データベース・システム 79－3
プログラミング言語 26－3
（1990．9．7）

分散データデースの木質問の最適化

李 紅、 佐藤 洋
(株)リコーソフトウェア事業部、電気通信大学情報工学科

1990年9月7日

本論文では分散データベースの木質問処理の準結合による解法の最適スケジュールを求める問題について考察した。全回線使用時間を最小化するものとしている。Yu、Ozsoyogolu、Lamは最適スケジュールに対するいくつかの必要条件を導き、最適でない多くのスケジュールを探索空間から除去することに成功している。本論文ではデータベースの大きさを比較することによって得られる最適性の必要条件がいくつか導かれ、定理の形で述べられている。これらの結果は最適スケジュールの性質について興味深い示唆を与え、また最適スケジュールを得るための探索空間を更に減少させる。

OPTIMIZATION OF TREE QUERY IN DISTRIBUTED DATABASE SYSTEMS

Hong LI
Software Division, RICOH Company,Ltd.
1-17 Koishikawa 1-Chome,Tomin-Bldg Bunkyo-ku,Tokyo,112, Japan

Hiroshi SATO
Department of Computer Science, University of Electro-Communications
1-5-1 Chofugaoka, Chofu-shi,Tokyo,182,Japan

In this paper the problem of funding an optimum strategy of semi-joins for solving tree queries is studied under the objective of total time minimization. Yu,Ozsoyogolu and Lam obtained a set of necessary conditions for the optimal schedules and succeeded in reducing the search space for finding the optimum by eliminating strategies that can never be the optimum. Some necessary conditions obtained by comparing sizes of database are obtained in this paper. These results give suggestive knowledge for the optimal schedule and will be useful to further reduce the search space.

1　前書き

多くのサイトに分散配置されたデータベースに関する質問処理を実行するため、各サイトに置かれた関係を他のサイトに転送して結合する操作を続けて行うことになるが、その順序を上手に選択して、ある基準に従って最適なスケジュールを求める研究が盛んに行われてきている[1]-[9]。

本論文では、問題を木質問に限り、任意の関係を簡約する(reduce)ための準結合の系列より成るスケジュールを考え、全回線使用コストを最適化する最適化問題について考察する。これに関しては特にYu、Ozsoyogolu、Lamによる研究があり[1](今後これをYOLと呼ぶ)、動的計画法の手法により最適スケジュールを求めるアルゴリズムが提案されている。本論文では関係やその結合のサイズに関する性質を用いていくつかの最適性の必要条件を求め、１つにはこの問題の本質の理解を試み、１つには上記のアルゴリズムを更に簡単化する可能性を探りたい。

以下、第２章では問題の定式化を行い、第３章ではYOLが発現した最適条件を述べ、第４章では２属性のみの問題について新しい最適条件を求め、第５章ではこれをより多い属性に適用した結果を述べる。第６章では種種の議論を行っている。

2　問題の定式化

本論文では、分散データベースの木質問処理の全回線使用時間を最小とする最適化問題について考察する。質問に関連するすべての関係は異なるサイトにあり、これらの任意のサイトの間には回線によって直接結ばれているとする。また、一方のサイトから他方のサイトへある関係を転送するコストはその関係の全サイズの線形関数であるとする。また、各サイトで行われる結合等の処理コストは無視できるものとする。YOLが仮定したように結合の対象となる任意の２つの関係間には複数の共通属性はなく、結合はつねに一属性について行われるものとする。結合は等結合に限るとする。また、YOLにならって多くの関係中の重要な一関係を二の質問に関して準結合の系列によって簡約する(reduce)という質問に限って考察する。木質問について関係を準結合によって簡約することが可能であり、これが関数の数の程度の準結合で実現できることが知られており、またこれが質問の答を得るための最も重要なステップとなったいることも周知の事実である。

