

RDBアクセス経路選択アルゴリズムの評価

水摩正行 日吉茂樹 渡辺正信
(日本電気(株) C&C システム研究所)

1. はじめに

関係データベース(RDB)システムの性能は、そのシステムに組み込まれるアクセス経路最適化機構(オプティマイザ)によるところが大きい。

著者らは、高性能なオプティマイザアルゴリズムの開発を目的として、種々の形態のシステムのアクセス経路のコストを評価するシミュレータ ROSE を開発し、アクセス経路の定量的評価を行ってきた。^[1]そして、その評価結果を反映し、実用システムとして十分適用できるオプティマイザアルゴリズムの開発を進めてきた。

オプティマイザアルゴリズムは、PARTV^[2]、SQUIRAL^[3]のように、条件式やアクセス機構に対し、先験的に優先度を与える発見的方式と、INGRES^[4]、SYSTEM R^[5]のようにシステムが維持する統計情報を用い、アクセス経路のコスト評価によって決定するコスト評価方式に大別される。しかしながら、その実現方式は様々で、システムの設計目標、対象規模、内部処理機構の形態等に依存する。

本稿では、索引機構とビットマップ処理機構を有するRDBシステムにおけるコスト評価に基づく一つの最適化方式を提案し、その妥当性をROSE上で評価した結果について述べる。

コスト評価方式は、与えられた問合せに対し、可能な全てのアクセス経路のコスト評価により、モデル上での真の最適解を得る方式と、ヒューリスティックなルールを一部導入し、可能な限り真の最適解に近づく準最適化方式に分けられる。本方式はその後者に位置付けられる。

2. システムモデル

本稿で議論するRDBシステムのモデルを示す。本システムの特徴は、複数の索引を利用する処理に対し、ビットマップ操作を基本としていることである。即ち、索引の有効利用をはかため、制限/結合演算処理に対し複数の索引を利用し、索引上で得られたタプル識別子(Tid)の集合をビットマップ形式で表現し、その上での集合演算により対象タプルを限定する。ここで言うビットマップとは個々のタプルに与えられるTid番号の順に一つのビットを割当てたものである。

Tid集合をビットマップ形式で表現する利点として、論理演算が行い易いこと、情報の圧縮効果により主メモリ内のめみでの操作の可能性が強いこと、Tid番号によるソーティングが自然に出来上がり、タプルアクセス時におけるバッファリング効果が最大限に生かせること等が挙げられる。

2.1 記憶構造とタプルアクセス機構

リレーションを構成するタプル集合には各々その識別子(Tid番号)が割当てられ、そのTid番号に従って、順次論理記憶空間の基本単位であるページに格納される。索引はB*-treeで構成され、キー値とTid番号を対として保存される。また、タプルへのアクセス機構として次の方式がある。

(1) リレーションスキャン: Tid番号に従ってタプルを連続的にアクセスし、目的タプルを得る。

(2) 索引を介したアクセス: マトリキュート値をキーとして、索引によりTid番号を得、目的タプルをアクセス

する。

(3) ビットマップアクセス: ビットマップ上に表現された Tsid 集合に従って、順次タプルをアクセスする。

2.2 処理方式

2.2.1 制限処理方式

制限処理方式には、次のものがある。

(1) リレーション走査 (SCN)

物理ファイル上のリレーションを走査して、制限条件をチェックする。

(2) 制限索引利用方式 (RNX)

制限条件の付されたアトリビュートに対する索引 (制限索引) により、条件を満たすタプルの Tsid を得、タプルアクセス時に残りの制限条件のチェックを行う。

(3) ビットマップ処理 (BMP)

複数の制限索引により得られた結果の Tsid 集合をビットマップ上に反映し、AND 条件や OR 条件をビットマップ間の集合演算により実行する。

2.2.2 結合処理方式

一つのリレーションにおける制限、結合演算の順序関係に注目した場合、次の方式に分けられる。

(1) 結合-制限方式 (JIX)

結合条件に付されたアトリビュートに対する索引 (結合索引) を利用して結合条件を満たす Tsid を求め、その Tsid を使ってタプルをアクセスし、タプル上で制限条件の検査を行う。

(2) 制限-結合方式 (SMJ)

制限索引または、ベースリレーション走査により制限処理を行った後で、結合キーに順してソートし、結合処理を行う。この場合、前述の制限処理方式の違いにより次の3つの方式に分けられる。

