

継承データに対する高速アクセス法の一考察

4 A a - 8

西岡 秀一 小西 史和

NTT 情報通信研究所

1 はじめに

近年、オブジェクト指向モデル等の複雑な構造のデータを管理する方式として、RDBMS と OODBMS の機能を併せ持つ ORDBMS [1] が着目されている。著者らは、交換処理向けに開発した高速 DB サーバ [2] のメモリアーキテクチャ技術を基盤に、リレーショナルモデルからオブジェクト指向モデルへの拡張を行うことで、高速 ORDBMS の実現を目指している。

オブジェクト指向モデルの特徴の一つとして、型の継承がある。高速アクセス法としてインデックスを用いる方法があるが、型継承により定義された継承データに対するインデックス検索では、検索対象となるクラスが、単一の場合及び複数の場合があり、いずれの場合に対しても効率的なアクセスを行えるインデックス方式は確立されていない。

本稿では、継承データに対する効率的なアクセス法を提案し、その有効性を示す。

2 継承データに対する検索

継承データとは、型継承で定義されたクラスのデータである。

図 1 において motorized_vehicle クラスは、vehicle クラスの属性(price, name)を継承しており、rotary_vehicle クラスは、motorized_vehicle クラスが継承している属性と、motorized_vehicle クラスで新たに定義された属性(power)を継承している。

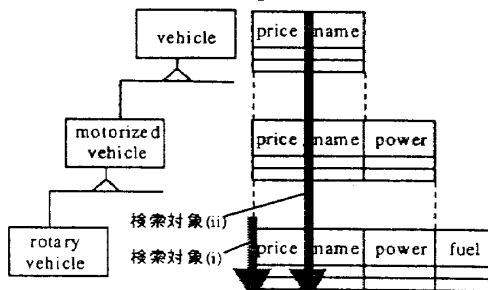


図 1 : 継承データ

継承データに対する検索には、大きく以下の 2 つのパターンが存在する。

A High Speed Access Method to Inheriting Data Structure
Shuichi Nishioka(nishioka@dq.isl.ntt.co.jp)
Fumikazu Konishi(konishi@dq.isl.ntt.co.jp)
NTT Information and Communication Systems Laboratories

(i) クラス内検索

検索対象のデータがクラス内に閉じている検索。

例、「rotary_vehicle クラス内で 200 万円(price=2000000)の乗り物名(name)は?」

(ii) クラス配下検索

検索対象のデータが属性を継承しているサブクラスにまたがる検索。

例、「vehicle クラス配下で 200 万円(price=2000000)の乗り物名(name)は?」

2.1 従来方式における問題点

従来のインデックスを用いて検索の高速化を行う場合、以下に示す 2 つの方法がある。

① クラス毎にインデックスを作成する方法

インデックスキーは、キーのシングルカラムとする。

② 継承関係を有するクラス間にまたがってインデックスを作成する方法

インデックスキーは、「キー」+「クラス識別子」のマルチカラムとする。

①の方法では、クラス内検索はインデックスと検索対象が一致するため効率的に行えるが、クラス配下検索は検索対象インデックスを特定する処理が更に必要となるため効率的に行えない。

②の方法では、クラス内検索はインデックスと検索対象が一致しないため効率的に行えない。また、クラス配下検索は検索対象と一致するが、マルチカラムにしたことによりインデックスサイズが増加し、効率的に行えない。

以上のように、いずれの方法においても、検索の種類によって検索対象クラスの範囲が異なるため、全ての検索パターンに有効なインデックスが確立されていない。

3 継承データアクセス法

本稿で提案するアクセス法は、メモリデータベースのインデックス方式で有効な T-木[3]を基盤としている。以下に、その提案方法を示す。

3.1 クラス毎のインデックスの拡張

クラス毎のインデックスを用いてクラス配下検索を行う場合、複数存在するインデックスの中から特定していたが、各クラスのインデックスを継承関係に従った結合(リンク)により、検索の高速化を図

