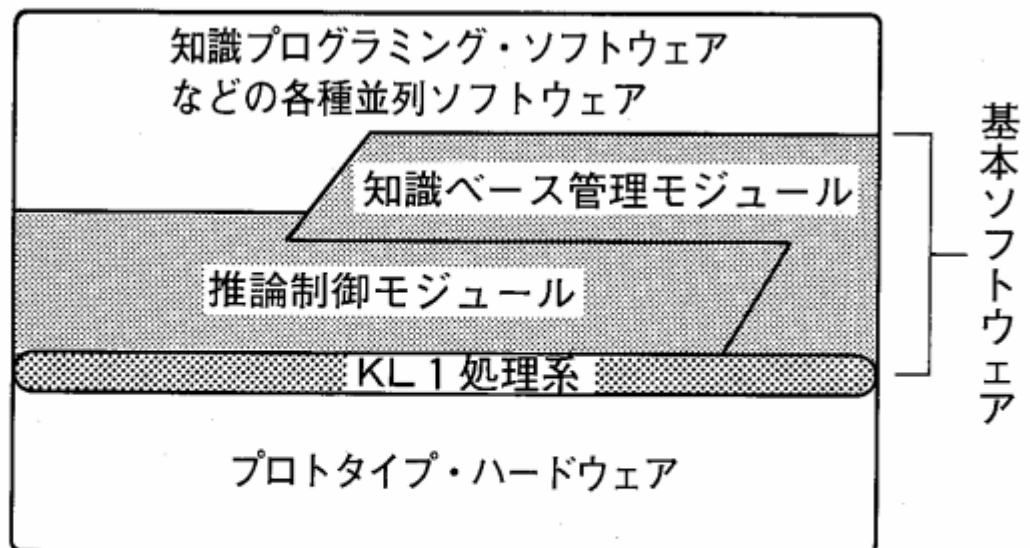


マルチポートページメモリ
(関係ページプール)

中期計画から後期計画への橋渡し



後期目標の5Gプロトタイプ・システムの構成



KL1 処理系とPIMOS

目標

- 並列推論処理方式の確立
- 並列ソフトウェア研究開発のベースを提供

構成

- 核言語KL1
 - ソフトウェアとハードウェアの接点
- マルチPSI
 - 並列推論マシンの実験機
 - 並列ソフトウェア開発ツール
- マルチPSI上のKL1 言語処理系
 - KL1 言語処理方式の確立
- オペレーティングシステムPIMOS
 - 並列推論の制御方式の確立

