Database

第4回：論理設計と正規化

上智大学理工学部情報理工学科
高岡詠子

No reproduction or republication without written *permission*.
許可のない転載、再発行を禁止します

©2011 Eiko Takaoka All Rights Reserved.

Schedule

	日程	内容
第1回	10月6日	ガイダンス，データベースとは？
第2回	10月13日	三層スキーマ，データモデル，リレーショナルデータベースの基礎
第3回	10月20日	概念設計：概念モデルとERモデル，論理設計へ
第4回	10月27日	論理設計と正規化
第5回	11月10日	正規化，物理設計
第6回	11月17日	物理設計
第7回	11月24日	SQL言語（データベース定義）
第8回	12月1日	SQL言語（データベース操作：射影制限，結合，合併除の各演算）
第9回	12月8日	SQL
第10回	12月15日	SQL言語（ビュー定義など）
第11回	12月22日	データベース管理システム：トランザクション処理
第12回	1月5日	データベース管理システム：同時実行制御，排他制御
第13回	1月12日	同時実行制御，排他制御，デッドロック
第14回	1月19日	データベース技術動向，リレーショナル代数，まとめ

©2011 Eiko Takaoka All Rights Reserved.

今日の授業

- 論理設計
- 正規化
- Today's Quiz

©2011 Eiko Takaoka All Rights Reserved.

論理設計

- 論理データモデルの作成
- データ間の関係を処理しやすい形式（表に）にする
- データの重複をなくす
- 手順
  - エンティティを「リソース系エンティティ」と「イベント系エンティティ」に切り出す
  - 属性の抽出
  - 重複のない無駄のないデータ管理を実現するために「正規化」を行う

©2011 Eiko Takaoka All Rights Reserved.

データベース設計と3層モデル

対象世界
企業全体

①概念設計（データモデリング）

概念データ
モデル

②論理設計

論理データ
モデル

③物理設計

物理データ
モデル

外部スキーマ

論理性データ独立

概念スキーマ

物理性データ独立

内部スキーマ

データベース設計：実世界をコンピュータシステムに乗せるためのプロセス

©2011 Eiko Takaoka All Rights Reserved.

例題　一意識別子

- 販売会社のエンティティ，属性，一意識別子を考えよう

顧客
- #＊顧客番号
- ＊顧客名
- ＊住所
- 性別
- 嗜好

受注
- #＊受注番号
- ＊顧客番号
- ＊受注日

受注品目
- #＊受注番号
- #＊商品番号
- ＊受注個数

商品
- #＊商品番号
- ＊商品名
- ＊通常価格
- 割引価格
- セット価格

顧客 ―― 受注 ―― 受注品目 ―― 商品

1　N　1　N　N　1

©2011 Eiko Takaoka All Rights Reserved.

例題　履修について

- 履修のエンティティ，属性，一意識別子を考えよう

学部
＊学部コード
＊学部名

学生
#＊学生番号
＊氏名
＊学部コード
＊学科コード
サークル

履修
＊履修年度
＊学生番号
＊科目コード
＊登録日時
成績

科目
＊科目コード
＊科目名
担当教員
教室

学部 —— 学生 —— 履修 —— 科目

1　N　1　N　N　1

©2011 Eiko Takaoka All Rights Reserved.

一意識別子とキー

- 一意識別子＝　　となる属性
  - 重複するデータを一切持たない属性
  - データを一意に識別できる属性
- キーは複数あってもよいので　　とも呼ぶ
- 候補キーのうち最も適切なものが
- 2つ以上の属性から成るキーは

©2011 Eiko Takaoka All Rights Reserved.

学生リレーションの属性

学生

履修

©2011 Eiko Takaoka All Rights Reserved.

候補キー

- 学生リレーションの属性：
  - 表中のデータを識別可能にする属性
  - 氏名や生年月日：候補キーにならないものをという
- 履修：履修年度，学生番号，科目コード，成績，登録月日
  - 学生番号と科目名が二つ揃わなければ成績も登録月日も一意に特定できないから，この場合の候補キーは{　　}

©2011 Eiko Takaoka All Rights Reserved.

