

知識ベースにおける部分評価

4M-4

坂間 千秋 森田 幸伯 伊藤 英則

(財) 新世代コンピュータ技術開発機構

1. はじめに

知識ベースが大規模化されるにつれて、知識ベースシステムにおける問い合わせ処理の効率化が重要になってくる。知識ベースの構造化、種々の探索戦略などはいずれも問い合わせ処理における検索の効率化を狙ったものである。

プログラムにおける計算の効率化手法の一つに部分評価がある[三村 83]。これはプログラム中のある入力データを固定することによって繰り返し計算における効率向上を狙ったものであるが、以後ロジックプログラミング[竹内 85]、演繹データベース[Miyazaki 86]などの分野に応用されてきた。

本稿では、Prolog のようなホーン節で表現された知識ベースを想定する。そして、そのような知識ベースの部分評価の概念を導入し、問い合わせ処理の効率化手法として適用した場合の方法とその意義について述べる。

2. 知識ベースの部分評価

2.1 概念

いま、知識ベース kb をホーン節からなる知識の集合とする。このとき、次のような項の集合 k を考える。

$$k = \{ k_i \mid k_i \in \text{Term}, \exists x_j \in \text{Term}', k_i \sim x_j \}$$

(但し、 Term は全ての項の集合、 Term' は kb における項からなる集合を表し、 $x \sim y$ は x と y がユニファイ可能であることを表す。)

即ち、 k_i は kb におけるある項 x_j とユニファイ可能な項である。

このような k による kb の部分評価とは、 kb における k からの演繹処理による特殊化である。

k による kb の部分評価の概念を下図に示す。



ここで、 kb_k は k によって特殊化された知識ベースである。

いま、 kb における k からの演繹は、関数 Ded

$$\text{Ded} : \text{Horn-set} \times \text{Term} \rightarrow \text{Horn-set}$$

とすると、

$$kb_k = \text{Ded}(kb, k) \quad (1)$$

と表される。ここで、 kb_k は kb において k からの演繹処理を行った結果の知識ベースである。

2.2 方法

次に、このような部分評価の概念を知識ベースにおける問い合わせ処理に適用した場合を考える。

いま、ある問い合わせの集合 $q (= \{q_i\}, (i=1, 2, \dots, n), q_i \in \text{Term})$ とすると、 q に対してある $k \in \text{Term}$ が存在し、

$$k = \sqcup q \quad (\sqcup : \text{上限}) \quad (2)$$

である。即ち、 k は q の Least-general なものである[吉田 85]。ここで、このような k を q のメタキュアリと呼ぶことにする。

さて、このような k によって kb を特殊化する場合を考える。

まず、(2)式を(1)式に代入して、

$$kb_k = \text{Ded}(kb, \sqcup q) \quad (3)$$

を得る。ここで、 kb_k は q を処理する環境に特殊化された kb である。

いま kb, kb_k における q の処理をそれぞれ $\text{Ded}(kb, q), \text{Ded}(kb_k, q)$ と書くと、(3)式から

$$\begin{aligned} \text{Ded}(kb_k, q) \\ = \text{Ded}(\text{Ded}(kb, \sqcup q), q) \end{aligned} \quad (4)$$

を得る。

ここで、 kb_k をある問題解決向きに再定義された論理的ビューと考えると、(4)式は kb と q が与えられたときの、メタキュアリによるビュー定義とそこの問い合わせ処理を表している。

このような、知識ベースのビューの概念はデータベースにおける概念に対応するものであり、問題に応じたそれぞれの処理環境への特殊化による、知識の効率的利用を可能にするという立場から意義があると考えられる。

いま

$$\begin{aligned} \text{Ded}(kb, \sqcup q) \subseteq \text{Ded}(kb, q) \\ \therefore \text{Ded}(\text{Ded}(kb, \sqcup q), q) \\ \subseteq \text{Ded}(kb, q) \end{aligned} \quad (5)$$

次に、 $\forall q_i, q_j \in q$ について

$q_i \sqsubseteq q_j$ のとき

$$\text{Ded}(kb, q_i) \subseteq \text{Ded}(kb, q_j)$$

(但し、 $a \sqsubseteq b$ は $a, b \in \text{Term}, \exists \theta \in \text{Sub}(\text{Sub}$ は置換の集合), $a = b\theta$ を意味する。)である。

Partial Evaluation of Knowledgebase

Chiaki SAKAMA Yukihiko MORITA Hidenori ITO

ICOT Research Center

いま, $\forall i \quad q_i \subseteq \cup q$ だから
 $\forall i \quad \text{Ded}(k b, q_i)$
 $\subseteq \text{Ded}(k b, \cup q)$

ここで,
 $\text{Ded}(k b, q) = \cup_i \text{Ded}(k b, q_i)$
 $\therefore \text{Ded}(k b, q) \subseteq \text{Ded}(k b, \cup q)$
 $\therefore \text{Ded}(\text{Ded}(k b, q), q)$
 $\subseteq \text{Ded}(\text{Ded}(k b, \cup q), q)$

また,
 $\text{Ded}(\text{Ded}(k b, q), q)$
 $= \text{Ded}(k b, q)$
 $\therefore \text{Ded}(k b, q)$
 $\subseteq \text{Ded}(\text{Ded}(k b, \cup q), q) \quad (6)$

(5),(6)より
 $\text{Ded}(\text{Ded}(k b, \cup q), q)$
 $= \text{Ded}(k b, q) \quad (7)$