① SMJ1 (SCN + SORT)

② SMJ2 (RNX + SORT)

③ SMJ3 (BMP + SORT)

(3) 制限結合並列方式 (IJR)

制限と結合演算をそれぞれを索引上で行い、ビットマップ上でそれらの結果を突合わせ、制限と結合条件の両方を満たすタプルのみをアクセスする。

なお、結合方式については、結合キーに順するソートマージ方式を採用している。他の方式としてはネストループ方式^[6]なども考えられるが、ROSE を用いた評価により、ソートマージ方式がほとんどの場合に優れていることが確認された。

2.2.3 複数の結合処理方式

n 変数向合せ ($n > 3$) の処理は、基本的には2変数向合せの積上げで行われる。積上げの際の中向結果は、ベースリレーションのアクセス回数を最小限に留めるため、最終結果に含まれるアトリビュート、および以降の処理で必要とされるアトリビュートの値を中向ワークとして保持する。

3. アクセス経路の選択方式 (OPT)

2章のシステムを対象としたオペティマイザのアクセス経路選択アルゴリズム OPT について説明する。

本オペティマイザは、結合順序の決定に際しては発見的手法を採用し、個々のリレーションに対する制限/結合処理の決定には、簡略化した計算式によるコスト評価方式をベースとしている。

ここで使用する変数の説明を表1に示す。

3.1 結合演算の順序決定

n 変数向合せにおける可能な結合順の数は次数 $O(n!)$ であり、可能な全ての中間結合結果リレーションをロードとする結合トリーを全探索して、最適結合順を得るには多大な計算オーバーヘッドを要する。なお、ここでは、ソートマージ型結合方式に限定している

ので、中間結果リレーションのスペース量の和を最小とする結合順序が最適となる。この結合順序を決定するために、次のような発見的な手法を採用する。

まず、 α_1 の結合は結合結果リレーションのスペース量が最小となるものを選ぶ。ここに、結合の対象リレーションを R_i, R_j としたとき、結合結果リレーションのスペース量 V_k は、

$$V_k = N_i d_i N_j d_j T_{ij} e_{ij}$$

で与えられる。また、 α_2 番目の結合は、 α_1 結合により得られた中間結果リレーションも含め、次に可能な結合の中から結合結果リレーションのスペース量が最小となるものを選択する。この操作を順次繰返す。

3.2 制限/結合方式の決定

制限演算と結合演算の処理方式の決定は、I/Oコスト計算に基づいて行われ、但し、コスト計算は制限選択度 d_i , β_i およびセミ結合選択度 T_i に基づいて一次近似的による簡略化をはかる。以下にコスト式を示す。

なお、ROSEにおけるコスト計算式については文献[1]を参照せよ。

(1) 単位処理とそのコスト

$$R(d_i, \alpha_i) : \text{制限索引アクセス} \\ = h_i + N_i d_i \alpha_i / p$$

$$A(n) : \text{タプルアクセス} \\ = \begin{cases} N_p & \text{ベ-スリレーション走査} \\ \min\{n, N_p\} & \text{ビットマップを介してアクセス} \end{cases}$$

但し、 N_p はベ-スリレーションベ-ツ数
 n は対象タプル数

$$B(n) : \text{ビットマップアクセス} \\ = \min\left\{\left\lceil \frac{n \cdot \gamma_b}{p} \right\rceil, N_b\right\}$$

但し、 N_b は全ビットON時のビットマップベ-ツ数

$$W(n, w_i) : \text{ワークデータアクセス} \\ = n \cdot w_i / p$$

表1. 変数定義

変数名	意味
N_i	タプル数
γ_b	タプルサイズ
α_i	索引エントリサイズ
p	ベ-ツサイズ
e_i	最終結果エントリサイズ
d_i	最小の制限選択度
δ_i	d_i 以外の制限選択度
β_i	d_i 以外の制限選択度の積 ($= \prod \delta_i$)
T_i	セミ結合選択度
γ_b	ビットマップレコードサイズ
h_i	索引トリ-レベル数
T_{ij}	結合選択度
w_i	中間結果エントリ-サイズ
v_i	異なるキー値の数

