

具体的な講義の項目

- イン트로ダクション
- データベースのモデル
- 関係データベースの操作(関係代数、関係論理)
- 正規化
- 問い合わせ言語
- トランザクション処理
- データベースの内部構造
- データベースの問い合わせ処理
- その他(オブジェクト指向DB、XML/RDF-DB等々)

2020/5/7

データベース (©横田)

37

関係の操作

- 関係代数と関係論理
- いずれも E. F. Codd によって提案
- 関係代数(Relational Algebra)
 - 関係に対する操作の手続き的記述
 - 関係を入力とし、関係を出力とする演算
 - 集合演算と専用演算
- 関係論理(Relational Calculus)
 - 出力としてほしい関係の仕様の論理式による記述

2020/5/7

データベース (©横田)

38

関係代数演算の種類

- 集合演算(Set Operations)
 - 和、積、差、直積
- 専用演算
 - 射影、選択、制約、結合、割り算
- 別の分類
 - 単項演算
 - 射影、選択、制約
 - 二項演算
 - 和、積、差、直積、結合、割り算

2020/5/7

データベース (©横田)

39

和両立(Union Compatible)条件

- 二項演算の二入力について
 - 属性数と対応する属性の定義域が一致
- 条件成立が必要な演算
 - 集合和、集合積、集合差
- 条件成立が必要でない二項演算
 - 直積、結合、割り算

2020/5/7

データベース (©横田)

40

和両立の関係の例

秋葉原支店関係 R_a

商品番号	販売価格
GDC10	150,000
GLC20	90,000
FLC33	70,000
HIP06	30,000
TLCD70	40,000

日本橋支店関係 R_b

商品番号	販売価格
GDC10	140,000
GLC20	90,000
FLC33	70,000
FLP44	50,000
HLP05	60,000
TLCD70	40,000

2020/5/7

データベース (©横田)

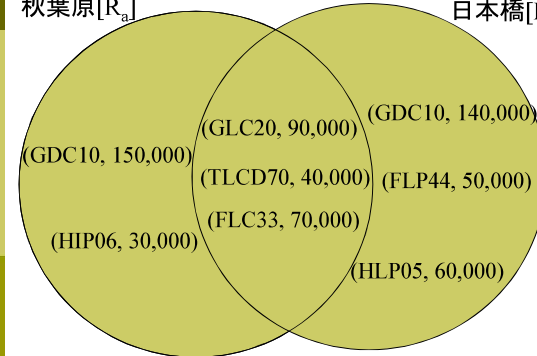
41

集合和 (Union) 演算 $R_a \cup R_b$

秋葉原 R_a

日本橋 R_b

二支店全商品関係



商品番号	販売価格
GDC10	140,000
GDC10	150,000
GLC20	90,000
FLC33	70,000
FLP44	50,000
HLP05	60,000
HIP06	30,000
TLCD70	40,000

- 注: 重複の除去が必要 Cardinality $5 + 6 \neq 8$

2020/5/7

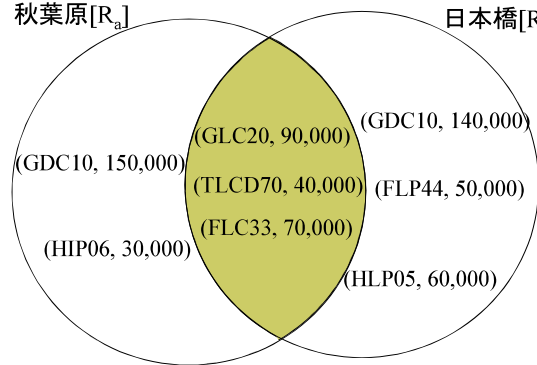
データベース (©横田)

42

集合積 (Intersection) 演算 $R_a \cap R_b$

秋葉原 R_a

日本橋 R_b



二支店共通商品関係

商品番号	販売価格
GLC20	90,000
FLC33	70,000
TLCD70	40,000

2020/5/7

データベース (©横田)

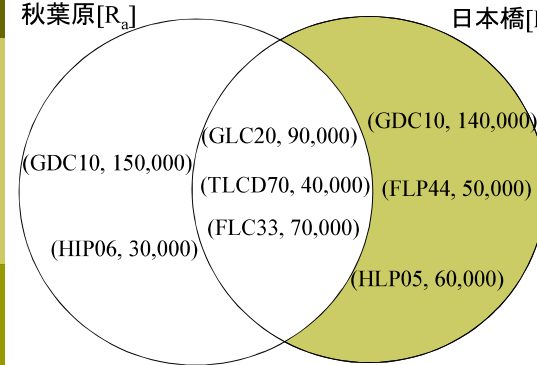
43

集合差 (Difference) 演算 $R_b - R_a$

秋葉原 R_a

日本橋 R_b

日本橋支店のみ商品関係



商品番号	販売価格
GDC10	140,000
FLP44	50,000
HLP05	60,000

$R_a - R_b$ なら
秋葉原支店のみ商品関係

商品番号	販売価格
GDC10	150,000
HIP06	30,000

2020/5/7

データベース (©横田)

