

「高度応用のための情報ベースモデルとその実現技術」 を目指して

牧之内 顕文*, 天野 浩文†, 岩井原 瑞穂‡, 岩間 一雄*,
茨木 俊秀§, 大堀 淳¶, 喜連川 優||, 掛下 哲郎**, 江 允††

*... 九州大学工学部情報工学科
†... 九州大学大学院総合理工学研究科
‡... 京都大学数理解析研究所
¶... 佐賀大学理工学部情報科学科

†... 九州大学大型計算機センター
§... 京都大学大学院工学研究科数理工学専攻
||... 東京大学生産技術研究所
††... 倉敷芸術科学大学産業科学技術学部

平成 8 年度から 3 年間, 科学研究費重点領域研究「メディア統合および環境統合のための高機能データベースシステムの研究開発」が実施される。本報告は, その内の A 班の研究内容の概略をまとめたものである。

Toward the Information Base Model and Its Implementation Techniques for Advanced Applications

Akifumi Makinouchi*, Hirofumi Amano†, Mizuho Iwaihara‡, Kazuo Iwama*,
Toshihide Ibaraki§, Atsushi Ohori¶, Masaru Kitsuregawa||, Tetsuro Kakeshita**, Yun Jiang††

*... Dept. Comp. Sci & Comm. Eng., Kyushu Univ.
†... Dept. Inform. Syst., Kyushu Univ.
‡... Res. Inst. for Math. Sci., Kyoto Univ.
§... Dept. Inform. Sci., Saga Univ.
¶... Coll. Sci. & Ind. Tech., Kurashiki Univ. Sci. & the Arts

†... Comput. Center, Kyushu Univ.
§... Dept. Applied Math. & Physics, Kyoto Univ.
||... Inst. of Ind. Sci., Univ. Tokyo

Under the support from a Grant-in-Aid for Scientific Research on Priority Areas, a three-year research project will be launched in 1996. The whole project is dedicated to advanced database systems for media and environment integration. This report gives an overview of subprojects to be carried out by Group A.

1 はじめに¹

データベースを取り巻く環境が大きく変わりつつある。一つは、計算機環境の変化である。並列計算機の普及と高速ネットワークの出現は分散・並列計算環境を実用的なものにしつつある。二つ目は、マルチメディアデータのニーズの高まりである。さらに、ディスクアレイなどの二次記憶装置の変革がある。そして、計算機応用の高度化が変化に拍車をかける。データベースシステムは、これら環境の変化に追随し、新しい要請に的確に応えなければならない。

本報告では、A 班の構成研究員が以上の要求に応えるべく、次の 3 年間で何を研究するのかについて述べている。

2 超並列計算機による高度データベース処理の研究²

2.1 研究の背景

データベースに対する要求が高度化・多様化するため、さまざまな形態の大量のデータに対する複雑な処理を極めて高速に処理する能力が必要になる。近年、多数の CPU を搭載して全体の演算性能の向上を目指す超並列計算機が広まりつつある。超並列計算機は、演算性能の他、主記憶容量の大きさ、並列入出力など、データベース処理の高速化に有利な性質を持っている。

並列計算機をデータベース処理、特に関係の結合演算の並列化に利用するシステムの多くは、クライアントの処理要求に応えるためのデータベースマシンとして実現されている。この方式では、並列計算機は外部からの要求を処理するためだけに用いられており、アプリケーション自体は並列に動作しない。しかし、今後データベースシステムのアプリケーションが高度化・複雑化していくと、それ自体も同じ超並列計算機の上で並列に処理しなければならない可能性がある。

この場合、アプリケーションを構成する多数のプロセスから処理要求が並列に発せられ、それらがデータベース管理システムを構成する多数のプロセスによって並列に処理されることになる。しかし、従来の並列データベース処理技術はこのよう

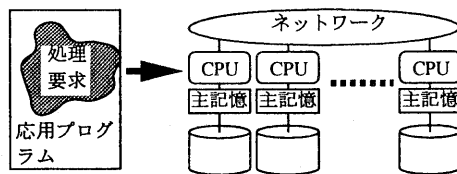


図 1: 従来のデータベースマシン

な状況を想定していない [1]。そこで、本研究では、高度で多様な要求に応える超並列データベースシステムの技術基盤の確立を目指す。

2.2 超並列オブジェクト指向永続プログラミング言語

1つのアプリケーションを構成する多数のユーザプロセスが超並列計算機で動作している場合、ユーザプロセスからの処理要求も多数同時に発生する。このため、単一の処理要求から並列実行できる部分を抽出して処理の高速化を図るよりも、多数のプロセスから多数のプロセスに対して発せられる大量の要求をいかにスムーズに伝達し、それをいかにそのまま並列に処理するかが重要になる。

