1000-9825/2002/13(08)1488-06 ©2002 Journal of Software 软 件 学 报 Vol.13, No.8

相应于无冲突依赖的规范化对象模式森林*

吴永辉[1,2,3], 周傲英[1,2]

[1](复旦大学 计算机科学与工程系,上海 200433);
[2](复旦大学 智能信息处理开放实验室,上海 200433);
[3](中国科学院 软件研究所 计算机科学重点实验室,北京 100080)
E-mail: slwu@fudan.edu.cn
http://www.cs.fudan.edu.cn

摘要: 首先概括对象依赖、无冲突对象依赖集合、规范化对象模式森林和复杂对象模式规范化设计算法的基本概念和性质;然后给出并证明相应于无冲突对象依赖集合 M 的规范化对象模式森林 F 的性质:$P(F)$是惟一的、不可分解的规范化对象模式森林的路径集合;$M\Leftrightarrow OD(F)\Leftrightarrow\bowtie P(F)$;$P(F)$是无$\alpha$环的.这对于面向对象信息系统的开发有一定的意义.

关 键 词: 对象模式;对象依赖;无冲突;森林;路径
中图法分类号: TP311 文献标识码: A

随着 Internet 和 Web 的发展和普及,诸如数字图书馆等存储海量信息的系统应运而生.这些系统的开发存在一系列的问题.在复杂对象模式规范化设计的研究中有代表性的有 Z.Tari 等人[1]和 W.Y.Mok 等人[2]的工作.Z.Tari 等人的工作仅假设用户的描述不存在冲突,而没有讨论无冲突用户描述在规范化设计中具有的特性;W.Y.Mok 等人的工作则没有考虑冲突的情况.

我们参考了已有的工作[1,2]和关系数据库理论[3],也在复杂对象模式规范化设计这方面进行了探索[4~7]:定义对象依赖表示对象间的语义关系;定义规范化对象模式森林作为对象范式;给出复杂对象模式的规范化设计算法——MIMI 算法,并且证明,如果 F 是由 MIMI 算法获得的相应于对象依赖集合 M 的规范化对象模式森林,则 M 蕴涵 $OD(F)$;并且 $OD(F)$蕴涵$\bowtie(P(F))$;若以规范化覆盖 M^*作为 MIMI 算法的输入,则获取相应于 M^+(M 的闭包)的路径不可分解的规范化对象模式森林;若以 M^-(M 所蕴涵的所有化简的对象依赖集合)作为 MIMI 算法的输入,则获取相应于 M^+的叶结点不可分解的规范化对象模式森林.

本文在已有工作的基础上,给出并证明了相应于无冲突对象依赖集合的规范化对象模式森林的性质.

1 基本概念和性质的概括

本节对在文献[4~7]中给出的基本概念和性质进行概括,作为全文的基础.

定义 **1**(对象依赖(**OD**)). 设 O 是一个对象集,$X,Y\subseteq O$,并且 $Z=O-XY$.在 X 和 Y 之间存在聚集关系或关联关系,使得对于在 X 中的每个实例 x,在 Y 中一定存在一组相应的实例 y 可以被 x 访问,并且对于在 XZ 中的任意实例 xz,在 Y 中访问到的是相同的一组相应的实例 y,则在 O 上 X 和 Y 之间存在着对象依赖(object dependency,简记为 OD),表示为 $X\rightarrow Y$,称 X 决定 Y,或 Y 依赖于 X.

根据 OD 的定义,如果 $X\rightarrow Y$ 在 O 上成立,$X\subseteq V\subset O$,则在 V 上,$X\rightarrow Y\cap V$ 成立.此外,设定 OD 不以$\varnothing$作为左式

* 收稿日期:2001-06-22; 修改日期:2002-01-08
基金项目:国家自然科学基金资助项目(60003008);中国科学院软件研究所计算机科学重点实验室资助项目(SYSKF0202)
作者简介:吴永辉(1966-),男,浙江温岭人,博士,副教授,主要研究领域为数据库,数字图书馆,面向对象技术;周傲英(1965-),男,安徽郎溪人,博士,教授,博士生导师,主要研究领域为 P2P 对象管理,WEB/XML 数据管理.

(LHS).设 $X,Y,Z\subseteq O$.OD 的推导公理如下:

O1. 自反公理.$Y\subseteq X\subseteq O$ 蕴涵 $X\to Y$.

O2. 增广规则.$X\to Y$ 蕴涵 $XZ\to Y$.

O3. 并规则.$X\to Y,X\to Z$ 蕴涵 $X\to YZ$.