本論文では関係のサイズに関する考察が重要となる。関係中のタプルの構成については通常行われている次の統計的な仮定をとることとする。即ち

『ある関係のある列(column)の異なる値はその領域(domain)内で一様分布をとり、１つの関係内の異なる列の値も互いに独立であるとする。』

ある単一属性Aをもつ関係R(A)の選択度pを次のように定義しておく

$$p = |R(A)|/D_a \qquad (2.1)$$

ここで $|R(A)|$ は関係R(A)のタプル数(単にサイズということもある)。D_a は属性Aの領域の大きさ、即ち、属性のとりうる異なる値の総数をいう。

ある１つの関係を簡約するためのある準結合の系列をスケジュールと呼び、これをその関係を最終的な頂点とする有向木によって表現することができる。枝は関係の属性の転送を表す。同一の関係が異なる節点として複数回現れ得る。

例えば、$R_1(A)$、$R_2(A)$、$S(B)$、$U(A,B)$ の4つの関係を準結合して、U(A,B)を簡約するスケジュールの例を3つ2a)図とab)図と2c)図に示す。枝の上に必要に応じて属性名を表示する。

図 2a):

$$\begin{array}{l} R_1(A) \rightarrow R_2(A) \rightarrow U(A,B) \\ \qquad\qquad\quad S(B) \nearrow \end{array}$$

図 2b):

$$\begin{array}{l} R_1(A) \rightarrow U(A,B) \xrightarrow{A} R_2(A) \rightarrow U(A,B) \\ S(B) \nearrow \end{array}$$

図 2c):

$$\begin{array}{l} S(B) \rightarrow U(A,B) \xrightarrow{A} R_2(A) \rightarrow U(A,B) \\ \qquad\quad R_1(A) \nearrow \end{array}$$

3　潜在的に最適なスケジュール

YOL は前章で定式化した問題について、全回線使用コストの最適化という観点に立って、多くの効率の悪いスケジュールの除去を試み、除去されずに残っているスケジュールを潜在的に(potentially) 最適なスケジュールと称し、そのための条件を列挙している。例えば、前章の 2b) と 2c)図の2つのスケジュールを比べるとき中間にある U(A,B)の属性の A のサイズが 2b)の場合に比べて 2c)の場合の方が大きいことが明らかであるため、2c)は除去される。

YOL の条件の主なものを次に述べよう

(1) 前述のようにスケジュールは簡約する関係を最終頂点とする有向木となる。

(2) 上記の有向木のある節点で分岐が生じ、複数の入力枝をもつのは複数属性をもつ関係を表す節点の位置であり、その属性の数が s 個の時 s 個分岐があり、分岐の下の s 個の部分木を見るとき、同一の関係が異なる部分木に現れることはない。

(3) 単一属性をもつ関係は有向木の上のある一つの有向路(path)上に大きさの順に(小から大)ただ1回現れる。

(4) 複数属性をもつ関係の節点への入力枝が1本のとき、入力枝と出力枝の属性は一般に異なっている。ただし、有向木の葉から始まる分岐のない直線上の単一属性のみの転送から成る路があるとき、その路の上に複数属性をもつ関係がその単一属性に射影され、入力、出力枝がその同一の属性となることがありうる。

上記の条件(3)は次章で述べる補助定理(Hevner-Yao)によるものである。

前節で述べた最後の例、$R_1(A)$、$R_2(A)$、$S(B)$、$U(A,B)$ について、上記の条件を満たすスケジュールを列挙すると 2a)図、2b)図の他につぎの3種となる。ただし、$|R_1(A)| < |R_2(A)|$ とする。

図 3a):

$$\xrightarrow{(U)} S(B) \rightarrow U(A,B) \rightarrow R_1(A) \rightarrow R_2(A) \rightarrow U(A,B)$$

図 3b):

$$R_1(A) \xrightarrow{(U)} R_2(a) \rightarrow U(A,B) \xrightarrow{A} S(B) \rightarrow U(A,B)$$

図 3c):

$$\xrightarrow{(U)} R_1(A) \rightarrow U(A,B) \rightarrow S(B) \rightarrow U(A,B) \rightarrow R_2(A) \rightarrow U(A,B)$$

