演繹データベースにおける制約伝播の実現方式

阿比留 幸展*， 羽生田 博美*， 宮崎 収兄*， 森田 幸伯**， 伊藤 英則**

*:沖電気工業（株）　　**:（財）新世代コンピュータ技術開発機構

1．はじめに

演繹データベースの研究において，制約優先による処理効率化は重要な課題である．特に，データが大量のときや再帰のルールがあるときの制約優先の実現は，処理効率化に大きく貢献する．

演繹データベースの制約優先に関する研究は，[Ceri86, Kife86]などがあるが，これらは，制約優先を実現するために関係代数式やグラフの変形を行うものであった．我々の方法は，ホーン節で表現されたルール集合に対して等価な定数代入が可能かどうかを判定するものである．等価な定数代入が可能な場合には，制約優先が実現できる．この方法は，ルールに対して，直接制約を導入することができるため，容易に実現できる．

本稿では，まず，我々の問い合せ処理の方式述べる．次に，等価な定数代入がどのような場合に可能かを述べ，その判定アルゴリズムを述べる．なお，以下で扱うルールは，否定（not）及び関数のない純ホーン節を仮定し，ルール中で扱う述語は，述語名が同じで引数の数が違う述語はないものとする．

2．問い合せ処理方式

我々の問い合せ処理方式は，データベースに存在する関係側から部分解の関係を順次組み立てていき，最終解を得るボトムアップ方式とする[宮崎87]．これを実現するために，ルール集合に対して強連結成分分解[羽生87]を行う．成分間には，半順序関係がついており，成分Aから成分Bへのパスが存在するとき，AはBより上（BはAより下）の成分であるという．この強連結成分分解された下の成分から順次解を求める．これにより，成分内の述語をヘッドとするルールのボディ中に含まれる述語は，成分内の述語，下の成分の述語，データが外延データベースに存在する述語のいずれかであるため，各成分ごとに処理効率化を考えることができる．

また，制約条件については，各成分の検索を実行する前に，上の成分から順に下の成分へ，トップダウンに制約条件の伝播が可能かどうかを判定する．

3．制約優先と制約伝播

本章では，どのような場合に制約優先が実現できるかを述べ，その判定アルゴリズム，および他成分への制約の伝播について述べる．

3．1　制約優先による検索空間の縮小

我々の目的は，データの検索空間の縮小による処理の効率化である．たとえば，q(X,c)を計算するとき，q(X,Y)の解集合(X,Y)をすべて求めてから，Y=cであるXの値を求めるより，q(X,Y)の代りにq(X,c)により定数代入を行ったルール集合により解集合を求めるほうが，もしそれが等価なら，効率的である．以下で，定数代入および等価な定数代入の可能性を定義する．

定義　specific, general　述語名が同じ述語PとP'について，PがP'よりspecific(PがP'よりgeneral)であるとは，σ(P')=Pとなる置換σが存在する場合をいう．
定義　定数代入　述語$P(X_1,\ldots,X_{i-1},c,X_{i+1},\ldots,X_n)$(cは定数)により，ルール集合に定数代入を行うとは，Pをヘッドとするルールのヘッドの第i引数に定数cを代入することをいう．
定義　等価な定数代入が可能
$P(X_1,\ldots,X_{i-1},c,X_{i+1},\ldots,X_n)$(cは定数)について等価な定数代入が可能とは，定数代入を行っていないルール集合により求めた解と$P(X1,\ldots,X_{i-1},c,X_{i+1},\ldots,X_n)$による定数代入を行ったルール集合により求めた解が一致していることをいう．

一般に，ルールに再帰が存在しない場合は，等価な定数代入が可能である．しかし，ルールに再帰が存在するときは，等価な定数代入が不可能な場合がある．そこで，どのような場合に等価な定数代入が可能かを以下の定理で述べる．以下の議論では，簡単化するために，問い合せをq(X,c)とおく．一般形への拡張は容易である．

補題1　問い合せをq(X,c)とする．このとき，q(X,c)と定数代入される前のルール集合より構成されるSLD木[野口86]の根から葉への各パスにおいて，初めて出現する述語qが，すべてq(X,c)よりspecificであるとき，q(X,c)によるルール集合への等価な定数代入が可能である．
証明　SLD木中で根から葉への各パスについて最初に現われるqの形がq(X,c)ならば，その節点以下の部分木に現われるqの形は，すべてq(X,c)の形になっている．したがってSLD木中に現われるqの形はすべてq(X,c)の形になっている．このとき，ルール集合中のq(X,Y)をすべてq(X,c)に置換えたルール集合は，問い合せq(X,c)に関して，明らかにもとのルール集合と等価である．定数代入したルール集合は，全部置換えられたルール集合よりgeneralであるため，もとのルール集合と等価である．
■
補題2　問い合せをq(X,c)とする．このとき，qのSLD木の根から葉への各パスについて，最初に現われるqと同一成分の述語をP0,P1...,Pnとする．ここでP0はqである．各Piを根とする部分木T_iを考えるとき，Tiの根から葉への各パスについて最初に現われるPiと同じ述語名のPi'がPiよりspecificならば，Piについて等価な定数代入が可能である．

Constant Propergation in Deductive Database
Yukihiro ABIRU*, Hiromi HANIUDA*, Nobuyoshi MIYAZAKI*,
Yukihiro MORITA**, Hidenori ITOH**
*Oki Electric Industry Co., Ltd.
**Institute for New Generation Computer Technology