核言語KL1

基本部分 = Flat GHC

- 論理型言語
 - プログラムを論理式と解釈できる
- データフロー並列
 - 暗黙の並列性と同期機構

機能拡張部分

- 実行制御機能
- メタレベル制御機能
- 組込み機能の充実
- ← 本格的な並列処理システムには不可欠

機能拡張点

実行制御機能

- 実行プロセサの指定
 - 並列プロセス間の優先度指定
 - 選択肢間の優先度指定
- 並列処理の効率化の研究に不可欠

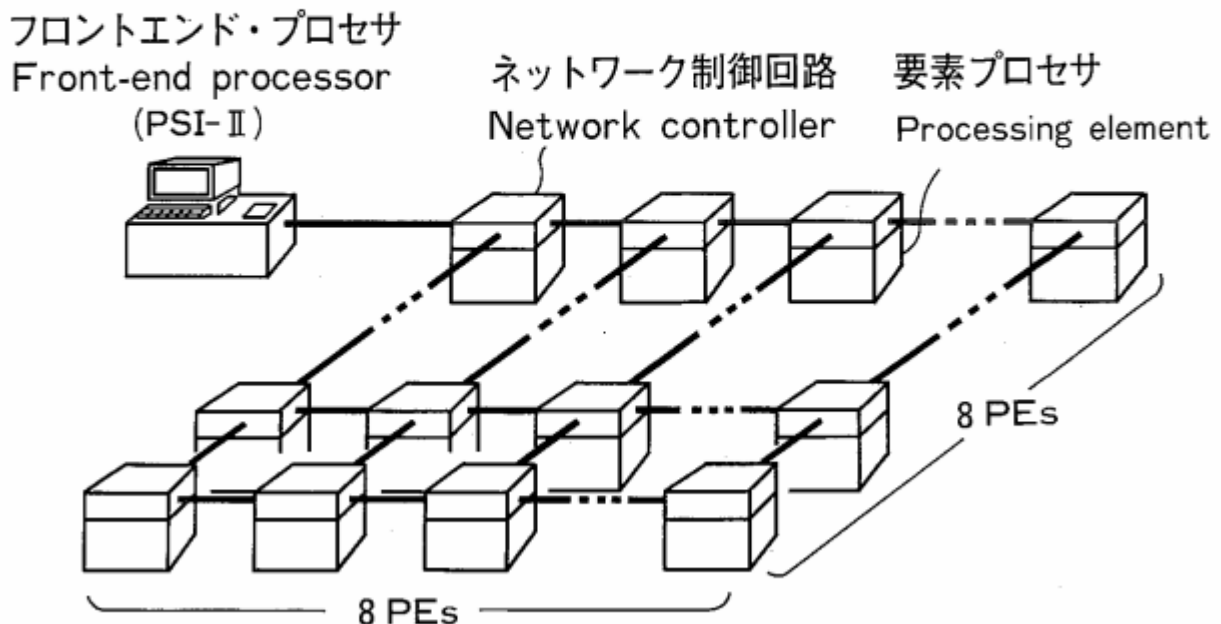
メタレベル制御機能（荘園機構など）

- OSの記述に不可欠(OSはメタプログラム)
- 複雑な応用プログラムの記述に便利

組込み機能の充実

- 一定の手間で更新できる配列
 - 組込みのストリームマージャ
- セマンティクスを保存したまま性能向上

マルチPSIシステムの概要



マルチPSI上のKL1言語処理系

単一プロセサ上での高効率実行

- 抽象機械語命令セットKL1-B
- 参照数管理 (MRB)
 - インクリメンタルGC
 - 一定の手間で更新できる配列

複数プロセサに渡る管理の効率化

- プロセサ間参照管理
 - 大域・局所の二重アドレス系(輸出入表)
 - 重み付き参照カウントによるGCの局所化
 - 分散実行の管理
 - 里親方式による管理の局所化
 - 重み付き参照カウントによる通信量削減
- ☆ できる限り局所的に管理 → 通信コスト低減

擬似マルチPSI

一台のPSI-IIでマルチPSIを模擬

- プロセサ → プロセス
- 広範に利用可能

実機と同じマイクロコード

→ 実機の単一プロセサと同等の速度

PIMOSの設計

方針

- すべて論理型言語 KL1 で記述
- 疎結合並列計算機を想定
 - 通信コストを意識
 - 集中管理を排除

特徴

- 副作用なし・暗黙の同期機構を利用
- システム全体を木構造に管理
 - 局所的な情報で管理可能
 - 通信コスト低減・負荷集中回避

PIMOSの開発

開発過程

1. エミュレータ (PDSS) 上で開発
2. 擬似マルチPSIに移植
3. マルチPSI実機に移植

開発を通して

- 高水準な記号処理言語で記述
 - デバッグが容易
- メタレベル制御機構
 - 低レベルな処理を処理系に任せる
- データフロー同期機構
 - 同期のバグは皆無

核言語の高水準化

○ OSの記述が容易に

× 処理系の負担増

逐次システムでは：

- ふたつの異なる概念を同じ逐次処理で実現
本質的な逐次性(データ依存関係)
効率上の処理の優先関係

→ 並列システムでは両者を区別しないと効率低下

- メタレベル処理をOSに依頼(低効率だが低頻度)