2a)は最終分岐スケジュール、2b)は中間分岐スケジュールと呼ぶことができよう。3a)、3b)、3c)は直線スケジュールである。$\xrightarrow{(U)}$ の記号は U を単一属性に射影した関係が(3)の意味で適切な位置に現れうることを示している。

次に U(A,B)、V(B,C)、R(A)の列、S(B)の列、T(C)の列を準結合により結合し、U(A,B)を簡約する潜在的に最適なスケジュールの典型的な例を 3d)、3e)、3f)に示す。R(A)、S(B)、T(C)の列と分割した列を図上では便宜上、同じ記号で書いてある。

3e)は一部直線スケジュールが現れる。3f)は全面的な直線スケジュールである。

次章では、上記の条件(3)に加えて、関係やその結合のサイズの大きさを考慮に入れた最適性の条件について述べることにする。

図 3d):

$$\begin{array}{lll} & & \xrightarrow{R(A)} \searrow \\ & \xrightarrow{S(B)} V(B,C) \xrightarrow{S(B)} & U(A,B) \\ \xrightarrow{T(C)} \nearrow & & \end{array}$$

図 3e):

$$\begin{array}{lll} & & \xrightarrow{R(A)} U(A,B) \xrightarrow{S(B)} U(A,B) \\ & V(B,C) \xrightarrow{S(B)} \nearrow & \\ \xrightarrow{T(C)} \nearrow & & \end{array}$$

図 3f):

$$\xrightarrow{R(A)} U(A,B) \xrightarrow{S(B)} V(B,C) \xrightarrow{T(C)} V(B,C) \xrightarrow{S(B)} U(A,B)$$

4　最適スケジュールの条件１

本章では属性が A、B の２つに限られる場合を考察する。これが次章で述べるより属性数の大きい場合の基本となる。

始めに Heven-Yao の定理を補助定理として述べる。m 個の単一属性 A を持つ関係 $R_j(A), j = 1, 2, ...m$ がそのサイズの順(小から大へ) に番号がついているとする。

【補助定理】(Hevner-Yao)
m 個の単一属性の関係 $R_j(A)(j = 1, 2, ...m)$を結合した結果を $R_j(A)$ のサイトで実現する最適スケジュールは次のいずれかである。

$$R_1 \rightarrow R_2 \rightarrow ...R_{j-1} \rightarrow R_j \rightarrow R_{j+1} \rightarrow ...R_m \rightarrow R_j \rightarrow$$
$$R_1 \rightarrow R_2 \rightarrow ...R_{j-1} \rightarrow R_{j+1} \rightarrow ...R_m \rightarrow R_j \rightarrow$$

□

この結果はかなり明らかであるから、証明は省略する。

これから後は属性 A をもつ単一属性の関係 $R_j(j = 1, 2, ..., m)$ と属性 B の単一属性の関係 $S_k(B)(k = 1, 2, ..., n)$ と A 、B を属性としてもつ１つの関係 $U(A, B)$の結合を $U(A, B)$ のサイトに実現する問題を考察する。R_j 、S_k はそれぞれサイズの大きさの順に番号がついているものとする。また、R_j、S_k の選択度を

p_j、q_k とし、$R_1 * R_2 * ... * R_j * U$ を B に射影した関係の選択度を q'_j、$S_1 * S_2 * ... * S_k * U$ を A に射影した関係の選択度を p'_k とする。明らかに次の関係が成り立つ。

$$p_1 \le p_2 \le \dots \le p_m$$
$$q_1 \le q_2 \le \dots \le q_n \tag{4.1}$$
$$p'_1 \ge p'_2 \ge \dots \ge p'_n$$
$$q'_1 \ge q'_2 \ge \dots \ge q'_m \tag{4.2}$$

