

## データベース概説(第4回)

2008.05.19  
宇陀則彦

1

## リレーショナルデータモデル

直観的には表(テーブル)

### 支店

支店名	所在地	代表電話番号
みなと	横浜	045-123-4567
なにわ	大阪	06-555-6666
おえど	東京	03-3444-8888
はかた	福岡	092-333-7777

2

## リレーショナルデータモデル

数学に基づく形式的定義

### 定義

- リレーションスキーマ  $R(A_1, \dots, A_n)$  が与えられたとき、 $D_1 \times \dots \times D_n$  の有限部分集合  $r \subseteq D_1 \times \dots \times D_n$  をリレーションと呼ぶ。

注: E-RモデルのRelationshipとリレーショナルデータモデルのRelationは異なる概念

3

## 直積

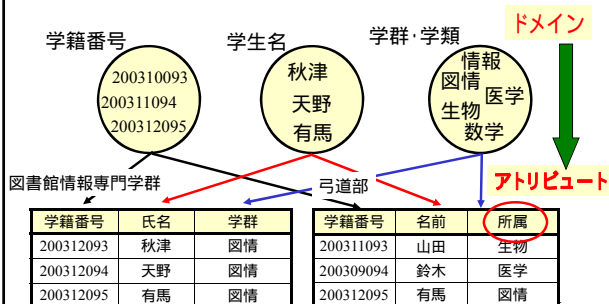
### 練習

- $\{a, b, c\} \times \{1, 2\} \times \{A, B\}$  の直積を求めよ。

$\{(a, 1, A), (a, 1, B), (a, 2, A), (a, 2, B),$   
 $(b, 1, A), (b, 1, B), (b, 2, A), (b, 2, B),$   
 $(c, 1, A), (c, 1, B), (c, 2, A), (c, 2, B)\}$

4

## 関係(リレーション)



## ドメイン、アトリビュート、スキーマ

- ドメイン(定義域)の集合の要素は、「シンプルな値」であること。
- アトリビュート(属性)の名前は、そのリレーション(関係)を適切に表現した(現実世界の意味を反映した)名前にしてよい。
- リレーションスキーマ  $R$  は、アトリビュートの組  $(A_1, A_2, A_3, \dots, A_n)$  によって記述される。

6

## 整合性制約(一貫性制約)

- 定義域制約(ドメイン制約)
  - タブルの値はドメインの値であること  
(図書館情報専門学群の学生は筑波大学の学生)
  - ドメインの条件に従っていること  
(学生の年齢が2歳とか500歳などでないこと)  
(100点満点の成績は0から100までの値であること)

7

## 整合性制約(一貫性制約)

- キー制約
  - 候補キーを構成する属性値は唯一であること
  - 主キーを構成する属性値は唯一であり、かつ空値(null値)でないこと
- スーパーキー(super key)
- 候補キーまたは単にキー(candidate key)
- 主キー(primary key)

8

## スーパーキー

- リレーションのタブルを一意に識別できる属性または属性の組合せをスーパーキーという。
  - (支店名、所在地、代表電話番号)
  - (支店名、所在地)
  - (支店名、代表電話番号)

9

## 候補キー

- スーパーキーのうち、極小(一意性を保つために除けないもの)の属性をいう。
    - (支店名、所在地、代表電話番号)
    - (支店名、所在地)
    - (支店名、代表電話番号)
- これらスーパーキーのうち、
- (支店名)
  - (代表電話番号)
- が候補キーになりうる。(条件(状況)による)

10

## 主キー

- 候補キーのうち、そのリレーションスキーマに最もふさわしい属性をいう。
    - (支店名)
    - (代表電話番号)
- これら候補キーのうち、
- (支店名)
- が主キーとなる。
- (空値となりえない。)
  - (「支店」というリレーションに最適)

11

## データ操作

- 質問(query): 「情報検索」との違い
  - 条件に合うタブルの集合を取り出す。
  - 条件に合う仮想リレーションを作り出す。
  - queryの結果もリレーション(再帰的操作が可)
- 更新(update)
  - 追加(insertion)
  - 削除(deletion)
  - 変更(modification)