主キー

- 候補キーの中の代表的なキー
  - 一つの表(リレーション)内でただ一つ
  - 空ではありえない
- 学生リレーションの属性：
- 主キー候補
  - 学生番号は?同じものはありえないし空ではない
  - 学内メルアドは？同じものはありえないが登録しなくてもよいから主キーにはなりえない
- 実世界でも主体的な役割を果たすもの
- 管理性のよいもの

©2011 Eiko Takaoka All Rights Reserved.

外部キー

- エンティティ間の関連を表現する属性となる.
- エンティティを正規化していく過程で活用される.
- 1:Nの場合：N(子供エンティティ)側に設置
- 1:1の場合
  - 必須：任意の場合⇒任意側
  - 任意：任意⇒データが後から発生する側

学部コード	学部名
100	神学部
200	理工学部

1:N

学生番号	氏名	学部コード
A0807700	上智	100
A0807720	紀尾井	200

©2011 Eiko Takaoka All Rights Reserved.

ここまでの演習問題

- 2つのエンティティ間にM:Nの関係があるときそれらを1:Nの関係にする
- 手順1: 間に新しいエンティティを入れてみる
- 手順2：カーディナリティの設定

©2011 Eiko Takaoka All Rights Reserved.

ここまでの演習問題

- 手順3：属性，一意識別子の設定

©2011 Eiko Takaoka All Rights Reserved.

ここまでの演習問題

- 主キーと外部キーはどれ？

©2011 Eiko Takaoka All Rights Reserved.

論理設計

- エンティティを「リソース系エンティティ」と「イベント系エンティティ」に切り出す
- 属性の抽出
- 重複のない無駄のないデータ管理を実現するために「正規化」を行う

©2011 Eiko Takaoka All Rights Reserved.

正規化を行う目的

- を抑制することができる
- データを部品化できる＝＞ がなくなる
- ユーザにわかりやすい
- アプリケーションプログラムに依存しないDBを構築することができる
- 一つの事実を一意に表現する

多くても第　正規形にすれば十分

©2011 Eiko Takaoka All Rights Reserved.

正規化の話

受注番号	顧客ID	顧客名	日付	商品名	数量	単価
123	A21	SJハウス	2010/10/28	ワイン	10	1000
				チーズ	5	500
				フォアグラ	3	3000
				パン	7	200
124	B22	Bacchano	2010/10/28	パン	2	400
				ワイン	6	1500
				チーズ	10	200
				ビール	5	100

正規化　　状態の一例

©2011 Eiko Takaoka All Rights Reserved.

第1正規形(1NF)
2次元のフラットな形式で表現される表

- を分離
- 導出属性(導出される属性－関係従属性とは異なる－)は除去する
  - フラットな表にする方法
    エンティティを分割せず，繰り返しの存在する分だけデータを増やす
  - 繰り返しグループを分離する
    繰り返し項目の属性を別エンティティとして分離する

©2011 Eiko Takaoka All Rights Reserved.

第1正規形(1NF)

- リレーションRが第1正規形であるとは・・・
  - リレーションの各属性値が単純な値(繰り返しを使っていない)である状態
- 第1正規形ではない関係
  - 繰り返し項目を含む
  - 集合を値として持つ
  - 入れ子になっている

©2011 Eiko Takaoka All Rights Reserved.

第1正規形(1NF)

受注番号	顧客ID	顧客名	日付	商品名	数量	単価
123	A21	SJハウス	2010/10/28	ワイン	10	1000
123	A21	SJハウス	2010/10/28	チーズ	5	500
123	A21	SJハウス	2010/10/28	フォアグラ	3	3000
123	A21	SJハウス	2010/10/28	パン	7	200
124	B22	Bacchano	2010/10/28	パン	2	400
124	B22	Bacchano	2010/10/28	ワイン	6	1500
124	B22	Bacchano	2010/10/28	チーズ	10	200
124	B22	Bacchano	2010/10/28	ビール	5	100

©2011 Eiko Takaoka All Rights Reserved.

第2正規形(2NF)
主キーに非キー属性が完全関数従属している

- 複数の実態が混在している可能性ある
- ⇒冗長なデータ記録やデータの更新時異常が発生する
- 主キーが複合キーでない場合は第　正規形を満たしていると考えてよい
- 主キーが複合キーならば第　正規化する！(他の要素がないようにする)
- 複合キーの一部が他の非キーを決定できてはならない。

©2011 Eiko Takaoka All Rights Reserved.

