

ダウンロード版

データベース問題 (2)

学籍番号

氏名

学年

1. 直積集合の定義を書け

2. ドメインの定義を書け

3. リレーショナルデータモデルにおけるリレーションの定義を書け

4. リレーションスキーマ $R(A_1, \dots, A_n)$ について.

(1) R は何を表わすか

(2) A_1 や A_n は何を表わすか

2. リレーションスキーマ $R(A_1, \dots, A_n)$ について,

(1) このリレーションスキーマに対するリレーションの定義を書け.

(2) このリレーションスキーマに対するタプルの定義を書け.

5 リレーションスキーマとリレーションの違いを説明せよ

<必須の用語>

■ データモデリング

概念モデルや論理モデルを記述する行為

■ 直積集合 (「直積」とも呼ぶ)

n 個の集合 $S_1, S_2, \dots, S_n$ が与えられたとき, $\{(X_1, X_2, \dots, X_n) \mid X_1 \in S_1, X_2 \in S_2, \dots, X_n \in S_n\}$ を $S_1, S_2, \dots, S_n$ の直積集合と呼び, 「 $S_1 \times S_2 \times \dots \times S_n$ 」と記す. 例えば,

$$S_1 = \{1, 2\}, S_2 = \{a, b, c\}$$

とすると

$$S_1 \times S_2 = \{(1, a), (1, b), (1, c), (2, a), (2, b), (2, c)\}$$

※ $S_2 \times S_1 = \{(a, 1), (b, 1), (c, 1), (a, 2), (b, 2), (c, 2)\}$ となり, $S_1 \times S_2$ とは違う集合である.

■ ドメイン

属性が取りえる値の集合のこと. 例えば, 属性「年齢」のドメインは, 整数の集合 $\{0, 1, 2, \dots, 120\}$ のようになるだろう.

■ リレーション (「関係」とも呼ぶ)

$D_1, D_2, \dots, D_n$ をドメインとすると, $D_1, D_2, \dots, D_n$ 上のリレーションとは, 直積集合「 $D_1 \times D_2 \times \dots \times D_n$ 」の任意の有限部分集合を言う.

■ タプル (「タプル」とも呼ぶ)

次の2つの意味を持つ

- (1) 直積集合の各要素のこと.
- (2) リレーションの各要素のこと.

■ 属性名 (「属性」とも呼ぶ)

リレーションを, 2次元の表の形で書き表したとき, 属性名は, その表の各列に付けられた名前になる.

※ リレーションのタプルが更新 (例えば, リレーションに新しいタプルが挿入されたり, リレーションからタプルが削除されたり, タプルの値が変化したり) されても, **属性名は不変である** (変化しない).

■ リレーション名.

個々のリレーションに付けられた名前のことで, リレーション全体の意味を表す.

※ リレーションのタプルが更新 (例えば, リレーションに新しいタプルが挿入されたり, リレーションからタプルが削除されたり, タプルの値が変化したり) されても, **リレーション名は不変である** (変化しない).

■ リレーションスキーマ

リレーションにはリレーション名と属性名を付与するのが普通である (リレーションとは, 直積集合「 $D_1 \times D_2 \times \dots \times D_n$ 」の任意の有限部分集合のこと, リレーション名も属性名も無い. このままでは不便だということ). このとき, リレーションスキーマを「 $R(A_1, A_2, \dots, A_n)$ 」のように書く. R はリレーション名で, $A_1, A_2, \dots, A_n$ は属性名である. リレーションスキーマは, リレーションの枠組みを記述している. リレーションのタプルが更新されても, リレーション名と属性名は不変であるから, リレーションスキーマも不変である.

■ リレーションスキーマのインスタンス

(「リレーションスキーマのインスタンス」のことを, 「リレーション」や「関係」とも呼ぶ)

リレーションスキーマ $R(A_1, A_2, \dots, A_n)$ について (リレーション名は R であり, 属性名は $A_1, A_2, \dots, A_n$ である), これら属性のドメインを $D_1, D_2, \dots, D_n$ とする. $D_1, D_2, \dots, D_n$ 上のリレーションは, 直積集合「 $D_1 \times D_2 \times \dots \times D_n$ 」の任意の有限部分集合のことである. このリレーションのことを, 「リレーションスキーマ $R(A_1, A_2, \dots, A_n)$ のインスタンス」と呼ぶ (あるいは「リレーション」や「関係」とも呼ぶ).

※ 「ドメイン」とは属性が取りえる値の集合のこと.

※ リレーシヨンスキーマ $R(A_1, A_2, \dots, A_n)$ のインスタンスのことを, 「リレーシヨン $R(A_1, A_2, \dots, A_n)$ 」と呼んだり「リレーシヨン R 」と呼ぶことが多い. 同じ記号「 $R(A_1, A_2, \dots, A_n)$ 」と「 R 」を, リレーシヨンスキーマの意味で使ったり, リレーシヨンの意味で使ったりすることがあるので注意が必要. リレーシヨンスキーマとリレーシヨンの違いを理解していれば, 混乱はしないだろう.

※ 上記で説明したように, 属性名とリレーシヨン名は不変であるから, リレーシヨンスキーマ $R(A_1, A_2, \dots, A_n)$ は不変である. 一方で, リレーシヨン $R(A_1, A_2, \dots, A_n)$ は変化する.

■ 濃度

リレーシヨン R のタプルの総数を, R の濃度という.

■ 次数

リレーシヨン R が定義されているドメインの個数のことを, R の次数という.

<参考>

SQL の詳細は, いずれ説明する.

SQL の簡単な例

```
sqlite> create table hoge (  
...>   id INTEGER PRIMARY KEY,  
...>   price INTEGER NOT NULL,  
...>   name VARCHAR(100) NOT NULL );  
sqlite> insert into hoge values ( 1, 200, 'orange' );  
sqlite> insert into hoge values ( 2, 150, 'apple' );  
sqlite> insert into hoge values ( 3, 980, 'melon' );  
sqlite> select * from hoge;  
1|200|orange  
2|150|apple  
3|980|melon
```

SQL によるソート

```
sqlite> select * from hoge order by price;  
2|150|apple  
1|200|orange  
3|980|melon  
sqlite>
```