

データベース(第4回)

情報工学科 木村昌臣

リレーショナルデータベース設計 理論





【復習】第一正規形(1/2)

- リレーションのあるべき姿
- タップルの属性値(レコードのフィールド値)に繰り返しがあってはいけない。
- 繰り返すなら、いっそ新しいタップル(レコード)を作るべし！

【復習】第一正規形(2/2)

社員番号	社員名	趣味
L001	木村	ドライブ、カラオケ
L002	坂本	テニス、音楽、山登り

フィールド内で繰り返すのではなく、
レコードとして繰り返す！

社員番号	社員名	趣味
L001	木村	ドライブ
L001	木村	カラオケ
L002	坂本	テニス
L002	坂本	音楽
L002	坂本	山登り

正規化



更新時異常(1/4)

- 第一正規化だけでは、実際にリレーショナルデータベースを設計・運用するときに不都合がでる場合がある。
- どんなときか？
 - タップルの挿入時
 - タップルの削除時
 - タップルの修正時

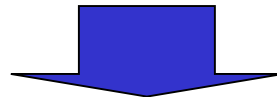
タプル挿入時異常(2/4)

注文

主キー

顧客名	商品名	数量	単価
A商店	テレビ	3	198,000
Bマート	テレビ	10	198,000
Bマート	パソコン	5	59,800
C社	ゲーム機	1	29,800

リレーション注文は、商品名と単価の属性をもつので
新製品である電子レンジ(単価74,800円)の単価情報を持たせたい



t=(null, 電子レンジ, null, 74,800, null) を挿入することになるが、
顧客名は主キーなのでこれは**挿入できない!!!**

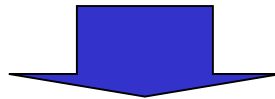
タプル削除時異常(3/4)

注文

主キー

顧客名	商品名	数量	単価
A商店	テレビ	3	198,000
Bマート	テレビ	10	198,000
Bマート	パソコン	5	59,800
C社	ゲーム機	1	29,800

リレーション注文から、C社のゲーム機注文記録を削除したい。



C社の注文記録を削除すると、ゲーム機の単価情報もなくなってしまう！！

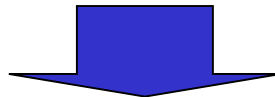
タプル修正時異常(4/4)

注文

主キー

顧客名	商品名	数量	単価
A商店	テレビ	3	198,000
Bマート	テレビ	10	198,000
Bマート	パソコン	5	59,800
C社	ゲーム機	1	29,800

テレビの単価が間違っていた。修正しなければ。



テレビの単価が変わったという**単一の事象が起こっただけなのに**、
テレビ注文**関連のタプルのすべて**(A商店、Bマート)を修正しなければならない

更新時異常が起こる原因(1/3)

注文

主キー

顧客名	商品名	数量	単価
A商店	テレビ	3	198,000
Bマート	テレビ	10	198,000
Bマート	パソコン	5	59,800
C社	ゲーム機	1	29,800

商品がいくらするか

顧客がどの商品をどれだけ注文したか

ひとつのリレーションに複数の事象
が含まれている！！



更新時異常を解決するには(2/3)

ひとつのリレーションには
ひとつの事象しか格納しないようにするべし

つまり……？

更新時異常を解決するには(3/3)

……事象ごとにリレーションを分ける

顧客名	商品名	数量	単価
A商店	テレビ	3	198,000
Bマート	テレビ	10	198,000
Bマート	パソコン	5	59,800
C社	ゲーム機	1	29,800

顧客がどの商品を
どれだけ注文したか

顧客名	商品名	数量
A商店	テレビ	3
Bマート	テレビ	10
Bマート	パソコン	5
C社	ゲーム機	1

商品がいくらするか

商品名	単価
テレビ	198,000
パソコン	59,800
ゲーム機	29,800

情報無損失分解 (1/4)

事象ごとにリレーションを分けても
もとのリレーションを再現できる

注文

顧客名	商品名	数量	単価
A商店	テレビ	3	198,000
Bマート	テレビ	10	198,000
Bマート	パソコン	5	59,800
C社	ゲーム機	1	29,800

自然結合

注文[顧客名,商品名,数量]

