

TM-0510

知識ベースサブシステムの研究開発状況

伊藤英則

May, 1988

©1988, ICOT

ICOT

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

(03) 456-3191~5
Telex ICOT J32964

Institute for New Generation Computer Technology

----- 知識ベースサブシステムの研究開発状況 -----

このメモは昭和63年5月31日～6月1日九段会館で開催される第6回 第5世代コンピュータに関するシンポジウムで報告する知識ベースサブシステムの研究開発状況について記述したものである。

1. 標題 知識ベースサブシステム

引き続きまして知識ベースサブシステムについてご報告いたします。

2. 開発経緯とこれからの計画

これは、知識ベースサブシステムの研究開発の経緯とこれからの研究開発計画を示したものです。これまでの知識ベースサブシステムの研究開発は先ず、前期では関係形のデータを事実形の知識であるとして関係データベースサブシステムの実験機を開発いたしました。

引き続きまして中期では、知識ベースサブシステムにおきましても、知識の集合を高度に扱うためには推論機能が不可欠でありますので、前期に開発した逐次形推論サブシステムをベースとして、前期で開発した関係データベースサブシステムから得られた知見を吸収統合して、知識ベースサブシステムの実験機と、ESP言語による分散知識ベース管理プログラム、また、GHCにより並列知識ベース管理のための実験プログラムを開発してまいっています。

本日はこれらを中心にしてご報告いたします。

なお、後期では、並列推論サブシステムと知識ベースサブシステムを融合し、並列推論クラスタと知識ベースクラスタから構成される5Gプロトタイプハードウェア上に並列知識ベース管理プログラムを研究開発して行く計画であります。

3. 知識ベースサブシステムハードウェアアーキテクチャ概念図

これは、中期における知識ベースサブシステムの実験機のハードウェアアーキテクチャの概念を示したものです。これは先程、ご紹介いたしました5Gプロトタイプハードウェアの構成要素にあたるものであります。5Gプロトタイプハードウェアは並列推論クラスタと、知識ベースクラスタと、これらを結合するネットワークから構成することになっています。

知識ベースクラスタの構成要素としては、ワレン風の抽象命令セットを基本とする8個の知識ベースプロセッシングエレメントと、共有メモリと、クラスタ制御ネットワークと、マルチポート制御ネットワークとから構成します。

知識ベースクラスタ内の共有メモリでは、知識ベースを蓄積する目的から管理単位を比較的大きくしていますが、8個の知識ベースプロセッシングエレメントは知識ベースクラスタ内のマルチポート制御ネットワークを介して共有メモリに接続されますので、これらの知識ベースプロセッシングエレメントは、共有メモリの同一論理空間の各々のサブエリアに同時にアクセスが可能となります。

また、並列推論クラスタと知識ベースクラスタとは5 Gプロトタイプハードウェアの上位階層のネットワークで接続されます。

4. 知識ベースクラスタの実験機（写真）

このスライドは知識ベースクラスタ構築に向けた実験機であります。左半分が8個の知識ベースプロセッサ部であり、右半分の上が共有メモリ部で、右半分の下側がマルチポート制御ネットワーク部であります。現在は、知識ベースプロセッサの並列制御プログラムとマルチポート制御ネットワークのための制御プログラムを開発して、知識ベースプロセッサの並列動作確認および評価を続けながら、また、共有メモリ内に実験用に文献データを入れて、演繹検索の並列制御の実験を行っています。

今後は、8個の知識ベースプロセッサ部の並列化のためにスタック、トレール制御機能を追加してゆきまして、5 Gプロトタイプハードウェアの構成要素として発展させてゆく計画であります。

5. 知識ベースクラスタ内プロセス構成概念図

つぎに、並列推論クラスタ内の推論プロセスと知識ベースクラスタ内の知識ベースプロセスの構成概念とこれらの関係についてお話しします。

並列推論クラスタ内には並列に動作する推論プロセスが存在し、知識ベースクラスタ内には知識ベースプロセスが存在します。知識ベースプロセスは推論プロセスへは知識ベースクラスタ内に知識としてクローズ集合があるように見えます。知識ベースプロセスはさらに、並列管理プロセスと検索実行プロセスからなります。並列管理プロセスは推論プロセスからのトリガーを受けて、大量の知識を効率良く手分けして検索するために、複数の検索実行プロセスを生成し、これを管理します。

ですから、知識ベースプロセスは推論プロセスの処理の（たとえば、定理証明等の）アクセラレータとして位置付けることになります。

後程、知識ベース管理プログラムのところでご説明いたしますが、複数の検索実行プロセスを8個のプロセッシングエレメントの負荷が均等になるように割り付けたり、自分の外のクラスタに処理依頼する制御が必要であります。

