

# リレーショナルデータベース設計 (その2)

## E-R モデルと正規化

- E-Rモデル:
  - エンティティ(リレーション)とエンティティ(リレーション)の関連を示す.
- 正規化:
  - リレーション自身(テーブル)の正しさをチェックする.

## 正規化とは

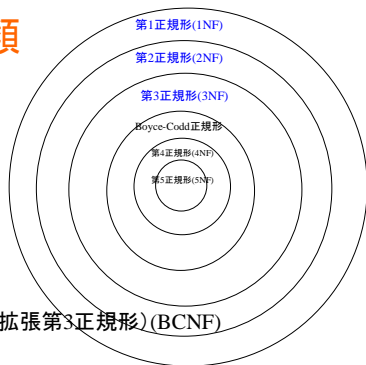
- 何らかの問題をもつリレーションを, 問題をもたない複数のリレーションに変換するプロセス
- リレーション(テーブル)のデータ(列の値:属性)間の意味的な関係に注目し, できる限り不必要なデータ項目の重複がないようなリレーションに分割していく

## 正規化の目的

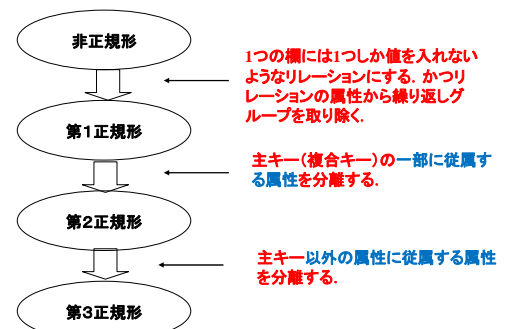
- データの重複を減少させること
  - データ項目を正しいキーと結びつけることにより, データの冗長度を排除すること
- データ操作を, 内容の矛盾を起こさずに, 正確に実行できるようにすること
  - リレーション内の行ごとのカラムが単一のデータ項目だけを含み, 繰り返しグループがないようにデータを整える

## 正規化の種類

- (非正規形)
- 第1正規形(1NF)
- 第2正規形(2NF)
- 第3正規形(3NF)
- Boyce-Codd正規形(拡張第3正規形)(BCNF)
- 第4正規形(4NF)
- 第5正規化(5NF)



## 正規化の手順



## 非正規リレーションと正規リレーション

非正規リレーション

科目番号	科目名	担当者
001	実験2	高田
		任
002	実験3	任
		吉田

正規リレーション

科目番号	科目名	担当者
001	実験2	高田
001	実験2	任
002	実験3	任
002	実験3	吉田

7

## 第1正規形(1NF)

第1正規形の条件

- (1) 1つの平坦な2次元のテーブルにすること。つまり2次元のテーブルに並び、1つの欄には1つしか値を入れないリレーション。
- (2) リレーションの属性から繰り返しグループを取り除いたリレーション。

注文番号	注文日	顧客コード	商品コード1	数量1	単価1	商品コード2	数量2	単価2
TM001	2000年9月1日	010045	S0001	2	¥1000	H0001	2	¥1000
TM003	2000年8月1日	991016	S0003	3	¥500	H0005	1	¥5000
TM005	2001年6月1日	981226	S0001	1	¥2000			

注文番号	注文日	顧客コード
TM001	2000年9月1日	010045
TM003	2000年8月1日	991016
TM005	2001年6月1日	981226

注文番号	商品コード	数量	単価
TM001	S0001	2	¥1000
TM003	S0003	3	¥500
TM005	S0001	1	¥2000
TM001	H0001	2	¥1000
TM003	H0005	1	¥5000

8

## 第2正規形(2NF)

第2正規化はリレーションから関数従属しているデータ項目を取り出して別の新しいリレーションを作ること

学籍番号	課題番号	課題名	評価点
KC98001	1	A	90
KC98001	2	B	70
KC98002	1	A	50
KC98003	3	C	80
KC98004	4	D	75