O4. 投影规则.$X\to Y,X\to Z$ 蕴涵 $X\to Y\cap Z,X\to Y-Z$.

O5. 传递规则.$X\to Y,Y\to Z$ 蕴涵 $X\to Z-Y$.

O6. 补规则.$X\to Y$ 蕴涵 $X\to O-XY$.

O7. 伪传递规则.$X\to Y,YW\to Z$ 蕴涵 $XW\to Z$.

用$\bowtie$标记对象联接依赖,OD 和 OJD 之间的转换规则如下:

OD-OJD: $X\to Y$ 蕴涵$\bowtie(XY,X(O-XY))$,其中 $X\to Y$ 在 O 上成立;

OJD-OD: $\bowtie(X,Y)$蕴涵 $Y\to Y'$,其中 $X\subseteq Y'\subseteq Y$.

以后 O 表示给出的全体对象的集合,M 表示 O 上的 OD 集合.

定义 **2**(依赖基). 给出对象集合 $X,X\subseteq O,X$ 的依赖基 $DEP(X)$定义为 $O-X$ 的一个划分$\{Y_1,\ldots,Y_n\}$,即 $DEP(X)=\{Y_i|X\to Y_i$ 并且对任何一个 $W\subset Y_i,X\to W$ 在 O 上不成立;$i=1,\ldots,n\}$,$DEP(X)$中的元素被称为 X 的依赖元.$RDEP(X)$表示相应于 M 的 X 的化简的依赖基;$EDEP(X)$表示相应于 M 的 X 的必要的依赖基.

定义 **3**(无冲突 **OD** 集合). 如果存在 $X,R\subseteq O$,如果 X 有两个依赖元 V_1 和 V_2,使得 $V_1\cap R$ 和 $V_2\cap R$ 都不为空,则称 X 分裂 R.如果在 $LHS(M)$中至少有一个 X 使得 X 分裂 R,则称 M 分裂 R.如果 M 不分裂在 $LHS(M)$中的任意 X,并且$(DEP(X)\cap DEP(Y))\subseteq DEP(X\cap Y)$,则 M 是无冲突的.

定义 **4**(对象模式树). 如果用 $H(O)$表示 O 上嵌套关系的层次[4,5],相应于 $H(O)$的对象模式树 T 递归定义为:(1) 如果 $H(O)=O$,则 T 是根为对象集 O 的单结点树;(2) 如果 $H(O)=A(O_1)^*\ldots(O_n)^*$(即 A 与 $O_1,\ldots,O_n$ 构成嵌套层次),T 的根为对象集 A,根的儿女是对象模式树 T_i 的根,T_i 是 O_i 相应的对象模式树,$1\le i\le n$.

设 T 是 O 上的一棵对象模式树且 $e=(u,v)$是 T 中的一条边.定义标识:$F(v)$表示 v 的父亲结点,即,结点 $u.A(v)$为包括 v 在内的 v 的所有祖先结点的集合.$D(v)$为包括 v 在内的 v 的所有后继结点的集合.$OS(T)$是在对象模式树 T 中的所有对象的集合.由对象依赖的定义,$OD(e)$是由边 e 表示的 OD,记为 $OD(A(u),D(v),OS(T))$,表示在 $OS(T)$上成立的一个 OD:$A(u)\to D(v)$.$OD(T)$是由 T 中的所有的边表示的 OD 集合.

定义 **5**(路径). 设 T 是对象模式树,并且 $u_1,\ldots,u_n$ 是 T 的所有的叶结点,则 T 的路径集合 $P(T)=\{A(u_1),\ldots,A(u_n)\}$.对于叶结点 $u,A(u)$是在 T 中由 T 的根到 u 的路径上对象的集合.

显然,在对象模式树对应的化简超图中,$A(u)$是一条超边.路径也表示一个语义单位.

对象模式树 T 具有如下这些性质:$P(T)$是无α环的;$OD(T)\Leftrightarrow\bowtie(P(T))$;并且 $OD(T)$是无冲突的.

在消除了对象模式树中存在的冗余和不规则之后,就可以给出能很好地表示对象间的语义关系的规范化对象模式树和规范化对象模式森的定义[4,5].复杂对象模式规范化设计的目标是产生规范化对象模式森林[4,5].

引理 **1**. 设 $X\to W$ 是在 M^+中的一个未化简的 OD,那么在 M^+中存在一个化简的 OD $X'\to W'$,使得 $X'\subset X$ 和 $X'W'\subseteq XW$ 成立;并且如果 $W\in DEP(X)$,那么 $W'=HW$,其中 $H\subseteq(X-X')$.