$$S(n, w_i) : \text{ソート}$$

$$= 2n \cdot w_i / p$$

$$J_1(\alpha_i) : \text{結合索引を利用した結合演算} \\ = N_i \alpha_i \delta_i / p$$

但し、 δ_i はソートマ-ツ時のアクセス効率

$$J_2(n, w_i) : \text{タプルアクセス後の結合演算} \\ = n_i w_i \delta_i / p$$

(2) 制限と結合の処理コスト

$$C_{SCN} : \text{ベ-スリレーション走査} \\ = A(N_i)$$

$$C_{RNX} : \text{制限索引利用} \\ = R(d_i, \alpha_i) + 2B(n_i) + A(n_i)$$

但し、 $n_i = N_i d_i$

$$C_{BMP} : \text{ビットマップ処理} \\ = \sum_j \{R(d_{ij}, \alpha_{ij}) + 2B(n_{ij})\} \\ + 2B(n_i) + A(n_i)$$

但し、 $n_{ij} = N_i d_{ij}$, $n_i = N_i \prod_j d_{ij}$

$$C_{JEX} : \text{結合-制限方式}$$

$$= J_1(\alpha_i) + 2B(n_i) + A(n_i) \\ + W(n_i, w) + S(n_i, w) + J_2(n_i, w)$$

但し、 $n_1 = N_i T_i$, $n_2 = N_i T_i \prod_j d_{ij}$

$$C_{SMT} : \text{制限-結合方式}$$

$$= C_{RES} + W(n_i, w) + S(n_i, w) + J_2(n_i, w)$$

但し C_{RES} は C_{SCN} , C_{RNX} 又は C_{BMP}

C_{ITR}: 制限結合並列方式

$$= \sum_j \{R(d_{ij}, x_{ij}) + 2B(n_{ij})\} + 2J(x_2) + 2B(n_2) + A(n_3) + W(n_4, w) + S(n_4, w) + I_2(n_4, w)$$

但し、 $n_{ij} = N_i d_{ij}$, $n_2 = N_i T_i$
 $n_3 = N_i T_i \prod_j d_{ij}$, $n_4 = N_i T_i d_i$

3.3 利用制限索引の選択

制限処理方式としてビットマップ処理(BMP)方式を考慮したとき、利用可能な制限索引のうちどれを使用するか決定する必要がある。その決定は次のように行う。利用可能な制限索引を制限選択度 d_i の値の小さい順に並べたものを X_1, X_2, \dots, X_k とする。このとき当然 X_1 から順に採用してゆくが、 k 番目の索引 X_k の採用基準は

$$d_{ik} < 1 - \frac{x_i + h_i p / N_i}{x_i + p \prod_{m=1}^{k-1} d_{im}} \quad \dots (1)$$

で与えられる。なお

$$\prod_{m=1}^{k-1} d_{im} \leq \frac{h_i}{N_i} \quad \dots (2)$$

なる場合には、(1)式の右辺 ≤ 0 となるため d_{ik} の値にかかわらず索引 X_k を不採用とする。(1)式の採用基準は、BMP方式のコスト式から導かれるものであるが、制限選択度に基づく簡明な索引選択を可能にする。

以上述べたように、OPTアルゴリズムでは、結合順決定、制限/結合処理方式決定、および利用制限索引の選択の3つの最適化要素を独立化し、階層的に評価することにより効率的最適化を図っている。

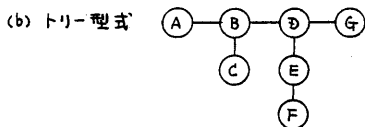
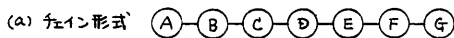


図1. 結合形式タイプ

4. アクセス経路選択方式の評価

前章で述べたアクセス経路選択方式(OPTアルゴリズム)の妥当性をシミュレータROSEを使って評価する。

4.1 結合順決定アルゴリズムの評価

結合トリーの全探索を基本とするF*アルゴリズムと発見的手法に基づくH*1およびH*2のアルゴリズムを設定し、シミュレーションにより前記OPTアルゴリズムと比較評価を行った。以下に各方式の説明と評価結果を示す。

4.1.1 比較対象アルゴリズム

(1) F*アルゴリズム

結合数 $=n$ のとき、全ての順列数は $n!$ 通りある。それらの結合順の必要総ワークスペース量を計算し、その値が最小の結合順を得る。

(2) H*1アルゴリズム

オ1段の中間結果リレーションのスペース量 V_k の小さい結合から順に実行する。この場合、トリーの構成により独立した結合が出来上がるため、複数の中間結合結果リレーションを同時に保持する必要性が生じる。