次のような中間分岐スケジュールを考える。

$$\begin{array}{l} R_1 \to \dots \to R_j \to U \to R_{j+1} \to \dots \\ \qquad\quad S_k \nearrow \\ \qquad \dots \\ S_1 \nearrow \end{array} \tag{S1}$$

ただし、$j = 1, 2, ..., m-1; k = 1, 2, ..., n; m \ge 2; n \ge 1$ とする。

【定理 1】 中間分岐スケジュール(S1)が最適ならば、次の式が成り立つ。

$$p_j \le p'_k \le p_{j+1} \tag{4.3}$$

【証明】

(S1)のような中間分岐スケジュールでは、Uが R列のどこに位置つけるかに関係なく、$S_1 \to S_2 \to ... S_k \to U$ の転送コストが同じ、また、$(S_1 * S_2 * ... * S_k * U)(A)$の結合の結果も同一である。従って、上記の結合は$U'(A)$という属性 A だけを持つ関係とみなすことができ、即ち、属性 A だけを転送する質問に還元することができる。補助定理にとって、式(4.3)が成り立つ。

□

【注 1】 $j+1 = 1$ に相当する次のスケジュール

$$S_1 \to ... S_k \to U \to R_1 \to ...$$

が最適である時、$p'_k \le p_1$ は必ずしも成立しない(定理 3 参照)。

【注 2】 $j = m$ に相当する次のスケジュール

$$\begin{array}{l} R_1 \to \dots \to R_m \to U \to S_{k+1} \to \dots \\ \qquad\quad S_k \nearrow \\ \qquad \dots \\ S_1 \nearrow \end{array}$$

または

$$\begin{array}{l} R_1 \to \dots \to R_m \to U \to \\ S_1 \to \dots \to S_n \nearrow \end{array}$$

が最適であるとき、$p_m \le p'_k; p_m \le p'_n$ は必ずしも成立しない(定理 4 参照)。

次に次のような形の直線スケジュールを考える。

$$R_1 \to ... \overset{(U)}{\to} R_j \to U \to S_1 \to ... \to S_k \to U \to R_{j+1} \to ... \tag{S2}$$

ただし、$j = 1, 2, ..., m; k = 1, 2, ..., n$ とする。

【定理 2】 直線スケジュール(S2)が最適ならば、次式が成り立つ。

$$p'_k \le p_{j+1} \tag{4.4}$$

特に $j = m$ のとき(このときは必然的に $k = n$ となる)

$$p_m \le p'_k \tag{4.5}$$

また、特に $q'_1 = q'_2 = ... = q'_m$ のとき (4.4)は

$$p_j \le p'_k \le p_{j+1} \tag{4.6}$$

となる。

【証明】

$p_1 \le p_2 \le ... \le p_l \le p'_k \le p_{l+1} \le ... \le p_m$ とする、$U \to S_1 \to ... \to S_k \to U$ が R_l より前に挿入されないことについて証明する。$S_1 \to S_2 ... \to S_k$ を G と書くことにする。R_l の直後と直前に $U \to G \to U$ を挿入された時のコストを比較し、前者の方がよいことを説明する。

$$R_1 \to ...R_{l-1} \to R_l \xrightarrow{c1} U \xrightarrow{c2} G \xrightarrow{c3} U \to R_{l+1} \to ... \tag{1}$$

$$R_1 \to ...R_{l-1} \to U \xrightarrow{d1} G \xrightarrow{d2} U \xrightarrow{d3} R_l \to R_{l+1} \to ... \tag{2}$$

上記の両スケジュールにおいて、異なる転送コストは $c1, c2, c3$ と $d1, d2, d3$ 及び G 自身のコストである。明らかに

$$c2 < d1, c3 < d2$$

が成り立つ。同様な理由で前者の G の掛かったコストは後者より低い。また、

$$c1 = D_a p_1 p_2 ... p_{l-1} p_l$$

$$d3 = D_a p_1 p_2 ... p_{l-1} p'_k$$

仮定の $p_l \le p'_k$ より、$c1 \le d3$ 。

従って、スケジュール(1)より(2)が損する。式(4.4)が成り立つ。

これから $j = m$ の場合について証明する。この時$(S2)$は次のようになる。

$$R_1 \to R_2 \to ... \to R_m \to U \to S_1 \to ... \to S_n \to U \to$$

(4.4)の特例として、(4.5)式が成り立つ。

$q'_1 = q'_2 = ... = q'_m$ の場合には、Rをもって Sを簡約することができないため、$U \to G \to U$の掛かるコストは挿入位置によらず、同じである。補助定理によって式(4.6)が成り立つ。

