

4章「データモデルとデータ設計」

(三層スキーマ・第一正規形・関数従属・決定項・キー・第三正規形)

中島康彦

§4.1 データモデル

データモデルとは

- ▶ データベースは、各自がデータを用意するのではなく、共通のデータを用意し、様々な利用者の要求に応えるもの。
- ▶ このためには、実世界の情報を自然に正確に把握するためのモデルが必要。
- ▶ 例)文献データモデル
文献の実体は紙の集合であるが、一般に必要なものは「紙」に関する情報ではない。文献を探すために何が必要かといった基本概念。

§4. 1 データモデル(続き)

階層モデル

- ▶ 以前に説明した「多分木」による表現
実世界が親子関係により自然に表現できる場合
組織内の階層構造など

ネットワークモデル

- ▶ 以前に説明した「グラフ」による表現
階層構造に多対多の関係がある場合(複数の親に所属)
互いに対等な関係である場合(空港および航空路線など)

関係モデル(現在の主流)

- ▶ 以前に説明した「リスト」による表現
表構造(1行が1件のデータを表す)
多分岐やグラフよりも扱いやすい

§4. 2 論理設計(データ設計)

論理設計(データ設計)とは

- ▶ データモデルに基づき、実世界からデータベースに必要な情報を抽象整理すること。
 - ▶ この段階では、使用する計算機やソフトウェアは考慮しない。
 - ▶ 例)文献データ
著者名, 題名, 発行者, ページ数, 発行日, 概要,
ISBN, 価格といった名前による抽象的な枠組。
-

§4. 3 概念スキーマ

概念スキーマとは

- ▶ 抽象した結果を記述したもの.
- ▶ 文字列や数値による記述.

著者名	題名	発行所	ページ数	発行日	概要	ISBN	価格
文字列	文字列	文字列	整数	文字列	文字列	文字列	整数

データモデルに基づく論理設計
実世界 → 概念スキーマ

§4. 4 内部スキーマ

内部スキーマとは

- ▶ 概念スキーマを計算機の上に具体化するために記述し直したもの.
- ▶ プログラミング言語のレベル

```
struct 文献データ構造体 {
    char * 著者名;
    char * 題名;
    char * 発行所;
    int   ページ数;
    char * 発行日;
    char * 概要;
    char * ISBN;
    int   価格;
} 文献データ [総数];
```

§4. 5 外部スキーマ

外部スキーマとは

- ▶ 概念スキーマのうち、各々の利用者が必要な部分だけを記述したもの.
- ▶ 各々の利用者は、概念スキーマの一部に興味を持つ.

著者名	題名	発行所	ISBN
文字列	文字列	文字列	文字列

実世界 → 一つ概念スキーマ → 複数の外部スキーマ

§4. 6 関係モデル

関係モデルとは

- ▶ 実世界を一つまたは複数の表として把握する考え方.
- ▶ 構造体
一行が一つの関係を表現
- ▶ 構造体のならび
複数の行が「関係表」を構成
行の順番には意味がない

§4.7 まず表にしてみる

正規化前の表

著者番号	著者名	生年月日	題名	発行所番号	発行所名	発行所連絡先	ページ数	発行日	概要	ISBN	価格
1	京大太郎	T14/1/1	大正史	100	京大出版	吉田本町	140	1926/1/1	大……	ISBNX-1234	10000
			昭和史				640	1989/1/1	昭……	ISBNX-5678	18000
			平成史	202	吉田印刷	吉田神楽岡町	110	1999/1/1	平……	ISBNX-9012	10000
2	経済花子	S14/2/2	人物金	150	近衛文庫	東山近衛	110	1999/1/1	人……	ISBNX-3456	9000
			情報	100	京大出版	吉田本町	110	1999/1/2	情……	ISBNX-7890	12000

- ▶ 1行が単純なデータ型ではない
- ▶ 取り扱いが難しい

§4.8 第一正規形

正規化後の表(第一正規形)

著者番号	著者名	生年月日	題名	発行所番号	発行所名	発行所連絡先	ページ数	発行日	概要	ISBN	価格
1	京大太郎	T14/1/1	大正史	100	京大出版	吉田本町	140	1926/1/1	大……	ISBNX-1234	10000
1	京大太郎	T14/1/1	昭和史	100	京大出版	吉田本町	640	1989/1/1	昭……	ISBNX-5678	18000
1	京大太郎	T14/1/1	平成史	202	吉田印刷	吉田神楽岡町	110	1999/1/1	平……	ISBNX-9012	10000
2	経済花子	S14/2/2	人物金	150	近衛文庫	東山近衛	110	1999/1/1	人……	ISBNX-3456	9000
2	経済花子	S14/2/2	情報	100	京大出版	吉田本町	110	1999/1/2	情……	ISBNX-7890	12000

- ▶ 各行が単純なデータ型であり、全データに重複がない

問題点

- ▶ 発行所名が間違っていた場合、複数箇所の修正が必要
- ▶ 京大太郎の出版物を削除した場合、吉田印刷に関するデータも消える

§4.9 関数従属・決定項・キー

観察してわかること

著者番号	著者名	生年月日	題名	発行所番号	発行所名	発行所連絡先	ページ数	発行日	概要	ISBN	価格
1	京大太郎	T14/1/1	大正史	100	京大出版	吉田本町	140	1926/1/1	大……	ISBNX-1234	10000
1	京大太郎	T14/1/1	昭和史	100	京大出版	吉田本町	640	1989/1/1	昭……	ISBNX-5678	18000
1	京大太郎	T14/1/1	平成史	202	吉田印刷	吉田神楽岡町	110	1999/1/1	平……	ISBNX-9012	10000
2	経済花子	S14/2/2	人物金	150	近衛文庫	東山近衛	110	1999/1/1	人……	ISBNX-3456	9000
2	経済花子	S14/2/2	情報	100	京大出版	吉田本町	110	1999/1/2	情……	ISBNX-7890	12000