これは並列推論クラスタでの制御と同様であります。知識ベースプロセスの並列制御におきましては、処理対象が大量であることから、知識を集合処理する機能と、また、その中から全部の解を探索集取することが要求されます。これを全解探索と呼びますが、このための試行錯誤処理機能などが主な特徴であると言えます。

先程の講演で紹介がありましたが、推論プロセスの並列制御のために並列論理形核言語GHCが開発されています。このGHCの潜在能力をうまく引出して、知識ベースクラスタ内の知識ベースプロセスの並列制御に活かすことと、推論プロセスと知識ベースプロセスとの整合の取れたインタフェースを設定することが研究課題

であります。現在、私どもはこのための実験プログラムを作成し、評価検討中であります。

これについて、これからご説明いたします。

6. 知識の表現言語と基本的処理機能

知識ベースサブシステム内に、オブジェクトをどのような言語で表現し、どのように整理して蓄積して置き、また、その動的検索処理のためにどのような基本機能を準備しておくかの考慮が必要であります。

知識ベースサブシステムを構築するためには、並列論理型核言語 G H C の上位に高度な知識表現、知識獲得のための言語と、知識ベース管理のための言語を設定する必要があります。また一方では、知識を共有化するために、プリミティブな知識表現にする必要がありますが、このために並列論理型核言語 G H C のなかに、または、さらに下位に、原始的な共通知識表現言語を設定する必要があります。

ここではその基本を確定節すなわちディフニットクローズまたはその構成要素である項すなわちタームであると仮定しています。

ここで、高度に表現された知識から、共通利用のために原始的に表現された知識への知識コンパイル処理が必要になります。

また、上位の知識表現言語としては人工知能のユーズに柔軟に、しかも、高度に対応するために、従来からオブジェクト指向プログラミング言語、とか、オブジェクト指向知識ベース管理言語、さらに、知識ベース問い合わせ言語の研究が盛んでありますがこれらは、知識コンパイル処理を一例といたしましても、同一論理体系の中で扱えれば大変便利であります。この要求条件を備えた上位の知識集合がコンパイルされて下位に大量に蓄積されることが想定されます。このため、大量の知識集合を下位で効率よく扱えるように、知識ベースサブシステムの基本機能を考慮しております。

なお、機能の高度化と処理効率の問題は、常にトレードオフの関係にありますので、これらの研究開発には種々の実験を重ねて、特に、並列化処理に重点をおいて、研究開発をすすめております。

知識ベースクラスタ内で知識として扱われるものは高度でありますので、論理的な構造を持っています。たとえば、記号列／項／リストであり、さらには、これらの集合であります。また、これらの処理の効率化のためにはソート／クラスティング／トライ化／重ね合わせ等により、上位プログラムの使用目的に合わせて知識を動的に整理整頓して蓄積管理することが必要であります。これらの方法の中に単一化演算を意識した幾つかの新しいアイデアを導入したことは既に前回ご報告いたしました。

なお、これらの論理構造体の間の関係付けにはタプル／テーブル／非正規形／ネステッドリレーション／ネットワーク等が考えられますが、これらも上位プログラムのそれぞれの知識モデルに依存します。ですから、今後は、この後の講演で報告される実証プログラムとの結合を密にとりトータルで評価してゆきます。

7. G H C による並列知識ベース管理実験プログラム

これは、知識ベース指向の応用プログラムと並列知識ベース管理プログラムと並

列検索実行管理プログラムとからなるシステムの全体イメージを示した図であります。

知識ベース管理プログラムは応用プログラムと並列検索実行管理プログラムの中間に存在して、知識ベース指向の応用プログラムのために並列に問題解決を実行する基本機構を構築するのが狙いがあります。

このために、クローズデータの管理機能の代表的なものとしまして、調節／均衡化／同化についてGHCによる実験プログラムにより、並列度と知識粒度の相関についての評価を実施しております。

後期におきましては、これからご説明いたしますこれまでの実験で得られた知見とこれまでに開発してきた知識ベース管理プログラムからの知見から、本格的にオブジェクト指向の考え方を取り入れた並列知識ベース管理実験プログラムをGHCにより開発する計画であります。

8. 並列検索実行管理の実験プログラム、位置付け

並列検索実行管理プログラムは並列知識ベース管理プログラムの下位に位置しまして、知識ベースクラスタ内の8個のプロセスエレメントと共有メモリを物理資源としてプロセスエレメントを並列に制御して、知識ベース検索の並列実行管理、並列資源管理を行います。つまり、先程もご説明いたしましたが、有限のプロセッサに検索プロセスを動的に割り付ける制御をおこないます。