引理 **2**. 如果 X 是非必要关键字,则 X 有惟一的化简依赖元;如果 X 是必要关键字,则 X 的化简依赖元多于一个.

引理 **3**. 设 X,Y 是对象集,$V_X\in DEP(X),V_Y\in DEP(Y)$.如果 $V_X\cap Y=\varnothing,V_Y\cap X=\varnothing$,并且 $V_X\cap V_Y\neq\varnothing$,那么 $V_X=V_Y$.

引理 **4**. 设 Z 是 M 的一个关键字,则对于任意的 $W\in DEP(Z),Z\to W$ 是非平凡的和不可转换的.

引理 **5**. 设 Z 是 M 的一个关键字,并且 $X\subset Z$,则存在一个 $V\in DEP(X)$使得

(1) $Z\subset XV$(即 X 不分裂 Z);

(2) 对于每个 $W\in RDEP(Z),W\subset V$;

(3) $X\to V$ 是左部化简的,并且如果 X 是一个关键字,则 $V\in RDEP(X)$.

引理 **6**. 如果 M 是无冲突的 OD 集合,则如果 M 蕴涵 OD $X\to V$ 和 $Y\to V$,并且 V 与 X 和 Y 都没有交集,则 $X\cap Y\to V$ 成立.

引理 **7**. 如果 M 是无冲突的 OD 集合,X,Y 和 V 是对象集,则

(1) 如果 $X\to V,Y\to V$ 是左部化简的和右部化简的,则 $X=Y$;

(2) 如果 X 是关键字,$V\in RDEP(X)$,并且 Y 是必要关键字使得 $Y\cap V\neq\varnothing$成立,则 Y 分裂 XV;

(3) 如果 X 是非必要关键字,对 X 的非化简依赖元 V,有必要关键字 $X'\subset X$ 使得 $X'\to V$;

(4) 如果 X 分裂 V,则存在一个必要的关键字 $X'\subseteq X$ 使得 X'分裂 V;

(5) 如果 M 没有分裂 X,并且在 M^+中 $X\to V$ 是左部和右部化简的 OD,则对于任意的分裂 XV 的必要关键字 $Y,Y\cap V\neq\varnothing$.

引理 **8**. 设 T 是相应于 OD 集合 M 的规范化对象模式树,P 是 T 中的路径,L 是 P 的叶结点.如果有关键字 $X\subset P$ 使得 X 分裂 P,则 $L\subseteq X$.

引理 **9**. 设 $W=W_0W_1...W_k$ 是相应于 M 的 X 和 Y 的依赖元.那么,如果 $W_i\in DEP(XW_0)$,则 $W_i\in DEP(YW_0)$,其中 $1\le i\le k,W_i,X$ 和 Y 是 O 的子集.

过程 $DECOMP(O)$返回一棵不可分解的半规范化对象模式树 T(叶结点不可分解).过程 NEWTREE 是通过删除一棵半规范化对象模式树的冗余结点来构造两个半规范化对象模式树.详细步骤参见文献[5].

算法 **1**. MIMI 算法.

输入:对象集合 O 以及 O 上的 OD 集合 M.

步骤:

求出 M 的所有关键字;

$T_1:=DECOMP(O);j:=1;$

设 $X_1,...,X_n$ 是 M 的关键字的一个偏序的次序;

For i:=1 To n Do

Begin If $X_i\subset OS(T_k)$,其中 $1\le k\le j$ Then

  Begin $V:=\varnothing$;

    For 每个在 T_k 中的满足$\exists Z\in DEP(X_i)$和 $D(W)=Z\cap OS(T_k)$的结点 W Do

    If (W 相应于 X_i 是冗余的) or ($Z\in EDEP(X_i)$ and $Z\notin EDEP(A(F(W)))$) Then $V:=V\cup\{W\}$;

    If $V\neq\varnothing$ Then Begin $j:=j+1$; $NEWTREE(X_i,V,T_k,T_j)$ End

  End

End

输出:规范化对象模式森林 $F=\{T_1,...,T_j\}$.

定理 **1**. 如果 T 是不可分解的半规范化对象模式树,则由 $NEWTREE(X,V,T,T_{new})$产生的 T 和 T_{new} 都是不可分解的半规范化对象模式树.

定理 **2**. 如果 F 是由 MIMI 算法获得的相应于 M 的 O 的规范化对象模式森林,则 M 蕴涵 $OD(F)$;并且 $OD(F)$ 蕴涵$\bowtie(P(F))$.其中 $P(F)$是规范化对象模式森林的路径集合.