(7)式は, $k b$ における q の解集合と, $k b_k$ における q の解集合が一致することを表している。

なお, 問い合わせ処理におけるこのような部分評価が有効である場合は k の要素の数に対して, q の要素の数が十分多いとき, 言い替えると, ある共通なパターンをもった問い合わせが繰返し行われる場合である。

(これは, 部分評価の本来的思想でもある。)

3. 具体例

いま, ホーン節集合 $k b$ に対して問い合わせの集合 q が以下のように与えられたとする。

$k b$
 $p(X, Y, Z) :- q(X, Y), r(Y, Z).$
 $q(X, Y) :- w(X, Z), q(Z, Y).$
 $q(X, Y) :- w(X, Y).$
 $r(X, Y) :- s(X, f(Y)), t(Y).$
 $u(X, Y) :- m(X, Y), t(Y).$
 $m(X, Y) :- n(c, X, Y).$
 $l(X) :- t(X), v(X).$

$s(b, f(g(c))).$	$t(a).$	$v(a).$
$s(d, f(a)).$	$t(b).$	$v(b).$
$s(f(a), e).$	$t(f(b)).$	$v(f(c)).$
$s(c, f(f(b))).$	$t(f(g(b))).$	
$s(g(a), f(f(g(b)))).$	$t(g(c)).$	
$s(f(d), f(c)).$	$n(c, a, b).$	
$w(f(a), b).$	$n(c, a, f(g(b))).$	
$w(f(a), c).$	$n(c, a, g(c)).$	
$w(f(f(c)), d).$	$n(c, b, g(d)).$	
$w(g(b), c).$	$n(b, c, f(d)).$	

$q = \{ p(f(X), Y, Z), p(f(X), Y, X), p(f(f(X)), g(X), Y),$
 $u(a, b), u(a, c),$
 $r(a, f(X)), r(X, f(a)) \}$

このとき, q のメタキュアリ k は

$k = \{ p(f(X), Y, Z), u(a, X), r(X, f(Y)) \}$

となり, k によって特殊化された知識ベース $k b_k$ は

$k b_k$
 $p(f(a), b, g(c)). \quad u(a, b). \quad r(c, f(b)).$
 $p(f(a), c, f(b)). \quad u(a, f(g(b))). \quad r(g(a), f(g(b))).$
 $p(f(f(c)), d, a). \quad u(a, g(c)).$

となる。即ち, $k b_k$ は k の全解の集合である〔竹内85〕。

また, $k b$ をルール(idb) とファクト(edb) の集合からなる演繹データベースと見ると, ホーン節変換〔Miyazaki 86〕を使って $k b_k$ は

$k b_k$
 $p(f(X), Y, Z) :- q(f(X), Y), s(Y, f(Z)), t(Z).$
 $q(X, Y) :- w(X, Z), q(Z, Y).$
 $q(X, Y) :- w(X, Y).$
 $r(X, f(Y)) :- s(X, f(f(Y))), t(f(Y)).$
 $u(a, Y) :- n(c, a, Y), t(Y).$
 $s(b, f(g(c))). \quad t(a).$
 $s(d, f(a)). \quad t(b).$
 $s(c, f(f(b))). \quad t(f(b)).$
 $s(g(a), f(f(g(b)))) \quad t(f(g(b))).$
 $w(f(a), b). \quad t(g(c)).$
 $w(f(a), c). \quad n(c, a, b).$
 $w(f(f(c)), d). \quad n(c, a, f(g(b))).$
 $w(g(b), c). \quad n(c, a, g(c)).$

となる。ここで, ルールに対しては k による具体化, 及び節変換が行われ, ファクトに対しては節変換の結果ルールのボディ部に表われるファクトの述語のうち, Least-general な項とユニファイ可能なファクトの選択が行われている。

4. おわりに

ホーン節からなる知識ベースにおける部分評価の概念と, 問い合わせ処理に適用した場合のメタキュアリによるビュー定義方法とその意義について述べた。

また, 本稿で述べた部分評価は知識ベースにおける特殊化変換であるが, 例えば頻繁に使われる知識を予めコンパイルしておく方法や, 最適化のための等価変換〔Tamaki 83〕なども広義の部分評価と考えることができよう。

現在, 我々は幾つかのモデルの下に知識ベースシステムの研究, 開発を行っており, そこでの問い合わせ処理における最適化制御として, 本稿で述べた部分評価手法をメタプログラムとして実現することを考えている。

これまでの議論の問題点として, 本文中にも述べたがメタキュアリがキュアリと殆ど変わらないときは, ここで述べた手法では不十分である。また, 発展として自然言語処理への応用, あるいは知識ベースにおいて知識が段階的に獲得されるような場合, 既知データによって知識ベースを前向きに特殊化していくことなどが考えられるが, これらは今後の研究課題としたい。

謝辞

KC 会議メンバ, 研究室の皆様へ感謝します。

参考文献

- 〔二村 83〕 二村良彦: プログラムの部分計算法, 電子通信学会誌, Vol 66, No 2, 1983.
- 〔竹内 85〕 竹内彰一, 他: 部分計算のメタプログラミングへの応用, 情報処理学会研究会資料, 1985.
- 〔Miyazaki 86〕 N.Miyazaki, et al: Compiling Horn Clause Queries in Deductive Databases, ICOT-TR, No183, 1986.
- 〔吉田 85〕 吉田幹, 他: 導出原理における置換の束論に基づく定式化, 電子通信学会研究会資料, 1985.
- 〔Tamaki 83〕 H.Tamaki, et al: A Transformation System for Logic Programs which Preserves Equivalence, ICOT-TR, No18, 1983.