44

直積 (Cartesian Product) $R_x \times R_y$

製造元国籍 [R_x]

製造元	国籍
GELL	米国
Fepson	日本
Hanon	日本

販売担当 [R_y]

販売員	担当地域
田中	千代田区
鈴木	新宿区
佐藤	渋谷区

製造元	国籍	販売員	担当地域
GELL	米国	田中	千代田区
GELL	米国	鈴木	新宿区
GELL	米国	佐藤	渋谷区
Fepson	日本	田中	千代田区
Fepson	日本	鈴木	新宿区
Fepson	日本	佐藤	渋谷区
Hanon	日本	田中	千代田区
Hanon	日本	鈴木	新宿区
Hanon	日本	佐藤	渋谷区

$$\text{cardinality}(R_x \times R_y) = \text{cardinality}(R_x) \times \text{cardinality}(R_y)$$

ちょっと考えてみよう(2-1)

□ 和両立の必要性

- いずれも二項演算なのに、集合和、集合積、集合差で和両立が必要で、直積で和両立が不要なのはなぜか考えてみよう

射影 (Projection): π

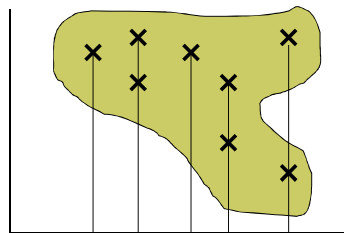
□ 指定した属性 ($X_1, X_2, \dots, X_n$) の切り出し

■ $\pi_{X_1, X_2, \dots, X_n}(R)$

- 例: $\pi_{\text{商品名}}(\text{製造元関係})$

商品名
Desktop
Laptop
Laser Pt.
Inkjet Pt.
LCDsplay

製造元



商品名

- 重複除去しない射影はデルタ射影 (Delta Projection) と呼ばれる

- 詳細は後半で説明

選択 (Selection): σ

□ R内の条件 ($X=a$) に合うタプルの検索

■ $\sigma_{X=a}(R)$

- 例: $\sigma_{\text{商品名}=\text{Laptop}}(\text{製造元関係})$

商品番号	商品名	製造元
GLC20	Laptop	GELL
FLC33	Laptop	Fepson

□ θ 選択 (θ -Selelction): $\sigma_{X\theta a}(R)$

- 条件を一般化
- $\theta: =, \neq, <, \leq, >, \geq$
- 例: $\sigma_{\text{在庫} < 4}(\text{秋葉原支店関係})$

商品番号	販売価格	在庫
FLC33	70,000	3
TLCD70	40,000	2

□ 制約 (Restriction): $\sigma_{Xa\theta Xb}(R)$

- 属性間の条件で検索

□ 条件を $\wedge$ (and) や $\vee$ (or) で結合することも可能

関係代数の組み合わせの例

- 自然言語による問い合わせ
 - 秋葉原支店で、販売価格が100,000円以下で、在庫が10以上である商品の商品番号と販売価格と在庫のデータが欲しい場合、関係代数で表現せよ。

書き方の例

- 秋葉原支店で、販売価格が100,000円以下で、在庫が10以上である商品の商品番号と販売価格と在庫のデータが欲しい場合、関係代数で表現せよ。

秋葉原支店で販売価格が100,000円以下で在庫が10以上の選択演算

$R1 = \sigma_{\text{販売価格} \leq 100,000 \wedge \text{在庫} \geq 10}(\text{秋葉原支店})$
全部の属性を出力したい場合(問題の場合)、このまま
商品番号だけ欲しい場合には、射影演算を行い

$R2 = \pi_{\text{商品番号}}(R1)$
一つの式にまとめて

$R2 = \pi_{\text{商品番号}}(\sigma_{\text{販売価格} \leq 100,000 \wedge \text{在庫} \geq 10}(\text{秋葉原支店}))$