2 相应于无冲突 OD 集合的规范化对象模式森林的性质

MIMI 算法产生的相应于某个 OD 集合 M 的规范化对象模式森林 F 并不总能保持 OD 集合.

引理 **10**. 设 $F=\{T_1,T_2,\dots,T_J\}$是由 MIMI 算法产生的规范化对象模式森林,$e=\langle u,v\rangle$是 T_i 中的一条边,$1\le i\le J$,并且 v_1 是 T_1'中相应于 v 的结点,$A(u)$和 $D'(v_1)$分别为定义在 T_i 和 T_1'中的 $A(u)$和 $D(v_1)$,则 $A(u)\to D'(v_1)$,并且它是左部和右部化简的.

证明:设 $T_{k1},T_{k2},\dots,T_{kn}$ 是在 F 中的一个规范化对象模式树的序列,其中 $T_{k1}=T_1,T_{kn}=T_i$,并且 T_{kj} 由 $T_{k(j-1)}$构造,$2\le j\le n$.对于 $1\le j\le n$,设 T_j'是 T_{kj} 刚被构造时的树的情形,$e_j=\langle u_j,v_j\rangle$是 T_j'中的边,并且 v_j 相应于 v,即,$v_j=v.A(u_j),D(v_j)$和 $A'(u_j),D'(v_j)$分别在 T_{kj} 和 T_j'中定义,则显然,$A(u_j)=A'(u_j),D(v_j)\subseteq D'(v_j)$,并且 $D(v_j)=D'(v_j)\cap OS(T_{kj})$.

首先通过对 j 的归纳,证明对于 $1\le j\le n,A(u_j)\to D'(v_1)$是右部化简的.

当 $j=1$ 时,$A(u_1)\to D'(v_1)$显然是右部化简的.

假设 $j=n-1$ 时命题成立,即 $A(u_{n-1})\rightarrow D'(v_1)$是右部化简的.设 T'' 是 $T_{k(n-1)}$在通过执行 $NEWTREE(X,V,T'',T_n')$构造 T_n'前的形式,则 $e_{n-1}=\langle u_{n-1},v_{n-1}\rangle$也在 T'' 中.设 $D''(v_{n-1})$是在 T'' 中定义的 $D(v_{n-1})$,则 $D''(v_{n-1})=D'(v_n)\subseteq D'(v_{n-1})$.基于定理 1,$T''$ 是半规范化对象模式树,因此 $OD(A(u_{n-1}),D''(v_{n-1}),OS(T''))$是右部化简的.因为 T_n'由 T'' 构造,基于引理 9,$OD(A(u_n),D''(v_{n-1}),OS(T''))$也是右部化简的.这两个 OD 是由 M 蕴涵的 OD 在 $OS(T'')$上的投影,即,存在 $D_{n-1}\in DEP(A(u_{n-1}))$以及 $D_n\in DEP(A(u_n))$使得 $D''(v_{n-1})=D_{n-1}\cap OS(T'')$以及 $D''(v_{n-1})=D_n\cap OS(T'')$.基于引理 3,$D_{n-1}=D_n$.由归纳假设,$D'(v_1)\in DEP(A(u_{n-1}))$成立,并且 $D_{n-1}\cap D'(v_{n-1})\neq\varnothing$,则 $D'(v_1)=D_{n-1}=D_n$,所以 $D'(v_1)\in DEP(A(u_n))$,即 $A(u_n)\rightarrow D'(v_1)$是右部化简的.

所以,对于 $1\le j\le n$,$A(u_j)\rightarrow D'(v_1)$是右部化简的.

T_{kn} 是规范化对象模式树,因此 $OD(A(u_n),D(v_n),OS(T_{kn}))$是左部和右部化简的,并且显然,这一 OD 是 $A(u_n)\rightarrow D'(v_1)$在 $OS(T_{kn})$上的投影.所以 $A(u_n)\rightarrow D'(v_1)$是左部化简和右部化简的. □

引理 **11**. 设 M 是无冲突 OD 集合,Z 和 V 是对象集合,$Z\rightarrow V$ 是左部和右部化简的,M 不分裂 Z,V 上的基本关键字集合 $FK(V)\neq\varnothing$,并且 $K=\{Y_i|Y_i$ 是必要关键字,并且 $Y_i\cap V\neq\varnothing\}$.在 K 上定义二元关系“<”如下:

$Y_i<Y_j$ 当且仅当 ZY_i 被 Y_j 分裂,但没有被在 K 中的任何 Y_j'分裂,其中 $Y_j'\subset Y_j$,则二元关系“<”是反自反的、反对称的和传递的.

