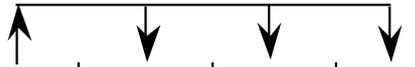


正規化について (承前)

(a) 種々の情報が混在する表



通知番号	学生番号	氏名	通知日	授業科目 番号	受講年度	成績	科目名	教官名
2001	1011	新井 一	2009.3.1	201	1989	70	電子工学	夏目素石
2001	1011	新井 一	2009.3.1	203	1990	65	英語	Peter Smith
2002	1015	井上太郎	2009.3.5	202	1991	90	文学	唯野耽二
2002	1015	井上太郎	2009.3.5	203	1992	50	英語	Peter Smith
2003	1123	宇野三太	2009.3.5	202	1992	80	文学	唯野耽二
2004	1238	大村花子	2009.3.7	203	1991	65	英語	Peter Smith
2005	1350	加藤雄一	2009.3.7	201	1993	81	電子工学	夏目素石

このような表があったとする。各人に一つの通知を発行し、通知番号を付すものとする。これはマス目 ( 1 箇の欄 ) に入るものが一つだけであるから第 1 正規化されている。

通知番号が定まると学生番号、氏名、通知日は一意に定まる。成績は通知番号、授業科目番号、受講年度 ( の属性の集まり ) を定めてはじめて定まる。

この表では特定のタプル ( 行 ) を識別するのに、通知番号、授業科目番号、受講年度を用いればよい、すなわちこれらの組み合わせをプライマリキーとすればよい ( 図ではプライマリキーに対応する属性名に下線を付している。以下の図でも同様 ) 。

ある属性の値が定まると他の属性値が定まるとき、関数的に従属している ( または単に従属している ) という。学生番号、氏名、通知日は通知番号に従属する。科目名と教官名は授業科目番号に従属する。成績は通知番号、授業科目番号、受講年度の属性の複合したものに従属する。

(a) において通知番号、授業科目番号、受講年度が定まるとき、通知日は定まるが、通知日を特定するには通知番号だけ定まれば足りる。これに反し成績はこの 3 つの属性が同時に定まらないと定まらない。このとき成績は 3 つの属性の複合したものに完全従属しているという。また、通知日はこの 3 つの属性の複合したものに完全従属はしていない。

ここでは複合属性を定めてはじめて定まるもの（成績）と、複合属性の一部（通知番号）を定めれば定まるもの（学生番号、氏名、通知日）が混在している。このような混在したままの表を用いると同じデータをくり返すことになり無駄や矛盾が多くなる。そこでこのような従属関係を整理して元の表を分割すると次のようになる。

なお、この過程で同一のタプルがくり返し発生する場合は一つに統合する（例 (c) で 201 電子工学 夏目素石 (c') 参照）。このような整理はデータの重複による記憶領域の無駄やデータの矛盾が減ることを意味する。

(b)

通知番号	学生番号	氏名	通知日
2001	1011	新井 一	2009.3.1
2002	1015	井上太郎	2009.3.5
2003	1123	宇野三太	2009.3.5
2004	1238	大村花子	2009.3.7
2005	1350	加藤雄一	2009.3.7

(c) 授業科目表

授業科目番号	科目名	教官名
201	電子工学	夏目素石
202	文学	唯野耽二
203	英語	Peter Smith

(d) 成績表

通知番号	授業科目番号	受講年度	成績
2001	201	1989	70
2001	203	1990	65
2002	202	1991	90
2002	203	1992	50
2003	202	1992	80
2004	203	1991	65
2005	201	1993	81

(c') 参考図 ((c) の前段階)

授業科目番号	科目名	教官名
201	電子工学	夏目素石
203	英語	Peter Smith
202	文学	唯野耽二
203	英語	Peter Smith
202	文学	唯野耽二
203	英語	Peter Smith
201	電子工学	夏目素石

このように、与えられた第1正規化された表をプライマリキーとそれに完全従属する属性だけからなる表に分解したとき第2正規化された（第2正規形）という。第2正規形は同時に第1正規形でもある。つまり第1正規形の条件を満たして同時に第2正規形の条件を満たすものを第2正規形と呼ぶから当然そうなる。

ここで表(b)について考える。通知番号が定めれば学生番号、氏名、通知日は定まるが、そのうち氏名については内在的に学生番号が定めれば定まる。通知番号が定まると学生番号が定まり、学生番号が定まることによって氏名が定まると考えるべきである。このような関係を氏名は通知番号に推移従属するという。

第2正規形の表から、推移従属の部分を独立させて分離するとき、第3正規化（第3正規形）という。次の図に、このように(b)を第3正規化したものを示す。なお、第2正規形で、第3正規化すべきものがなければそれは第3正規形といってよい。

(b') (再掲)

通知番号	学生番号	氏名	通知日
2001	1011	新井 一	2009.3.1
2002	1015	井上太郎	2009.3.5
2003	1123	宇野三太	2009.3.5
2004	1238	大村花子	2009.3.7
2005	1350	加藤雄一	2009.3.7

(e) 通知表

通知番号	学生番号	通知日
2001	1011	2009.3.1
2002	1015	2009.3.5
2003	1123	2009.3.5
2004	1238	2009.3.7
2005	1350	2009.3.7

(f) 学生表

学生番号	氏名
1011	新井 一
1015	井上太郎
1123	宇野三太
1238	大村花子
1350	加藤雄一

結局、元の表 (a) (種々の情報が混在する表) は (e) (通知表)、(f) (学生表)、(c) (授業科目表)、(d) (成績表) に分解された。全体として第3正規化されている。

このような分解によって正規化する操作は、日常的な操作というよりはデータベースを設計する段階で種々の考察の一環として行うべきものである。

仮に、(e)、(f)、(c)、(d) のような表があったとすると、これらは独立に維持管理することができる。日常業務としては (a) のような表がほしいことがあるが、これはビューとして扱うのがよい。必要があれば自然結合によって (a) を (ビューとして) 作ることがで

きる。(a)のような表だけで (a) を実表として) 教務関係のデータを処理するより、(e)、(f)、(c)、(d)のような表を実表として処理する (a) は必要なときビューを作る) 方がはるかに合理的である。なお、この例は人為的なものなので実際上の妥当性は不明である。