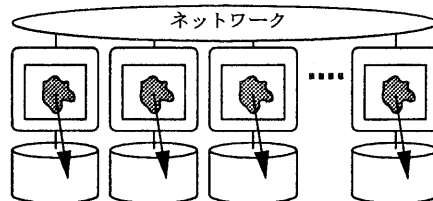


図 2: 超並列データベースシステム

本研究では、ユーザプログラムとシステムとを完全に分けて構築するよりも、主記憶中のデータと二次記憶中のデータを全く同様に操作できる永続プログラミング言語 [2] の考え方を超並列システムに応用する立場を採る。

また、超並列計算機は大規模科学技術計算のために用いられることが多いため、プログラミング言語も行列演算の高速化だけを目指したものが多い。ところが、複雑な構造を持つデータの処理や条件を満たすデータのみに対する操作が必要となるデータベースシステムにおいては、行列を主たるデータ構造と考えるのは適切ではない。したがって、複雑な構造を持つオブジェクトの集合を基本的なデータ構造と考えるような、超並列オブジェ

¹執筆分担：牧之内頼文

²執筆分担：牧之内頼文，天野浩文，江 允

クト指向永続プログラミング言語の研究を行う予定である。すでに、主記憶常驻型データベースを対象にその予備的な研究を開始している [3]。

3 データの論理的解析と知識獲得³

客観的判断のためのデータ収集の重要性がさげばれ、種々の事象に関するデータの蓄積が加速度的に進行している。しかし、これらの大量のデータから意味のある知識を抽出するための科学的手法については、知識獲得 (knowledge acquisition) あるいはデータ発掘 (data mining) という名称の下にその研究が盛んになりつつあるものの、ともすれば情報の海に呑み込まれてしまうという状況が生じている。

そこで、本研究では、ある事象に関する正データの集合と負データの集合 (たとえば、ある病気をもつと診断された患者の医療データと、そうではないと診断された患者のデータ) が与えられたとして、これらを区別するための論理的判断規則を抽出することを、論理関数の構成の問題として捕え、計算の複雑さおよびアルゴリズムの開発の立場から検討する。また、この種の問題の多くが NP 困難、あるいは co-NP 困難であることから、効率よくかつ実用的に有用なヒューリスティック・アルゴリズムについても研究する。

A. 問題の定式化 まず、すべてのデータが n 次元 0-1 ベクトルとして与えられている場合を考える。例として、上記の医療データをとると、一つのデータ $x = (x_1, x_2, \dots, x_n)$ は 1 人の患者の n 種の項目についての検査結果を表している。すなわち、 x_i は、項目 i の結果 (血圧が正常値より高い (= 1)、高くない (= 0); 体温が正常値より高い (= 1)、高くはない (= 0) など) である。正データの集合 T は、その病気をもつと診断された患者の検査ベクトルを集めたものであり、負データの集合 F は、そうではない患者の検査ベクトルの集合である。 (T, F) からの知識獲得とは、検査ベクトルのどのような特長に基づいて診断が行なわれたかを明らかにすることに相当する。本研究では、これを T と F を区別する論理関数 f を見出すことであると考へ、 f の具体的表現を求め。ただし、 f は通常一意的に定まるわけではないので、できるだけ簡単な f (Occam のカミノリ)、

あるいは単調性や分解性など特殊な構造をもつ f を見出すことに興味をもっている。

これらの問題を数学的に定式化すると以下のよう書かれる。 n 次元 0-1 ベクトルの集合 $\{0, 1\}^n$ の部分集合 T および F が与えられたとき、 (T, F) を部分定義ブール関数という。一方、(完全定義) ブール関数は写像 $f: \{0, 1\}^n \rightarrow \{0, 1\}$ であり、真ベクトルの集合 $T(f) = \{x \in \{0, 1\}^n | f(x) = 1\}$ と偽ベクトルの集合 $F(f) = \{x \in \{0, 1\}^n | f(x) = 0\}$ をもつ。 f が $T(f) \supseteq T, F(f) \supseteq F$ をみたすとき、部分定義ブール関数 (T, F) の拡大であるという。 C をある性質をもつブール関数のクラス (たとえば、単調関数のクラス) とする。

EXTENSION(C)

入力: 部分定義ブール関数 (T, F) 。

出力: (T, F) の拡大 $f \in C$ が存在するならば f , 存在しなければ no。

一般に、与えられた (T, F) に対し、拡大 $f \in C$ が存在するとは限らないので、その場合、次の問題が重要である。

BEST-FIT(C)