→ 並列システムでは高頻度

並列マシンでは応用プログラムも並列動作

→ 従来OSだけに必要だった機能が応用プログラムでも必要

→ 言語処理系で用意し、共通に利用するのが得策

研究開発の現況

- マルチPSI上で処理系が動作

- OSの基本機能を完成

- FGCS'88などでデモ

- 並列ソフトウェア開発研究に利用開始

今後の研究開発

KL1 言語仕様

- 利用されない機能の削除による簡素化
- 必要な機能の拡張・強化

KL1 言語処理系

- 機能強化(ソフトウェア開発支援機能など)
- 最適化の推進
- 分散処理技術をPIM上の処理系に適用

PIMOS

- ユーザインタフェースの整理
- 記述言語をさらに高水準化
- プログラム開発支援機能の強化
- 入出力機能の強化

並列応用実験プログラム

開発目的 — 並列プログラムが組めるための技術の蓄積

並列アルゴリズム

プログラミング・パラダイム

応用領域指向のユーザ言語

→ ユーザ向け技術

— 効率の良い並列実行制御のための技術蓄積

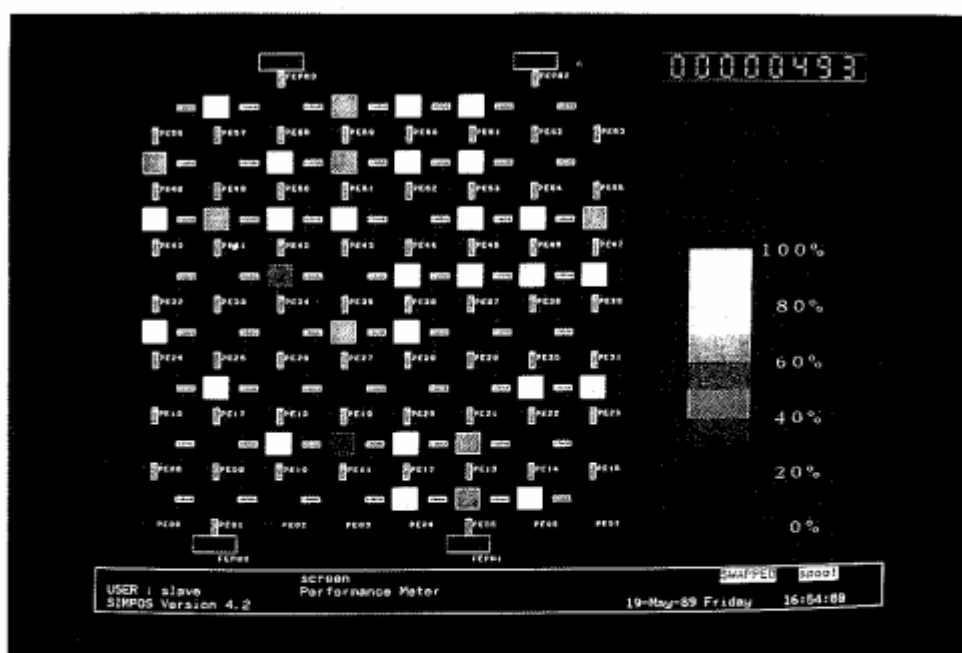
負荷分散方式，通信の局所化方式

→ システムに反映すべき技術

問題領域：疎結合大規模MIMDマシン上の非定形記号
処理

4つの並列応用実験プログラム

1. 自然言語構文解析プログラム：PAX
 - 談話理解システムDUALSの構文解析機能を並列化
 - レイヤード・ストリーム法による全解探索問題
 - 通信量削減を主眼とした静的負荷分散
2. 詰め碁(囲碁対局の部分問題)
 - ゲーム木探索($\alpha\beta$ 枝刈り法)の並列化
プライオリティを用いた並列アルゴリズム
 - 動的負荷分散, 静的負荷分散の実験
 - プロセッサ台数の平方根に比例する性能向上
3. 最適(最短)経路問題
 - 1万ノードのネットワークにおける最短経路探索
 - 最良解探索問題の並列アルゴリズム
 - 分布するプロセス間にローカルな通信がある場合の負荷分散
4. 詰め込みパズル(ペントミノ)
 - “OR並列型”の全解探索
 - 64プロセッサで45~50倍の性能向上