実際のデータベースを使う際には、質問と更新の両方が行えるデータベース言語SQLを使う。

12

## データ操作体系

- データ操作 (query) に関して理論的基盤を与える。
- リレーショナル代数 (relational algebra)
  - 手続き的 (集合論を基礎とした体系)
- リレーショナル論理 (relational calculus)
  - 宣言的 (一階述語論理を基礎とした体系)
- リレーショナル代数とリレーショナル論理は同等の記述力を持つ。

13

## リレーショナル代数

- 集合演算
  - 和集合演算 (union)\*
  - 差集合演算 (difference)\*
  - 共通集合演算 (intersection)
  - 直積集合演算 (cartesian product)\*
- リレーショナル代数に特有の演算
  - 選択演算 (selection)\*
  - 射影演算 (projection)\*
  - 結合演算 (join)
  - 商演算 (division)

原理的には  
\*がついた5つ  
で十分

14

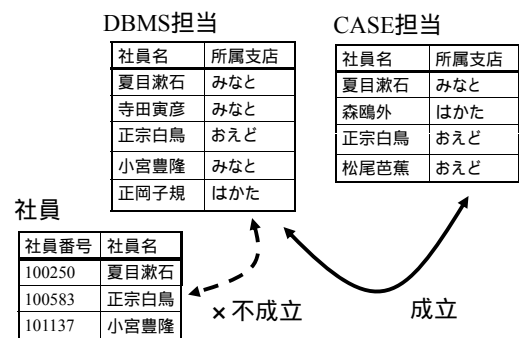
## 和両立条件 (union compatibility)

二つのリレーション  $R(A_1, \dots, A_m)$  と  $S(B_1, \dots, B_n)$  において、和両立条件が成立するとは、

1.  $m = n$ 
  - $R$  と  $S$  の次数が同じ
2.  $\text{dom}(A_i) = \text{dom}(B_i)$ 
  - 対応するアトリビュートのドメインが同じを満たすことである。

15

## 和両立条件の例



16

## 和集合演算 (union)

DBMS担当

社員名	所属支店
夏目漱石	みなと
寺田寅彦	みなと
正宗白鳥	おえど
小宮豊隆	みなと
正岡子規	はかた

- $R \cup S = \{t \mid t \in R \vee t \in S\}$
- 和両立条件が成立
- タブルの集合の和
- 重複タブルは除去

DBMS担当 CASE担当

社員名	所属支店
夏目漱石	みなと
寺田寅彦	みなと
正宗白鳥	おえど
小宮豊隆	みなと
正岡子規	はかた
森鷗外	はかた
松尾芭蕉	おえど

CASE担当

社員名	所属支店
夏目漱石	みなと
森鷗外	はかた
正宗白鳥	おえど
松尾芭蕉	おえど

17

## 差集合演算 (difference)

DBMS担当

社員名	所属支店
夏目漱石	みなと
寺田寅彦	みなと
正宗白鳥	おえど
小宮豊隆	みなと
正岡子規	はかた

- $R - S = \{t \mid t \in R \wedge t \notin S\}$
- 和両立条件が成立
- タブルの集合の差
- $R - S \neq S - R$

DBMS担当 - CASE担当

社員名	所属支店
寺田寅彦	みなと
小宮豊隆	みなと
正岡子規	はかた

CASE担当

社員名	所属支店
夏目漱石	みなと
森鷗外	はかた
正宗白鳥	おえど
松尾芭蕉	おえど

18

## 差集合演算 (difference)

DBMS担当

社員名	所属支店
夏目漱石	みなと
寺田寅彦	みなと
正宗白鳥	おえど
小宮豊隆	みなと
正岡子規	はかた

- $R - S = \{t \mid t \in R \wedge \neg t \in S\}$
- 和両立条件が成立
- タブルの集合の差
- $R - S \neq S - R$

CASE担当 - DBMS担当は？

CASE担当

社員名	所属支店
夏目漱石	みなと
森鷗外	はかた
正宗白鳥	おえど
松雄芭蕉	おえど

19

## 差集合演算 (difference)

DBMS担当

社員名	所属支店
夏目漱石	みなと
寺田寅彦	みなと
正宗白鳥	おえど
小宮豊隆	みなと
正岡子規	はかた

- $R - S = \{t \mid t \in R \wedge \neg t \in S\}$
- 和両立条件が成立
- タブルの集合の差
- $R - S \neq S - R$