全解探索時の全ての場合の組み合わせをOR並列処理するため生成される検索プロセス数の爆発を抑えるためにベクターファースト検索、つまり、その時点で最もよさそうな物の幾つかを検索実行させる実験プログラムを試作して、並列プロセッサ実験機上でその振舞について評価いたしました。

ここで、多少細かなお話になりますが、検索プロセスからのレスポンスが集合でありますのでGHC、並列論理形核言語で定義された並列推論プロセスと検索プロセスとのインタフェースには、集合処理のために特殊な処置を施しています。

ただいまご説明いたしました、並列知識ベース管理プログラムと並列検索実行管理プログラムの基本機能の評価については、明日の技術講演の“知識ベース並列検索実行管理について”と題しまして報告することになっております。

9. 知識ベース管理プログラムの開発経緯とこれからの計画

この図は知識ベース管理プログラムの開発経緯とこれからの計画について示したものです。

これまでのお話は、この図の一番外側の並列知識ベース管理の実験プログラムについてでありましたが、これは、後期において5Gプロトタイプハードウェア上に本格的に試作開発するものであります。

これまでに単一プロセッサ上に、つまり、PS1上に、知識ベース管理プログラムとして開発した、KAISERでは関係データベース、つまり、ファクトベース知識管理として知識の同化／調節／均衡化等の手法と、ルールからファクト検索コマンドを生成するアルゴリズムを提案しました。

また、このKAISERの発展形態としてKAPPAを、また、複数のPSIからなる分散プロセッサ上で、演繹機能を取り込んだ、分散知識ベース管理プログラムPHIをSIMPOS上に開発してまいりました。

特に、KAPPAは自然言語処理システムにおける大容量辞書検索システム、各種の専門家支援システムにおける大容量知識検索システム、プログラム検証、合成、修正支援システムにおけるプログラム部品の検索および対象プログラムの正当性の証明支援システムなどのアプリケーションを対象して現在稼働し始めています。

また、これはデータモデルを非正規系として、問合わせ言語はこれに整合のとれた、コンプレックスオブジェクトを意識したものとなっておりまして、オブジェクト指向の基本的概念を、既に一部取込んでおります。

このようにして、これまでの試作開発を通して得られた知見と、先程ご説明いたしました並列知識ベース管理プログラムに要求される機能についての、GHCによる実験プログラムでの評価と基礎的研究とから得られた知見を基にしまして、後期では、5Gプロトタイプハードウェア上に本格的オブジェクト指向の並列知識ベース管理プログラムを開発してゆく計画であります。

なお、明日の技術講演で“KAPPAにおける非正規形演繹データベース”と題しまして報告することになっております。

10. 部分計算の概念を問合わせ評価に導入

つぎに、知識ベースにおける幾分基礎的な研究成果の2、3について、ご説明します。

まず、部分計算の概念をクローズデータベースの問合わせ評価に導入いたしました。もともと部分計算とは繰返し計算の効率化のために考え出されたものでありますが、これを論理プログラムに導入して、たとえば、遅延評価処理に応用したり自己適用可能なコンパイラ生成機の研究が進められていたりしています。クローズデータベースを対象とした知識ベースの演繹処理に部分計算の概念を導入して問合わせ処理の効率化が図れることについて明らかにしています。

たとえば、部分計算を繰返し実行して知識ベースのサーチ範囲をトップダウンとボトムアップの双方向から縮退させてゆくことにより、検索を効率化できる双方向検索アルゴリズムを提案しました。このアルゴリズムは先程ご説明いたしました分散知識ベース管理プログラムPHIの基本部に開発しました。

また、知識ベースのサーチ範囲に共通部を持つ問合わせの集合間に部分計算を施すことにより、類似条件問合わせ処理で繰返す無駄な検索をしなくて済むアルゴリズムを提案しました。これと双方向検索アルゴリズムを組み合わせれば一層効率的であると言えます。

11. 否定が扱える層状知識ベース管理について

つぎに、否定を扱える知識ベース管理についての研究成果をご紹介します。

確定ホーン節では本体に負のゴールを持たないために、否定を一般的には扱えないのですが、このために、なんらかの拡張を行い、否定も扱えるようにして知識表現、管理能力を向上させたいわけですが、これに答えるための一つの方法として述語問

に半順序関係を設けて、どの述語の否定をどの述語が使えるかを層状に定義します。この概念は既に存在していましたが、このような知識集合への問い合わせ評価手法については強力なものはありませんでした。