プロセッサ稼働率の表示画面

並列応用実験プログラムの開発状況

前記4種のプログラムの初版をFGCS'88にデモ出展
 今回改訂版をデモ実演中

今後の課題

- 新たな問題，より大規模な問題の並列プログラム化
- 応用技術者と並列処理技術者の密な連携
- 並列プログラミング人口の拡大，技術の普及
- 個別プログラム開発の工夫点から一般性を抽出
 類型化とテクニックの標準化
 負荷分散の一部システムサポート

並列推論マシン(PIM)

- 大規模な並列推論処理を高速実行するプロトタイプ・ハードウェア
- 並列ソフトウェア技術を改良・蓄積してゆく実験環境アプローチ：

基本となる並列推論方式—マルチPSIでほぼ検証
↓
高性能化，大規模化の方式—多くの不明点と可能性
↓
複数方式について実験評価(複数モジュールの開発)
↓
成果の相互フィードバック(処理系レベルへ還元)
↓
複数モジュールを実験環境として接続・統合

PIMの後期目標

1. プロトタイプ・ハードウェア・モジュール群
 - 異なる方式を追及した複数ハードウェアモジュールを開発し実験評価，成果を相互還元
 - 1モジュール当たり
数100台規模の要素プロセッサ(PE)を接続
推論性能30～300M LIPS (512PE 構成時)
2. 複数モジュールを疎結合により統合した実験環境
全体で1000PE程度，計算能力100～1000M LIPS
3. PIMOSおよび複数の並列応用プログラムが稼動

PIM/p (推論処理モジュール)の構成と特徴

1. 要素プロセッサ

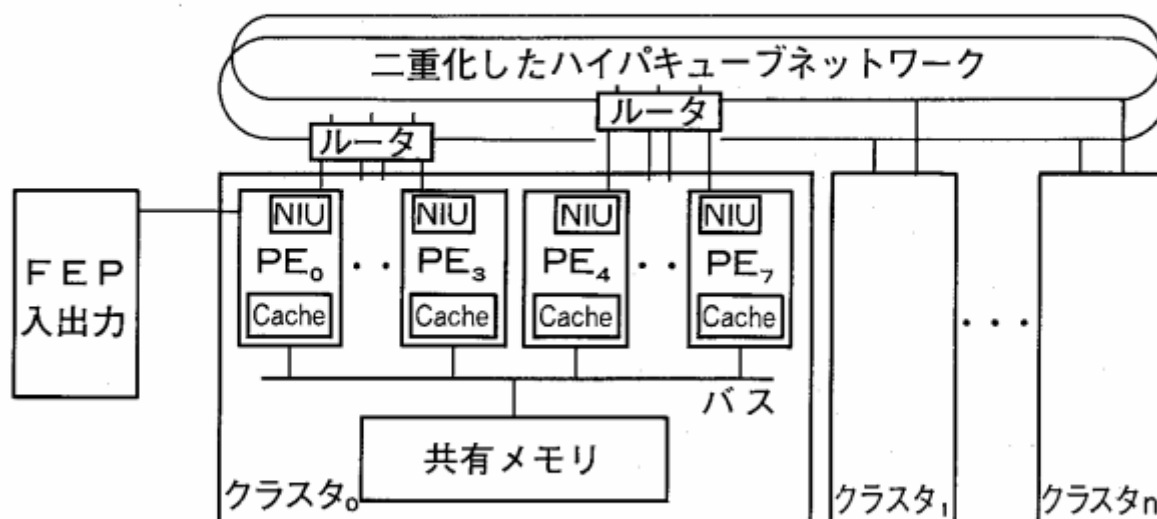
- 抽象機械語命令セットKL1-B*とRISC命令の融合
- MRB方式による実時間GC*
- タグアーキテクチャの採用*とパイプライン処理