CASE担当 - DBMS担当は？

CASE担当

社員名	所属支店
夏目漱石	みなと
森鷗外	はかた
正宗白鳥	おえど
松雄芭蕉	おえど

社員名	所属支店
森鷗外	はかた
松尾芭蕉	おえど

20

## 共通集合演算 (intersection)

DBMS担当

社員名	所属支店
夏目漱石	みなと
寺田寅彦	みなと
正宗白鳥	おえど
小宮豊隆	みなと
正岡子規	はかた

- $R \cap S = \{t \mid t \in R \wedge t \in S\}$
- 和両立条件が成立
- $R \cap S = R - (R - S)$

DBMS担当  $\cap$  CASE担当

社員名	所属支店
夏目漱石	みなと
正宗白鳥	おえど

CASE担当

社員名	所属支店
夏目漱石	みなと
森鷗外	はかた
正宗白鳥	おえど
松雄芭蕉	おえど

21

## 和、差、共通集合の練習

司書科目履修者

学生名	学年	クラス
若月	3	A
手塚	4	D
吉田	3	C
萩尾	3	B
筒井	4	A

司書教諭科目履修者

学生名	学年	クラス
紫堂	3	C
若月	3	A
萩尾	3	B
成島	4	A

1. 司書科目と司書教諭科目のどちらかをとっている学生は？
2. 司書科目だけをとっている学生は？
3. 司書教諭科目だけをとっている学生は？
4. 司書科目と司書教諭科目の両方をとっている学生は？

22

## 直積集合演算 (cartesian product)

- $R \times S = \{t * u \mid t \in R \wedge u \in S\}$
- $t * u$  はタブル  $t$  とタブル  $u$  の連結

支店名	所在地	代表電話番号
みなと	横浜	045-123-4567
なにわ	大阪	06-555-6666

社員番号	社員名	所属支店
100250	夏目漱石	みなと
100580	寺田寅彦	みなと
100583	正宗白鳥	おえど

支店名	所在地	代表電話番号	社員番号	社員名	所属支店
みなと	横浜	045-123-4567	100250	夏目漱石	みなと
みなと	横浜	045-123-4567	100580	寺田寅彦	みなと
みなと	横浜	045-123-4567	100583	正宗白鳥	おえど
なにわ	大阪	06-555-6666	100250	夏目漱石	みなと
なにわ	大阪	06-555-6666	100580	寺田寅彦	みなと
なにわ	大阪	06-555-6666	100583	正宗白鳥	おえど

## 選択演算 (selection)

ある条件を満たすタブルを取り出す操作

選択条件

- = (等しい)
- $\neq$  (等しくない)
- $>$  < (大小)

$\sigma$ 社員名 = '夏目漱石' (社員)

社員番号	社員名	所属支店
100250	夏目漱石	みなと

これらを AND OR NOT で組み合わせることができる。

社員

$\sigma$ 社員番号 > 100300 (社員)

社員番号	社員名	所属支店
100250	夏目漱石	みなと
100583	正宗白鳥	おえど
101137	小宮豊隆	みなと

社員番号	社員名	所属支店
100583	正宗白鳥	おえど
101137	小宮豊隆	みなと

24

## 射影演算 (projection)

指定した属性を残し、他の属性を削除する操作  
その際、重複したタプルは削除

社員

社員番号	社員名	所属支店
100250	夏目漱石	みなと
100583	正宗白鳥	おえど
101137	小宮豊隆	みなと

π所属支店 (社員)

所属支店
みなと
おえど
<del>みなと</del>
みなと
おえど

(重複削除)

25

## 結合演算 (join)

- 直積演算は全ての組合せだが、結合演算は条件に合ったもの
- 支店  $\bowtie$  支店名-所属支店 社員

支店名	所在地	代表電話番号
みなと	横浜	045-123-4567
なにわ	大阪	06-555-6666

社員番号	社員名	所属支店
100250	夏目漱石	みなと
100580	寺田寅彦	みなと
100583	正宗白鳥	おえど

支店名	所在地	代表電話番号	社員番号	社員名	所属支店
みなと	横浜	045-123-4567	100250	夏目漱石	みなと
みなと	横浜	045-123-4567	100580	寺田寅彦	みなと

