

知識ベースマシン *Mu-X* (1) - 項集合に対する検索言語 -

物井 秀俊、森田 幸伯、伊藤 英則(新世代コンピュータ技術開発機構)
柴山 茂樹、酒井 浩(東芝)、沖 健治(東芝ソフトウェアエンジニアリング)

1. はじめに

ICOTでは、知識処理システムの実現を目指し、論理型言語に基づいた推論マシンと知識ベースマシンを研究開発中である。特に、知識ベースマシンの一つのモデルとして、大量の項集合に対する高速な検索処理を実現するため、検索等を実行する複数のプロセッシングエレメントと大容量の共有メモリとしてマルチポートページメモリ(MPPM: Multiport Page-Memory)^[3]を用いる、知識ベースマシン *Mu-X* を開発中である^[2]。

知識ベースマシン *Mu-X* では、Prolog等の論理型言語が操作対象とする項(term)を(単純な文字列としてではなく)構造データのまま、格納及び検索が可能な関係型知識ベースモデル(RKB: relational knowledge base)を提案している^[1]。RKBでは、データベースにおける関係モデルの拡張であり、関係の定義域を単純な値の集合から項の集合に拡張した項関係(term relation)と、項関係に対する操作言語として関係代数の等号条件を単一化操作に拡張したRBU演算を導入している。

本稿では、論理型言語から項集合をアクセスする時の、問い合わせ言語(query language)について述べる。この問い合わせ言語では、論理型言語から知識ベース内の項集合を無理なくアクセスできるように、単一化による項の操作が比較的柔軟に記述できるようにしている。

2. 関係型知識ベースと演えき処理

大量の知識を対象とする演繹システムとして、関係データベースシステム(RDB: relational database)の検索機能を演繹処理の中に取り込んだ、演繹データベースシステム(DDBS: deductive database system)が盛んに研究されている^[4]。

DDBSでは、RDB内に定義された関係を、関係名を述語名とするような事実(fact)の表明とみなし、この事実を用いた演繹規則(deductive law)をデータベースの一部として定義する。そして、元々の関係だけでなく、この演繹規則により定義された仮想的な関係(演繹規則と関係として定義された事実から推論できる事実の集合)も問い合わせることができるようにする。DDBSは、RDBを対象とするため、関数子を含むような事実集合を格納したり、ルールの適用を知識ベース内の集合操作として実現することはできない。

関係型知識ベースは、項集合を属性の定義域とす

る項関係の集合として定義される。項関係は、以下のように定義される。即ち、一つの項関係をTで示し項の集合を K_i で表わすと、項関係Tは

$$T \subset K_1 \times K_2 \times \dots \times K_n$$

と定義される。ここで、各 K_i の要素となる項は以下のように再帰的に定義される。

(1)定数と変数は項である。(2) f は n 変数の関数とし $t_1, \dots, t_n$ を項とすると、 $f(t_1, \dots, t_n)$ も項である。(3)以上の規則を用いて生成されるものだけが項である。

項関係により、RDBでは扱えないルールや関数子を含む知識表現データを知識ベースの中で直接扱うことが可能となる。例えば、よく使われる先祖データベースをDDBSで扱う場合、fatherとmotherに関する表明をRDBの関係として定義し、

```
ancestor(X,Y) :- parent(X,Y).
ancestor(X,Y) :- parent(X,Z), ancestor(Z,Y).
parent(X,Y) :- father(X,Y).
parent(X,Y) :- mother(X,Y).
```

という演繹規則を追加することにより、先祖に関する検索が可能となる。しかし、項関係では、項を直接格納し単一化により項を操作できるから、

```
gene(ancestor(X,Y), [parent(X,Y)]).
gene(ancestor(X,Y), [parent(X,Z), ancestor(Z,Y)]).
gene(parent(X,Y), [father(X,Y)]).
gene(parent(X,Y), [mother(X,Y)]).
gene(father(a,b), []).

-----
gene(mother(h,u), []).
```

という項関係geneを定義すれば、推論システムを介せず、RKB内の演算のみでancestor(X,Y)を求めることができる。また、演繹規則そのものも、同じ枠組みにより検索することができる。

以上のように、項関係を対象とすることで、知識ベースシステムの中だけで、ルールを適用したり構造体を操作したりすることができる。システム的には、DDBSのように演繹規則を実行する推論システムは必要なく、大量の事実に対してルールを適用したりルールの量が多くなった場合に、専用システムを用いて効率的な処理が可能となる。

Knowledge Base Machine *Mu-X* (1) - Query Language for Term Relations -
Hidetoshi MONOI, Yukihiko MORITA, Hidenori ITOH (ICOT Research Center),
Shigeki SHIBAYAMA, Hiroshi SAKAI (Toshiba R & D Center), Kenji OKI (TSE)

3. 項集合に対する問い合わせ言語

3.1 関係論理に基づく枠組み

本稿で述べる問い合わせ言語は、関係スキーマに格納された項集合に対する操作言語となる。項関係に対する操作言語としては、検索専用プロセッサのアーキテクチャ検討を目的として、関係代数を拡張したRBU演算を既に提案している。しかし、以下に述べる理由により関係代数より関係論理に基づく言語の方が望ましい。

