リレーショナル代数

リレーションに格納されたデータを操作するデータ操作言語

リレーショナルデータベースの操作

- リレーショナル代数 (relational algebra)
  - リレーションに対する基本的操作を行う代数演算子を用いてデータ操作を記述する
  - 今回の授業のメインである
- リレーショナル論理 (relational calculus)
  - 一階述語論理を基礎とした論理系を用いてデータ操作を記述する
  - この授業では簡単に触れるのみ

リレーショナル代数

- 1つまたは2つのリレーションから1つのリレーションを作り出す操作を数式で表現するもの
- 演算子
  - 4つの集合演算
  - 4つの固有の演算

4つの集合演算

- 和集合演算 (union) $R \cup S$
- 差集合演算 (difference) $R - S$
- 共通集合演算 (intersection) $R \cap S$
- 直積演算 (direct product または) $R \times S$
  (Cartesian product)

和両立

- 和集合演算，差集合演算，共通集合演算を行うための条件
- 定義
  - 2つのリレーションの次数が同じかつ対応する属性のドメインが同一でなければならない
- 名前は一致していなくても良い
  - 結果の表の属性名は，片方の名前に合わせてもいいし，別の名前を付けても良い

和集合演算

- 2つのリレーションがあり，少なくともどちらかのリレーションに含まれるタップルからなるリレーションを得る
- 定義

  R と S を和両立なリレーションとする

  $$R \cup S = \{t \mid t \in R \vee t \in S\}$$

R

A	X	U
1	x	u
2	y	v
1	z	u

S

A	X	U
1	x	t
2	y	v

$R \cup S$

A	X	U
1	x	u
2	y	v
1	z	u
1	x	t

青字はRとSに共通に含まれている

差集合演算

- 2つのリレーションがあり，オペランドの左のリレーションに含まれ，右のリレーションに含まれないタップルからなるリレーションを得る
- 定義

  R と S を和両立なリレーションとする

  $$R-S = \{t \mid t \in R \wedge \neg t \in S\}$$

R

A	X	U
1	x	u
2	y	v
1	z	u

S

A	X	U
1	x	t
2	y	v

$R-S$

A	X	U
1	x	u
1	z	u

青字はRとSに共通に含まれている

共通集合演算

- 2つのリレーションがあり，両方のリレーションに共通に含まれるタップルからなるリレーションを得る
  - 1番目の参考図書では 共通部分 と表記されている
- 定義

  R と S を和両立なリレーションとする

  $$R \cap S = \{t \mid t \in R \wedge t \in S\}$$

- 以下の表現も可能である

  $$R \cap S = R - (R - S)$$

R

A	X	U
1	x	u
2	y	v
1	z	u

S

A	X	U
1	x	t
2	y	v

$R \cap S$

A	X	U
2	y	v

青字はRとSに共通に含まれている

直積演算

- 2つのリレーションがあり，両方のリレーションのタップルの連結のすべての組み合わせからなるリレーションを得る
- 定義

  R と S をリレーションとする

  $R \times S = \{(r, s) \mid r \in R \wedge s \in S\}$　　(r, s) はタップル r と s の連結

  属性名はそのまま引き継がれるが，衝突する場合はドット表記で区別する

R

B	X	Y
1	x	t
2	y	v

S

A	X	U
1	x	t
2	y	v

$R \times S$

A	$S.X$	U	B	$T.X$	Y
1	x	t	1	x	t
1	x	t	2	y	v
2	y	v	1	x	t
2	y	v	2	y	v

ドット記法

- 同じ属性名を含む2つのリレーションの直積の結果には同じ名前の属性が含まれる

↓

- 属性名は同じでも出所は異なる

↓

- それらを区別するために，出所のリレーション名とその属性をドットで結んで表す

$R(A_1, \cdots, A_n)$，$S(B_1, \cdots, B_n)$ で A_i とB_j が同じ名前の場合，
$R.A_i$, $S.B_j$ と表記する

4つの固有の演算

- 射影演算 (projection)　$R[X]$

  1番目の参考図書では $\pi_X(R)$

- 選択演算 (selection)　$R[A_i \, \theta \, A_j]$

  1番目の参考図書では $\sigma_F(R)$

  Fは選択条件

- 結合演算 (join)　$R[A_i \, \theta \, B_j]S$

  1番目の参考図書では $R \bowtie_F S$

  Fは結合条件

- 商演算 (division)　$R \div S$

射影演算

- リレーションが持つ属性のうち，指定した属性のみを取り出す
- 定義

$R(A_1, \cdots, A_n)$ をリレーション，

X を R の全属性集合 $\{A_1, \cdots, A_n\}$ の

部分集合 $\{A_{i_1}, \cdots, A_{i_k}\}$ $\{1 \leqq i_1 < \cdots < i_k \leqq \mathrm{n}\}$ とするとき，

