

大規模KBM

ICOT研究所

第3研究室

物井秀俊

第5世代コンピュータプロジェクトにおける 知識ベースシステムの研究

前 期 (57° ~ 59°)

関係データベースマシン (Delta 1)

中 期 (60° ~ 63°)

知識ベースマシン

●パイロットモデル

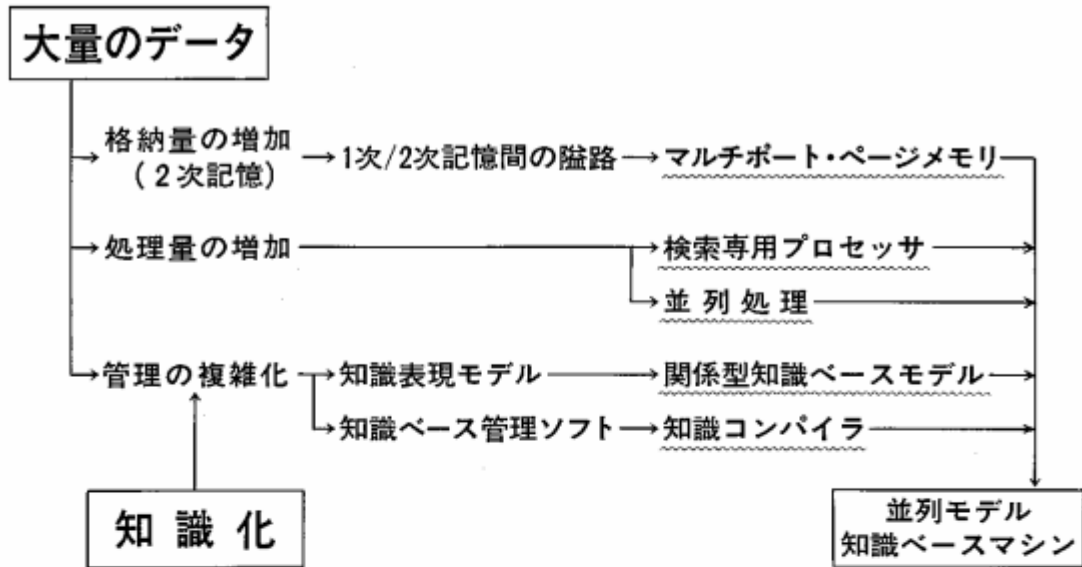
●分散モデル

●**並列モデル**

- 並列アクセス制御メカニズム
- 並列資源管理制御メカニズム
- 演算特性・処理粒度

〈ハードウェア実験機〉

大量知識処理の要求条件と解決策



関係型知識ベースモデル

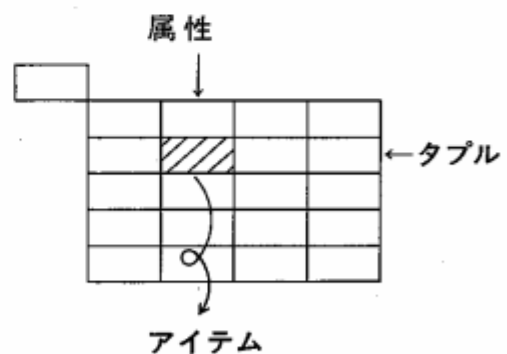
背景

- 大量の知識に対する高速な検索機能
- 各種の知識表現が扱える柔軟な操作対象
- 形式的な取り扱いに耐え得るモデル
- 関係データベースで得た技術を利用



関係型知識ベースを提案

項関係



〈定数〉 関係データベース



〈項〉 関係型知識ベース

関係型知識ベースの操作

項関係に対する単一化を用いた検索

RBU (Retrieval By Unification)

単一化結合 \longrightarrow $T_a \underset{\text{att}_a \diamond \text{att}_b}{\bowtie} T_b$
Unification-Join

単一化制約 \longrightarrow $\sigma_{\text{att} \diamond \text{condition}} T$
Unification-Restriction

射影 \longrightarrow $\pi_{\text{att}_1, \dots, \text{att}_n} (T)$
Projection

◇：単一化操作を示す。

項関係の例

an (X, Y) : -par (X, Y)
an (X, Y) : -par (X, Z),
 an (Z, Y)

par (smith, clark)
par (clark, turner)