第　正規形

		顧客名	日付	商品名	数量	単価
123	A21	SJハウス	2010/10/28	ワイン	10	1000
123	A21	SJハウス	2010/10/28	チーズ	5	500
123	A21	SJハウス	2010/10/28	フォアグラ	3	3000
123	A21	SJハウス	2010/10/28	パン	7	200
124	B22	Bacchano	2010/10/28	パン	2	400
124	B22	Bacchano	2010/10/28	ワイン	6	1500
124	B22	Bacchano	2010/10/28	チーズ	10	200
124	B22	Bacchano	2010/10/28	ビール	5	100

- 更新時異常
  - 主キーが　と　だから　をもらっていない顧客は登録できない
- A21の顧客名が変更になったら：
  - 4行変更しなくてはならない．データを部品化していれば1か所変更すれば済む

顧客ID	顧客名
A21	SJハウス
B22	Bacchano

©2011 Eiko Takaoka All Rights Reserved.

関数従属性

- 関数従属性　X -> Y
  - が決まると，顧客ID・顧客名は必然的に決まるとき
  - 顧客IDは受注番号に　する
  - 顧客名は　に関数従属する
  - 数量は受注番号と　に関数従属する
  - つまり，どちらかの属性値が決まればもう片方の属性値が決まる場合のこと

受注番号	顧客ID	顧客名	日付	商品名	数量	単価
123	A21	SJハウス	2010/10/28	ワイン	10	1000
123	A21	SJハウス	2010/10/28	チーズ	5	500

©2011 Eiko Takaoka All Rights Reserved.

完全関数従属性

- 関数従属性　X -> Y
  - 　　が決まると，顧客ID・顧客名は必然的に決まるとき
  - 顧客IDは受注番号に　　する
  - 顧客名は　　に関数従属する
- 部分関数従属性：キーの一部に非キー属性（キー以外の属性）が関数従属している状態
  - 　　と　　が主キー
  - しかし，顧客ID・顧客名は受注番号が決まると決まる．
- 完全関数従属性（関数従属性から部分関数従属性を取り払った状態
  非キー属性はキー属性がすべて決まって初めて1つに決まる
  - {　　}->顧客ID
  - {　　}-> 数量

©2011 Eiko Takaoka All Rights Reserved.

第2正規形(2NF)

- リレーションRが第2正規形であるとは？
  - 第1正規形から，部分関数従属を取り払った状態
  - リレーションRが第1正規形で，かつ，すべての非キー属性が主キーに対して　　従属するとき，Rは第2正規形であるという．

©2011 Eiko Takaoka All Rights Reserved.

第　正規形

受注番号	顧客ID	顧客名	日付	商品名	数量	単価
123	A21	SJハウス	2010/10/28	ワイン	10	1000
123	A21	SJハウス	2010/10/28	チーズ	5	500
123	A21	SJハウス	2010/10/28	フォアグラ	3	3000
123	A21	SJハウス	2010/10/28	パン	7	200
124	B22	Bacchano	2010/10/28	パン	2	400
124	B22	Bacchano	2010/10/28	ワイン	6	1500
124	B22	Bacchano	2010/10/28	チーズ	10	200
124	B22	Bacchano	2010/10/28	ビール	5	100

部分関数従属を取り払った状態

第　正規形

受注番号	顧客ID	顧客名	日付
123	A21	SJハウス	2010/10/28
124	B22	Bacchano	2010/10/28

©2011 Eiko Takaoka All Rights Reserved.