問い合わせの評価手法として既に、佐藤／玉木らのOLDT反駁法が提案されていましたが、これに否定処理のために、知識ベースから証明できない事実に関しては、その否定が成立すると解釈する閉世界仮説演算を追加することにより、否定も扱える強力な評価手法を提案することができました。

つまり、OLDT反駁処理はOLD (orderd linear resolution for definite clauses)反駁処理に表すなわち、テーブルを使うことにより計算結果の再利用と無限ループの抑止などを可能としたものです。

このOLDT反駁処理を使って、下位の層から順に閉世界仮説演算により問い合わせに答えれば良いことを示したわけです。

これによりまして、否定情報が扱える知識表現とその管理能力を向上させることができました。

より詳しく厳密なお話は、明日の技術講演で“演繹データベースにおける問い合わせ処理”と題しまして報告することになっております。

12. 制約付き性質継承における並列アルゴリズムと並列論理型言語

最後に、もう一つ知識ベース管理についてのトピックをご紹介します。

一般に、知識間には関係が存在していて、これらの関係に制約条件を付けて知識を表現すれば、ナチュラリティーをキープしていても便利で、知識の表現能力を向上させることができます。また、これを知識の集合間の関係の表現にも拡張できます。

つまり、知識の表現言語としてis-a 関係、is-not-a 関係等に制約を付加するわけです。知識集合間の関係も同一表現手法が可能でありますので、これはオブジェクト指向の知識ベース管理に親和しているとも考えられます。

これまでに、例外事象を含んだ知識集合の非単調な性質継承における並列アルゴリズムについては既に提案されていましたが、これに制約の概念を導入して、より複雑な知識の管理を高度に、しかも、並列におこなえることを示しました。さらに、制約付き性質継承における新しい並列アルゴリズムは並列論理型言語GHCで表現記述できることを示しました。また、知識の量にたいして線形のオーダーで処理可能な知識ベースのクラスも明らかにしています。