入力: 部分定義ブール関数 (T, F) 。

出力: (T, F) に対し最小数の誤りをもつ $f \in C$ 。

これらについては、さまざまなクラス C についてその計算の複雑さが明らかにされている [4]。

B. 不完全データの取扱い 現実のデータでは、ベクトルのいくつかの要素に不定値 * が存在するのが普通である。* 値の取扱いには、* の値が 0 と 1 のどちらであっても正データあるいは負データとしての役割は変化しないと考へるか、* の値を正しく定めることによってはじめて正データあるいは負データとして利用できるかと考へるかによって異なる。したがって、関数のクラス C が与えられたとき、* の値に関わらず拡大 $f \in C$ が存在するかを問う問題、* の値をうまく定めることによって拡大 $f \in C$ が存在するかどうかを判定する問題、さらに、最少数の * の値を固定することによって、他の * の値は自由にしたまま拡大 $f \in C$ の存在を保証する問題などが考へられる。

C. 数値データ データとして、0-1 ベクトルでなく実数値を要素とするベクトルを扱う場合もある。このようなデータは、各要素ごとについて

³執筆分担: 茨木俊秀

のカットポイントを導入し、その値より大きい小さいかによって0と1の2値化をすれば、ブール関数による論理的扱いが可能となる。この場合、導入すべきカットポイントの最少数およびその位置の決定が基本的な問題である。

D. その他の研究 論理関数 f の表現には DNF (disjunctive normal form) あるいは CNF (conjunctive normal form)、さらに決定木などが考えられる。これらの表現の最簡形の定義にも種々の提案があり、それに応じてアルゴリズムも異なる。また、本研究のテーマと、データをつぎつぎと加えて行くことで真の f を見出そうとする学習との間には互いに双対とも言える関係がある。さらに、 n 次元ベクトルのデータだけでなく、遺伝子の処理にみられるような1次元ストリングのデータとの関係も興味深い。

4 高機能処理の高速化のためのアルゴリズムの研究⁴

新しい情報ベースモデルは様々な個別事例に統合的に対応できる性質が要求されている。従って、その上でのデータ処理アルゴリズムに関しても統合的性質が求められるのは当然で、従来のような個々の問題に最適な個別アルゴリズムをアドホックに開発する手法では有効な対応ができないことは明らかである。幸いなことに、近年、難易度の高い組合せの性質の強い問題 (NP 完全問題等) に対して、例えば10年前では想像もできなかった性能を発揮する統合的性質を有したアルゴリズム的手法 (メタヒューリスティックス) がいくつか発見され、NP 完全問題は手に負えないという従来の常識を覆しつつある。

本研究では、そのようなメタヒューリスティックスとして局所探索と多項式時間近似図式を取り上げ、新情報ベースモデルに基づいた統合的データ処理アルゴリズムの開発に応用する。さらに、そのようなアルゴリズムの性能評価のためのベンチマーク問題を効率良く且つ現実性をできるだけ損なわない形で生成できるアルゴリズムを開発し、新情報ベースへの組み込みの可能性を探る。

(1) 局所探索. 局所探索法は巡回セールスマン問題のような組合せ的性質が比較的弱い問題に対しては従来から積極的に利用されてきた。しかし

最近それが、例えば論理式の充足可能性問題 (SAT) ように極端に組合せ的性質の強い問題に対しても、バックトラック等の従来法をはるかに凌駕する性能を示すことが発見された。その理由は、局所法の本質的課題である局所解からうまく抜け出す方法がいくつか発見されたからである。特に局所解に落ちた時に問題に多少の操作を加えることによってその局所解を解消してしまえるようになった。このことにより、例えばSATでは 10^4 変数規模の論理式まで解けるようになって来ている。

本研究では、この局所探索法を一般の制約充足問題に適用することを試みる。データ検索を始めとしてデータベース関連の組合せ問題のかかなりの部分が制約充足問題と考えることができる。局所解からの脱出に関しては我々が最近開発した制約式変形法 [5] の応用を試みる。

(2) 多項式時間近似図式. 近似アルゴリズムは古くから研究が盛んであるが、その性能 (最悪の近似度) を保証できるアルゴリズムの研究が進んでいる。あるタイプの問題に対しては比較的良い (例えば $1+\epsilon$ 倍) 性能保証が得られるが、また別の問題ではそれが非常に難しいといった近似アルゴリズムに対する各種問題の親和性が判明してきた。特に時間をかければかけるほど良い近似度が保証される多項式時間近似図式が存在する問題が数多く発見されている。近似アルゴリズムの多くはグリーディ手法を中心とした単純なものが多く実用性が非常に高いという利点もある。

本研究では、グラフの分割問題の近似アルゴリズムの構築を行なう。クラスター化を始めとして、最適化問題の多くがある種のグラフ分割問題に帰着できることが分かっており、本研究の目的にも合致する。我々は既にグラフから枝の多い部分 (相互の関連性が高い部分) を抜き出す単純なグリーディアルゴリズムがかなり良い性能保証を持つことの証明に成功している [6, 7]。

(3) ベンチマーク生成. 性能評価のためのテスト例題の重要性は誰もが認めるところである。従来固定のベンチマーク集合が使われることが多かったが、一般性に欠けるという批判や、ベンチマーク集合を意識したアルゴリズム開発というある種の不正の存在も根強く噂されてきた。さらにはベンチマーク集合が重要な企業秘密になり、公的な研究機関が利用することはほとんど絶望的になってきている。そこで自然に考えられるのがテ

