

模糊空值环境下扩展关系模型的冗余处理^{*}

王艳亮 孙 生 赵红梅

(燕山大学管理学院工业工程系 秦皇岛农业银行科技处 秦皇岛066000)

摘 要 本文将模糊空值环境下的关系模型扩展为扩展关系模型,并借用模糊语义距离给出清除扩展模型上冗余信息的基本方法。

关键词 模糊空值,扩展关系模型,冗余信息

The Expanding Relational Model under Fuzzy Null Value Enviroment

WANG Yan-Liang SUN Sheng ZHAO Hong-Mei

(School of Management, Yanshan University Qinhuangdao Agriculture Bank, Qinhuangdao, 066000)

Abstract This paper expands the basic relational model to the expanding one under fuzzy null value enviroment. And if also gives out the methods how to clear the redundant information in the expanding relational model with using the defininition of fuzzy distance.

Keywords Fuzzy null values, Expanding relational model, Redundant information

1 引言

为了表述不确定、模糊值,关系数据库引入模糊空值^[1]。模糊空值环境下数据库更新也包括插入、删除和修改三种操作,但由于模糊空值语义的不确定性,使得元组对更新条件存在不同确定程度的满足,因此引入取值 $[0,1]$ 之间的模糊逻辑值作为更新条件判定的逻辑结果。模糊空值环境下,插入操作与模糊空值关系不大,而删除、修改则要依据元组对更新条件的不同满足程度施以不同的处理。确定满足条件的元组可直接进行删改操作,一定程度满足条件的元组要在更新的同时保留原数据备份,以防把可能不满足条件的元组删改后不能恢复而破坏关系一致性。为此,引入扩展关系模型^[2]进行模糊空值环境下关系数据库的更新处理。本文介绍模糊空值环境下的扩展关系模型及模型上冗余信息的清除。

2 扩展关系模型

模糊空值环境下的关系模式^[1]可以表示为 $R(U,D)$,其中 U 为含有模糊空值的有限属性集合, D 为属性集合对应的含有模糊值的域。将这一模式扩展为 $R(U,U,D)$ ^[2]的形式,对应的二维表就是由元组对形成的扩展元组,再将更新条件项加入扩展二维表中,就构成模糊空值环境下更新操作的扩展模

式。

扩展关系模型中引入基本表(B 表)、修改基本表(MB 表)、删除附加表(DA 表)和删除修改附加表(DMA 表)来记录关系中因为更新操作产生的不同性质的信息。一个含模糊空值的二维关系,最初元组在 B 表中, MB 表、 DA 表和 DMA 表都是空表。插入操作只涉及 B 表;而进行删除和修改时,在满足关系约束的条件下,不确定满足删除条件的元组附带删除条件进入 DA 表;不确定满足修改条件的元组,修改前和修改后的元组附带修改条件一并进入 MB 表; MB 表中不确定满足删除条件的扩展元组及 DA 表中不确定满足修改条件的都进入 DMA 表。 B 表、 MB 表、 DA 表和 DMA 表的结构分别如下所示。

	A_1	...	A_n
	a_1	...	a_n
$B:$		...	
	b_1	...	b_n

	A_1	...	A_n	A_1	...	A_n	$Conn$
	a_1	...	a_n	a_1	...	a_n	m_1
$MB:$		...			...		...
	b_1	...	b_n	b_1	...	b_n	m_s

^{*} 本课题得到河北省教育厅基金资助。王艳亮 硕士,研究方向为关系数据库理论。孙 生 主要从事网络技术研究。赵红梅 燕山大学教务处。

DA:	A_1	$\cdots$	A_n	Cond
	a_1	$\cdots$	a_n	d_1
	$\cdots$	$\cdots$	$\cdots$	$\cdots$
	b_1	$\cdots$	b_n	d_b

DMA:	A_1	$\cdots$	A_n	A_1	$\cdots$	A_n	Cond	Conm
	a_1	$\cdots$	a_n	a_1	$\cdots$	a_n	d_1	m_1
	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
	b_1	$\cdots$	b_n	b_1	$\cdots$	b_n	d_b	m_b