キャラクタ入力での書き方

- キャラクタ入力での書き方(ローカル)
 - 集合和: u(関係名1, 関係名2)
 - 集合積: i(関係名1, 関係名2)
 - 集合差: -(関係名1, 関係名2)
 - 直積: x(関係名1, 関係名2)
 - 射影: p[属性名](関係名)
 - 選択: s[条件](関係名)
 - 条件: 属性1 >= 属性2
 - 論理演算: and / or / not

キャラクタ入力での書き方

- 秋葉原支店で、販売価格が100,000円以下で、在庫が10以上である商品の商品番号と販売価格と在庫のデータが欲しい場合、関係代数で表現せよ。

秋葉原支店で販売価格が100,000円以下で在庫が10以上の選択演算

$R1 = s[\text{販売価格} \leq 100,000 \text{ and } \text{在庫} \geq 10](\text{秋葉原支店})$
全部の属性を出力したい場合(問題の場合)、このまま
商品番号だけ欲しい場合には、射影演算を行い

$R2 = p[\text{商品番号}](R1)$
一つの式にまとめて

$R2 = p[\text{商品番号}](s[\text{販売価格} \leq 100,000 \text{ and } \text{在庫} \geq 10](\text{秋葉原支店}))$

ちょっと考えてみよう(2-2)

- 自然言語による問い合わせの関係代数で表現
 - 秋葉原支店か日本橋支店で、販売価格が100,000円以上の商品の商品番号と販売価格のデータが欲しい場合、キャラクタ入力での関係代数で表現してみよう。

関係代数演算の種類

- 集合演算 (Set Operations)
 - 和、積、差、直積
- 専用演算
 - 射影、選択、制約、結合、割り算

結合演算 (Join) : $\bowtie$

- 2つの関係の間の属性同士のつき合わせ
 - θ 結合 (θ -Join)
 - $R_a \bowtie_{x_a \theta x_b} R_b$ ($\theta: =, \neq, <, \leq, >, \geq$)
 - 等結合 (Eq-Join)
 - θ が $=$
 - $R_a \bowtie_{x_a = x_b} R_b$
 - 自然結合 (Natural-Join)
 - θ が $=$ で、結果関係の中の結合属性が1つ (なぜ?)
 - 正規化で利用

結合演算の例: 結合対象

製造元関係			秋葉原支店関係		
商品番号	商品名	製造元	商品番号	販売価格	在庫
GDC10	Desktop	GELL	GDC10	150,000	5
GLC20	Laptop	GELL	GLC20	90,000	12
FLC33	Laptop	Fepson	FLC33	70,000	3
FLP44	Laser Pt.	Fepson	HIP06	30,000	4
HLP05	Laser Pt.	Hanon	TLCD70	40,000	2
HIP06	Inkjet Pt.	Hanon			
TLCD70	LCDisplay	Tenao			

自然結合演算の例: 結合結果

製造元関係 $\bowtie$ 商品番号=商品番号 秋葉原支店関係

商品番号	商品名	製造元	販売価格	在庫
GDC10	Desktop	GELL	150,000	5
GLC20	Laptop	GELL	90,000	12
FLC33	Laptop	Fepson	70,000	3
HIP06	Inkjet Pt.	Hanon	30,000	4
TLCD70	LCDisplay	Tenao	40,000	2

Cardinality ?

結合が入った問い合わせの例

□ 自然言語による問い合わせ

- 商品分類がComputerに属するGELLの商品で、秋葉原支店か日本橋支店で販売価格が 100,000円以下で売られている商品の商品名と商品番号を出力することを考える。
 - 製造元関係 (商品番号、商品名、製造元)
 - 商品分類関係 (商品分類、商品名)
 - 秋葉原支店関係 (商品番号、販売価格、在庫)
 - 日本橋支店関係 (商品番号、販売価格、在庫)

関係代数式の例

- 商品分類がComputerに属するGELLの商品で、秋葉原支店が日本橋支店で販売価格が 100,000円以下で売られている商品の商品名と商品番号

Π 商品名、商品番号 (

$(\sigma_{\text{販売価格} \leq 100,000} (\text{秋葉原支店}) \cup$

$\sigma_{\text{販売価格} \leq 100,000} (\text{日本橋支店})) \bowtie \text{商品番号} = \text{商品番号}$

$(\sigma_{\text{商品分類} = \text{Computer}} (\text{商品分類関係}) \bowtie \text{商品名} = \text{商品名})$

$\sigma_{\text{製造元} = \text{GELL}} (\text{製造元関係}))$

結合演算の補足

- θ 結合は、直積演算と制約演算で表現可能

- 結合演算における法則

- 可換則 (Commutative Law)

$$R \bowtie S = S \bowtie R$$

- 結合則 (Associative Law)

$$R \bowtie (S \bowtie T) = (R \bowtie S) \bowtie T$$

ちょっと考えてみよう(2-3)