⁴執筆分担：岩間一雄

スト例題のランダム生成である。ランダム生成は比較的簡単な方法で多くの例題を生成でき、それらを用いて、アルゴリズムの平均動作時間や、例題のサイズが大きくなるにつれてどの程度動作時間が増大するかを知ることができる。しかし強い批判もあり、その代表的なものは「ランダム例題は人工的過ぎて現実の世界を反映していない」というものである。

本研究では、この批判に耐えられるランダム例題生成技術を研究し新情報ベースへの適用を目指したい。基本的方針は、ランダム生成でありながら、生成例題の各種属性を制御できるようにすることである。我々はこのような方針の例題生成系を既にいくつか開発しており [8, 9], 新情報ベースモデルへ適合させることもそれほど困難ではないと考えている。

5 データベースシステムに於ける高性能化技法の研究⁵

データベース管理システムはシステムソフトウェアの中でも最も巨大なものの一つとなった。データベースシステムの重要性は言うまでもなく、その高性能化は不可欠であり不断の努力がなされている。ここでは2つの側面から高性能化へのアプローチについて述べる。

5.1 高並列処理方式の研究

関係データベースシステムはユーザーに対し非手続的インタフェースを提供することを可能にするとともに、柔軟性を向上させたが、反面、その支援の為、結合演算等の処理負荷は従来のナビゲーションシステムに比べ、大幅に増加した。勿論、ソフトウェアの効率化に関し、たゆまぬ努力が払われ、現在では比較的使いやすいシステムに育ったものの、巨大データベースをその操作対象とする場合には性能面で大きな問題がある。1995年10月25日サイベース社は128台のプロセッサからなる分散メモリ並列マシンSP-2上で並列データベースのデモンストレーションを行いその動作を確認した。恐らく、オープン環境下でのこの規模の実験は始めてであろう。実際に並列データベースが必要とされる時代に入りつつあるという事実を示している。我々は過去、並列

データベースシステムに関し処理アルゴリズム負荷分領域、コンパイラ等に関し研究を進めてきたが [10, 11, 12], 今後も重要な研究領域になろう。又、OLAPやDATA MININGなる名称も定着しつつある。とりわけ小売業を中心としたバスケットアナリシスの研究は活発になされている。我々も Associate Rule のマイニングに関する効率の高い並列処理方式を開発するとともに AP1000 上に実装し有効性を確認している [13]。

5.2 高並列記憶方式の研究

近年計算機エレメントのコモディティ化が進んでおり、プロセッサ、メモリ、ディスクいずれをとっても、一部のアプリケーションを除き専用化は影をひそめつつある。ディスクは従来、大型、ミニコン、パソコンとそれぞれ別の製造ラインにより製造されていた。これらを廉価なパソコンディスクで一元化しコストを下げるとともに信頼性を冗長記憶により維持する方式が定着しつつある。これが RAID と呼ばれるディスクアレイシステムである。ディスクアレイの基本的アイデアは IBM によって提案され、その後カリフォルニア大学パークレイ校のパターソンらによって広く RAID としてその有効性が明らかにされた。RAID 0, 1, 3, 5 なる名称は、ワークステーション等のカタログによく目にとまる様になった。通常最も良く利用されるのは RAID-5 であるが、ブロック書き込みに関し性能が 1/4 に低下することが従来大きな問題とされてきた。これに関し、STK社は LFS を用いた ICEBERG なる製品を長い期間をかけて開発に成功し、高性能化を図っている。我々は Virtual Striping と呼ぶ方式を提案し [14], ほぼ同等の性能を達成出来ることを示し、更に、Hot Block Clustering なる手法を提案し [15], これにより LFS の Cost Benefit 手法より高い性能を達成可能であることを示した。更に最近提案した Hot Mirror なる手法を用いることでディスク故障時のリビルドに於いても殆ど性能低下を隠蔽することが可能となる [16]。この様に高度な記憶管理手法を導入することにより二次記憶システムの性能を大幅に向上させることが可能であることが明らかになったが、この代償としてディスクコントローラの制御ソフトウェアは従来に比べ大幅に複雑化する傾向にありそのトレードオフを見極めることが必要であろう。二次記憶の

⁵執筆分担：喜連川優

みならず我々は三次記憶システムに関しても、スケーラブルアーカイバなるコモディティロボティクスを利用した拡張柔軟性に富む大規模三次記憶システムを開発している [17, 18]。本システムではアクセス頻度分布に応じたカセットマイグレーション方式を採用し極めて有効であることを確認している。