证明:

(1) “<”是反自反的,即 $Y_i<Y_i$ 不成立.因为 M 不分裂 Z,所以 Y_i 不分裂 Z,Y_i 不分裂 Y_iZ.

(2) “<”是反对称的.假设 $Y_i<Y_j$,并且 $Y_j<Y_i$ 也成立,即 Y_i 分裂 ZY_j 也成立.因为 $Y_i<Y_j$,所以 Y_j 分裂 ZY_i,又因为 M 无冲突,所以存在 $V_j\in DEP(Y_j)$使得 $Z_j=V_j\cap Z\neq\varnothing$,并且 $V_j\cap Y_i=\varnothing$.同理,因为 Y_i 分裂 ZY_j,所以存在 $V_i\in DEP(Y_i)$,使得 $Z_i=V_i\cap Z\neq\varnothing$,并且 $V_i\cap Y_j=\varnothing$.现在证明 $Y_j-Y_i\neq\varnothing$成立,假设 $Y_j-Y_i=\varnothing$成立,则 $Y_j\subset Y_i$.因为 M 不分裂 Z,所以 Y_i 不分裂 Z,Y_i 不分裂 Y_jZ,则导致矛盾.所以 $Y_j-Y_i\neq\varnothing$成立,因而有 $Y_i\cap Y_j\subset Y_j$.现在证明 $Z_i\cap Z_j\neq\varnothing$.因为 Y_i 不分裂 $Z,Z_i=V_i\cap Z$,所以有 $Z-Z_i\subseteq Y_i$ 成立,如果 $Z_i\cap Z_j=\varnothing$,则 $Z_j\subseteq Z-Z_i\subseteq Y_i$,所以 $Z_j\subseteq Y_i$,因为 $V_j\cap Y_i=\varnothing$,所以 $Z_j\not\subseteq Y_i$,则导致矛盾.所以 $Z_i\cap Z_j\neq\varnothing$,则 $V_i\cap V_j\neq\varnothing$成立.所以,由引理 3 和引理 6,$V_i=V_j\in DEP(Y_i\cap Y_j)$,则 $Y_i\cap Y_j$ 分裂 ZY_i,由引理 7(4),存在必要关键字 $Y'\subseteq(Y_i\cap Y_j)$,使得 Y'分裂 ZY_i;因为 M 无冲突,$Z\rightarrow V$ 是左部化简和右部化简的,$Y_i\cap V\neq\varnothing$,所以 $Y_i\subseteq ZV$;亦即 Y'分裂 ZV;由引理 7(5),$Y'\cap V\neq\varnothing$,所以 $Y'\in K$,则与 $Y_i<Y_j$ 矛盾.所以 $Y_j<Y_i$ 不成立.

(3) “<”是传递的,即 $Y_i<Y_j,Y_j<Y_k\Rightarrow Y_i<Y_k$.$Y_i<Y_j$ 蕴涵 $DEP(Y_j)=\{V_{j1},V_{j2},\dots\}$,其中 $Z\subseteq Y_jV_{j1},Z\cap V_{j1}\neq\varnothing,Y_i\subseteq Y_jV_{j2}$,并且 $Y_i\cap V_{j2}\neq\varnothing$.$Y_j<Y_k$ 蕴涵 $DEP(Y_k)=\{V_{k1},V_{k2},\dots\}$,其中 $Z\subseteq Y_kV_{k1},Z\cap V_{k1}\neq\varnothing,Y_j\subseteq Y_kV_{k2}$ 并且 $Y_j\cap V_{k2}\neq\varnothing$.