(3) H*2アルゴリズム

まずオ1の結合は結合結果リレーションのスペース量 V_k が最小となるものを選出する。オ2番目以降の結合は、隣接する結合のうち結果リレーションのスペース量の小さい方を選出結合する。これを順次繰返す。この方式は、ベースリレーションを一つづつ付加してゆく積上げ型の結合形態である。

4.1.2 シミュレーションの環境条件

(1) 結合形式は図1に示す2つのタイプを対象とした。

(2) リレーションの数 $=3 \sim 8$ 個

(即ち、結合数 $=2 \sim 7$ 個)

(3) タプル数 $N_i = 10000$

(4) 結合選択度 \bar{d}_{ij} と制限選択度 d_i の値はともに平均0.01の指数分布に

後から。

図2は結合数2~7について、F*, H*1, H*2, OPTの各アルゴリズムを1000回づつ試行したときの平均コスト(中間結果総テーブル件数)を示している。但し、結合形式は図1(a)のチェーン形である。また、AVGは結合順をランダムに結合した場合のコストである。図2のようにOPTアルゴリズムはH*1, H*2より優れた特性を示す。改善度は、

$$\frac{C_{AVG} - C_{H*}}{C_{AVG} - C_{MIN}} \times 100$$

にて表わしている。

また、計算オーバーヘッドはF*アルゴリズムが次数 $O(n!)$ であるのに対し、OPT, H*1, H*2の各アルゴリズムは $O(n)$ である。

一方、結合形式が図1(b)のトリーの場合のコスト特性を表2に示す。

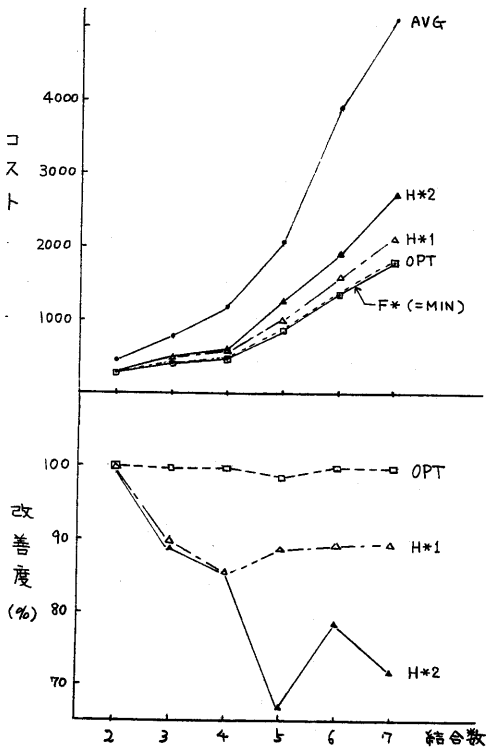


図2. 結合順序決定アルゴリズムの評価

この場合にもOPTアルゴリズムは良好な解を与える。

表2. トリー形式結合の評価結果

アルゴリズム	平均コスト(C_{H*})	コスト比($\frac{C_{H*}}{C_{MIN}}$)	改善度
OPT	1448	1.02	99.2%
H*1	1726	1.15	93.8
H*2	1625	1.22	90.8
MIN	1418	—	—
AVG	3804	3.40	—

4.2 制限/結合処理方式の評価

制限, 結合処理方式の決定に際してOPTアルゴリズムは線形近似により簡略化したコスト計算式を採用しているが、ここでは、その近似精度の観点からROSEのモデルと比較する。標準的な評価環境パラメータ値を表3に示す。

4.2.1 制限処理方式の決定基準の比較

4個の制限条件をそれぞれ変数向合せを例にOPTとROSEの制限処理方式の選択判定の差異を調べる。向合せ中の4個の条件式は選択度 d の条件が1つと選択度 γ ($>d$)の条件式が3つから成るAND結合のものとする。

図3は $N_i = 10^5$, $\gamma = 0.2$ におけるコスト評価値の比較を示している。破線はOPT, 実線はROSEの評価値を示す。

OPTとROSEの各アクセス経路のコスト推定値を比較した場合、RNX方式の d 値0.02~0.03付近で差がみられ、他はほぼ等しい。しかしながら、最適解を得るとする観点からみると、上記のコスト推定値の誤差が直接影響

