

# 知識ベース管理システムKAPPAの構想

—PSIの環境と設計方針—

5M-5

横田一正 (ICOT)      内田俊一 (ICOT)      溝口徹夫 (三菱電機)

## 1. はじめに

Kappa (Knowledge Application Oriented Advanced DBMS/KBMS)は、逐次型推論マシンPSI上の知識処理システムのためのデータベース機能と知識ベース機能を提供することを目的としている。とくに自然言語システムと定理証明支援システムはそれらの機能を緊急に必要としていること、またそれらは将来のPSIの大ユーザに成長することが予想されるので、Kappaの設計にその要求を反映している。本発表では、それらの応用システムが要求する機能と、それを踏まえた設計方針、ならびに現在開発中の試作システムの概要を説明する。

## 2. PSIの環境と要求される機能

自然言語処理のために、大規模な電子化辞書と概念辞書の開発が進められている。自然言語の本格的なシステムはこれらを効率よく利用しなければならない。電子化辞書は複雑にネスト化したテーブル構造として定義されており、数10万語の単語辞書を何種類か抱えている。概念辞書(シソーラス)は概念間が木構造あるいはネットワーク構造をしており、そのリンクを高速にたどる必要がある。さらに機械翻訳では概念構造の変換規則、談話理解ではそのときそのときの(複雑なデータ構造で表現される)状況を操作する必要がある。これらはいずれも変数と構造を含んだもので、従来のデータベース・システムそのままではその処理はひじょうにむずかしい。

一方定理証明支援システムでは、数学の定義、定理、証明などをテキスト、木構造、項などの形で格納し、それらを必要に応じて(とくに項は単一化で)検索しなければならない。証明の検証中には、仮定の追加、更新が頻繁に必要となる。項の扱いをどうするか、膨大な数学知識をいかに効率的に扱うか、知識間の制約をいかに制御するか、など従来のデータベース・システムでは考慮されていなかったデータ構造と操作が必要となる。さらにこのシステムは知的プログラミング・システムの中核的な役割をもっているので、将来は仕様やプログラムや書換え規則を含む多種のデータの統一的な扱いも考えておかねばならない。

そしてシステム環境は第5世代コンピュータ研究のためのツールとして開発されたPSIであり、パーソナル・ワークステーションとしてのパーソナル型の使用が中心になる

と考えられるので、中央集中型とは異なった使用環境を想定しなければならない。またPSIネットワークに接続された分散機能も当然必要になる。実装言語も利用者の使用言語も論理型言語ESPであり、そのインタフェースも重要である。また現在予定されているPSI2は、ワークステーションながら主記憶を320MBもっている。つまり従来の二次記憶データベースの他に、大規模主記憶を利用した主記憶データベースと、ESP中のプログラム・データベースの共通のインタフェース、統合化も考慮しておかねばならない。

以上のKappaの環境を図示すると図1.のようになる。

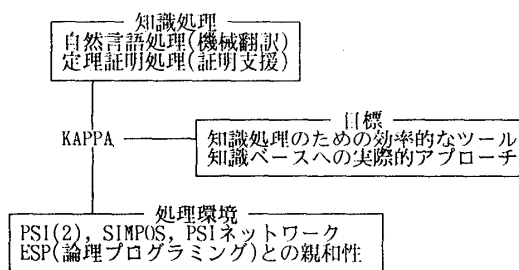


図1. Kappaの位置づけ

## 3. 設計方針

このような知識(データ)の量、処理効率、構造の保持、稼働環境、などを考慮して、次の設計方針を設定した:

- (1) 特殊なハードウェアの使用は前提としない。
- (2) 処理効率を重視し、ツールとしての役割も果たす。
- (3) オブジェクト指向をソフトウェア工学的な面だけではなくシステムの基礎概念としても導入する。
- (4) ワークステーション環境に適合したシステムとする。
- (5) ネスト化されたデータや構造型データを基本的なテーブル型として導入する。
- (6) 項(ターム)を基本データ型の1つとし、単一化による検索を可能とする。
- (7) 現在想定できる応用システムの中で、汎用的な推論機構を装備する。
- (8) その他の知識処理のための一般的な知識表現言語と推論機構も備える。

そしてそのためには、単に関係モデルに基づいたシステムではなく、その拡張と考えられる非正規型(NF2)モデルに基づいた、項の扱いを含んだシステムを下層にするのがふさわしいと判断した。非正規型モデルはこれまでその表現

力の豊かさから概念モデルとしてよく使われてきているが、内部モデルとしてもディスク資源、I/O 回数、結合演算の数などから効率的に優れていると考えられる。これまでもIBM のAIM プロジェクト、Darmstadt 工科大学のDASDBSシステム、あるいはINRIAのVersoマシン、などさまざまなシステムが提案されている。このモデルに論理型言語での実装には困難が予想されたが、ESP で一部を試作した結果は上々であり、さらにファームウェアのサポートも約束されており、実用的なシステムとしても問題ないと判断した。

もっとも検討を要したのが項の取扱いである。一般的に項の取扱いは効率上ひじょうに重い処理となることが予想されるが、項の非正規型テーブルの展開、二次記憶上の処理と主記憶上の処理の分離などを検討した結果、充分使用に耐えられると判断した。