証明　補題１より明らか。■

定理　問い合せをq(X,c)とする。このとき，q(X,c)による定数代入を行い，qのボディ中に現われた同一成分の述語の引数に定数が代入されていれば，その述語についても定数代入を行う。この作業を繰り返してすべてのルールについて定数代入を行ったあとのルール集合について，ボディ部に現われるPiと同じ述語名のPi'はすべて，Piをヘッドとするルールのヘッドよりspecificならば，qと同一成分の述語P0,....,Pnすべてのルール集合について等価な定数代入が可能である。

証明　定理の定数代入の方法は，補題２において述べているSLD 木の実行を模擬したものである。定理の条件は，他の部分木に現われる述語の形まで制限したものであり，明らかに補題２の条件の'部分木における節点が根よりspecificである'という条件を満たしている。したがって，補題２よりすべてのP0,...,Pnについて等価な定数代入が可能である。■

３．２　制約優先の実現可能性判定アルゴリズム

前章で述べた定理をもとに，定数代入が可能かどうかを判定するアルゴリズムを例により示す（表記法はPrologに準拠する）。

(1) 再帰述語が唯一つの場合

```
ancestor(X,Y):-parent(X,Y).              (1.1)
ancestor(X,Y):-parent(X,Z),
               ancestor(Z,Y).            (1.2)
?-ancestor(X,taro).
```

アルゴリズム：

① 各ルールのヘッドの第二引数の変数をすべて'taro'に置換える。

```
ancestor(X,taro):-parent(X,taro).        (1.1)'
ancestor(X,taro):-parent(X,Z),
                  ancestor(Z,taro).      (1.2)'
```

② ヘッドと同じ述語が，すべて，第二引数が'taro'になっていれば3.1の補題の仮定の条件を満たしている。上記の例では，(1.2)'のボディ中のancestor述語が，ancestor(Z,taro)となり，定理の条件を満たしている。したがって，ancestor(X,taro)について等価な定数代入が可能である。

これに対し，問い合せがancestor(taro,X)の場合は定理の条件を満たさない。

(2) 再帰が複数の場合

```
p(X,Y):-a(X,Y).             (2.1)
p(X,Y):-q(X,Z),b(Z,Y).      (2.2)
q(X,Y):-b(X,Y).             (2.3)
q(X,Y):-p(X,Z),a(Z,Y).      (2.4)
?-p(c,Y).
```

アルゴリズム：

① pをヘッドとするルールの第一引数の変数をすべて'c'に置換える。

```
p(c,Y):-a(c,Y).             (2.1)'
p(c,Y):-q(c,Z),b(Z,Y).      (2.2)'
```

② (2.2)'のボディにq(c,Z)があるので，成分内のもう一つの述語qについて，同様のことを行う。

```
q(c,Y):-b(c,Y).             (2.3)'
q(c,Y):-p(c,Z),a(Z,Y).      (2.4)'
```

③ p,q両方の述語について，pは第一引数がすべてcであり，かつqも第一引数がすべてcであるかどうかをチェックする。この場合，両方とも成りたっているため，pおよびqについて等価な定数代入が可能である。

３．３　成分間の制約条件の伝播

成分間の制約の伝播は，２．１で述べたようにトップダウンに行う。制約条件は，成分内の述語に対応するルール集合に定数代入が可能で，かつ別の制約を表現したパスが存在しないときに，上の成分から下の成分へ伝播できる。制約を伝播できる例を示す。

```
goal(X,Y):-friend(X,Y).                 (3.1)
goal(X,Y):-friend(X,Z),
           ancestor(Z,Y).               (3.2)

ancestor(X,Y):-parent(X,Y).             (3.3)
ancestor(X,Y):-parent(X,Z),
               ancestor(Z,Y).           (3.4)
?-goal(X,taro).
```

アルゴリズム

①goalを含む成分について，定数代入が可能かどうかを調べる。

```
goal(X,taro):-friend(X,taro).           (3.1)'
goal(X,taro):-friend(X,Z),
              ancestor(Z,taro).         (3.2)'
```

goalは再帰でないため，等価な定数代入が可能である。

②下の成分へのパスの形を調べる。この場合下の成分はancestorで，制約を表現したパスの形が(3.2)'のancestor(Z,taro)のみなので，下の成分への制約の伝播が可能である。

③ancestor(Z,taro)を問い合せとみて，等価な定数代入が可能かどうかを調べる。前章の例の(1)より，ancestor(Z,taro)についても等価な定数代入が可能である。

４．おわりに

３章で述べた方法の一部を先に報告したPHI/K²[阿比87]で実現し，現在インプリメント中のKBMS PHIにおいて改良を行っている。この方法は，ルールに対し，ルールからグラフへの変換などを行わずに定数を代入することが可能であり，論理型言語のユニフィケーション機能をそのまま取り込めるため，従来の方式と比較して実現が容易である。今後，KBMS PHIにおいて評価を行っていく予定である。

《参考文献》

[Ceri86]　Ceri, S., et al., "Translation and Optimization of Logic Queries: The Algebraic Approach", VLDB86, pp.395-402, 1986

[Kife86]　Kifer, M., et al., "A Framework for an Efficient Implementation of Deductive Databases", Proc. of the 6-th Advanced DB Symposium, pp109-116, 1986

[宮崎87]　宮崎他,"KBMS PHI:問い合せ処理の効率化",情報処理学会第34回全国大会,3K-6,1987

[羽生87]　羽生田他,"KBMS PHI:問い合せ処理モデル",情報処理学会第34回全国大会,3K-5,1987

[野口86]　野口,滝沢,"知識工学基礎論",オーム社, 1987

[阿比87]　阿比留他,"演繹データベース実験システムPHI/K²",情報処理学会第34回全国大会,3K-8,1987