表3. 標準パラメータ設定値

N_i	10^5	γ	128
γ_i	200	ν_i	10^4
α_i	10	T	400
p	4K	M	5
分布	一様分布		

注)
T: n-1内索
引エントリ数
M: マシオンク

することはなく、OPTは α 値0.0185を境に、より小さい所ではRNX, 大きい所ではBMP方式と選択、その解はROSEの結果とほぼ一致する。なお、OPTがROSEと異なった方式と選択区域はOPTとROSEが算定する各トロードオフ点により挟まれる部分で、図中互で示されている。記号 $\Phi[A, B]$ は、アクセス経路Aと選択すべきところをOPTがアクセス経路Bと選択する誤判定区域を示す。表4にタプル件数 N と γ が異なる場合の誤判定区域 Φ の詳細値を示している。表中 α は γ に対応する制

限選度 α の区間、 Γ は最悪の場合の処理損失コスト量を示している。 Γ の分子はOPTが誤った経路と選択した時の最適経路に対するコスト差の最大値で、分子は最適経路のコストを示している。

4.2.2 結合処理方式の選択基準の比較

2変数向合せを例に、OPTとROSEのアクセス経路選択判定基準の比較を行う。この向合せでは、セミ結合選度 τ の結合条件が1個と各リレーションに対し、制限選度 α の制限条件が1個、 γ の制限条件が3個与えられている。

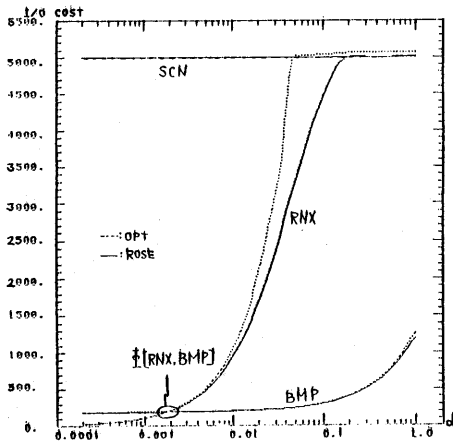


図3.1変数向合せにおけるアクセス経路のコスト評価値の比較 ($\gamma=0.2$)

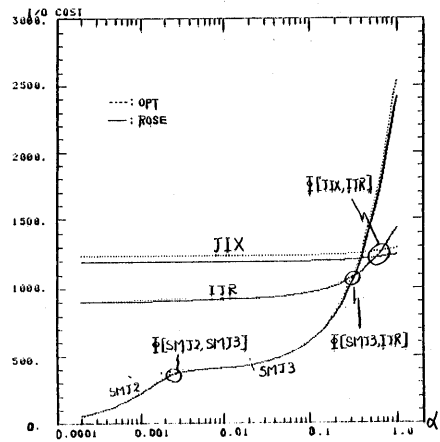


図4.2変数向合せにおけるアクセス経路のコスト評価値の比較 ($\gamma=0.2$)

表4.1変数向合せにおけるOPTのアクセス経路誤判定領域

条件			$\Phi[RNX, BMP]$	$\Phi[BMP, SCN]$
N	γ			
10^5	0.2	α	$[1.85 \times 10^3, 1.88 \times 10^3]$	—
		Γ	3/191	—
	0.5	α	$[4.68 \times 10^3, 4.90 \times 10^3]$	$[39 \times 10^1, 79 \times 10^1]$
		Γ	45/459	1318/3706
5×10^3	0.2	α	$[4.6 \times 10^3, 4.6 \times 10^3]$	—
		Γ	0/29	—
	0.5	α	$[7.9 \times 10^3, 8.0 \times 10^3]$	$[33 \times 10^1, 66 \times 10^1]$
		Γ	1/42	68/182