顧客名	商品名	数量
A商店	テレビ	3
Bマート	テレビ	10
Bマート	パソコン	5
C社	ゲーム機	1

顧客がどの商品を
どれだけ注文したか

注文[商品名,単価]

商品名	単価
テレビ	198,000
パソコン	59,800
ゲーム機	29,800

商品がいくらするか

自然結合演算 (2/4)

社員

社員番号	社員名	給与	所属コード
E3059	城島 茂	22万円	100
E3015	宇多田 ヒカル	22万円	200
E0201	桑田 圭祐	57万円	200
E4119	島谷 ひとみ	18万円	100
E0304	子門 正人	51万円	300

部門

部門コード	部門名
100	人事部
200	営業部
300	経理部

$\theta = "="$ の場合

社員*部門

社員番号	社員名	給与	所属コード	部門名
E3059	城島 茂	22万円	100	人事部
E3015	宇多田 ヒカル	22万円	200	営業部
E0201	桑田 圭祐	57万円	200	営業部
E4119	島谷 ひとみ	18万円	100	人事部
E0304	子門 正人	51万円	300	経理部

社員と部門
の自然結合

共通部分は
片方だけ残す

情報無損失分解 (3/4)

リレーションを分解することにより、
更新時異常が解決

注文[商品名,単価]

商品名	単価
テレビ	198,000
パソコン	59,800
ゲーム機	29,800

商品がいくらするか

注文[顧客名,商品名,数量]

顧客名	商品名	数量
A商店	テレビ	3
Bマート	テレビ	10
Bマート	パソコン	5
C社	ゲーム機	1

顧客がどの商品をどれだけ注文したか

1. タップル挿入時異常→単価情報を持たせたい→左側のリレーションを更新
2. タップル削除時異常→注文記録を削除すると単価情報が消える
→右側のリレーションのタップルのみ削除
3. タップル修正時異常→単一の事象が起こっただけなのに関連タップルを
全部修正しなければならない
→左側のリレーションを修正

解決!

解決!

解決!

情報無損失分解(一般論)(4/4)

勝手に分解するともとのリレーションが再現できない

R

顧客名	商品名	数量	単価
A商店	テレビ	3	198,000
Bマート	テレビ	10	198,000
Bマート	パソコン	5	59,800
C社	ゲーム機	1	29,800

R1

顧客名	商品名
A商店	テレビ
Bマート	テレビ
Bマート	パソコン
C社	ゲーム機

R2

商品名	数量	単価
テレビ	3	198,000
テレビ	10	198,000
パソコン	5	59,800
ゲーム機	1	29,800

$$R1 * R2 \neq R$$

顧客名	商品名	数量	単価
A商店	テレビ	3	198,000
A商店	テレビ	10	198,000
Bマート	テレビ	3	198,000
Bマート	テレビ	10	198,000
Bマート	パソコン	5	59,800
C社	ゲーム機	1	29,800

Rにない
タプルができて
いる
(結合の罫)



多値従属性 (1/4)

- 今までの情報無損失分解は、タプルレベルでの議論
- 本来、情報無損失分解は、リレーションスキーマレベル(要するにテーブル設計)で議論されるべき



多値従属性 (2/4)

リレーションスキーマ $R(A_1, A_2, \dots, A_l, B_1, B_2, \dots, B_m, C_1, C_2, \dots, C_n)$ について R に多値従属性(MVD)が存在するとは

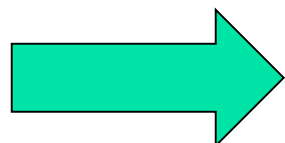
**B_1, B_2, \dots, B_m の情報 (の集合) が
 C_1, C_2, \dots, C_n に属する情報に依存せず
 A_1, A_2, \dots, A_l の情報のみで決定すること**

$$A_1, A_2, \dots, A_l \twoheadrightarrow B_1, B_2, \dots, B_m$$

多値従属性(3/4) 具体例-1

社員	子供	本人スキル
高田健二	タロウ	ネットワークエンジニア
高田健二	ユキコ	ネットワークエンジニア
吉川裕子	フウタ	データベースエンジニア

社員が決まれば、その社員の子供の**集合が求まる**



子供には社員に対する多値従属性がある
社員→→子供

多値従属性(4/4) 具体例-2

商品がいくら
するか

↓
単価は商品名に
多値従属

↓
だれがいくつ
買ったかに
依存しない

注文

顧客名	商品名	数量	単価
A商店	テレビ	3	198,000
Bマート	テレビ	10	198,000
Bマート	パソコン	5	59,800
C社	ゲーム機	1	29,800