□

【注 3】定理 2 はスケジュール(S2)の左端が R_1 から始まる必要はない。また、S の列も S_1から始まる必要はない。S の列が S_iから始まり、S_k で終るときは $S_i * S_{i+1} * ... * S_k * U$ を A の射影した関係の選択度を p'_{ik} とかくならば、(4.4)、(4.5)、(4.6)の p'_k を p'_{ik} で置き換えればよい。

【注 4】(4.4)式は S の列が R の列内のどこの位置に置かれるかの下限を与えている。上限についてもある近似の下で見当をつけることができる。(略)

次に再び直線スケジュール(S2)を考える。

(S2)を修正し、常に(S2)より小さくないコストをもつ次のスケジュール(S3)を考える。

$$U \to R_1 \to ...R_j \to U \to S_1 \to ...S_k \to U \to R_{j+1} \to ... \tag{S3}$$

ただし、$j = 1, 2, ..., m; k = 1, 2, ..., n$ とする。

【定理 3】スケジュール(S3)が(S2)を除いて最適ならば、次の式が成り立つ。

$$q'_j \le q_1 \tag{4.7}$$

【証明】

結論の式(4.7)が成り立たない、即ち $q'_j > q_1$ とすると、次のスケジュール

$$S_1 \to U \to R_1 \to ...R_j \to U \to S_2 \to ...S_k \to U \to R_{j+1} \to ...$$

はコストの計算により、$(S3)$よりコストが低い。これは$(S3)$が最適なスケジュールとの仮定に反することである。従って、(4.7)が成り立つ。

□

【注 5】定理 2、定理 3 は直線スケジュール内で属性が交替する位置に関する定理である。定理 2 は R 列内に中間に挿入された S の列の位置を定め、定理 3 は S の列の左に入る R_1より始まる R の列の大きさに関する条件である。定理 2 は直線スケジュールでの中間挿入位置 に関するもの、定理 3 は直線スケジュールの 左端の列の配置に関するものと言えよう。

【注 6】定理 3 においては、(S2)という現実的なものをそのまま使えず、(S3)という人工的のなものを考えざるを得なかった。問題の困難さを示すものといえよう。しかし、完全な最適性でなく、準最適な解を求めるという立場であれば、定理 3 の(S3)を(S2)に置き換えることも許されよう。

次に中間分岐スケジュールについての定理 1 で $j = m$ の場合、即ち注 1 で述べたスケジュール(S4)について考察する。

$$\begin{array}{lll} R_1 \to ... \to R_m \to U \to S_{k+1} \to ... & & \\ \qquad\qquad S_k \nearrow & & \\ \qquad\quad ... & & \\ \quad S_1 \nearrow & & \qquad\qquad\qquad \text{(S4)} \end{array}$$

ただし、$k = 1, 2, ..., n-1$ とする。

この場合にも(S4)よりコストが小さくない(人工的な) スケジュール(S5)を考える。

$$\begin{array}{lll} R_1 \to ... \to R_m \to U \xrightarrow{*} U \to S_{k+1} \to ... & & \\ \qquad\qquad S_k \nearrow & & \\ \qquad\quad ... & & \\ \quad S_1 \nearrow & & \qquad\qquad\qquad \text{(S5)} \end{array}$$