6 超並列オブジェクトベース言語の設計理論⁶

高性能大規模オブジェクトベース構築のための基本技術の開発は、本研究の重要なテーマである。この目的を達成するための一つの鍵は、データベース技術を、最近急速に進歩している超並列技術と統合することである。このためには、データベース処理に適した並列ハードウェアの研究開発や、データベース問い合わせ処理のための並列処理アルゴリズムの開発が重要であり、これらが本研究の重要なテーマとなっていることはすでに述べた通りである。しかしながら、高機能、高性能の並列データベースシステム構築のためには、それら基盤技術を使いこなすためのプログラミング言語の開発も重要な課題である。この認識のもとに、本研究では、超並列オブジェクトベース言語構築のための基礎理論およびその実用化技術に関する研究も行う予定である。

オブジェクト指向データベースはそれ自体複雑なシステムであり、その超並列化は、従来の関係データベースの並列化技術の延長上には存在し得ない。望ましい超並列オブジェクトベース用プログラミング言語構築のためには、複雑な機能の諸性質を宣言的に記述し、それらに内在する並列性を抽出することを可能にする理論的基礎が必須である。本研究では、ラムダ計算やプログラミング言語の理論研究で発展した型理論を基礎とし、それを、データベース処理に適した超並列処理機構を取り入れ拡張することによって、超並列オブジェクトベース用プログラミング言語のための基礎理論の構築をめざす。

並列計算の分野では種々の並列モデルが提案されているが、大量データを扱うデータベース処理にもっとも適していると思われるものは、データ並列モデルである。このモデルは、近ごろ実用化

されつつある超並列計算機のための主要な並列モデルでもあり、超並列データベースを超並列計算機上に実現する上でも重要な意味を持つ。このモデルの基礎となっているアイデアは、均一な構造をもつ集合データの処理を、各データエレメントに対して同一の演算を並列実行することによって行う、というものである。この構造から容易に想像される通り、データ並列処理は、相互に依存関係が少ない均一な構造の大量データの処理に最も適したものであり、集合演算を中心とした関係データベースの処理と親和性が高い。現在、データ並列の考えに基づく並列関係データベースシステムが実用化されつつある。これは、関係データベースが、均一の構造をもつタブルの集合からなり、タブル間には複雑な参照関係は存在しないという性質に基づく。しかしながら、本研究の主要な研究対象であるオブジェクト指向データモデルにおいては、データベースは、均一の構造をもつほぼ独立なデータエレメントの集合という構造はもっておらず、上記のデータ並列処理の考え方をそのまま適用することは不可能である。

現在のところ、オブジェクト指向データモデルのためのデータ並列処理方式は確立されていないと言ってよい。困難の原因は、オブジェクト指向データベースのデータが、オブジェクト識別子を媒介とした複雑な相互参照構造をもっていること、および、検索処理が、集合に対する一括操作のみならず、オブジェクト識別子を順次たどるナビゲーションをその重要な要素として含んでいることによる。データ構造のみに注目すれば、オブジェクト指向データベースは、ポインターで複雑に接続された入れ子型の巨大なレコード構造、言わばポインターで複雑に絡み合ったデータの塊である。従って、データ並列の考え方を、ポインターを含んだデータ構造に一般化し、さらにそれらをデータベース処理に必要な集合演算と統合することができれば、オブジェクト指向データベースのための超並列プログラミング言語の構築が可能になると期待できる。

我々は、この問題を、プログラミング言語の新たな評価モデルの構築の問題として捕え、ラムダ計算の型理論を基礎とし、ポインター構造を含んだ並列処理及びデータベース処理が表現可能な計算系の構築を目指す。我々の戦略は、[19]によって提案された並列ラムダ計算をデータベースの検索処理に必要な機能に拡張することである。[19]

⁶執筆分担：大堀淳

の基礎となっている考え方は、ポインタをふくんだデータ構造を、ポインタを変数集合とする等式系(すなわち連立方程式)と見なし、その等式系を、各等式に同一の処理を並列に実行することによって新たな等式系に変換する、という操作を繰り返すことによって並列処理を達成するというものである。これにより、ポインタを逐次的にたどる操作では不可能であった高度な並列化が達成可能になる。オブジェクト指向データベースも、オブジェクト識別子を変数とする等式系と見なせば、以上の考え方を適用することが可能ならずである。この見方のもとでは、検索処理は、等式系を別な等式系に変換する操作に相当する。この方式は、十分な表現能力のある計算系のもとでは、従来のナビゲーションの機能をも損なうことなく、オブジェクト指向データベース処理における高度な並列性を抽出する可能性を持つものである。さらにこの処理方式は、オブジェクト識別子を含んだデータに対する一般性のある代数を構成するとみなすことができ、したがってこの方式を基礎とし検索言語を設計するならば、従来のオブジェクト指向データベースでは通常得られない問い合わせ処理の関数閉包性(すなわち問い合わせの結果がもとのデータベースと同一の構造をもつという性質)を達成でき、関係データベースに比べ後述しているオブジェクト指向データベースのviewをも実現できる可能性も期待できる。我々は、以上の戦略のもとに、データ並列ラムダ計算を、集合操作及びデータ永続性を取り入れて拡張し、並列オブジェクト指向データベースのためのプログラミング言語構築を可能にする基礎理論とその実用化技術の確立を目指す。