さらに、言語の記述能力として以下のものが要求される。まず、(1)前節で示したような項集合を対象とする演算手順を記述するため、単一化による項の操作と項の定義が自由に書けること。また、論理型言語に組込むため、(2)構文的に論理型言語との差がないこと、(3)問い合わせ言語内の基本演算が論理型言語と同じであること。

知識ベースマシン *Mu-X* では、項関係に対する問い合わせ言語として定義域関係論理^[5]に基づくものを考える。まず、定義域関係論理式

$$\{x_1 x_2 \dots x_k \mid \psi(x_1, x_2, \dots, x_k)\}$$

において、論理式 ψ を構成する素論理式を、項集合を対象として以下のように拡張する。

1. $R(x_1, x_2, \dots, x_k)$:

ここで R は項数 k の関係で、 x_i は項(term)である。意味は、 R 内の任意の組を $\langle y_1, y_2, \dots, y_k \rangle$ とすると、 $R(y_1, y_2, \dots, y_k)$ が $R(x_1, x_2, \dots, x_k)$ と単一化できた R の組 $\langle y_1, y_2, \dots, y_k \rangle$ のみを抜き出すことを示す。

2. $x\theta y$:

x と y は項であり、 θ は次節で定義する項間の比較演算子である。意味は、 x と y が $x\theta y$ を満足しなければならないことを示す。

以上の定義に基づき、 X, Y を変数とするとき $r(X, Y)$ と問い合わせ言語の中で書いた場合は、関係名が r である2属性の項関係の全タプルを表わすことになる。さらに、 a を定数 f を関数子とするとき $r(f(a, X), Y)$ と書くことにより、項関係 r のタプルの中で、第1属性が $f(a, X)$ と単一化できたタプルのみを抜き出すことができる。

3.2 項間の比較演算

RKBでは、属性値として項を許し基本演算として項間の単一化操作を導入しているため、RDBのように要素もしくは定義域変数間に $\langle x=y \rangle$ 等の算術比較のみを定義しても、問い合わせ言語の記述能力としては充分とは言えない。

項間の比較演算として、新たに以下のような比較演算を問い合わせ言語の中で指定できるようにする。以下では、 x と y は共に任意の項を示すものとする。

(1) 単一化に基づく比較

- $x=y$: x と y を単一化する。
- $x\leq y$: x と y が単一化できることを調べる(単一化は行わない)。

(2) ジェネラリティに基づく比較

- $x<y$: x が y よりジェネラリティが低い
- $x<=>y$: x と y のジェネラリティが同等

但し、 x が y よりジェネラリティが高いとは、 x と y の間に代入 θ が存在し $x\theta y$ であるとき。

(3) 単純な等価比較

- $x=y$: x と y が文字列として等しい。

3.3 否定の扱い

項関係に対する否定は、差演算により定義する。このような定義を行う理由は、否定を考える項関係に対してスーパーセットを考えることにより、補集合の定義が明確になるためである。従って、項関係単独での否定、即ち項関係 $R(x_1, x_2, \dots, x_k)$ に対して $\neg R(x_1, x_2, \dots, x_k)$ は、定義されない。

項関係 X と Y 間の差演算は、項間の比較演算子を陽に指定して以下のように定義される。すなわち、関係 X と関係 Y の差は、関係 X のタプルで関係 Y の任意のタプルと、比較演算子 θ によって一致しないものを結果とする。ここで、関係 X と Y の属性数は等しくなければならない。また、 θ は $=, <, <=>$ のいずれかが指定できる。

4. おわりに

以上項集合に対する問い合わせ言語について述べた。本項で述べた検索言語により、これまでに提案しているRBU演算の拡張が必要となる。例えば、3.2節に述べた項間の比較演算の追加は、RBU演算の拡張を必要とする。我々は、これらの機能をマイクロプロセッサ上に実現し、単一化エンジンの拡張を検討する。

DDBSでは、データベースシステムにおける集合演算と推論システム側のタプル単位の演算という処理上のギャップが問題となり、このギャップの解消を目指した研究がある。また、推論システム側からの要求に対して、データベースシステム側の表現能力を上げる研究も、いくつか行われている。我々は、項集合を定義域とする項関係の利点を生かし、推論システムと知識ベースシステム間に、より柔軟でかつ密なインタフェースを構築することを目指し、問い合わせ言語の検討を行った。

[参考文献]

- [1] Yokota, H., et al, "A Model and Architecture for a Relational Knowledge Base", in *Proc. 13th Int. Symp. on Computer Architecture*, 1986, pp. 2-9
- [2] 伊藤, 他, 「大規模知識ベースマシンの開発」, 第33回情報全国大会予稿, 1986
- [3] Tanaka, Y., "A Multiport Page-Memory Architecture and Multiport Disk-Cache System", *New Generation Computing*, OHMUSHA, Feb. 1984
- [4] Gallaire, H., et al, "Logic and Database: A Deductive Approach", *ACM Computing Survey*, Vol. 16, No. 2, 1984, pp. 153-185
- [5] Ullman, D. J., "Principles of Database Systems", *Computer Science Press*, Maryland, USA, 1982