⋮

an : ancestor
par : parent



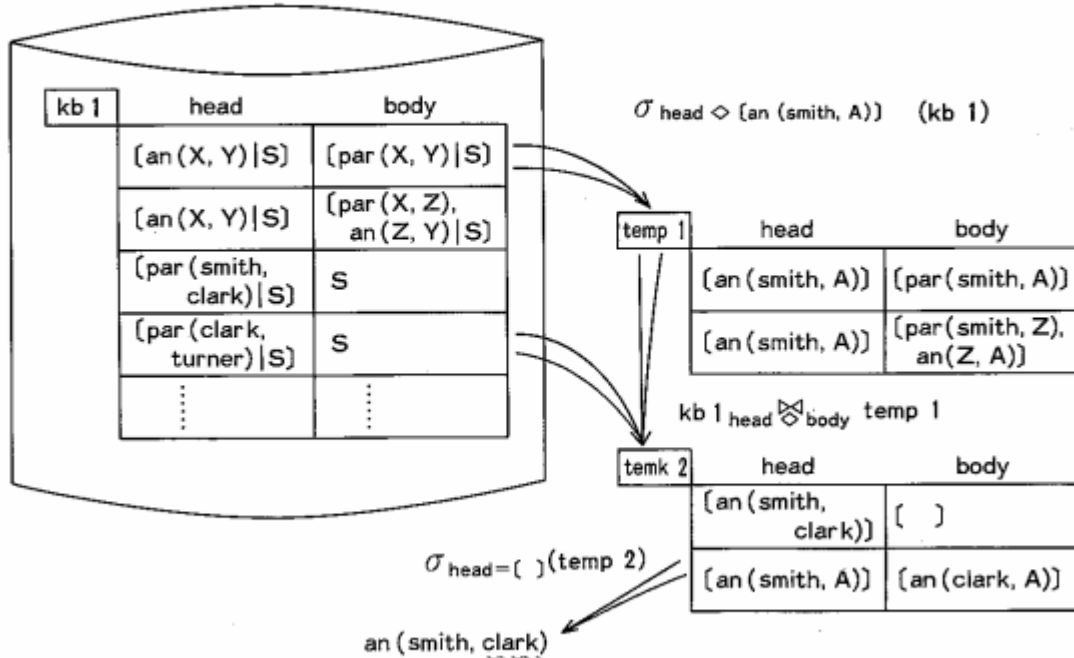
[an (X, Y) S]	[par (X, Y) S]
[an (X, Y) S]	[par (X, Z), an (Z, Y) S]
[par (smith, clark) S]	S
[par (clark, turner) S]	S
⋮	⋮

S : 2属性に共通な変数

RBUによる演繹の例

問：smithの祖先

〈関係型知識ベースの検索〉



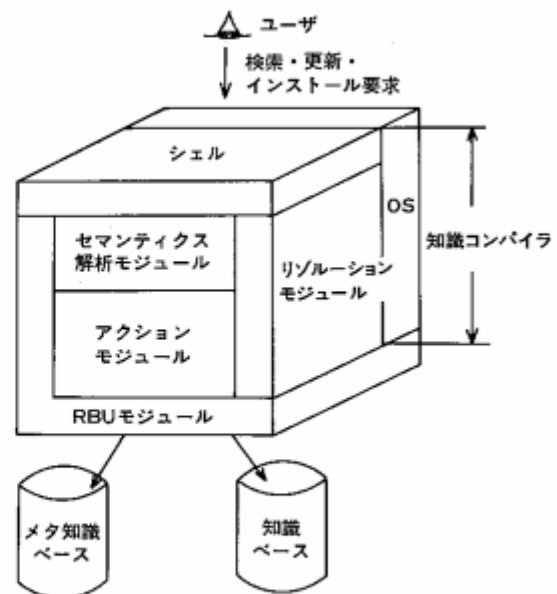
知識コンパイラ

目的

- プリミティブな格納形式とホスト側との論理的インタフェース。
- 知識ソース、メソッド、コンストレイント等の格納と管理。
- 検索、更新等の処理を知識に対する specific な操作として処理する。



メタプログラミングの枠組みで统一的に扱うインテリジェントな知識ベース管理システム。



並列モデル知識ベースマシンの目的

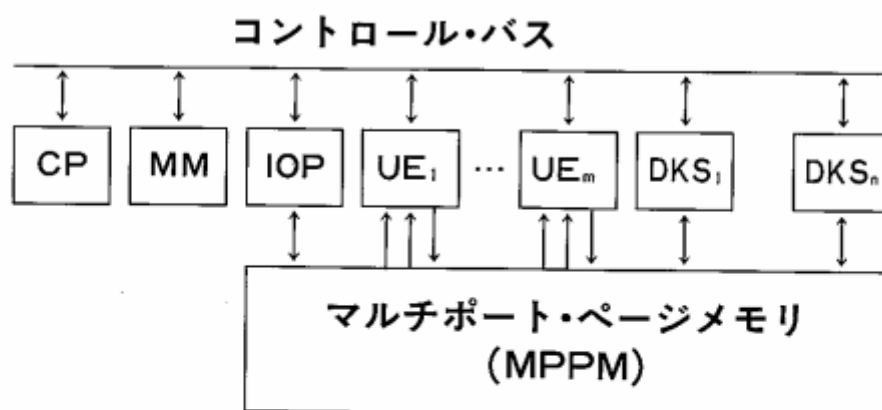
- 大量データに対する処理の高速化……………専用プロセッサ
並列処理
- 二次記憶と主記憶間の隘路の解消……………バス幅の拡張
- 二次記憶上の大量データに対する高速処理……………ストリーム処理
- RBU演算の高速実行
 - 単一化処理の高速化……………アルゴリズムのハードウェア化
 - 効率の良いデータ分配……………アクセス競合のない共有メモリ