R の X 上の射影 $R[X]$ は次のように定義される

$$R[X] = \{t[X] \mid t \in R\}$$

$t[X]$ はタップル t から属性集合 X のみを残したタップル

R

A	X	U
1	x	u
2	y	v
1	z	u

→

$R[A,U]$

A	U
1	u
2	v

選択演算

- リレーションが持つタップルから，指定した条件を満足するタップルを取り出す
- 定義

  $R(A_1, \cdots, A_n)$ をリレーション，

  A_i と A_j を θ-比較可能な属性とするとき，

  R の A_i と A_j 上の θ-選択 $R[A_i\,\theta\,A_j]$ は

  次のように定義されるリレーションである

$$R[A_i\,\theta\,A_j] = \{t \mid t \in R \wedge t[A_i]\,\theta\,t[A_j]\}$$

比較演算θは$=$，$<$，$>$，$\leqq$，$\geqq$，$\neq$ の何れか

1番目の参考図書 pp. 40～41

1番目の参考図書での選択条件の表現

選択条件としては通常以下を考える

① 属性 A_i の値と定数 c の θ による
比較条件$A_i\,\theta\,c$

② 属性 A_i と属性 A_j の θ による
比較演算 $A_i\,\theta\,A_j$

③ 上記①②の条件を論理和(∨), 論理積(∧), 否定(¬) を用いて
組み合わせたもの

R

A	X	U
1	x	u
2	y	v
1	z	u

→

$R[U=u]$

A	X	U
1	x	u
1	z	u

結合演算

- 直積したものから条件に合ったタップルを取り出す
- 定義

$R(A_1, \cdots, A_n)$ と $S(B_1, \cdots, B_m)$ を2つのリレーション，

A_i と B_j をθ-比較可能とするとき，

R と S のA_i と B_j 上のθ-結合 $R[A_i \,\theta\, B_j]S$ は

次のように定義されるリレーションである

$$R[A_i \,\theta\, B_j]S = \{(t, u) \mid t \in R \wedge u \in S \wedge t[A_i] \,\theta\, u[B_j]\}$$

自然結合

- 2つのリレーションで同じ名称の属性が存在し，それらの属性値で等しいという条件で結合し，重複する片方を削ったタップルを出力する操作で，

  $R * S$

  と表現する
  1番目の参考図書では $R \bowtie S$ と表記している

結合

R

A	X	U
1	x	u
2	y	v
1	z	u

S

B	X
0	w
2	z
1	x

$R[A=B]S$

A	$R.X$	U	B	$S.X$
1	x	u	1	x
1	z	u	1	x
2	y	v	2	z

$R[A \leqq B]S$

A	$R.X$	U	B	$S.X$
1	x	u	1	x
1	x	u	2	z
1	z	u	1	x
1	z	u	2	z
2	y	v	2	z

$R * S$

A	X	U	B
1	x	u	1
1	z	u	2

自然結合

商演算

- $R(A_1, \cdots, A_{n\text{-}m}, B_1, \cdots, B_m)$ をn次，$S(B_1, \cdots, B_m)$ をm次 ($m<n$) のリレーションとするとき，
  R を S で割った商は
  次のように定義されるリレーションである
- 定義

$$R \div S \;=\; \{\{t \mid t \in R[A_1, \cdots, A_{n\text{-}m}] \wedge (\forall u \in S)((t, u) \in R)\}$$

- $(R \times S) \div S = R$ が成り立つので商演算と名付けられた

R

A	B	X	W
1	1	x	u
1	1	y	v
1	1	x	v
1	2	y	v
1	2	x	u
2	1	x	u
2	1	y	u
2	2	y	v

S

X	W
x	u
y	v

$R \div S$

A	B
1	1
1	2

リレーショナル代数表現

1. リレーショナルデータベースの実リレーションRは表現である.
2. RとSを表現とするとき, RとSが和両立なら, $R \cup S$, $R - S$, $R \cap S$ は表現である. (和集合, 差集合, 共通集合)
3. RとSを表現とするとき, $R \times S$は表現である. (直積)
4. Rを表現とするとき, $R[X]$は表現である. ここに, XはRの属性をなす集合である. (射影)
5. Rを表現とするとき, $R[A_i \theta A_j]$は表現である. A_iとA_jはRの属性で, θ-比較可能とする. (選択)
6. RとSを表現とするとき, $R[A_i\ \theta\ B_j]S$は表現である. A_iとB_jはそれぞれRとSの属性で, θ-比較可能とする. (結合)
7. RとSを表現とするとき, $R \div S$は表現である. (商)
8. 以上の定義によって得られた表現のみがリレーショナル代数表現である.

リレーショナル論理

- リレーショナル代数
  - 集合としてのリレーションの性質に基づき、既存のリレーションから新たなリレーションを作成する演算体系
- リレーショナル論理
  - 一階述語論理を基礎としたデータ操作の体系
  - 目的とするリレーションを論理式により宣言的に記述
  - 2種類のリレーショナル論理
    - ✓タップルを変数として用いるタップルリレーショナル論理
    - ✓ドメインの要素を変数として用いるドメインリレーショナル論理
- 記述能力
  - リレーショナル代数、タップルリレーショナル論理、ドメインリレーショナル論理は同じ記述能力を持ち、本質的に同等である

（参考）論理式の問題点

- 問題点
  - 否定が入ると無限集合になる可能性がある
    $\{t \mid \neg R(t)\}$
  - 直積の場合でも巨大集合あるいは無限集合になることがある
- 解決法
  - タップル変数やドメイン変数の取りうる値 (ドメイン) を適切に定める (安全表現 (safe expression))
  - 例
    - ✓安全な例 : $\{t \mid R(t) \wedge \neg S(t)\}$
      t の取りうる範囲をリレーションR に含まれるものに限定する
    - ✓安全でない例 : $\{t \mid R(t) \wedge (\exists u)\ (u[A] > t[A])\}$
      u の取り得る範囲を限定できない