2. 接続方式

- 8 PEを密結合したクラスタ構成
- 一貫性を保つ並列キャッシュメモリ
- 二重ハイパーキューブによるクラスタ間接続
- クラスタ間実時間GC(WEC方式)*

* 共通仕様

PIM/pのモジュール構成



PIM/c, PIM/m, その他のモジュール

- クロスバ・スイッチ，2次元格子網，クラスタ内二重化バス，その他の接続方式による実験・評価
- 水平型マイクロ命令制御等を用いる要素プロセッサ
- クラスタ間自動負荷分散，ユーザプログラムとシステムの協調による負荷分散制御，等に焦点を当てた処理系改良実験

処理系開発システム

- 仮想PIM処理系仕様(KL1 処理系共通部)：VPIM
- VPIM記述言語：PSL = C言語風マクロ言語
各モジュール対応処理系中核部の共通化
共通化による開発工数低減
技術移転および仕様拡張・変更の容易化
- 仮想PIMのシミュレータ
VPIM仕様の早期確認 — (複雑な)高機能分散処理系の開発に必須

シミュレータ ← VPIM → 各PIMモジュール処理系

PIMの開発状況

- 各モジュールに付き，要素プロセッサ，ネットワーク制御等のLSIの初版を製造中（一部詳細設計中）
- 最初のモジュールは，元年度に数クラスタを接続したハードウェア試験システムを製造予定
- VPIM：クラスタ内処理系共通仕様の初版を開発
小規模プログラムによりVPIM仕様を検証
クラスタ間仕様の開発と全体仕様の改良を継続中
- 各モジュールに付き，ハードウェア試験，保守，処理系の低レベル・デバッグ等の支援機能を開発中

知識ベース管理ソフトウェア KBMS

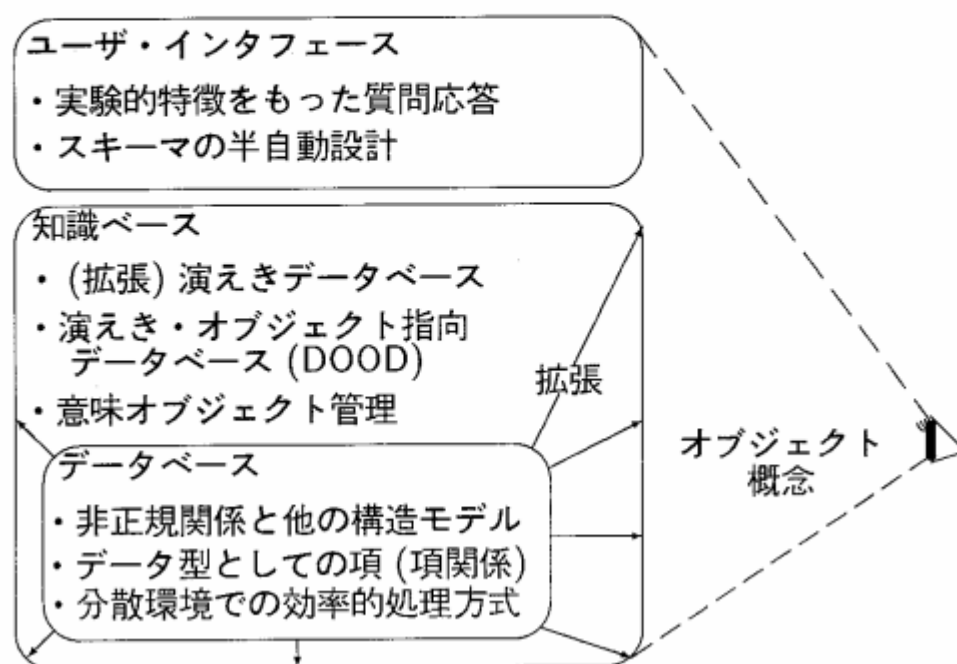
研究開発内容

- データベース / 知識ベース管理 (Kappa-II ⇒ Kappa-P, DO-I)
- 分散知識ベース処理 (PHI ⇒ DO-φ)
- 知識オブジェクト管理 (ETA ⇒ DO-η)
- DNA 知識ベース (処理) 実験 (KNOA ⇒ GIP)