7 動的データ再配置による負荷分散⁷

コンピュータシステムの効果的な高速化技法として並列処理が挙げられるようになって久しい。しかし、並列処理を行なう場合、要素間の負荷分散が図られなければ台数効果による速度向上は望めない。データベースシステムにおいては、低速な二次記憶アクセスがシステム中に含まれるため、負荷分散の必要性は一層高まる。

データベースマシンにおける代表的な負荷分散手法として、RAIDを用いたストライピングが挙

げられる。しかし、ストライピングはデータサイズが大きい場合を想定した技法であり、トランザクション処理システムのようにデータのグラニュラリティが極めて小さい場合には効率良く動作しない。

これに対して掛下らは、トランザクション処理用 Trinity アーキテクチャ(図3)の中で、ブロードキャストバスを活用してデータの動的再配置を行なう手法を提案している [21]。

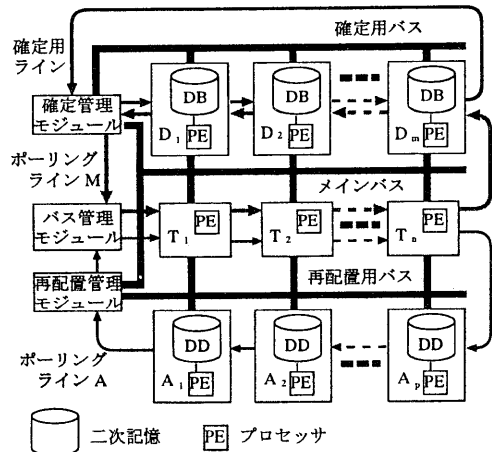


図3: Trinity アーキテクチャ

再配置管理モジュールはメインバスのモニタリングを行なっており、各ディスク D_i の負荷を把握している。最も負荷の大きいディスク D_{max} と CPU T_i の間で転送されるデータについては、負荷の最も小さなディスク D_{min} へも転送することによってデータの再配置を行なう。本方式は、(1) メインバスに余分な負荷をかけずにデータの再配置が可能になる、(2) アクセス頻度の高いデータほど再配置され易くなり、負荷分散効果が高まるといった性質を持つ。また、シミュレーションにより良好な性能を持つことも確認されている [22]。しかし、ディスク台数の増加とともにディスク間の負荷の差が線形に増加するなど、いくつかの課題も持つ。

本サブテーマでは、動的なデータ再配置方式を改良することで、上記の問題点を解決し、効率的な負荷分散を行なうための制御手法に関する研究を予定している。具体的には以下の項目が挙げられる。

- 平均以上の負荷を持つディスクから平均以下の負荷を持つディスクへの動的再配置による負荷分散効果の向上

⁷執筆分担：掛下哲郎

- ディスク間でのコピー分散を活用した負荷分散
- ディスク負荷の変動を推定する再配置制御方式
- 動的再配置に対応した並行処理制御方式、ディスク内での領域割り当て方式 など
- 動的再配置を行なうために必要とされるハードウェア要素の設計
- 動的再配置に伴うフラグメンテーション問題、バックアップ問題に関する考察
- 故障発生時における動的再配置の性能評価

本手法はブロードキャストバスを利用しているため、バスがボトルネックになる可能性がある。しかし、バス速度は飛躍的に向上しつつある [20]。また、ATM 技術を始めとするコンピュータネットワークの速度向上は著しく、通信媒体がボトルネックになる可能性は将来とも低下することが期待できる。また、動的再配置ではデータ転送量を削減するための種々の技術によりバスへの負荷を減らしている。これにより、100MB/sec 程度のバスであっても、0.5KB のデータを 1 秒間に 10 万個以上転送できるため、1000 台以上のディスクを接続できる可能性を持つ。

8 高機能情報ベースのための質問処理⁸

8.1 共有可能な索引構造

データベース/知識ベースにおいて、関係演算の結果が巨大になる場合の記憶コストを抑制する手法として、我々は二分決定グラフ (Binary Decision Diagram) に基づく共有可能な索引構造を提案している [23]。二分決定グラフは論理回路設計の分野で用いられている論理関数の表現法である。関係を論理関数に符号化することにより、二分決定グラフで表現することが可能であるが、この表現の長所は (1) 同じ関係はつねに同じ構造の二分決定グラフになる、(2) 1 つの関係に同じ部分集合が重複して含まれる場合は、1 つだけ記憶することにより、記憶サイズを削減する、などがある。(1) の性質を用いると、過去に行なった演算を記憶 (キャッシュ) しておくことにより、同じ関係演算が重複して実行されるのを防止でき、質問処理が高速化される。