$$\Gamma = (\text{最大損失コスト}) / (\text{最適経路コスト})$$

表5.2変数向合せにおけるOPTのアクセス経路誤判定領域

条件			$\Phi[SM12, SM13]$	$\Phi[SM13, IJR]$	$\Phi[JIX, IJR]$
γ	τ				
0.2	0.009	α	$[2.85 \times 10^3, 2.88 \times 10^3]$	$[3.23 \times 10^1, 3.33 \times 10^1]$	$[5.93 \times 10^1, 6.77 \times 10^1]$
		Γ	2/259	16/1067	41/1229
	0.5	α	$[1.85 \times 10^3, 2.88 \times 10^3]$	$[6.46 \times 10^1, 6.96 \times 10^1]$	—
		Γ	2/260	70/1640	—
0.5	0.008	α	$[4.68 \times 10^3, 4.90 \times 10^3]$	$[2.07 \times 10^2, 2.14 \times 10^2]$	$[1.64 \times 10^1, 1.0]$
		Γ	41/910	19/1364	232/1844
	0.5	α	$[4.68 \times 10^3, 4.90 \times 10^3]$	$[4.13 \times 10^2, 4.46 \times 10^2]$	—
		Γ	41/912	173/1842	—

$$\Gamma = (\text{最大損失コスト}) / (\text{最適経路コスト})$$

図4は、 $N=10^5$
 $\gamma=0.2$, $\tau=0.007$
 に設定した場合
 のOPTとROSE
 の各アクセス経
 路のコスト評価
 値と比較したも
 のである。

図に示される
 ように、コスト
 評価値には大き
 な差はみられず
 OPTはよい近
 似を与えている
 と云える。また
 ROSEの評価値
 から最適アクセ

ス経路は α 値の増大に従って、SMJ2
 SMJ3, IJR, JIXと変動しているこ
 とが判るが、OPTも同様の結果を示
 しており、トレードオフ点の誤差で表
 われる各誤判定区域互も狭い。

図5は、ROSEとOPTの各々にお
 いて、各アクセス経路が最適経路とし
 て選択される領域を (τ, α) 平面上で
 表わしたものである。図中、実線は、
 ROSE、破線はOPTの評価領域を示す。
 また、斜線で示される領域互はOPT
 のアクセス経路誤判定領域を示す。残
 りの領域は真の解を得る領域である。
 図(a)は $\gamma=0.2$, (b)は $\gamma=0.5$ に設
 定した場合を示している。また、表5は、
 図5上の特徴的な性質を示している τ
 の値を選り、OPTの誤判定領域互の
 詳細な値を調べた結果を示している。

図5に示されるように、OPTの誤
 判定領域互は非常に狭い。これは、本
 条件下の向合せに対して、OPTが真
 のアクセス経路を選択する確率が高い
 ことを示す。また、表5に示されるよ
 うに、たとえ誤ったアクセス経路を選
 びたとしても、その損失コストは最適
 解の処理コストに比べ、小さくOPT

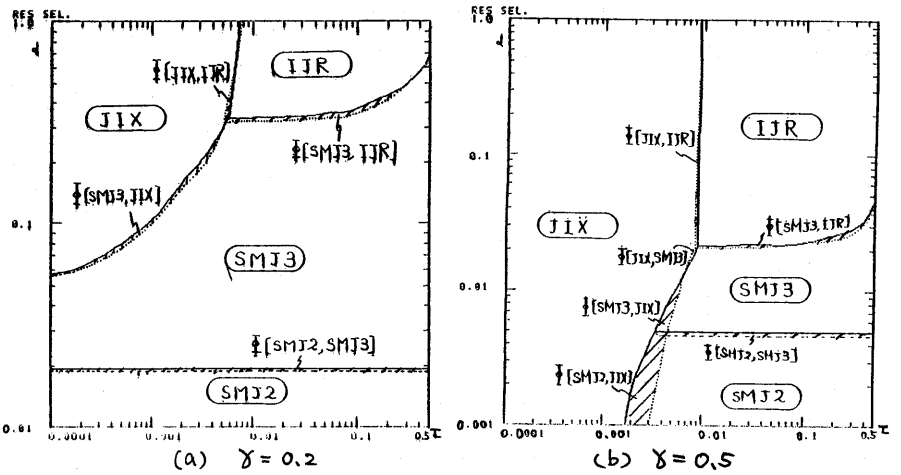


図5. 2変数向合せに対する (τ, α) 平面上での
 アクセス経路選択基準の比較
 (実線: ROSE 破線: OPT 斜: 誤判定領域)

は準最適の経路と選りおめと云える。

4.3 制限索引の選択

4つの制限条件を1つ1つ変数向合せ
 における利用制限索引の選択について
 評価した結果を図6に示す。1つの制
 限条件 C_1 の制限選択度を α , 残りの制
 限条件 C_2, C_3, C_4 の選択度を $\gamma=0.2$ とする。
 そして、 $C_1 \sim C_4$ に対応する制限索引を
 $X_1 \sim X_4$ とし、