- ▶ 著者番号 → 著者名, 生年月日
- ▶ 発行所番号 → 発行所名, 発行所連絡先
- ▶ 著者名から生年月日は決まらない (同姓同名).
- ▶ 生年月日は著者番号に「関数従属」
著者番号を「決定項」と言う
- ▶ ISBN → 全ての属性が一意に決まる.
1行を選択する最小の属性集合を「キー」と言う.

§4.9 関数従属・決定項・キー(続き)

別の例

カナ氏名	漢字氏名	学科番号	学部	学部記号	学科	入学年度	年度記号	学生証番号	偶奇	種別	学科名
キョウダ	京大太郎	0435000	04	e	30	11	x	1234	0	修士	経済学研究科経済システム分析
ケイザイ	経済花子	0435000	04	e	60	10	w	6789	1	博士	経済学研究科経済システム分析

- ▶ カナ氏名/漢字氏名 × 何も決まらない
 - ▶ 学科番号 → 学科名
 - ▶ 学科番号 × 学科
 - ▶ 学部 → 学部記号
 - ▶ 学科 → 種別
 - ▶ 入学年度 → 年度記号
 - ▶ 学生証番号 → 偶奇
- キーは(学部・学科・入学年度・学生証番号)の組

§4. 9 関数従属・決定項・キー(続き)

ということは,

学籍番号	カナ氏名	漢字氏名	学科番号	学部	学部記号	学科	入学年度	年度記号	学生証番号	偶奇	種別	学科名
0430111234	キョウダ	京大太郎	0435000	04	e	30	11	x	1234	0	修士	経済学研究科経済システム分析
0460106789	ケイザイ	経済花子	0435000	04	e	60	10	w	6789	1	博士	経済学研究科経済システム分析

- ▶ 学部・学科・入学年度・学生証番号から学籍番号を作る.
- ▶ 学籍番号 → 全ての属性が一意に決まる.
学籍番号がキーとなる

§4. 10 第二正規形

表を分割する(第二正規形)

文献表

著者番号	著者名	生年月日	題名	発行所番号	ページ数	発行日	概要	ISBN	価格
1	京大太郎	T14/1/1	大正史	100	140	1926/1/1	大……	ISBNX-1234	10000
1	京大太郎	T14/1/1	昭和史	100	640	1989/1/1	昭……	ISBNX-5678	18000
1	京大太郎	T14/1/1	平成史	202	110	1999/1/1	平……	ISBNX-9012	10000
2	経済花子	S14/2/2	人物金	150	110	1999/1/1	人……	ISBNX-3456	9000
2	経済花子	S14/2/2	情報	100	110	1999/1/2	情……	ISBNX-7890	12000

発行所表

発行所番号	発行所名	発行所連絡先
100	京大出版	吉田本町
202	吉田印刷	吉田神楽岡町
150	近衛文庫	東山近衛

- ▶ 各表にキーが存在し, 他の属性が一意に決まる
- ▶ 京大太郎の出版物を削除しても, 吉田印刷に関するデータは消えない

問題点

- ▶ キー以外に関数従属が存在する. 著者名⇒生年月日など

§6. 11 第三正規形

さらに分割(第三正規形)

文献表

著者番号	題名	発行所番号	ページ数	発行日	概要	ISBN	価格
1	大正史	100	140	1926/1/1	大……	ISBNX-1234	10000
1	昭和史	100	640	1989/1/1	昭……	ISBNX-5678	18000
1	平成史	202	110	1999/1/1	平……	ISBNX-9012	10000
2	人物金	150	110	1999/1/1	人……	ISBNX-3456	9000
2	情報	100	110	1999/1/2	情……	ISBNX-7890	12000

著者表

著者番号	著者名	生年月日
1	京大太郎	T14/1/1
2	経済花子	S14/2/2

発行所表

発行所番号	発行所名	発行所連絡先
100	京大出版	吉田本町
202	吉田印刷	吉田神楽岡町
150	近衛文庫	東山近衛

- ▶ キーが決定項であり、他に関数従属性が存在しない関係
- ▶ 文献表のキーは「ISBN」
- ▶ 著者表のキーは「著者番号」
- ▶ 発行所表のキーは「発行所番号」

§6. 11 第三正規形(続き)

別の例では、

学籍表

学籍番号	カナ氏名	漢字氏名	学科番号
0430111234	キョウダ	京大太郎	0435000
0460106789	ケイザイ	経済花子	0435000

学科名表

学科番号	学科名
0435000	経済学研究科経済システム分析

学部表

学部	学部記号
04	e
94	e

学科表

学科	種別
30	修士
60	博士
61	博士

年度表

入学年度	年度記号
11	x
10	w
01	a
02	a

- ▶ 学籍表のキーは「学籍番号」
- ▶ 学科名表のキーは「学科番号」
- ▶ 学部表のキーは「学部」
- ▶ 学科表のキーは「学科」
- ▶ 年度表のキーは「入学年度」

今日はここまで