4. 開発方針

Kappa は1985年9月に立ち上がったプロジェクトであり、ICOT計画の中期末1989年3月までの約3年半で上記の機能の実現を目指している。そこでKappa システムを試作システムと本格システムの2つに分けて順次開発をおこなうことにした。

試作システムは項の扱いと構造型データのサポートを含んだ高機能データベース・システムの実現に重点をおいて、本格システムではさまざまな推論を実行する知識ベース・システムの実現に重点をおいている。試作システムは開発期間が1年程度とひじょうに短いことから、下層に重点をおきつつ、機能のある程度絞りこんでいる。そして試作システムの開発と並行して、上位層の検討を現在おこなっており、来年度からの本格システムに実装する予定である。

5. 試作システムのモデル階層

Kappa の試作システムを論理的にみると、いくつかのモデルから構成されている。それを参照モデルに準じて考えると図2.のような階層にまとめることができる。

ユーザ・リクエスト	外部モデル (ビュー)	定理証明 シンタックス 日本語辞書	ユーザ情報 外部モデル 定義情報	意味ネット フレーム ルール	外部コマンド
ツール	通信モデル (モデル間の交換)	モデル展開	ネットワーク情報	分散データベース	コマンド変換 コマンド生成
データベース サービス (2)	基本モデル (NF2,関係,ファイル,構造,...)	木構造 ネットワーク ファイルライブラリ	基本モデル 定義情報 モデル 交換情報	Is_a木 IIs_a木 ファイルライブラリ	基本コマンド
データベース サービス (1)	原始モデル (NF2)	NF2レコード B+トリー ハッシュテーブル	NF2定義 制御情報 物理空間 制御情報	テーブル タブル一時 集合, RID	原始コマンド
基本 サービス (OS)	Kappa ファイル	ページ データ	ファイル, ユニーム 定義情報	ページ	Kappa ファイル コマンド
	データモデル	データ本体	メタデータ	基本要素	操作言語

図2. Kappa のモデル階層

ファイル層はデータベースとしての効率上の理由から新しく作り直すことにした。原始モデルでは非正規型に忠実なデータ構造とそれに対するプリミティブなデータ操作を

サポートし、基本的なデータ型としては項をもっている。基本モデルでは概念辞書のような構造型データと静的な結合関係をサポートする。通信モデルではローカル/リモートPSI(Kappa)の通信機能を持ち、ロケーションを意識しないアクセスが可能になっている。外部モデルでは、意味ネットワークのようなデータ構造とフレームのようなデータ操作を可能にしている。項とルールの変換もこの層でおこなう。

そして通信モデルを除き、各階層間のインターフェースはコマンドによって取られている。下の階層からファイル・コマンド、原始コマンド、基本コマンド、そして外部コマンドとなっている。外部コマンドはフレームに対応したユーザ・インターフェースをしており、構造型データはそれですべての操作をおこなう。上に述べたように、ユーザはすべてのコマンドを述語呼出しで実行することもできる。その他にKappa の内部状態を監視したり変更するための操作コマンド、ユーティリティのための内部コマンドがある。

通信モデル以下では項の扱いをもった高機能データベース・システムの役割を果たし、外部モデルで推論を含む知識ベースの機能をもたせることにしている。本格システムでは外部モデルの機能を大幅に強化する予定である。

6. SIGMAとKSI

Kappa と並行して、SIGMA (Set Inference Grounded Knowledge Base Management System) と KSI (Knowledge Representation Language for Set Inference) の開発を進めている [1,2,3]。これはKappa が応用システムとして設計に反映している自然言語システムと定理証明支援システム以外の、たとえば専門家システムなどの知識処理システムとのインターフェースとなるものである。

KSI は集合と要素という概念を基にした知識表現言語であり、数学的セマンティクスとフレーム的な表現が特徴である。SIGMA はKSI で表現した知識に対する知識エディタ、推論エンジン、推論コメントータなどからなるシステムで、Kappa とのインターフェースをもっている。SIGMA とKSI を含めたシステムを広義のKappa といい、そうでないものを狭義のKappa と位置づけている。これを図示すると図3.のようになる。

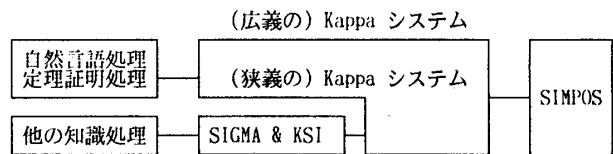


図3. "広義のKappa"と"狭義のKappa"

参考文献

[1] 三石,吉良,溝口, "知識ベース管理システムSIGMA の構想", 情報処理学会第32回全国大会, 1M-1, 1986.  
 [2] 吉良,三石,溝口, "集合の概念に基づく知識表現言語KSI ", 情報処理学会第32回全国大会, 4M-7, 1986.  
 [3] 吉良,三石,溝口, "集合の概念に基づく知識表現と推論", 情報処理学会第33回全国大会, 6M-1, 1986.