首先证明 Y_k 分裂 ZY_i.因为如果 $V_{k1}\cap Y_i=\varnothing$并且 $Y_i-Y_k\neq\varnothing$,则 Y_k 分裂 ZY_i.所以要证明 $V_{k1}\cap Y_i=\varnothing$和 $Y_i-Y_k\neq\varnothing$.如果 $Y_i-Y_k=\varnothing$,则 $Y_i\subseteq Y_k$成立,因为 Y_j 分裂 ZY_i,所以 Y_j 分裂 ZY_k,因为 $Y_j<Y_k$,所以 Y_k 也分裂 ZY_j.由上面(2)的证明可知,Y_j 分裂 ZY_k 和 Y_k 分裂 ZY_j 不能同时成立,导致矛盾,所以 $Y_i-Y_k\neq\varnothing$成立.现在证明 $V_{k1}\cap Y_i=\varnothing$.设 $Y_{i1}=Y_i\cap Y_j$ 并且 $Y_{i2}=Y_i\cap V_{j2}$,因为 $Y_i\subseteq Y_jV_{j2}$,所以 $Y_i=Y_{i1}Y_{i2}$.因为 $Y_j\cap V_{k1}=\varnothing$,所以 $Y_{i1}\cap V_{k1}=\varnothing$.因为 $Y_k\rightarrow Y_kV_{k2}$,由 $Y_j\subseteq Y_kV_{k2},Y_kV_{k2}\rightarrow V_{j2}$ 成立,则 $Y_k\rightarrow V_{j2}-Y_kV_{k2}$ 成立,即 $Y_k\rightarrow V_{j2}-V_{k2}$ 成立.由 $Y_k\rightarrow V_{k1}$,由投影规则,$Y_k\rightarrow(V_{k1}-(V_{j2}-V_{k2}))$,即 $Y_k\rightarrow(V_{k1}-V_{j2})$成立.因为 $V_{k1}\in DEP(Y_k)$,所以$(V_{k1}-V_{j2})=V_{k1}$,也就是说,$V_{j2}\cap V_{k1}=\varnothing$,因此 $Y_{i2}\cap V_{k1}=\varnothing$成立.所以可以导出 $Y_i\cap V_{k1}=\varnothing$,则 Y_k 分裂 ZY_i.

然后证明 K 中不存在 $Y'\subset Y_k$ 使得 $Y_i<Y'$.假设在 K 中存在 $Y'\subset Y_k$ 使得 $Y_i<Y'$.设 $DEP(Y')=\{V_1',V_2',\dots\}$,由 $Y_i<Y'$ 可以设 $Z\subseteq Y'V_1',Z\cap V_1'\neq\varnothing,Y_i\subseteq Y'V_2'$,并且 $Y_i\cap V_2'\neq\varnothing$.因为 $Y_j<Y_k$,由“<”的定义,所以 Y'不分裂 ZY_j,即 $ZY_j\subseteq Y'V_1'$,并且 $V_2'\cap Y_j=\varnothing$.同理,设 $DEP(Y_j)=\{V_{j1},V_{j2},\dots\}$,其中 $ZY_k\subseteq Y_jV_{j1},Z\cap V_{j1}\neq\varnothing,Y_i\subseteq Y_jV_{j2}$,并且 $Y_i\cap V_{j2}\neq\varnothing$.由 $ZY_k\subseteq Y_jV_{j1}$,$ZY'\subseteq Y_jV_{j1}$,所以 $V_{j2}\cap Y'=\varnothing$成立.现在证明 $V_2'\cap V_{j2}\neq\varnothing$.因为 $V_2'\cap Y_i\neq\varnothing$,并且 $V_{j2}\cap Y_i\neq\varnothing$,又因为 $Y_j\rightarrow V_{j2}$,并且 Y_j 不分裂 Y_i,所以 $Y_i\subseteq Y_jV_{j2}$.如果 $V_2'\cap V_{j2}=\varnothing$,则 $V_2'\cap Y_i\subseteq Y_j$ 成立,即 $Y_j\cap V_2'\neq\varnothing$,这与 $Y_jZ\subseteq Y'V_1'$矛盾,所以 $V_2'\cap V_{j2}\neq\varnothing$.因为 $V_2'\cap V_{j2}\neq\varnothing$,则由引理 3 和引理 6,$V_2'=V_{j2}\in DEP(Y_j\cap Y')$.因此 $Y_j\cap Y'$分裂 ZY_i,由引理 7(4),有一个必要关键字 $X\subseteq(Y_j\cap Y')$使得 X 分裂 ZY_i.由上述(2)的证明可知,X 分裂 ZV.因为 $Y_i<Y_j$ 以及 $Y_i<Y'$,由“<”的定义蕴涵 $Y'\not\subseteq Y_j$,即 $Y_j\cap Y'\subset Y'$.所以 $X\subset Y'$成立,并且由引理 7(5),X 在 K 中,这与 $Y_i<Y'$矛盾.所以 K 中不存在 $Y'\subset Y_k$ 使得 $Y_i<Y'$. □

引理 **12**. 设 M 是无冲突 OD 集合,Z,V 是对象集,并且 $Z\rightarrow V$ 是左部化简和右部化简的.如果 M 不分裂 Z 并且 $FK(V)\neq\varnothing$,则在 $FK(V)$中存在一个 X_0 使得 M 不分裂 ZX_0.