この成果は、より高度な知識表現言語自身の研究の橋懸かりとなることはもとより協調問題解決と並列処理のパラダイムを制約によって結合する研究へと発展が期待されます。

後半ご説明いたしましたこれらの基本的なアイデアについては、これからの知識ベースクラスタへ導入してゆくことにしています。

以上が中期に研究開発してまいりました成果の報告とこれから後期にむけた研究方針についての概要を述べました。

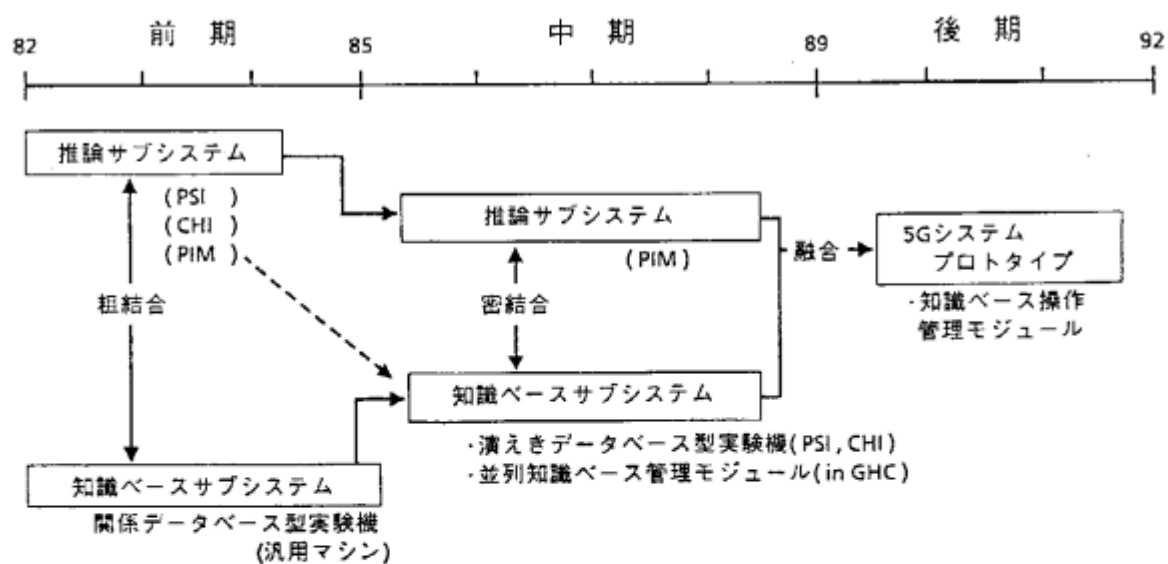
知識ベースサブシステム

ICOT 研究所

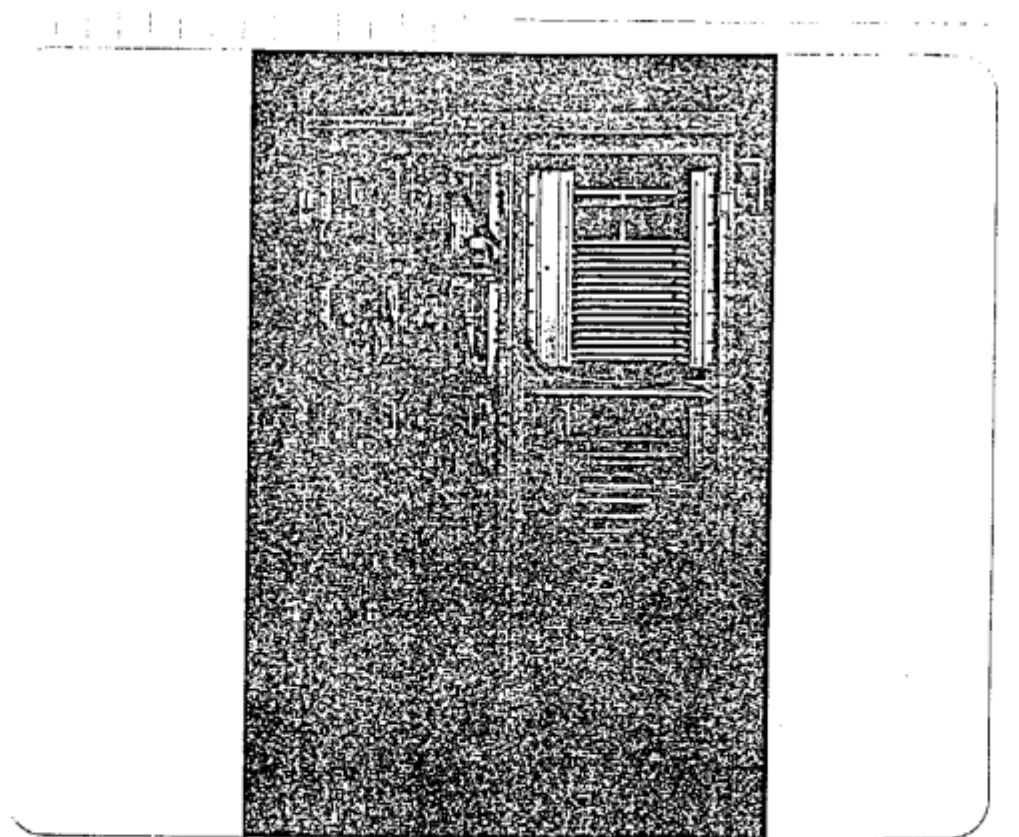
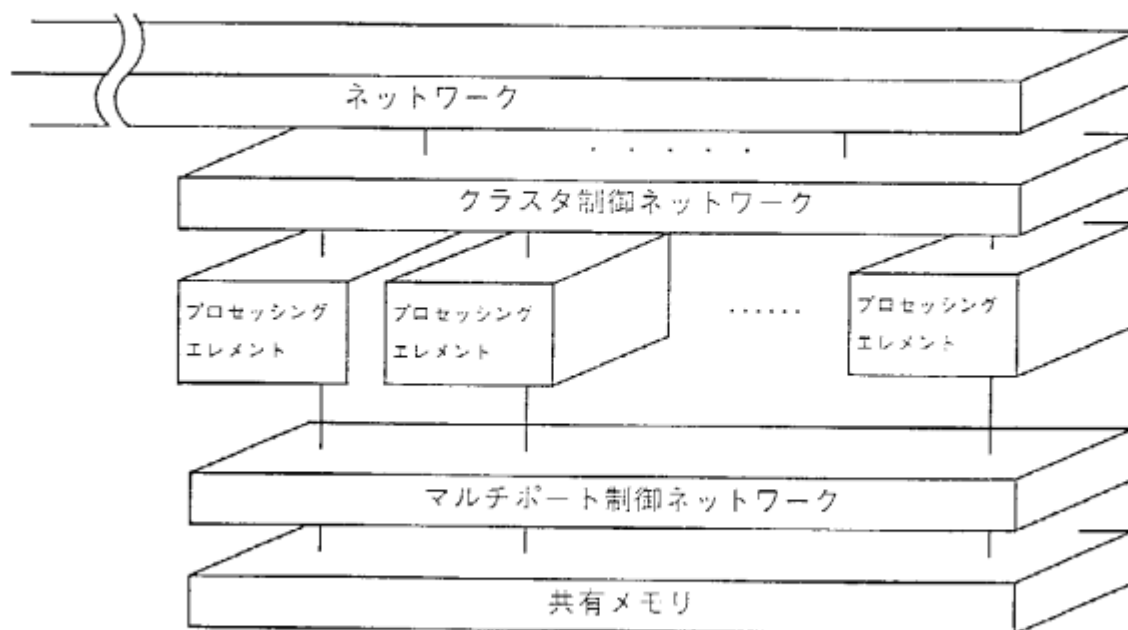
第3研究室