⁸執筆分担：岩井原瑞穂

上記のような特長がよく生かされる質問処理として、再帰質問のボトムアップ評価がある。我々はこれまでに、関係を二分決定グラフに符号化する方法として、対数符号化と線形符号化の 2 つを提案し、両者と既存の演繹データベースの 1 つを実験を通して比較を行なっている。その結果二分決定グラフを用いた方法が、従来のハッシュ結合に基づく方法よりも、密な関係が生じる場合に実行速度を大きく改善していることを示している。

これまでの我々の研究では主記憶に格納する索引を想定していたため、今後二次記憶への格納に適した共有可能な索引構造、および分散主記憶環境に適した索引構造の設計および実装を進める。そして、性能評価を通して共有可能な索引が有効となる条件を明らかにする。

また、以下のような用途にも応用してみる。

- 複合オブジェクトの格納, 集合演算, 包含判定
- 1 つのアプリケーションが同一の関係演算を冗長的に繰り返し発行する場合や、複数の利用者からの質問が同一の関係演算を多く含む場合など、演算のキャッシュが有効になる状況

以上の応用は、共有可能な索引が有効と考えられる。

8.2 遠隔協調設計作業を支援するためのデータベース機能

近年の計算機ネットワークの普及に伴ない、遠隔地を高速ネットワークで結び、協調作業を行なうことが可能になりつつある。本テーマでは、特に LSI 設計 CAD システムを対象として、その遠隔協調設計作業を支援するために必要な、新たなデータベース機能について検討する。

現在の LSI 設計開発においても、遠隔地間のコミュニケーションが必要となる場合として、以下のような多くの形態が存在する。

- (設計工程内) 同一の工程を設計チームが地理的に分散して設計している場合
- (設計工程間) システム設計, 論理回路設計, レイアウト設計, ファウンダリなどの設計工程がそれぞれ別の事業所で行なわれることが多い。上流工程から下流工程へ渡された設計データに問題が見つかった場合、設計者が実際に遠

隔地に赴き、ワークステーションの同じ画面を見ながら対話的に問題解決することが多い。

- (ユーザサポート—CAD ツールベンダ—ユーザ間の対話) ツールの初歩的な使い方の質問など、電話で対応できるものから、複雑な操作やバグのレポートなど、より高度なコミュニケーションが必要なケースがある。
- (ツールトレーニング) CAD ツールの使い方を教育する場合、複数の講習者を同じ部屋に集めてレクチャー形式で指導が行なわれることが多い。

上記のようなコミュニケーションが高速ネットワークを介して、遠隔地間で実現されることが有用である。そのためには、例えば以下のような機能が必要になると考えられる。

- リアルタイム協調設計のための排他制御、バージョン管理、バージョン統合。
- 設計データの効率的な分散配置。知的所有権の問題上すべての設計データを分散できない場合を考慮する必要がある。
- 通信容量を考慮したトランザクション管理。画面共有などの人対人の通信と、設計データの転送が円滑に行なわれるようなスケジューリングを行なう。

9 おわりに⁹

計算機科学は具体的な「計算機」を対象にしているせいか、実用指向が強い学問である。特に、データベース研究はその成り立ちから「実用的」であり続けた。現在の科研重点もそのような性格になりつつあるように思える。例えば、3年の重点研究の結果として、「目に見える」成果が期待される。実際、昨年度に完了した超並列関連の重点研究では、開発したハードを見せ、ソフトのデモを行って成果報告したと聞いている。一方、文部省による大学活性化施策の一つとして、「ベンチャービジネス」を起こせる人材の養成施策があり、重点大学に建物を含む設備予算がついた。

このような状況から見ると、重点研究の成果としては、単に論文を発表するだけでは不十分で、

⁹執筆分担：牧之内顕文

将来のベンチャービジネスの芽を出すような研究開発を行うことが期待されているのではなからうか。本報告中の研究から、将来世界を代表するユニークな「製品」あるいは、それを支えるアルゴリズム（「性能向上コンサルテーション」なるビジネスがあってもいい？）が生まれることを期待したい。