证明:设 $K=\{Y_i|Y_i$ 是必要关键字并且 $Y_i\cap V\neq\varnothing\}$.$K$ 上定义的二元关系“<”如下:

$Y_i<Y_j$ 当且仅当 ZY_i 被 Y_j 分裂,但没有被在 K 中的任何 Y_j' 分裂,$Y_j'\subset Y_j$.

由引理 11,在 K 中二元关系“<”是反自反的、反对称的和传递的,则在 K 中至少存在这样一个 Y 使得对于在 K 中的任意其他的 $Y',Y\not< Y'$,即 Y' 不分裂 ZY.所以 M 不分裂 ZY.因为 $Y\cap V\neq\varnothing$ 并且 $Y\cap V\in FK(V)$,设 $X_0=Y\cap V$,则 $X_0\in FK(V)$,因为 $ZY\supseteq ZX_0$,所以 M 不分裂 ZX_0. □

引理 **13**. 设 F 是由 MIMI 算法在 O 上获得的相应于 M 的规范化对象模式森林.如果 M 是无冲突的,则在 $P(F)$中的任何路径 p 相应于 M 是不可分解的.

证明:基于 MIMI 算法和过程 DECOMP.设 T_1'是 MIMI 算法中 $DECOMP(O)$产生的不可分解的半规范化对象模式树.引理 12 保证过程 DECOMP 的第 2 步[5]的选择条件总是满足的,则对于 T_1'中的边 $e=\langle u_1,v_1\rangle$,在 T_1'中存在 $X\subseteq A(u_1)$,使得 $X\to D(v_1)$是左部化简和右部化简的,并且对于任何的必要关键字 $Y,Y\cap D(v_1)\neq\varnothing,X$ 不会被 Y 分裂.设 $e=\langle u,v\rangle$是路径 p 中的任意边,则在 T_1'中存在边 $e=\langle u_1,v_1\rangle$使得 v_1 是 v 的相应结点,即 $v=v_1$.由引理 10,$A(u)\to D(v_1)$是左部和右部化简的,并且由引理 7(1),$A(u)=X$ 成立,并且 $D(v)\subseteq D(v_1)$成立,所以对于任意的必要关键字 $Y,Y\cap D(v)\neq\varnothing,A(u)$不会被 Y 分裂.设 L 是 p 的叶结点并且在 p 中 A 为 $A(F(L))$,假设 p 相应于 M 是可分解的,则由引理 7(4)和引理 8,存在必要关键字 $Z\supseteq L,Z$ 分裂 p,即 Z 分裂 A,导致矛盾.所以命题成立. □

参照文献[8]中的引理 8.7 及其证明可以给出并证明如下的引理 14.

引理 **14**. 设 M 是无冲突 OD 集合,$G(M)$是它的超图表示.对于任意一个关键字偏序次序,MIMI 算法产生的规范化对象模式森林 F 的路径集合 $P(F)$由 $G(M)$的最大团组成,而且 M 等价于对象联接依赖$\bowtie P(F)$.

在关系数据库理论中已经证明,如果 MVD 集合是无冲突的,则关系数据库模式有惟一的 4NF 分解[3];如果 MVD 集合是无冲突的,则在属性集合上有一个无分裂范式(SFNF),因为 SFNF 蕴涵 4NF,关系数据库模式是惟一的[9].基于此,可以给出并证明如下的引理 15.

引理 **15**. 设 M 是 O 上的无冲突 OD 集合,$P(F)$是 O 的规范化对象模式森林的路径集合,则 $P(F)$是惟一的和不可分解的.

相似于关系数据库理论中无α环数据库的特性的定理(文献[3]中的定理 18.2)及其证明,可以证明复杂对象模式也有相似的无α环复杂对象模式特性,如引理 16 所示.

引理 **16**. 对于规范化对象模式森林 F 的路径集合 $P(F)$,下列命题是等价的.

(1) $P(F)$是无α环的;(2) 对象联接依赖$\bowtie P(F)$等价于一个无冲突 OD 集合.

相应于无冲突 OD 集合的规范化对象模式森林具有如下特性.

定理 **3**. 设 F 是由 MIMI 算法获得的相应于对象集合 O 及其 OD 集合 M 的规范化对象模式森林,而且 M 是无冲突的,则

(1) $P(F)$是 O 的惟一的不可分解的规范化对象模式森林的路径集合;

(2) $M\Leftrightarrow OD(F)\Leftrightarrow\bowtie P(F)$;并且

(3) $P(F)$是无α环的.