顧客がどの商品を
何個いくら分注文したか

↓
商品名と数量は
顧客名に多値従属

↓
商品の単価に
依存しない

商品

商品名	単価
テレビ	198,000
パソコン	59,800
ゲーム機	29,800

発注

顧客名	商品名	数量
A商店	テレビ	3
Bマート	テレビ	10
Bマート	パソコン	5
C社	ゲーム機	1

関数従属性 (1/4)

注文

顧客名	商品名	数量	単価
A商店	テレビ	3	198,000
Bマート	テレビ	10	198,000
Bマート	パソコン	5	59,800
C社	ゲーム機	1	29,800

商品名がひとつ決まると単価が
ひとつ決まる!
(多値従属性の特別な場合)

商品

商品名	単価
テレビ	198,000
パソコン	59,800
ゲーム機	29,800

単価は、商品名にしか依存せず、
商品名が決まると単価がただひとつ決まる



関数従属性(定義) (2/4)

リレーショナルスキーマ $R(A_1, \dots, A_l, B_1, \dots, B_m, C_1, \dots, C_n)$ に
関数従属性

$$A_1 A_2 \dots A_l \rightarrow B_1 B_2 \dots B_m$$

が存在するとは、次の条件が成立することをいう。

R を R のインスタンスとするとき、

$$(\forall t, t' \in R) (t[A_1 A_2 \dots A_l] = t'[A_1 A_2 \dots A_l] \Rightarrow t[B_1 B_2 \dots B_m] = t'[B_1 B_2 \dots B_m])$$

候補キー

- リレーションスキーマ $R(A_1, A_2, \dots, A_n)$ の属性集合 K が候補キーであるとは、以下を満たすことをいう。
 - R を R の任意のインスタンスとする
 - ① $(\forall t, t' \in R)(t[K] = t'[K] \Rightarrow t = t')$
 - ② K のどのような真部分集合 H についても①は成り立たない

発注

顧客名	商品名	数量	金額
A商店	テレビ	3	594,000
Bマート	テレビ	10	1,980,000
Bマート	パソコン	5	299,000
C社	ゲーム機	1	29,800

$K = \{\text{顧客名}, \text{商品名}\}$

属性集合: 属性名のまとまり・集合



完全関数従属性

- 関数従属性 $X \rightarrow Y$ で、 X' を X の真部分集合とすると、 $X' \rightarrow Y$ が成立しない場合、 Y は X に完全関数従属しているという。

発注

顧客名	商品名	数量	金額
A商店	テレビ	3	594,000
Bマート	テレビ	10	1,980,000
Bマート	パソコン	5	299,000
C社	ゲーム機	1	29,800

$\{\text{顧客名}, \text{商品名}\} \rightarrow \{\text{数量}\}$
は成り立つが、
 $\{\text{顧客名}\} \rightarrow \{\text{数量}\}$
 $\{\text{商品名}\} \rightarrow \{\text{数量}\}$
は成り立たない。


$$X \rightarrow Y \Rightarrow X \twoheadrightarrow Y$$

- リレーションスキーマ ***R*** で,
 $X \rightarrow Y$ なら $X \twoheadrightarrow Y$ であるが、逆は成り立たない。
 - なぜなら、「 $X \rightarrow Y$ 」は X がきまると Y が一意にきまるという特殊な「多値従属性」であるから。
 - 逆に、一般に「多値従属性」があっても、 X がきまったからといって Y はひとつには決まらない。
- よって、関数従属性は多値従属性の十分条件
 - よって、 $A_1 A_2 \dots A_l \rightarrow B_1 B_2 \dots B_m$ はリレーションスキーマ $R(\{A\}, \{B\}, \{C\})$ を情報無損失分解するための十分条件である。