

3M-3

知識ベース管理システムKappa
—データベースから知識ベースへ—

溝口敬夫* 横田一正** 内田俊一**

*三菱電機(株) ** (財) 新世代コンピュータ技術開発機構

1. はじめに

知識ベースシステムあるいは知識ベース管理システムはデータベースからの類推で論じられることはあっても、内容として正確に表現されているとは言えない。

システムを作り上げるという観点で、「システムは何かの形で既に存在するものである」との認識に立てば、既に我々の持つシステムとしてのデータベースシステムや情報システムがなぜ存在するかということと共通の土俵にいるのだと言える。

Kappa の開発の動機はPSI マシンでの今後展開が予想される自然言語処理や定理証明処理のための環境作りである。最も分かりやすいものは自然言語辞書格納・検索である。自然言語辞書利用がいかなるものかの点からのシステム実現がなされている。

他方、何かが形になってくる背景には、枠組とそれを実現する要素が存在するであろうという観点がある。

知識ベースシステムの枠組みとは何かについて説得性のあるものを提案するのは困難である。最近情報システムについて分析・設計方法論が華やかである。知識ベースシステムの性格が何であるかを差別化できれば方法論、枠組みを設定できるかもしれない。しかし要素について言えば、枠組みのはっきりしない現在、システムの実現するのに必要な要素を用意することになる。

Kappa の開発においての要素に対する前提は、辞書における同一項目の繰り返し出現と可変長テキスト文や定理証明のための推論規則の格納・検索である。

これらを基に知識ベースシステムへの展開を図る。

2. Kappa システム構築の基盤

データベース管理を実現するについて、機能的、性能的に使用に耐えるもの、例えば従来のデータベース管理機能相当が提供されなければ本プロジェクトの意味は半減する。Kappa を開発するについての特殊性は、対象とする機種がPSI というワークステーションであり、言語は論理型言語ESP であるという2点である。この2点を

無視して有効なものは存在し得ない。データベースをユーザ言語で実現することの利点はデータベース機能と応用プログラムの境界を必ずしも限定せず、実現された機能の任意の階層に入り込めることであり、データベース機能と応用プログラムの違和感をなくせることである。身近な例では手続き言語の中に集合を基本とするデータベース言語が埋め込まれるというものがこれとは逆の形で存在する。Kappa はワークステーションの上のデータベース機能を提供するが、単一ユーザ環境のみをサポートしている訳ではない。Kappa は分散環境を提供することを前提としているので、その考慮も払われているからである。つまりサーバとしてのKappa の位置付けも含まれている。分散データベースというよりもリモートデータベースアクセスという性格が強い。

論理型言語による相当規模のソフトウェア開発において設計、デバッグ、テストをどのように計画し、実行すればよいのであろうか等は大に関心のあるところである。設計について言えば、手続き言語に比べればより高度な言語を用いたプログラム設計ではプログラム詳細仕様書の内容はプログラムそのものに非常に近いものになるし、それだけにプログラミング言語の持つ特性を熟知して詳細設計を実施することが肝要ということになる。

論理型言語でのデバッグはどのように行われているであろうかについてはまだ相異点を十分に理解は出来ない。テストについては従来の考え方と言えるプログラムパスをつぶすというよりも機能仕様・詳細仕様の検査ということになる。

これらのことは機能性等での進歩というよりは新たな環境への適応に関するものであるが、新たな問題への解決策を必要としている。

開発を推進する上での仕掛はほぼ従来のものを採用した。設計・コーディング・単体テスト・結合テスト・登録・版管理等がそれである。ユーティリティ系をも含めて80K 行にも及ぶ規模である。試作を通じてKappa が置かれている環境でのソフトウェア開発の知見を掴つつあ

Knowledge Base Management System Kappa

From Database Toward Knowledge Base

Tetsuo MIZOGUCHI* Kazumasa YOKOTA** Shunichi UCHIDA**

*Mitsubishi Electric Corp. **ICOT

る。例として検査に関する報告を参照願いたい[4]。

プログラムインタフェースとしてのデータベース機能の実現の上にエンドユーザインタフェースの提供が必要であることには変わりはない。対話機能についても以下で報告したい[7]。