参考文献

- [1] DeWitt, D. and Gray, J.: "Parallel Database Systems: The Future of High Performance Database Systems," *Comm. ACM*, Vol. 35, No. 6, pp. 85-98, June 1992.
- [2] 有次, 天野, 牧之内: "永続プログラミング言語 IN-ADA におけるビューの実現," 情報処理学会論文誌, Vol. 36, No. 4, pp. 215-224, 1995年4月.
- [3] Imasaki, K. et al.: "Design of a Database Language for a Massively Parallel Main Memory Database System and Its Performance Benchmark," *Proc. Int. Symp. on Parallel and Distributed Supercomputing*, pp. 198-205, September 1995.
- [4] E. Boros, T. Ibaraki and K. Makino: "Error-free and best-fit extensions of partially defined Boolean functions," *RUTCOR*, Rutgers University, RRR 14-95, 1995.
- [5] Cha, B., Iwama, K.: "Performance test of local search algorithms using new types of random CNF formulas," *Proc. IJCAI-95*, Montreal, pp.304-310, Aug 1995.
- [6] Asahiro, Y. and Iwama, K.: "Finding dense subgraphs," *Proc. ISAAC '95*, (LNCS 1004), pp. 102-111, Cairns, Dec. 1995.
- [7] 朝廣, 岩間, 玉木, 徳山: "Greedy finding a dense subgraph," 通信学会コンピュータシミュレーション研究会, 1996年1月.
- [8] Asahiro, Y., Iwama, K., Miyano, E.: "Random Generation of Test Instances with Controlled Attributes," *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, American Mathematical Society (in press).
- [9] Iwama, K., Hino, K.: "Random generation of test instances for logic optimizers," *Proc. 31st ACM/IEEE DAC*, San Diego, 430-434, San Diego, June 1994.
- [10] Masaru Kitsuregawa, Satoshi Hirano, Masanobu Harada, Minoru Nakamura, Takayuki Tamura, Mikio Takagi: "Overview of The Super Database Computer(SDC-I):(INVITED PAPER)," *IEICE TRANSACTIONS ON ELECTRONICS*, Vol.E77-C, No.7, pp.1023-1031, July 1994.

- [11] 中村稔, 平野聡, 田村孝之, 喜連川優, 高木幹雄: “スーパーデータベースコンピュータSDC-IIにおけるシステムソフトウェアの設計と実装,” 電子情報通信学会論文誌 D-I, Vol.J78-D-I, No.2, pp.129-141, 1995.2.
- [12] 中野美由紀, 新谷隆彦, 喜連川優: “並列データベースシステムにおける多重結合演算処理の最適化とその評価,” 情報処理学会計算機アーキテクチャ研究会 113-19, pp.145-152, 1995.8.
- [13] 新谷隆彦, 喜連川優: “データマイニングの並列化に関する一考察,” 電子情報通信学会データ工学研究会, 1995.12.
- [14] Kazuhiko Mogi, Masaru Kitsuregawa: “Dynamic Parity Stripe Recoganzitions for RAID5 Disk Arrays,” Proceedings of Third International Conference on PARALLEL & DISTRIBUTED INFORMATION SYSTEMS, pp.17 - 26, 1994.9.
- [15] Kazuriko Mogi, Masaru Kitsuregawa: “Hot Block Clustering for Disk Arrays with Dynamic Striping =exploitation of access locality and its performance analysis=,” Proceedings of the 21st International Conference on Very Large Data Bases, pp.90-99, 1995.9.
- [16] 茂木和彦, 喜連川優: “Hot Mirroring を用いたディスクアレイの基本性能評価,” 情報処理学会研究報告データベースシステム研究会 VOL.95, NO.65, pp.49-56, 1995.7.
- [17] Kazuhiko Sako, Toshihiro Nemoto, Masaru Kitsuregawa, Mikio Takagi: “Partial Migration in an 8mm Tape Based Tertiary Storage File System and its Performance Evaluation through Satellite Image Processing Applications,” Proc. of 6th International Conference on Information Systems and Management of Data (CISMOD '95), pp.15 - 17, November 1995.
- [18] Toshihiro Nemoto, Yasuhiko Sato, Kazuhiko Mogi, Ken-ichiro Ayukawa, Masaru Kitsuregawa, Mikio Takagi: “Performance evaluation of cassette migration mechanism for scalable tape arhivber,” Proc. of SPIE Photonics East '95, Digital Image Storage and Archiving System, Vol.2606, pp.23-26, October 1995
- [19] Nishimura, S. and Ohori, A.: “A calculus for exploiting data parallelism on recursively defined data,” *Proc. International Workshop on Theory and Practice on Parallel Programming*, Springer Lecture Notes in Computer Science 907, 1995.
- [20] 今井 他: “変革期のパソコン用周辺インタフェース - UltraSCSI, IEEE 1394, USB へ,” 日経エレクトロニクス No.647, pp.141-162, 1995年10月.
- [21] 掛下, 高橋: “ブロードキャストを利用した高効率トランザクション処理マシンの基本アーキテクチャ,” 情報処理学会 DB 研究会 97-8, 1994年3月.
- [22] 掛下 他: “ブロードキャストを利用したディスク間の動的データ再配置方式の性能評価,” 電子情報通信学会 CPSY 研究会 95-83, 1995年12月.
- [23] Iwaihara, M., and Y. Inoue: “Bottom-Up Evaluation of Logic Programs Using Binary Decision Diagrams,” *Proc. 11th Int. Conf. Data Engineering*, pp. 467-474, Mar. 1995.