□ 結合演算での可換則と結合則

- 結合演算で可換則と結合則が成り立つ理由と、そのメリットは何か考えてみよう

θ結合に関する補足

関係A		関係B		結果			
属性p	属性q	属性r	属性s	属性p	属性q	属性r	属性s
A	10	a	20	A	10	a	20
B	20	b	40	A	10	b	40
C	30	c	60	A	10	c	60
				B	20	a	20
				B	20	b	40
				B	20	c	60
				C	30	a	20
				C	30	b	40
				C	30	c	60

$$R_A \bowtie_{q \geq s} R_B$$

外結合(Outer Join): $\bowtie$

- 一方の関係に対応する属性値がない場合にnullを入れる結合

□ 例

R		S	
商品分類	商品名	商品名	製造元
Computer	Desktop	Desktop	GELL
Computer	Laptop	Laptop	Fepson
Printer	Laser Pt.	LCDisplay	Tenao

- 商品名による自然結合なら、結果の濃度は...

- 結果に含まれなくなる情報がある

左外結合、右外結合、全外結合

$R \bowtie S$ Left Outer Join

商品分類	商品名	製造元
Computer	Desktop	GELL
Computer	Laptop	Fepson
Printer	Laser Pt.	null

$R \bowtie S$ Full Outer Join

商品分類	商品名	製造元
Computer	Desktop	GELL
Computer	Laptop	Fepson
Printer	Laser Pt.	null
null	LCDisplay	Tenao

$R \bowtie S$ Right Outer Join

商品分類	商品名	製造元
Computer	Desktop	GELL
Computer	Laptop	Fepson
null	LCDisplay	Tenao

割り算 (Division) : $R_a[X] \div R_b[Y]$

- 関係 R_a のタプルの集合の中で、その属性 X の値として、関係 R_b の全タプルの属性の値を含むものを全て抽出する
- 直積の逆算
 - 結果の関係と関係 R_b の直積を取り、剰余を加えると関係 R_a になる

割り算の例(1)

販売実績関係 [R_s]

販売員	商品名
田中	Desktop
田中	Laptop
田中	Laser Pt.
田中	LCDisplay
鈴木	Laptop
鈴木	Laser Pt.
鈴木	LCDisplay
佐藤	Desktop
佐藤	Laptop
佐藤	Laser Pt.

販売員関係 [R_p]

販売員
田中
鈴木
佐藤

全販売員の扱った商品
 $R_s[\text{販売員}] \div R_p[\text{販売員}]$

商品名
Laptop
Laser Pt.

ちょっと考えてみよう(2-4)

- $R_s[\text{販売員}] \div R_p[\text{販売員}]$ の余り
 - 余りをテーブルで書いてみましょう
 - 余りを求める関係代数式を書いてみましょう
 - キャラクタ入力で割り算は $d(R_s[\text{販売員}], R_p[\text{販売員}])$

割り算の例(2)

販売実績関係 [R_s]

販売員	商品名
田中	Desktop
田中	Laptop
田中	Laser Pt.
田中	LCDisplay
鈴木	Laptop
鈴木	Laser Pt.
鈴木	LCDisplay
佐藤	Desktop
佐藤	Laptop
佐藤	Laser Pt.

商品項目関係 [R_i]

商品名
Desktop
Laptop
Laser Pt.

商品項目 R_i の全てを扱った販売員
 $R_s[\text{商品名}] \div R_i[\text{商品名}]$

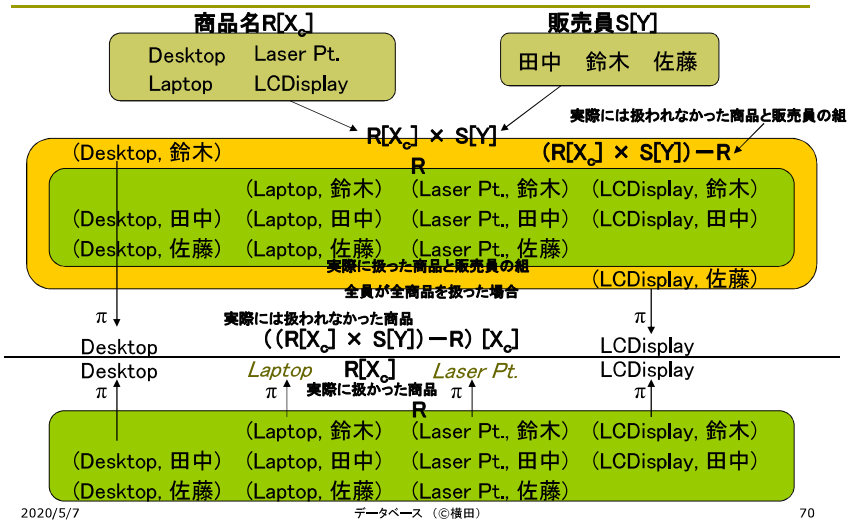
販売員
田中
佐藤

割り算の補足(1)