更にKappa が置かれている環境の特徴を出し、後で述べる知識ベースへの基盤作りとするために、主記憶データベース機能の導入は重要との認識である[6]。

3. データベースからの出発

データベースの持つ表現力を増すためにデータの持つ意味を表現できるデータモデルの充実というアプローチがある。意味データ表現、メタデータ管理、データ辞書はその出発点と言えるが、データモデルの提案・メタデータの活用でのメリットの模索がなされてきた。

もう1つのアプローチは格納されるデータの制約を緩和し、格納されたデータの持つ内容につれて、その利用面を拡大することと、利用者とシステム間の境界をよりシステムの充実に向けることを目指すものである。明らかに我々は後者のアプローチを採っている。しかしながらデータベースとしてのテーブル（ファイル）やレコードという概念からは出ていない。データモデルとして非正規形を採用することとした。動機は当面の対象である自然語辞書の扱いにおける下位レベルでの非正規データの格納・検索を基本的な動作とすることの必要があったためである。正規データを部分として含み、正規データの組合せで非正規データテーブルを実現可能であり、その選択をKappa では上位レベルで許しているが、非正規データを下位レベルで直接サポートすることで設計の容易さと性能の向上を目指すこととした[1]。

非正規データの定義はテーブル全体を支配するものと実現値（レコード）個別のものがある。データ表現の多様さを吸収する一歩である。実現値個別データ定義は実現値に近付けて置くことで処理の局所性を保つ試みがなされている。技術的には非正規データ項目の索引の扱い、非正規データ項目の操作に関する意味付けの明確化等の課題がある。

内部モデルでの非正規データ管理機能の上に、複数テーブルの結合やシソーラスに代表される概念分類データの扱いのための構造テーブル等の組立を可能にし、その基本的動作を提供できる[2]。

格納されるデータ型として項（ターム）を扱うことによって内容に関する検索・ユニフィケーションを行う機能を提供している[3]。

以上述べてきたように、Kappa では自然語辞書利用を手本にデータベース管理を拡大し、知識ベース管理への模索を行ってきた。

4. 知識ベースへ

データ処理でのデータの形式の大きな特徴は1つの定義情報に対する実現値が多数あることを前提としていることである。この前提は知識ベースを対象とするときは成り立たない。既に述べた非正規データによる表現の多様さは1つの試みであるがまだ充分な答になっていない。

知識を非定型、場当りのものとし、事実の1つの解釈と見たり、教科書や書籍の形に見られるように安定した内容を持つものと見たり、あるいは安定した形への到達過程が重要（または永久に安定には達しないものもある）と見たり等いくつかの解釈がありそうである。知識ベースと言うときにはその背景に安定したものの存在を仮定しているとするのが自然であり、その表現に用いられる表現として、論理の枠組みからの関係データとPrologを非正規データとCRLに拡張した演繹データベース[5]と自然言語処理等に用いられるフレーム型の言語の検討[6]を行っている。

5. おわりに

Kappa の開発ではデータベース機能およびその拡張機能を現実的に作ることで応用システム開発を有利に進められる一方、より上位での解放型設計による知識ベースシステム構築を可能にしようといくつかの試みをしている。この後者の試みはいくつかのシステム実現を経て洗練化されるものである。

参考文献

- [1] 小澤、金枝上、根本、川村、杉崎「Kappa:試作システムの内部モデル」情処35全大 3M-4
- [2] 根本、金枝上、宮地、椿野、合田「Kappa:試作システムの基本モデル」情処35全大 3M-5
- [3] 宮地、横田、椿野、丹羽「Kappa:非正規形モデルと項の処理」情処35全大 3M-6
- [4] 小林、河村、溝口「Kappa:論理型言語でのテスト」情処35全大 3M-7
- [5] 横田「Kappa:非正規形モデルと演繹データベース」情処35全大 3M-8
- [6] 金枝上、横田、河村、横塚「Kappa:大規模主記憶を利用した全体システム」情処35全大 3M-9
- [7] 河村、丸山、梶山「Kappa:利用者インタフェース」情処35全大 3M-10