第　正規形

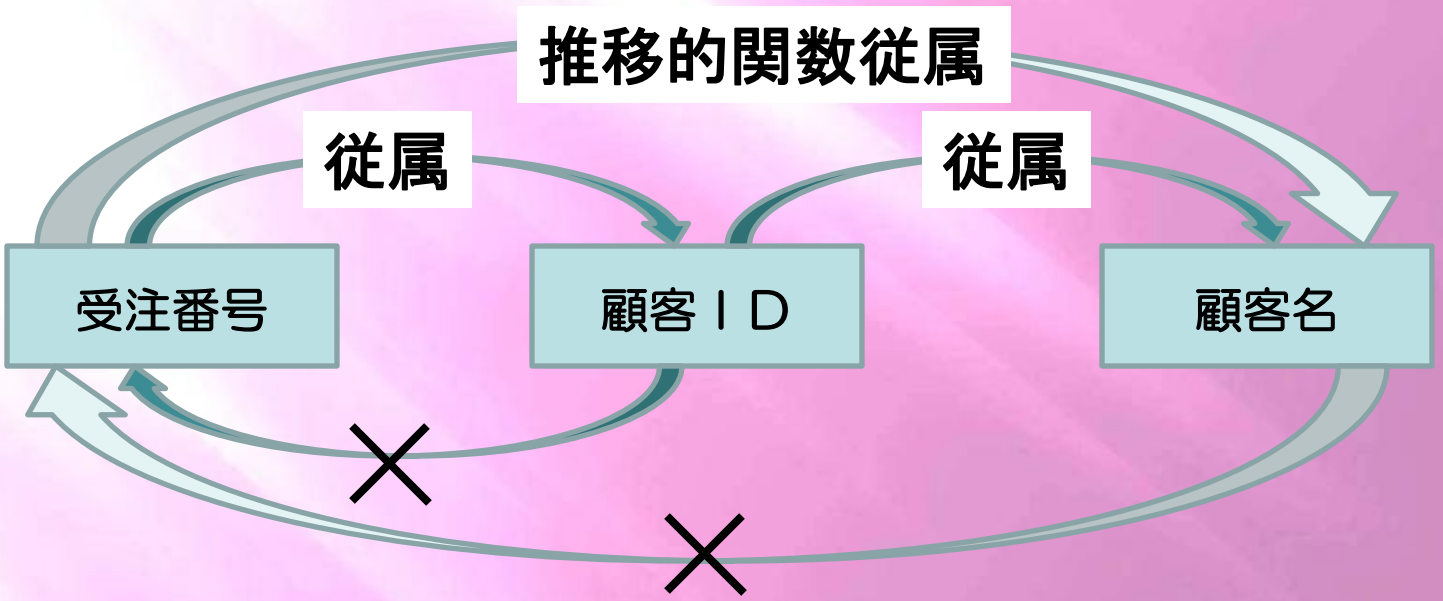
受注番号	顧客ID	顧客名	日付
123	A21	SJハウス	2010/10/28
124	B22	Bacchano	2010/10/28
125	B22	Bacchano	2010/10/30

- この場合の更新時異常
- 受注番号123の受注が　　されたら：
  - 顧客IDのA21,顧客名SJハウスという組み合わせの顧客履歴が　　してしまう！！
  - 受注が常時ない場合はデータが消えてしまう？？

©2011 Eiko Takaoka All Rights Reserved.

第3正規形(3NF)
主キー以外の列の値が他の列の値を決定しない

- があればそれらを分割する



逆は成り立たない

©2011 Eiko Takaoka All Rights Reserved.

第3正規形(3NF)

- リレーションRが第3正規形であるとは？
  - 第2正規形から，従属を取り払った状態
  - リレーションRが第2正規形で，かつ，リレーションに非キー属性があればそれらすべてが主キーに非推移的に関数従属するとき，Rは第3正規形であるという．

©2011 Eiko Takaoka All Rights Reserved.

第　正規形

[illegible]	[illegible]ID	顧客名	日付
123	A21	SJハウス	2010/10/28
124	B22	Bacchano	2010/10/28
125	B22	Bacchano	2010/10/30

この表をさらに正規化してみよう

受注番号	商品名	数量	単価
123	ワイン	10	1000
123	チーズ	5	500
123	フォアグラ	3	3000
123	パン	7	200
124	パン	2	400
124	ワイン	6	1500
124	チーズ	10	200
124	ビール	5	100

関数従属を
取り払った状態

第　正規形

[illegible]	顧客ID	日付
123	A21	2010/10/28
124	B22	2010/10/28
125	B22	2010/10/30

顧客ID	顧客名
A21	SJハウス
B22	Bacchano

これが正規化！！

©2011 Eiko Takaoka All Rights Reserved.

第1正規形にする 正規化の例

受注No	受注日	受注先	商品	数量	単価	合計金額
1	00/10/01	A	S	3	1,000	4,900
			T	2	950	
2	00/10/01	B	S	1	1,000	22,000
			U	10	1,200	
			V	5	1,800	
3	00/10/02	B	T	8	950	7,60
4	00/10/02	C	U	25	1,200	30,00

レコードに複数の情報が入っている

繰り返しのグループを分解する

©2011 Eiko Takaoka All Rights Reserved.