$X \rightarrow Y$  課題番号  $\rightarrow$  課題名

9

学籍番号	課題番号	課題名	評価点
KC98001	1	A	90
KC98001	2	B	70
KC98002	1	A	50
KC98003	3	C	80
KC98004	4	D	75

$X \rightarrow Y$  課題番号  $\rightarrow$  課題名

成績表

学籍番号	課題番号	評価点
KC98001	1	90
KC98001	2	70
KC98002	1	50
KC98003	3	80
KC98004	4	75

課題表

課題番号	課題名
1	A
2	B
3	C
4	D

ER図



10

## 第3正規形(3NF)

第3正規化とは、推移的な関数従属を排除し別の新しいリレーションを作ること

受注番号  $\rightarrow$  顧客番号  $\rightarrow$  (顧客会社名, 顧客担当者名)

受注表

受注番号	受注年月日	顧客番号	顧客会社名	顧客担当者名
------	-------	------	-------	--------

受注表

受注番号	受注年月日	顧客番号
------	-------	------

受注表

受注番号	受注年月日	顧客番号	顧客会社名	顧客担当者名
------	-------	------	-------	--------

顧客番号	顧客会社名	顧客担当者名
------	-------	--------

M

1

ER図



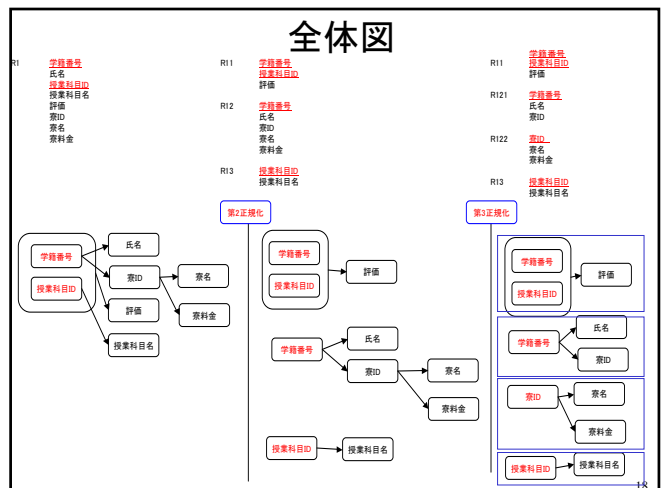
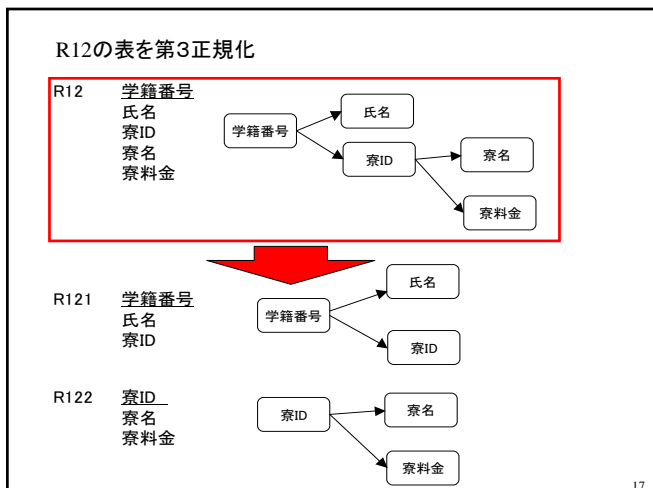
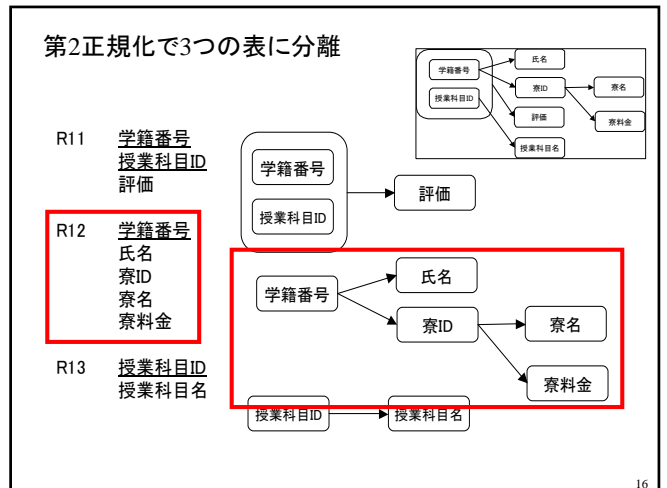
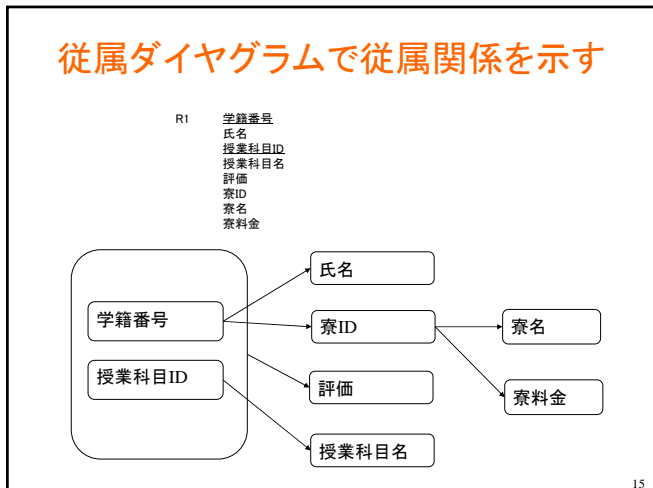
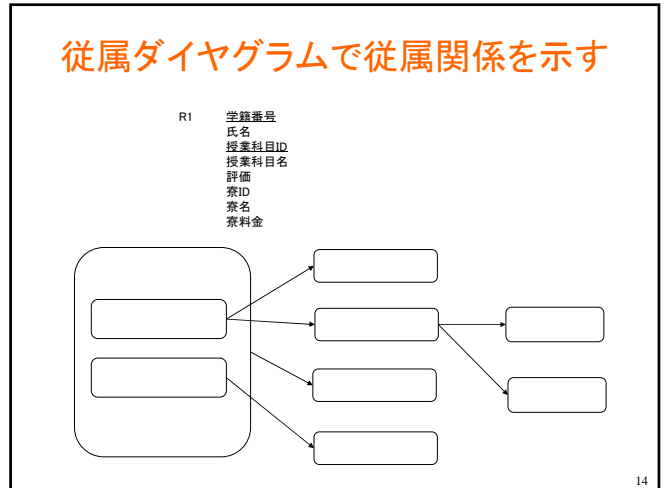
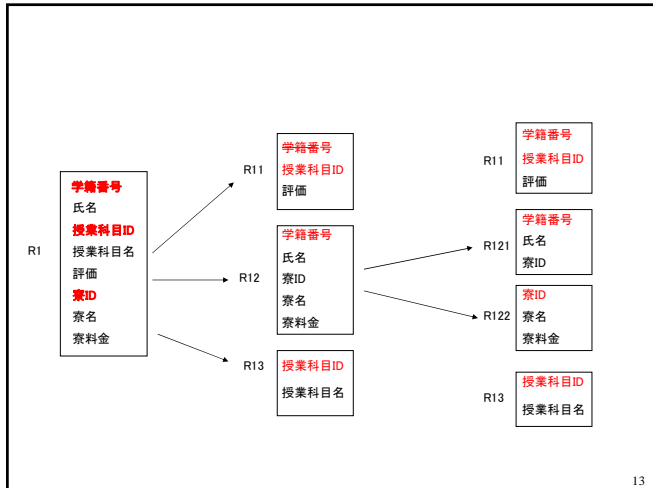
11