3 信息冗余处理

扩展关系的信息冗余存在两种情况:表内冗余和表间冗余。表内冗余指基本二维表内的两个元组相等、等价^[1]或是扩展二维表内两个扩展元组的左边元组和右边元组对应或交叉冗余。表间冗余则指二个表内的元组之间或元组与扩展元组的左边元组或右边元组之间相等或等价。表内冗余可通过消除互为冗余的元组或扩展元组中的一个而消除,表间冗余要根据不同冗余情况做出相应处理。需说明扩展二维表 MB 表和 DMA 表中两个扩展元组左部与左部不冗余,可能左部与右部冗余,只有当两个元组左部与右部对应或交叉均不冗余时,两个扩展元组才是不冗余的。

(1) 表内冗余消除

比较元组等价,需要依次比较元组每个属性值是否等价,因此给出 ϵ -相等概念。

定义1 设 $SD(x, y)$ 表示模糊集 x, y 之间的语义距离^[3], ϵ 是一个实小量且有 $\epsilon > 0$; 如果 $SD(x, y) < \epsilon$, 则称模糊集 x 与 y 为 ϵ -相等。记为: $x \stackrel{\epsilon}{=} y$ 。

算法1 表内冗余检查算法 $min-in(R, \epsilon)$

输入: 扩展关系 R , 冗余系数 ϵ

输出: 消除冗余元组后的结果扩展关系 R

```
PROC simply-min( $R, \epsilon$ )
BEGIN
   $G = \phi$ ;  $num = 0$ ;
  For  $i = 1$  to  $n$  { $n$  是关系  $R$  的元组数目}
    For  $j = i + 1$  to  $m$  { $m$  是关系  $R$  的属性数目}
      If  $SD(t_i, t_j) < \epsilon$  Then  $G = G \cup t_i$ ; {利用语义句里判定属性是否相等}
    Next  $j$ ;
  Next  $i$ ;
   $R_x = R - G$ ;
END;
```

ENDPROC.

{主程序} BEGIN

$simply-min(R_B, \epsilon); simply-min(R_{DA}, \epsilon);$ { B 表、 DA 表的处理}

$T_{MB} = \phi;$ { MB 表的处理}

For $i = 1$ to n { n 是 MB 表中扩展元组的数目}

For $j = i + 1$ to n

$R_1 = t_i$ 的左部 $\cup t_j$ 的左部; $R_2 = t_i$ 的右部 $\cup t_j$ 的右部;

$R_3 = t_i$ 的左部 $\cup t_j$ 的右部; $R_4 = t_i$ 的右部 $\cup t_j$ 的左部;

$simply-min(R_1, \epsilon);$

If R_1 中有一个元组 Then

Begin $simply-min(R_2, \epsilon);$

If R_2 中有一个元组 Then [$w = r_1$ 和 r_2 连串; $\{r_1 \in R_1, r_2 \in R_2\}$]

$T_{MB} = T_{MB} \cup t_i \cup t_j; W_{MB} = W_{MB} \cup w;$

Loop;

END;

$simply-min(R_3, \epsilon);$

If R_3 中有一个元组 Then

Begin $simply-min(R_4, \epsilon);$

If R_4 中有一个元组 Then [$w = r_3$ 和 r_4 连串; $\{r_3 \in R_3, r_4 \in R_4\}$]

$T_{MB} = T_{MB} \cup t_i \cup t_j; W_{MB} = W_{MB} \cup w; Loop;$

END;

Next j ;

Next i ;

$R_{MB} = R_{MB} - T_{MB}; R_{MB} = R_{MB} \cup W_{MB};$

{ DMA 表的处理与 MB 表的处理基本相同, 此处省略}

END.

$min-in$ 算法的时间复杂度为 $O(mn^2)$, 其中, m 为关系中属性个数, n 为关系中元组的数目。算法将关系中的冗余元组剔除, 使得关系一定程度满足无冗余条件。

(2) 表间冗余消除

定义2 设 $SD-r(t, s)$ 表示元组 x, y 之间的距离, ϵ 是一个实小量且有 $\epsilon > 0$; 如果对于每一个属性都有 A_i 都有 $SD(t[i], s[A_i]) < \epsilon$, 则称模糊集 x 与 y 为 ϵ -相等。记为: $x \stackrel{\epsilon}{=} y$ 。