*印では U から U へ属性 B のみを送る無駄な転送を行う。

【定理 4】スケジュール(S5)が(S4)を除いて最適ならば、次の式が成り立つ。

$$p_m \leq p'_k \qquad (4.8)$$

【証明】

結論の式(4.8)が成り立たない、即ち $q_m > q'_k$ とすると、次のスケジュール

$$\begin{array}{l} R_1 \to ... \to R_{m-1} \to U \to R_m \to U \to S_{k+1} \to ... \\ \qquad\qquad S_k \nearrow \\ \qquad\quad ... \\ \quad S_1 \nearrow \end{array}$$

はコストの計算により、$(s5)$よりコストが低い。これは$(s5)$が最適なスケジュールとの仮定に反することである。従って、(4.8)が成り立つ。

□

【注 7】(S4)では、定理 1 を用いて $q_k \leq q'_m \leq q_{k+1}$ は成立つのは勿論である。

【注 8】下図の最終分岐スケジュール(S6)

$R_1 \to ... \to R_m \to U \to$

$$S_1 \to \ldots \to S_n \nearrow \qquad\qquad (S6)$$

に関して右端の U でさらに U に属性 A のみ、B のみを運ぶ現実的でないスケジュール(S7)、(S7′)を考え、(S6)以外に (S7)、(S7′)のいずれかよりコストの低いスケジュールが存在しないとき、$p_m \leq p'_n$ 、$q_n \leq q'_m$ が成り立つことが言える。このような性質は余り現実的ではないが、このようなスケジュールの最適性についてある種の示唆をあたえてくれた。

5 最適スケジュールの条件 2

前章で述べたは最適スケジュールの必要条件に関する定理を属性数のより多い場合に拡張することができる。ここては定理 1 と定理 2 を拡張し、定理 5 と定理 6 とし、これを次に示す。定理 5 と 6 の中にある T_{j-1}、T_j、T_{j+1} は質問の部分木を表している。同様に定理 3 と定理 4 も拡張できるが、ここでは述べないことにする。

【定理 5】

$$\begin{array}{cccc} R_{j-1} \xrightarrow{A} & R_j \xrightarrow{A} & R_{j+1} \xrightarrow{A} & \\ \wedge & \wedge & \wedge & \\ \mathrm{T}_{j-1} & \mathrm{T}_j & \mathrm{T}_{j+1} & \end{array}$$

スケジュールの中に図のように属性 A を含む 3 つの関係 R_{j-1}、R_j、R_{j+1} が続いて並び、それぞれの関係から属性 A の転送が行なわれているとする。R_{j-1}、R_j、R_{j+1} は属性 A の転送を含まない部分木 T_{j-1}、T_j、T_{j+1} の頂点となっていてよい。ただし、T_{j-1}が空でないときは R_{j-1} に他の関係から属性 A が転送されているとする。T_{j-1}、T_j、T_{j+1} 内の関係を結合し、属性 A に射影して得られる関係の選択度を p_{j-1}、p_j、p_{j+1}とする。このスケジュールが最適である時、次の式が成立つ。

$$p_{j-1} \leq p_j \leq p_{j+1}$$

【定理 6】

$$\begin{array}{ccc} R_{j-1} \xrightarrow{A} \mathrm{G} \xrightarrow{A} & R_j \xrightarrow{A} & \\ \wedge & \wedge & \\ \mathrm{T}_{j-1} & \mathrm{T}_j & \end{array}$$

スケジュールの中に図のように属性 A を含む 2 つの関係 $R_{j-1}(A)$、$R_j(A)$と属性 A の転送を含まない部分スケジュール G とが続いて並び属性 A の転送が行なわれているとする。R_{j-1}、R_j は属性 A の転送を含まない部分木 T_{j-1} 、T_j の頂点となっていてよい。ただし、T_{j-1} が空でないときは R_{j-1}に他の関係から属性 A の転送がされているものとする。T_{j-1}、T_j、G 内の関係を結合し、属性 A に射影して得られる関係の選択度を p_{j-1}、p_j、p とする。このスケジュールが最適である時、次の式が成立つ。

$$p \leq p_j$$

また、G 内での転送のコストがスケジュール内の G の位置によらない時は、次の式が成立つ。

$$p_{j-1} \leq p \leq p_j$$

6 議論

本論文ではYOLの得た条件を押し進めて、あるスケジュールが最適であるための必要条件について調べた。十分条件が見つけることができれば、これを利用して、さらに多くの無用なスケジュールを除去できるはずであるが、これがかなり困難なように思われる。