る。リンクで結合する対象は、親インデックスのノードにおける最小値と最大値が、子インデックスにおいて部分木内に含まれるようなノードとした(方式α)(図2)。

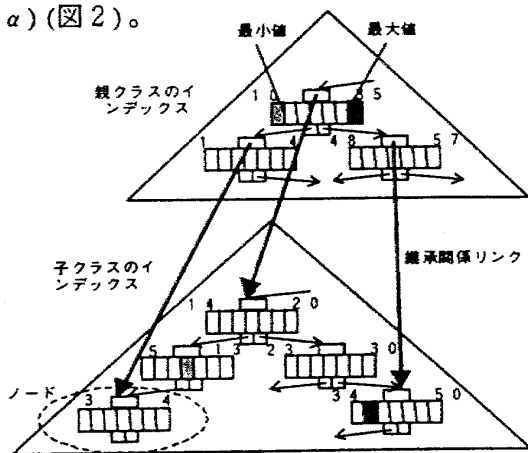


図2：方式αインデックス

このことにより、クラス配下検索は、検索開始クラスを特定した後、インデックスを特定せずにリンク情報をたどることにより可能となる。

3.2 クラス間にまたがるインデックスの拡張

クラス間にまたがるインデックスを用いて、クラス内検索を行う場合に、検索対象クラス外のデータがインデックスに含まれていた。このデータの冗長性を軽減するために、インデックスをキーのみで構成し、クラスが異なる同一キーについてはインデックス領域とは別にクラス識別領域を設けた(図3)。

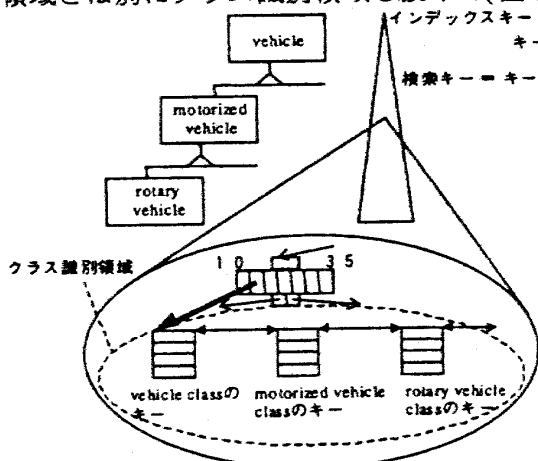


図3：方式βインデックス

このクラス識別領域は、一クラスにある同一キーをクラスタリングし、それを双方向リンクで結合することにより実現した(方式β)。

このことにより、クラス内検索は、キーのみで構成されたインデックスを用いることにより可能となる。また、クラス配下検索は、インデックスをキーのみで一度探索した後、該当クラスを検索すること

により可能となる。

4 提案方法を用いた実測

前述した2つのインデックスを、1クラスあたりのデータ数を1万件、最上位クラスから最下位クラスまでの段数(継承段数)を4段、1クラスから定義される子クラス数(分岐数)をパラメータとして、検索性能についての評価を行った(図4)。

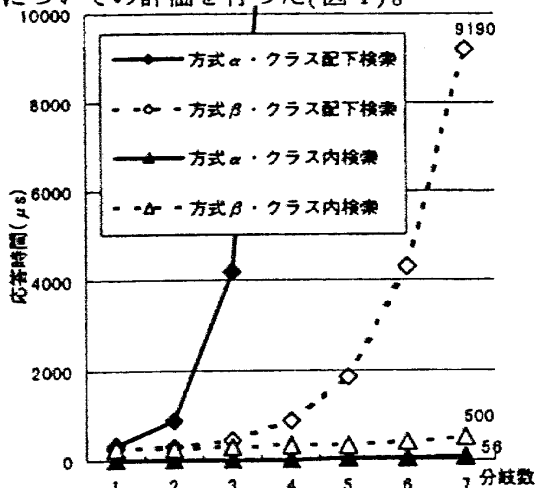


図4：継承段数=4における応答時間

図4より、クラス内検索における応答時間は、両インデックスとも分岐数の増加による顕著な変化は無い。一方、クラス配下検索における応答時間は、分岐数の増加に伴って指数関数的に増加する。増加の割合は、方式αが方式βより顕著である。

以上より、検索の種類に依存せずに効率的なアクセスが可能であるのは方式βのインデックスであると考えられる。

5 おわりに

本稿では、継承データに対する高速アクセス法を提案し、実装・計測によりその有効性を確認した。

今後は、更新系操作に関する定量評価を行った後、高速ORDBMSに組み込み、DBMS全体としての評価を継続して行う予定である。

参考文献

- [1] 太田 佳伸訳, Michael Stonebraker, "オブジェクトリレーショナルDBMSs," インターナショナル・トムソン・パブリッシング・ジャパン, 1996.
- [2] T.Honishi, etc., "Design and Implementation for an Enhanced Relational Database Management System for Telecommunication and Network Application," 18th PTC'96, T.1.7.4.
- [3] Lehman.T.J, "A Study of Index Structure for Main Memory Database Management," Proc. 12th Int. Conf. on VLDB, Aug 1986.