## 次のリレーションを正規化せよ

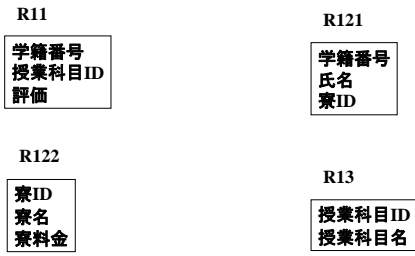
R1

学籍番号
氏名
授業科目ID
授業科目名
評価
寮ID
寮名
寮料金

12



問題: R11, R121, R122, R13のそれぞれの主キーまたは外部キーを示しなさい。



19

問題: E-Rモデル(E-R図)を示しなさい。

21

	R11	R121	R122	R13
R11	\	\	\	\
R121	○	\	\	\
R122	×	○	\	\
R13	○	×	×	\

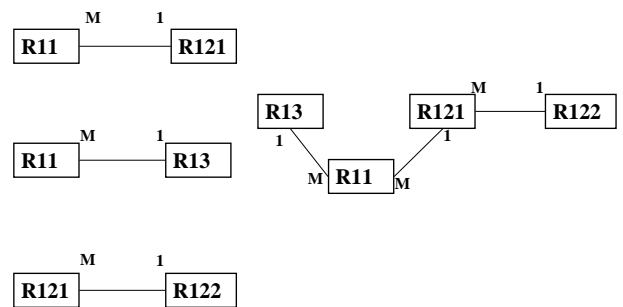
22

	R11	R121	R122	R13
R11	<b>R11</b>	\	\	\
R121	○	<b>R121</b>	\	\
R122	×	○	<b>R122</b>	\
R13	○	×	×	<b>R13</b>

23

	<b>R11</b>	<b>R121</b>	<b>R122</b>	<b>R13</b>
<b>R11</b>	○	\	\	\
<b>R121</b>	×	○	\	\
<b>R122</b>	○	×	×	\
<b>R13</b>	○	×	×	○

24



25

複合キーの一部であり、外部キーであり

X → Y 課題番号 → 課題名  
(X, Y) → Z (学籍番号, 課題番号) → 評価点

成績表

キー

学籍番号	課題番号	評価点
KC98001	1	90
KC98001	2	70
KC98002	1	50
KC98003	3	80
KC98004	4	75

主キー：学籍番号+課題番号

外部キー：課題番号

課題表

キー

課題番号	課題名
1	A
2	B
3	C
4	D

主キー：課題番号

学籍番号	課題番号	課題名	評価点
KC98001	1	A	90
KC98001	2	B	70
KC98002	1	A	50
KC98003	3	C	80
KC98004	4	D	75



複合キーの一部であり、外部キーであり

X → Y 課外活動 → 料金  
(X, Y) → Z (学籍番号, 課外活動) → 料金

学生・課外

学籍番号	課外活動
100	スキー
100	ゴルフ
150	水泳
175	スケート
175	スキー

学籍番号	課外活動	料金
100	スキー	200
100	ゴルフ	150
150	水泳	50
175	スケート	50
175	スキー	200

主キー：学籍番号+課外活動

外部キー：課外活動

課外活動

課外活動	料金
スキー	200
ゴルフ	150
水泳	50
スケート	50

主キー：課外活動



複合キーの一部ではない、外部キーであり

X → Y → Z  
学籍番号 → 寮 → 料金

学生・寮

学籍番号	寮
100	寮1
150	寮2
200	寮1
250	寮3
300	寮1

主キー：学籍番号

外部キー：寮

寮

寮	料金
寮1	1200
寮2	1100

主キー：寮

学籍番号	寮	料金
100	寮1	1200
150	寮2	1100
200	寮1	1200
250	寮3	1100
300	寮1	1200



設計のトレードオフ

- 正規化の度合いが高い場合に
  - 結合により、種々の見方が可能な柔軟なデータ構造である。その反面パフォーマンス(performance)が悪い。
  - データの正確性が高まり、整合性に問題がすくない。
  - テーブルの分割が多いため、行が非常に少ないテーブルができることがある。
  - 列の数が少ない。
- 正規化の度合いが低い場合に
  - テーブルの結合が減少するので、パフォーマンスが良い。
  - 物理アクセスが単純になる。
  - 行の多いテーブルができることがある。
  - 列の数が多。

顧客(顧客番号, 顧客名, 郵便番号, 市, 県)

?

顧客(顧客番号, 顧客名, 郵便番号)

郵便番号(郵便番号, 市, 県)