本論文で得られた定理1、2、5、6等に述べた必要条件によっても、かなり多くの無用なスケジュールを除去することが可能で、これをアルゴリズムに組み込むことが望ましい。これを最適アルゴリズムに利用しないとしても、これらの定理は最適スケジュールのもつ特性について重要な示唆を与えてくれる。もた、定理3、4や、第4章の注8に述べた内容は厳密な最適性を犠牲にするならば、効率のよいスケジュールの特性についての知見を与えるものである。非常に簡単な例だが、第2章で例示した $R_1(A), R_2(A), S(B), U(A, B)$ について 2a)図の最終分岐スケジュールが最適となる条件は注8を考察すれば、p' が p_1, p_2より十分大きく、$(p' \gg p_1, p_2)$ また同時に $q'_2 \gg q$ となることである。

第4章で扱った2属性に関する問題では、あらかじめ(4.1)、(4.2)の p'_k や q'_j 等を計算しておき、p'_k と p_k のおよび q'_j と q_j の大小関係を調べておくと、最適なスケジュールの可能性がかなり制限されることになり、極めて有用である。なお p'_k や q'_j を求めるのには、Bernstein 等の文献に述べられた Yao の公式を用いるのがよい。

本論文で述べた結果は、例えば Wong の SDD-1 に関して用いたような簡単な最適化法と組み合わせて用いることも有用であろう。即ち、簡単なアルゴリズムによって得られた結果が本論文で得られた条件を満たすかどうかを調べ、満たしていない場合には適切に修正するという方法である。

本論文においては、最適性の基準のコストを全回線使用時間(total time)としてきたが、応答時間(response time)の場合に対して、関係のサイズを考慮した最適条件を考察することが必要であると思われる。

文献

[1] C. T. Yu, Z. M. Ozsoyoglu, and K. Lam, "Optimization in Distributed Queries", J. Computer and System Sciences 29, pp.409-445,1984.

[2] A. R. Hevner and S. B. Yao,"Query processing in Distributed Database Systems", IEEE Trans. on Software Engineering,Vol. SE-5, No. 3, May 1979, pp. 177-187.

[3] P. M. G. Apers, A. R. Hevner and S. B. Yao,"Optimization Algorithms for Distributed Queries", IEEE Trans. on Software Engineering,Vol. SE-9, No. 1, January. 1983, pp. 57-68.

[4] P. A. Bernstein, N. Goodman, and E. Wong, et al., "Query Processing in a System for Distributed Dababase(SDD-1)", ACM Trans. on Database Systems,Vol. 6, No. 4, December.1981, pp.602-625.

[5] A. L. P. Chen and V. O. K. Li, "Improvement Algorithms for Semi-join Query Processing Programs in Distributed Database Systems", IEEE Trans. on Computers,Vol. C-33, No. 11, November. 1984, pp. 959-967.

[6] S. B. Yao, "Approximating Block Accesses in Database Oragnizations", Communication of the ACM, Vol. 20, No. 4, April 1977, pp. 260-261.

[7] W. Cellary, Z. Krolikowski, and T. Morzy, "Other Comments on 'Optimization Algorithms for Distributed Queries' ", IEEE Trans. on Software Engineering, Vol. 14, No. 4, April 1988. pp. 439-441.

[8] 李紅、佐藤洋、"分散データベースの質問処理最適化に関する考察"、データベース・システム研究会 64-6 (1988.3.15).

[9] 李紅、佐藤洋、"簡単な分散データベースの質問処理について"、情報処理学会第３６回全国大会 3E-10 (1988.3.17).