

2010年度「データ工学／データマネジメント／データベース」定期試験問題

1. 関係代数、SQL

以下に示すのは、あるオンラインミュージックストアのデータベースのスキーマである。このストアでは、アルバム単位でも楽曲単位でも購入が可能であるため、表「商品」にはアルバムも楽曲も含まれている。アルバムの場合には、「曲名」にアルバム名が、「アルバムID」にはNULLが入るものとする。また、楽曲の場合には「曲名」に楽曲名が、「アルバムID」には収録アルバムの商品IDが入るものとする。

商品(商品ID, 曲名, アーティスト, アルバムID, ジャンル, 価格)

ユーザ(ユーザID, 名前, 年齢)

購入履歴(ユーザID, 商品ID, 購入年, 購入月, 購入日)

このとき、以下の問合せを指定された方法で記述せよ。

(1) ジャンルが“Rock”で価格が2000円以上の商品の曲名とアーティスト名（関係代数）（10点）

π 曲名, アーティスト σ ジャンル="Rock" \wedge 価格 \geq 2000 商品

誤答が目立ったのは、 \wedge 記号を \cap や $,$ にしている、表名「商品」を書き忘れている、 π と σ を逆にしている、などでした。

(2) 商品を購入したことのあるユーザの名前一覧（関係代数）（10点）

π 名前 (ユーザ \bowtie 購入履歴)

誤答が目立ったのは、 π を適用する前に σ を適用する（例： σ ユーザ.ユーザID = 購入履歴.ユーザID）ものでした。 σ の書式の間違いもかなり見られました。

(3) ユーザ Alice が2008年に購入した商品の総額（SQL）（10点）

SELECT SUM(価格) FROM 商品, ユーザ, 購入履歴 WHERE 商品.商品ID = 購入履歴.商品ID AND ユーザ.ユーザID = 購入履歴.ユーザID AND 購入年=2008 AND ユーザ="Alice"

誤答が目立ったのは、結合条件をWHERE節に含めていない、SELECT節を単に「価格」とだけしている、などでした。

(4) アルバム“Not Too Late”に含まれる楽曲の曲名と価格。このデータベース中には、“Not Too Late”という名前のアルバムは1枚しかないものとする。（SQL）（15点）

SELECT t.曲名, t.価格 FROM 商品 t, 商品 s WHERE s.アルバムID=t.商品ID AND s.曲名="Not too Late"

(別解)

```
SELECT 曲名, 価格 FROM 商品 WHERE アルバムID IN (SELECT 商品ID FROM 商品 WHERE
  曲名="Not too Late")
```

自己結合を含む問い合わせですので、組変数を用いて自己結合にするか、入れ子問い合わせにするか、いずれかになります。目立った誤答としては、組変数を用いていない（例: SELECT 曲名, 価格 FROM 商品 WHERE アルバムID=商品ID AND 曲名="Not too Late"）、入れ子問い合わせの際に=を使っている（入れ子問い合わせの結果は組の集合なので、INを使うのが正しい）、などでした。

(5) 20歳代のユーザが購入した楽曲のジャンル毎総額を集計し、総額が100万円以上のジャンルについて、ジャンル名の辞書順に、ジャンル名と総額を出力せよ。（SQL）（15点）

```
SELECT ジャンル, SUM(価格) FROM 商品, ユーザ, 購入履歴 WHERE 商品.商品ID=購入履歴.商品ID
  AND ユーザ.ユーザID=購入履歴.ユーザID AND 年齢 BETWEEN 20 AND 29 GROUP BY
  ジャンル HAVING SUM(価格) >= 1000000 ORDER BY ジャンル ASC
```

目立った誤答としては、集計関数をSUM(価格)ではなくCOUNT(*)としている、結合条件がWHERE節に含まれていない、GROUP BY/HAVING/ORDER BYの順序が間違っている、HAVING節の条件がWHERE節にある、などでした。

2. キー、正規形

(1) 問1の3つの表に関して、外部キーの関係にある属性の組をすべて列挙せよ。（10点）

購入履歴.商品IDは「商品」の外部キー、購入履歴.ユーザIDは「ユーザ」の外部キー、商品.アルバムIDは「商品」の外部キー

3つ目は講義で説明していない例なので、抜けていても減点していません。

(2) 以下に示すスキーマを持つ表を第3正規形に分解せよ。完全関数従属性、推移的関数従属性など、分解の根拠を添えて答えよ。（20点）

履修(学籍番号, 学生名, 学科, 学年, 科目番号, 科目名, 教員番号, 教員名, 教室番号, 成績)

次に示す完全関数従属性、推移的関数従属性が成立していると考えられます。

学籍番号 →→ 学生名, 学科, 学年

科目番号 →→ 科目名, 教員番号, 教員名, 教室番号

学籍番号, 科目番号 →→ 成績

科目番号 → 教員番号 → 教員名

これらを基に表を分解すると、次の4つの表になります。これで第3正規形になっています。

R1(学籍番号, 学生名, 学科, 学年)

R2(科目番号, 科目名, 教員番号, 教室番号)

R3(教員番号, 教員名)

R4(学籍番号, 科目番号, 成績)

誤答は多岐に渡っていましたが、目立ったものを次に挙げてみます。

* 分解結果を自然結合しても元の表に戻らない: 解答例を見ると分かるように、分解後の関係間には属性の重複があります。例えば、R1の主キー「学籍番号」がR4に含まれる、R3の主キー「教員番号」がR2に含まれる、といった具合です。このような重複があることで、分解結果を自然結合して元に戻すことができるわけです。

* 学籍番号 →→ 成績が成立すると考え、R1(学籍番号, 学生名, 学科, 学年, 成績) としている: スキーマの意味から考えて、ここでいう「成績」は成績表全体ではなく、各科目ごとの点数もしくは評価であると考えべきです。従って、学籍番号 →→ 成績 ではなく、学籍番号, 科目番号 →→ 教員名 が成立すると考える方が自然です。

* R2(科目番号, 科目名), R4(学籍番号, 科目番号, 教員番号, 教室番号, 成績)としている: 科目が決まると担当教員と開講教室は一意に決まる、と考えるほうが自然です。もし、一つの科目に対して担当教員や開講教室が複数あることを認めるなら、次のように、R2をさらに分解するのが適切であろうと考えられます。

R2'(科目番号, 科目名)

R2''(科目番号, 教員番号)

R2'''(科目番号, 教室番号)

(3) (2)の答の表すべてについて、主キーを指摘せよ。(10点)

R1: 学籍番号、R2: 科目番号、R3: 教員番号、R4: 学籍番号, 科目番号

表と属性の対応を明確に書けていない答案が見られました。