↑ 属性名は異なるが、属性値は同じ ↑

26

## 自然結合演算 (natural join)

- 直積演算は全ての組合せだが、結合演算は条件に合ったもの
- 支店  $*$  支店名-所属支店 社員

支店名	所在地	代表電話番号
みなと	横浜	045-123-4567
なにわ	大阪	06-555-6666

社員番号	社員名	所属支店
100250	夏目漱石	みなと
100580	寺田寅彦	みなと
100583	正宗白鳥	おえど

所在地	代表電話番号	社員番号	社員名	所属支店
横浜	045-123-4567	100250	夏目漱石	みなと
横浜	045-123-4567	100580	寺田寅彦	みなと

27

## 直積集合演算の練習

学生

学生名	学年	クラス
石田	1	A
松林	1	B

科目

科目名	履修年次	担当教員
情報基礎	1	石塚
データベース	1	宇陀
情報検索演習	2	岩澤

この2つのリレーションの直積をとれ。

28

## 等結合演算の練習

学生

学生名	学年	クラス
石田	1	A
松林	1	B

科目

科目名	履修年次	担当教員
情報基礎	1	石塚
データベース	1	宇陀
情報検索演習	2	岩澤

等結合 (equi-join)

学生  $\bowtie$  学年=履修年次 科目  
を求めなさい。

29

## 自然結合演算の練習

学生

学生名	学年	クラス
石田	1	A
松林	1	B

科目

科目名	履修年次	担当教員
情報基礎	1	石塚
データベース	1	宇陀
情報検索演習	2	岩澤

自然結合 (natural-join)

学生  $*$  学年=履修年次 科目  
を求めなさい。

30

## 選択演算の練習

履修

学生名	学年	クラス	科目名	担当教員
石田	1	A	情報基礎	石塚
石田	1	A	データベース	宇陀
松林	1	B	情報基礎	石塚
松林	1	B	データベース	宇陀

- σ 担当教員=石塚 (履修) を行いなさい。

31

## 射影演算の練習

履修

学生名	学年	クラス	科目名	担当教員
石田	1	A	情報基礎	石塚
石田	1	A	データベース	宇陀
松林	1	B	情報基礎	石塚
松林	1	B	データベース	宇陀

- π 学生名, 科目名, 担当教員 (履修) を行いなさい。

32

## 商演算 (division)

担当

担当者	担当製品
夏目漱石	DBMS ●
夏目漱石	CASE ●
寺田寅彦	DBMS
正宗白鳥	DBMS ●
正宗白鳥	CASE ●
松雄芭蕉	CASE

製品

製品名
DBMS
CASE

から

の全ての製品を担当してる  
担当者を抜き出す操作。

結果は

夏目漱石 ●
正宗白鳥 ●

33

## 外結合演算 (outer join)

結合では、結合する相手がいないタプルは結果に残らない  
どうしても残したい場合に **ナル値を相手の値にして残す**。

支店

支店名	所在地	代表電話番号
みなと	横浜	045-123-4567
おえど	東京	03-3444-8888

社員

社員番号	社員名	所属支店
100250	夏目漱石	みなと
101137	小宮豊隆	みなと

「おえど」の相手が無いので、社員の属性に-が入る)

支店名	所在地	代表電話番号	社員番号	社員名	所属支店
みなと	横浜	045-123-4567	100250	夏目漱石	みなと
みなと	横浜	045-123-4567	101137	小宮豊隆	みなと
おえど	東京	03-3444-8888	-	-	-

34

## 外和集合演算 (outer union)

ナル値を使って和両立でないリレーションで和を取る。

- 両者の和両立になる属性について (普通の) 和をとる。
- 残りの属性について、値がないところには-を入れる。

社員名	所属支店	DBスキル	社員名	所属支店	職歴
夏目漱石	みなと	RDB	夏目漱石	みなと	10
寺田寅彦	みなと	NDB	森鷗外	はかた	8
正宗白鳥	おえど	HDB			

上のリレーションの  
外和集合 =

社員名	所属支店	DBスキル	職歴
夏目漱石	みなと	RDB	10
寺田寅彦	みなと	NDB	-
正宗白鳥	おえど	HDB	-
森鷗外	はかた	-	8

35