第　正規形

受注No	商品	数量	単価
1	S	3	1,000
1	T	2	950
2	S	1	1,000
2	U	10	1,200
2	V	5	1,800
3	T	8	950
4	U	25	1,200

主キーに非キー属性が
していない

{商品}->{　　}
{　　}->{数量}

単価は、主キーの
一部によって
決定される

主キーの一部が他の属性を決定できる

第　正規形

受注No	受注日	受注先	合計金額
1	00/10/01	A	4,900
2	00/10/01	B	22,000
3	00/10/02	B	7,600
4	00/10/02	C	30,000

©2011 Eiko Takaoka All Rights Reserved.

受注No	商品	数量	単価
1	S	3	1,000
1	T	2	950
2	S	1	1,000
2	U	10	1,200
2	V	5	1,800
3	T	8	950
4	U	25	1,200

そこで，主キーの一部{商品}->{単価}の部分を切り離す

関数従属を取り払った状態

受注No	商品	数量
1	S	3
1	T	2
2	S	1
2	U	10
2	V	5
3	T	8
4	U	25

商品	単価
S	1,000
T	950
V	1,800
U	1,200

推移的関数従属がないので
第　正規形

©2011 Eiko Takaoka All Rights Reserved.

主キー

外部キー

受注No	商品	数量
1	S	3
1	T	2
2	S	1
2	U	10
2	V	5
3	T	8
4	U	25

商品	単価
S	1,000
T	950
V	1,800
U	1,200

主キー以外の列の値が他の列の値を決定しないから

受注No	受注日	受注先	合計金額
1	00/10/01	A	4,900
2	00/10/01	B	22,000
3	00/10/02	B	7,600
4	00/10/02	C	30,000

すべて
第　正規形

©2011 Eiko Takaoka All Rights Reserved.

例題をやってみよう

(1)次のデータに関する説明として適切なものはどれか。

ア)正規化は行われていない。

イ)第1正規化だけ行われている。

ウ)第2正規化まで行われている。

エ)第3正規化まで行われている。

オ)第4正規化まで行われている。

商品コード	商品名	顧客コード	顧客名	単価	数量
001001	Aマウス	0015	東京商事	2,500	12
001002	Bマウス	0020	埼玉工業	3,900	38

©2011 Eiko Takaoka All Rights Reserved.

商品コード	商品名	顧客コード	顧客名	単価	数量
001001	Aマウス	0015	東京商事	2,500	12
001002	Bマウス	0020	埼玉工業	3,900	38

- 繰り返しはない→第　正規形
- 候補キー
  - 商品コード⇒　と　がきまる
  - 顧客コード⇒　が決まる
  - 商品コード＋顧客コード⇒
- 主キーの一部が他の属性を決定できる
- ということは第　正規形ではない

©2011 Eiko Takaoka All Rights Reserved.

学生は複数の授業を履修し，授業は複数の学生が履修する。さらに，どの学生も一つの授業を1回だけ履修する。正規化してみよう

学生コード	学生名	授業コード	授業名	履修年度	成績

©2011 Eiko Takaoka All Rights Reserved.

学生コード	学生名	授業コード	授業名	履修年度	成績

主キーに非キー属性が　　していないから
第　　正規形ではない

{　　}->{履修年度、成績}
{　　コード} -> {学生名}
{　　コード} -> {授業名}

正規化

©2011 Eiko Takaoka All Rights Reserved.

下記の表は第?正規形？

商品ID	名前	産地	国名	価格
11	いちじく	愛知	日本	500
12	ナツメヤシ	エルサレム	イスラエル	600
13	ライチ	上海	中国	400

©2011 Eiko Takaoka All Rights Reserved.

主キーに非キー属性が完全関数従属しているが
主キー以外のキーがほかの列を決定できてしまうので　　ではない

である

商品ID	名前	産地	国名	価格
11	いちじく	愛知	日本	500
12	ナツメヤシ	エルサレム	イスラエル	600
13	ライチ	上海	中国	400

©2011 Eiko Takaoka All Rights Reserved.