アーキテクチャ

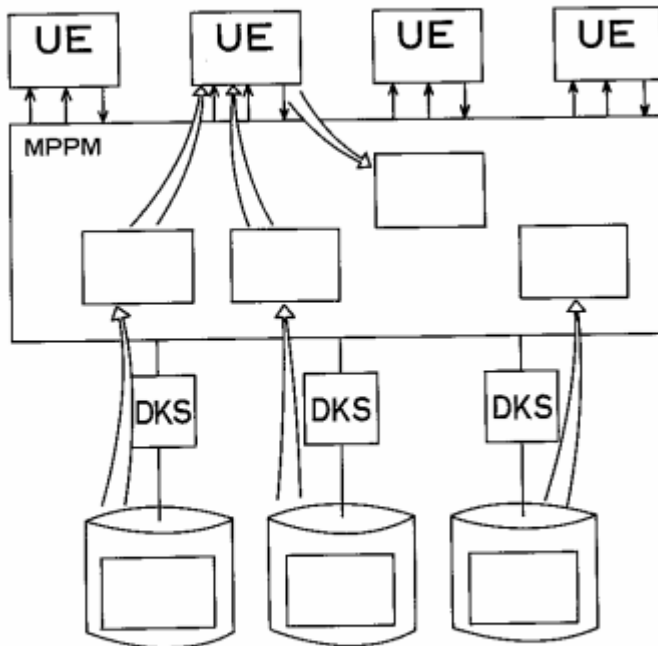
〈単一化検索専用プロセッサ〉+〈マルチポート・ページメモリ〉

並列モデル実験機の全体構成



- UE : 単一化エンジン (Unification Engine)
DKS : ディスクシステム (Disk System)
CP : 制御プロセッサ (Control Processor)
MM : 主メモリ (Main Memory)
IOP : 入出力プロセッサ (I/O Processor)