伊 藤 英 則

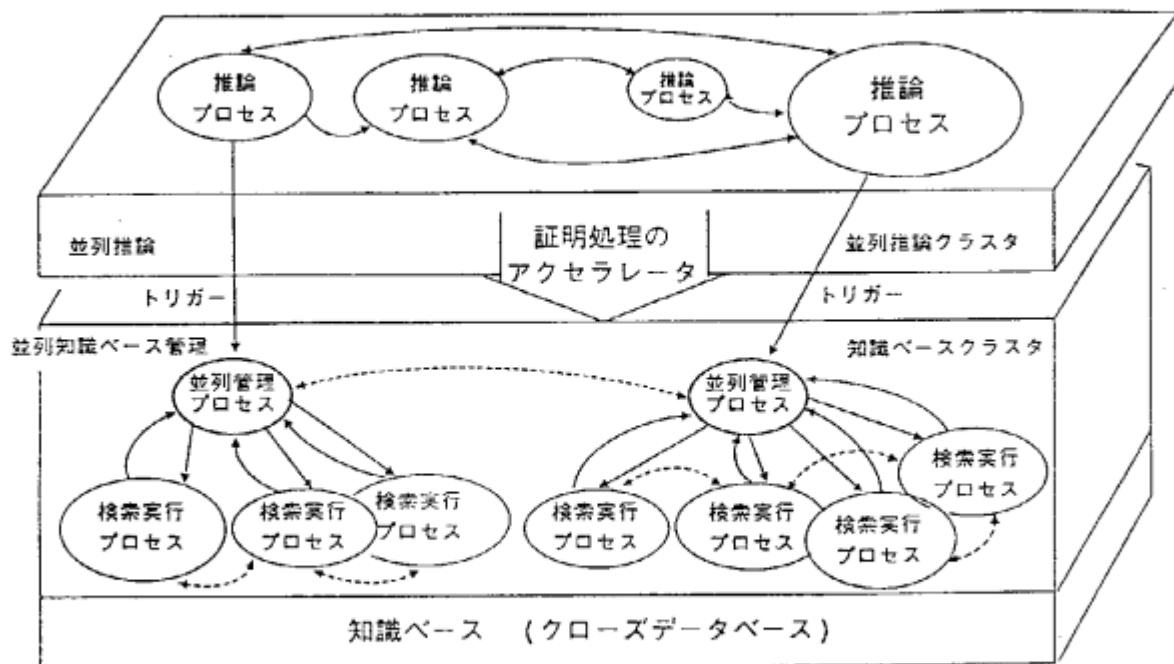
知識ベースサブシステム開発経緯



知識ベースクラスタ アーキテクチャ概念図



知識ベースクラスタ内プロセス構成概念図



知識 / 言語

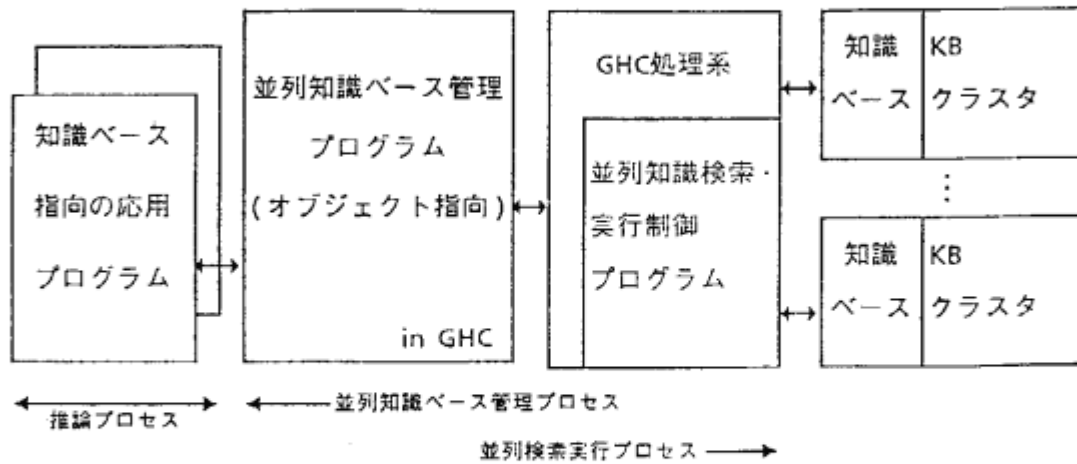
データ / 知識

- ・ 構造 データ / 項 / リスト、集合
- ・ 整理 ソート / クラスタリング / トライ化 / 重ね合せ
- ・ 関係 テーブル / 非正規形 / ネステッド / ネットワーク
- ・ 基本演算 関係演算 / パターンマッチング / 代入 / 単一化