- BM_1 : X_1 のみ利用, BM_2 : X_1, X_2 利用
 - BM_3 : X_1, X_2, X_3 利用, BM_4 : X_1, X_2, X_3, X_4 利用
- とする。

図6は $N=10^5$ におけるROSEとOPT
 のコスト評価値を示している。また表
 6はOPTで選択したアクセス経路と
 最適アクセス経路でのコスト差 (R) を示
 している。

利用制限索引の組合せ互によりかは
 りのコスト差が生じるが、OPTは互
 の中から良好なコストを与える制限索
 引の組合せを選択する。

なお、ビットマップ方式において全
 ての索引を利用する方式は BM_4 , 制限
 索引利用方式は BM_1 に対応する。OPT
 での制限索引カット手法により、 BM_1

BM4のみでなく、BM2, BM3も選択される。その効果が図のハッチング部に現われている。

5. おわりに

索引機構とビットマップ処理をベースとしたRDBシステムにおけるアクセス経路最適化アルゴリズムを提案した。本方式は、コスト評価をベースとしているが、現実性を考慮し、準最適化方式を採用している。そして、最適化要素と、結合順の決定、制限結合演算処理方式の決定、利用制限索引の決定の3つの独立した項目に分け、階層的評価をすることにより効率化を図っている。また、本提案アルゴリズムをシミュレータROSEを用いて評価した。その結果、かなり良好な結果が得られ、実用システムに十分適用できることが確認された。現在、我々は本方式に従ったオプティマイザを実システムに組み込み中である。

最後に、本研究の機会を与えて頂いた基本ソフトウェア開発本部原部長、並びに、C&Cシステム研究所三上部長、橋崎課長に、有意義な助言を頂いた基本ソフトウェア開発本部宮崎主任

表6. OPTアルゴリズムでの制限索引カットの精度

	Φ [BM1, BM2]	Φ [BM2, BM3]	Φ [BM3, BM4]
α の区間	$[7.9 \times 10^{-4}, 8.5 \times 10^{-3}]$	$[3.5 \times 10^{-3}, 3.6 \times 10^{-2}]$	$[1.7 \times 10^{-2}, 1.8 \times 10^{-1}]$
コスト差	$\frac{5}{94}$	$\frac{3}{152}$	$\frac{5}{213}$

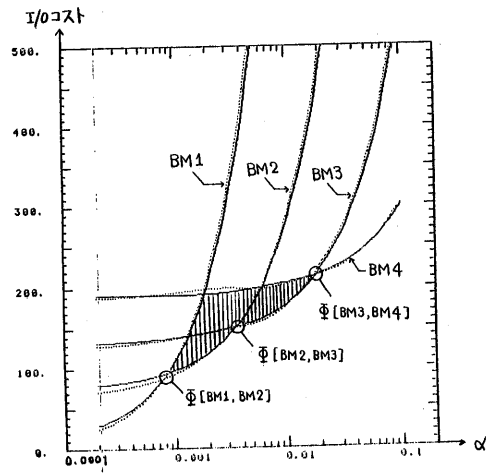


図6. BMP方式における制限索引カット手法の効果 (---: OPT, - ROSE)

C&Cシステム研究所牧野主任に、また、本評価に協力頂いた日本システムアプリケーション(株)小林氏に深く感謝の意を表します。

<参考文献>

- [1] 水廣, 日吉, 渡辺, 「関係データベースのアクセス経路の評価」情報処理学会DBMS研究会26-1 (1981.9)
- [2] S.J.P.Todd, "The Peterlee Relational Test Vehicle - a system overview", IBM Syst. J., No.4, 1976.
- [3] J.M.Smith, et al., "Optimizing the Performance of a Relational Algebra Database Interface", CACM, Vol.18, No.10, 1975.
- [4] M.Stonebraker, et al., "The Design and Implementation of INGRES", ACM TOD, No.3, 1976.
- [5] P.G.Selinger, et al., "Access Path Selection in a Relational Database Management System", Proc. ACM SIGMOD, 1979.
- [6] S.B.Yao, "Optimization of Query Evaluation Algorithms" ACM TOD, Vol.4, No.2, 1979.