证明:

(1) 基于定理 2 和引理 13,可得 $P(F)$是在 O 上的相应于 M 的规范化对象模式森林 F 的路径集合.由于 M 是无冲突的,由引理 15,$P(F)$是惟一的和不可分解的.

(2) 基于上面的(1)和引理 14,$M\Leftrightarrow\bowtie P(F)$.并且由定理 2,可以导出 $M\Leftrightarrow\bowtie OD(F)\Leftrightarrow\bowtie P(F)$;

(3) 由上面的(2)和引理 16,可以导出 $P(F)$是无α环的. □

3 结束语

本文论证了在无冲突的条件下产生的规范化对象模式森林可以更好地反映对象间的语义关系.今后,我们将把复杂对象模式规范化设计理论与 XML 语言、数字图书馆等相结合并加以改进.

References:

[1] Tari, Z., Stokes, J., Spaccapietra, S. Object normal forms and dependency constraints for object-oriented schemata. ACM Transactions on Database Systems, 1997,22(4):513~569.

[2] Mok, W.Y., Ng, Y.K., Embley, D.W. Using NNF to transform conceptual data models to object-oriented database designs. Data & Knowledge Engineering, 1998,24(3):313~336.

[3] Shi, Bai-le, He, Ji-chao, Cui, Jing. The Theory and Application of Relational Database. Zhengzhou: He'nan Science and Technology Press, 1989 (in Chinese).

[4] Wu, Yong-hui, Zhou, Ao-ying. A normal form for complex object schemes. Advances in Systems Science and Applications, 2000, 1(1):48~55.

[5] Wu, Yong-hui, Jiang, Wen-yun, Zhou, Ao-ying. Implementation and proof for normalization design of object-oriented data schemes. In: Chen, Jian, Chen, Ping, Bertrand, M., eds. Proceedings of the 36th International Conference on TOOLS. 2000. 220~227.

[6] Wu, Yong-Hui, Zhou, Ao-Ying. Research on properties of a set of object dependencies. Journal of Computer Research and Development, 2001,38(12):1491~1498 (in Chinese).

[7] Wu, Yong-Hui. Normalization design for complex object schemes [Ph.D. Thesis]. Shanghai: Fudan University, 2001 (in Chinese).

[8] Beeri, C., Fagin, R., Maier, D., *et al*. On the desirability of acyclic database schemes. Journal of the ACM, 1983,30(3):479~513.

[9] Beeri, C., Kifer, M. Comprehensive approach to the design of relational database schemes. In: Yuan, L., ed. Proceedings of the Conference on VLDB. Singapore: ACM Press, 1984. 196~207.

附中文参考文献:

[3] 施伯乐,何继朝,崔靖.关系数据库的理论及应用.郑州:河南科学技术出版社,1989.

[6] 吴永辉,周傲英.对象依赖集合的性质的研究.计算机研究与发展,2001,38(12):1491~1498.

[7] 吴永辉.复杂对象模式的规范化设计[博士学位论文].上海:复旦大学,2001.

The Normal Object Scheme Forest with Respect to Conflict-Free Dependencies*

WU Yong-hui[1,2,3], ZHOU Ao-ying[1,2]

[1](*Department of Computer Science and Engineering, Fudan University, Shanghai* 200433, *China*);

[2](*Laboratory for Intelligent Information Processing, Fudan University, Shanghai* 200433, *China*);

[3](*Key Laboratory of Computer Science, Institute of Software, The Chinese Academy of Sciences, Beijing* 100080, *China*)

E-mail: slwu@fudan.edu.cn

http://www.cs.fudan.edu.cn

Abstract: The properties for a normal object scheme forest with respect to a conflict-free set of ODs are shown in this paper. Firstly basic concepts and properties about object dependency, a conflict-free set of object dependencies, normal object scheme forest and the algorithm of normalization design for complex object schemes are summarized. Then the properties for a normal object scheme forest with respect to a conflict-free set of ODs are presented and proved: $P(F)$ is a unique split-free path set for a normal object scheme forest; $M\Leftrightarrow$OD$(F)\Leftrightarrow\bowtie P(F)$; and $P(F)$ is α-cyclic. There is a signification in the development for the object-oriented information systems.

Key words: object scheme; object dependency; conflict-free; forest; path

* Received June 22, 2001; accepted January 8, 2002

Supported by the National Natural Science Foundation of China under Grant No.60003008; the Foundation of the Key Laboratory of Computer Science, Institute of Software, The Chinese Academy of Sciences under Grand No.SYSKF0202