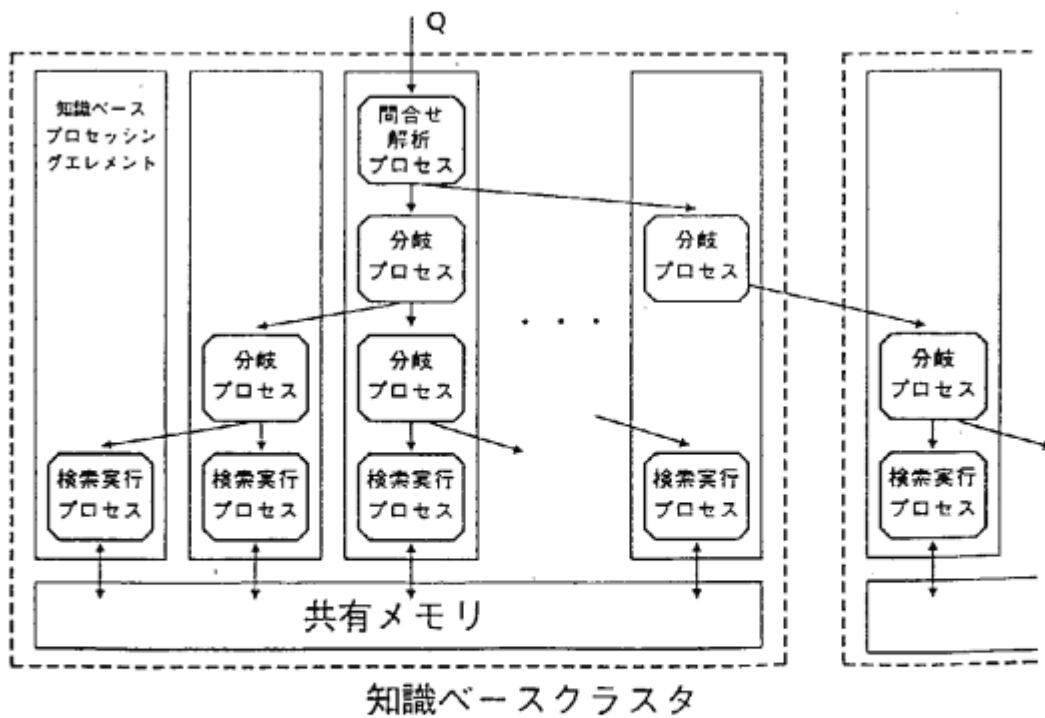
言語

- ・ 知識表現言語
- ・ プログラミング言語
- ・ 知識ベース管理言語
- ・ 問合せ言語

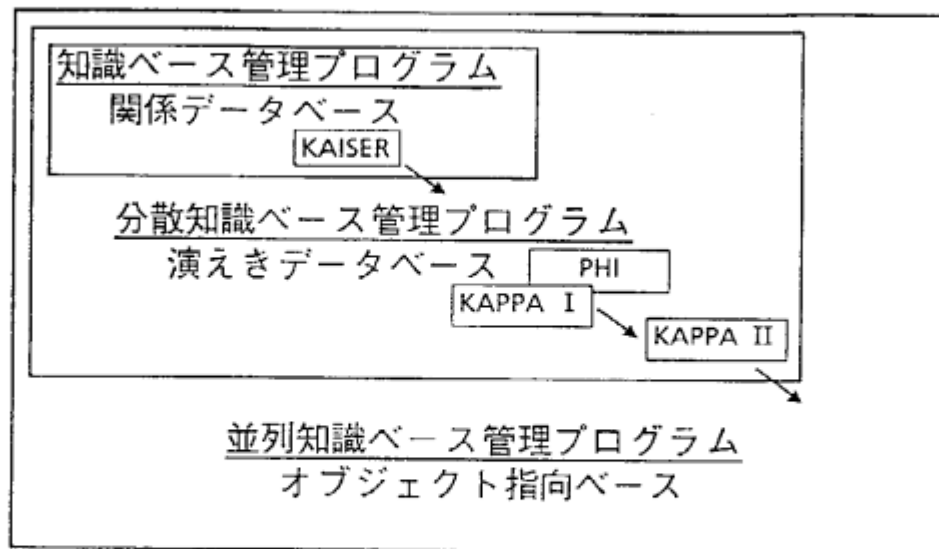
並列知識ベース実験システム構成



並列知識ベース管理プロセスと検索実行プロセス

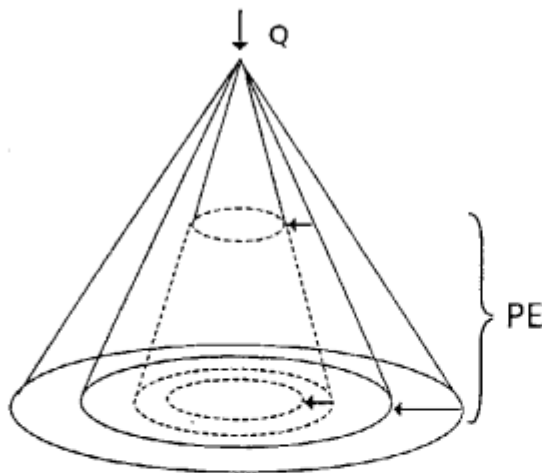


知識ベース管理プログラム