KBMS の研究開発の目的

- 大量のデータと知識を効率的に管理
 - (項を含む) 構造データの記憶モデル
 - ラージ・オブジェクト格納, 拡張非正規関係モデル, …
- データと知識の効率的な表現
 - 複雑なデータと知識に対する表現モデル
 - 拡張非正規関係モデル, 複合オブジェクト, 意味ネットワーク, …
- データと知識の効率的な処理
 - データ / 知識の分散化と質問処理の最適化 / 並列化
 - 演えきデータベース, オブジェクト・ベース, オブジェクト指向データベース, …
- 具体的な (並列) 知識処理による実証
 - PSI/Multi-PSI/PIM, SIMPOS/PIMOS, ESP/KL1 環境
 - 自然言語, 証明支援, CAD, 遺伝子情報, …

KBMS の基本的枠組



データベース管理 Kappa-II (DB, UI)

- PSI, SIMPOS, ESP 環境でのデータベース管理システム
- 構造データに対するモデリング (記憶モデル, 表現モデル)
 - 拡張非正規関係, 意味ネットワーク (ETA), ...
- ★ PSI 上のツール化 (⇒ Kappa)
知識処理応用で使用する (項を含む) さまざまなデータを効率的に処理
- 並列データベース管理システム (⇒ Kappa-P)
 - Multi-PSI/PIM, PIMOS, KL1 環境でのデータベース管理システム
 - 分散データベース + 主記憶データベース
 - 並列処理固有の最適化

知識ベース管理 Kappa-II (KB)

- 確定節 (関係モデル), CRL (非正規関係モデル) に基づいた演えきデータベースを Kappa-II (DB) 上に実装
- 一般マジック集合法 / セミナイーブ法による質問処理最適化 / 評価
- 演えき・オブジェクト指向データベース (⇒ DO-I)
 - 動的構造 (外延階層) を持った演えきデータベース
 - 演えきデータベースのオブジェクト指向パラダイムへの組込み
 - オブジェクト指向概念の演えきデータベースへの組込み

分散知識ベース処理 PHI

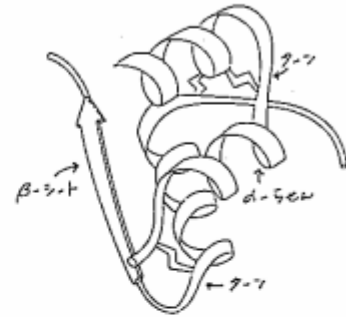
- 関係モデル (分散データベース管理システム), 確定節に基づいた演えきデータベースを実装
- 重ね合わせ符号 (SCW) を用いた高速索引方式を実装
- (HCT/P, HCT/S,) HCT/R とセミナイーブ法による質問最適化 / 評価
- 分散環境での質問処理最適化アルゴリズム
- 格納データのオブジェクト化 (\Rightarrow DO- ϕ)
 - ・演えき・オブジェクト・ベース
 - ・オブジェクト指向概念の取込み

知識オブジェクト管理 ETA (DO- η)

- 意味ネットワーク支援モジュール実装 (Kappa-II)
- ★ 汎化階層の意味ネットワーク群への組込み
- ★ 意味ネットワークを含む表現言語検討
- 技術文献検索での評価を予定

DNA 知識ベース (処理) 実験 KNOA

- 知識ベース処理機能の抽出
 - 知的検索と構造推論
 - 多重継承の階層型知識ベース
- DNA 配列検索による評価システム
 - 並列応用としての遺伝子情報処理 (⇒ GIP)



KBMS の研究開発計画