データの流れ



<各要素の動作>

UE

- MPPM上のデータをストリームとして入力し、MPPM上に出力。

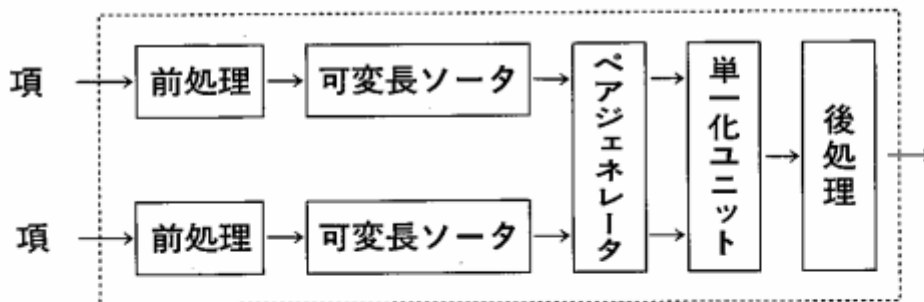
MPPM

- 二次記憶と UE 間のキャッシュメモリ
- 複数 UE の共有メモリ

DKS

- MPPM に対するステージング/デステージング

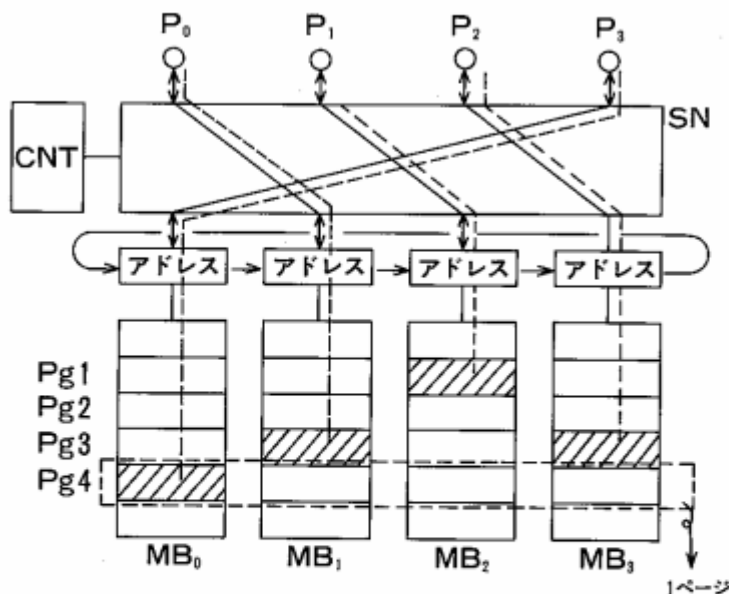
単一化エンジンの内部構成と動作



- 前処理部
：タプルからの属性の切り出し処理
- 可変長ソータ部
：項のジェネラリティによるソート処理
- ペアジェネレータ部
：単一化可能な項のペアの生成
- 単一化ユニット部
：mguの生成と置換操作
- 後処理部
：タプルの形成

— ストリームを対象とするアルゴリズムにより、パイプライン的な並列化を狙う。

マルチポートページメモリの構造とアクセス方式



CNT : ネットワークコントロール
SN : スイッチングネットワーク
MB : メモリバンク
P : ポート
Pg : ページ

特徴

- 複数のポートから同一のページを競合なしで同時にアクセス可能。
- アクセスはページ単位。
- 任意のポートから任意のページをアクセス可能。



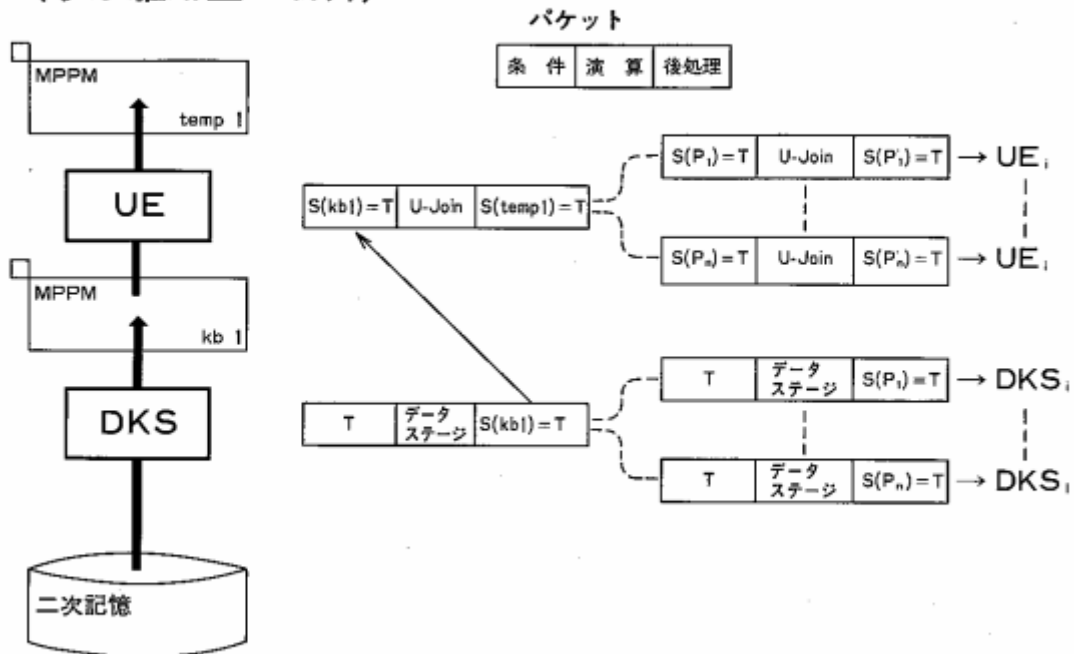
ページ単位の並列処理。

制御の目的

- 並列アクセス制御
 - 検索処理の並列化
 - 演算とデータ量に従う物理資源の割り当て戦略
 - 演算順序の最適化
 - アクセス競合の極小化
 - アクセス単位の最適化
 - プロセッサ間の同期処理
- 並列処理資源管理
 - プロセッサ割り当ての効率化
 - メモリ割り当ての効率化

制御方式の一例

(事象駆動型の制御)



開発の現状と今後の課題

- 〈現 状〉
- ソフトウェアシミュレータ
 - ・ 単一化エンジンのシミュレータ作成
 - ・ 並列処理制御部のシミュレータを検討
 - ハードウェアシミュレータ
 - ・ マルチポートページメモリの設計
 - ・ 単一化エンジンのハード化を検討
 - ・ プロセッサ結合方式の詳細化
- 〈今後の課題〉
- 並列処理制御方式の詳細化
 - ・ プロセッサの割り当て戦略
 - ・ メモリ管理方式
 - ・ 処理単位
 - RBU演算処理系の詳細化
 - 知識ベース管理方式の詳細化