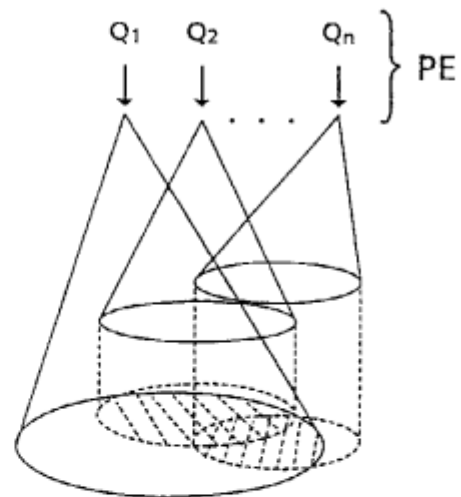


部分計算 (Partial Evaluation)

・ 双方向検索アルゴリズム



・ 問合せ集合処理



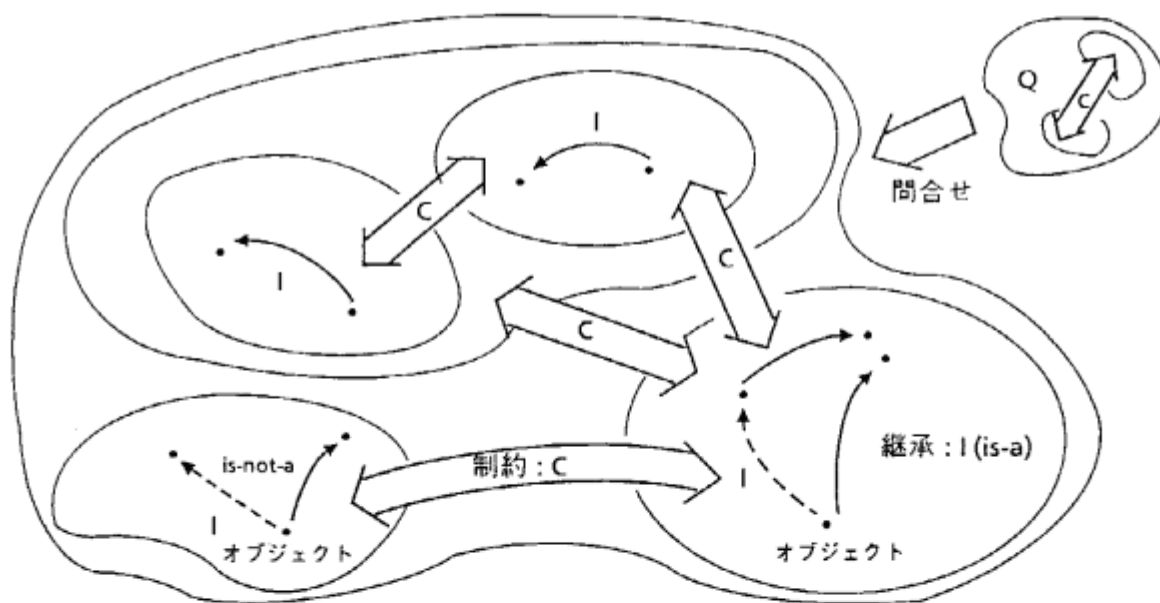
層状知識ベース管理

- ・ 否定情報の扱い
- ・ ルールの層状化(Stratified)
- ・ 問合せ評価

OLD 反駁+テーブル

+ Negation as Failure ルール

並列知識ベース管理 in GHC



制約、並列、協調