□ 他の演算による導出

- $R[X] \div S[Y]$
 $= R[X_c] - ((R[X_c] \times S[Y]) - R)[X_c]$
- ここで、 X_c は、関係Rの属性Xに対する補並び
 - つまり、関係 R の X 以外の属性の並び

割り算の補足(2): $R[X] \div S[Y]$

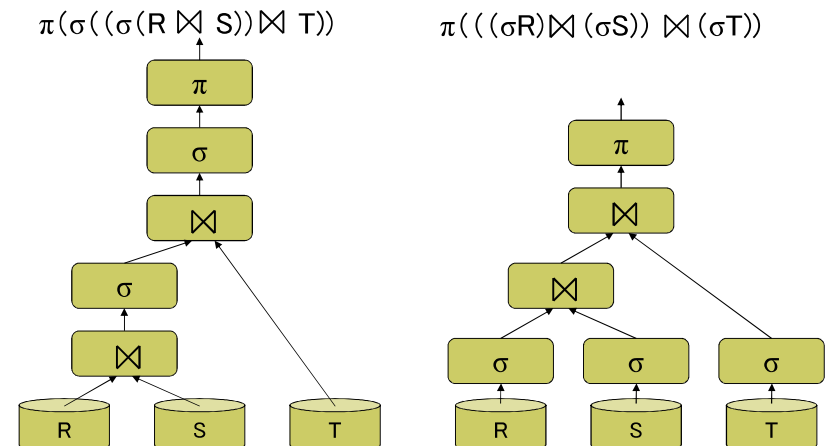


問い合わせ木 (Query Tree)

□ 一般の関係代数の問い合わせは、最終結果を根とする木を構成する

- ノードが関係代数演算(入力も出力も関係)
- 二項演算が分岐
- 結合演算の可換則、結合則によって最適化が可能
 - どのような構造が最適かは条件によって異なる
 - 各演算(特に結合演算)の実現アルゴリズム
 - 各関係の濃度

同じ問い合わせの異なる木の例



先ほどの例の場合

- 商品分類がComputerに属するGELLの商品で、秋葉原支店が日本橋支店で販売価格が 100, 000円以下で売られている商品の商品名と商品番号

Π 商品名、商品番号 (

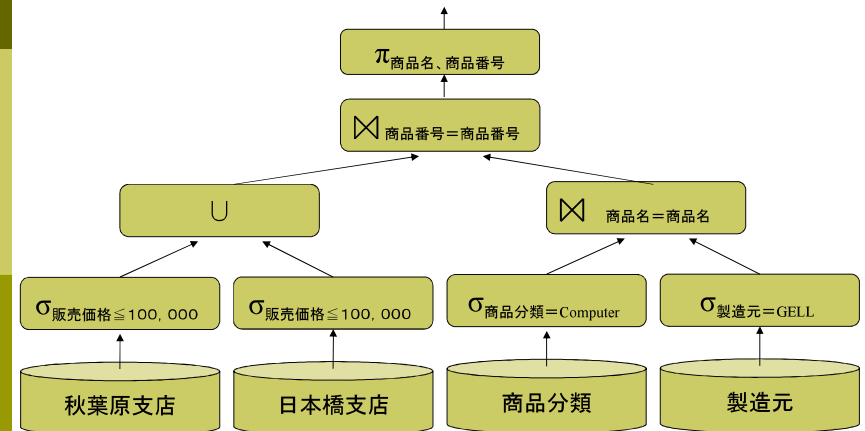
(σ 販売価格 $\leq$ 100, 000 (秋葉原支店) $\cup$

σ 販売価格 $\leq$ 100, 000 (日本橋支店)) $\bowtie$ 商品番号 = 商品番号

(σ 商品分類 = Computer (商品分類関係) $\bowtie$ 商品名 = 商品名

σ 製造元 = GELL (製造元関係))

例の場合の問い合わせ木



本日の講義の最後に

- 初めてのオンライン講義でこなれていませんが、コメントがあったら書き込んで下さい(オプションル)