表间元组冗余一般由更新操作产生, 依据表间元组之间的不同冗余情况进行不同的处理, 策略如下 (w_L 和 w_R 分别代表扩展元组的左部份和右部份):

如果元组 $t \in B$, 元组 $s \in DA$, 有 $t \stackrel{\epsilon}{=} s$, 则从 DA 中除去 s ;

如果元组 $t \in B$, 扩展元组 $w \in MB$, 有 $t \stackrel{\epsilon}{=} w_L$ 或者 $t \stackrel{\epsilon}{=} w_R$, 则从 MB 中除去 w ;

如果元组 $t \in B$, 扩展元组 $w \in DMA$, 有 $t \stackrel{\epsilon}{=} w_L$ 或者 $t \stackrel{\epsilon}{=} w_R$, 则从 DMA 中除去 w ;

如果元组 $t \in DA$, 扩展元组 $w \in MB$, 有 $t \stackrel{\epsilon}{=} w_L$ 或者 $t \stackrel{\epsilon}{=} w_R$, 则从 MB 中除去 w ;

如果扩展元组 $t \in DA, w \in DMA$, 有 $t \stackrel{\epsilon}{=} w_L$ 或者 $t \stackrel{\epsilon}{=} w_R$, 则从 DMA 中除去 w ;

如果扩展元组 $w \in MB, v \in DMA$, 有 w 和 v 左右部分对应等价或者交叉等价, 则从 DMA 中除去 v ;

算法2 表间冗余检查算法 $min-out(R, \epsilon)$

输入: 扩展关系 R , 冗余系数 ϵ

输出: 消除表间冗余元组后的结果扩展关系 R

BEGIN

{ B 表出发的处理} $G_{DA} = \phi; G_{MB} = \phi; G_{DMA} = \phi;$

For $i = 1$ to n { n, m, p, q 分别是 B, DA, MB 和 DMA 表的元组数目}

For $j = 1$ to m { t, w, v, u 分别是 B, DA, MB 和 DMA 表中的元组}

If $SD-r(t, w_j) < \epsilon$ Then $G_{DA} = G_{DA} \cup w_j;$

Next j ;

For $k = 1$ to p

If $SD-r(t, v_{kL}) < \epsilon$ or $SD-r(t, v_{kR}) < \epsilon$ Then $G_{MB} = G_{MB} \cup v_k;$

(下转第164页)

atorByTitle 和 BookNodeCreatorByKind。由于上下文是可嵌套的,为了避免复杂的继承关系,我们在这里使用了 Decorator 模式。上下文生成器将调用导航上下文接口(NavigationalContext)生成上下文相应的内容和链接,最终由导航上下文对象通过 NavigationalView 接口来访问 EJBs 以取得必要的数
据。

PSMs 自动生成以后,用户可以对其进行必要的改进和精化,然后可以利用 PSMs 生成系统的实现代码。限于篇幅,本文中省略了示例最终的代码。

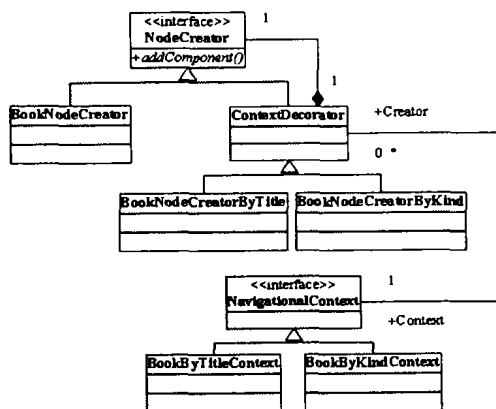


图6 图书馆管理系统的 Web 层模型局部

结束语 本文提出了一个基于 MDA 的 Web 应用程序开发方法 MDHDM。我们说明了应用 MDHDM 设计 Web 应用程序的三个步骤,并给出了与其相应的平台无关模型(PIMs)以及由 PIMs 映射而成的基于 J2EE 的平台相关模型(PSMs)。由于我

们将 MDA 和现有的 Web 应用程序开发方法相结合提出了 MDHDM, 因此 MDHDM 在吸收了现有开发方法的优点的同时, 利用 MDA 的特性也很好解决了它们所具有的缺陷。

目前 MDHDM 对于导航设计和展示设计中的行为建模的支持并不完善,这也是我们下一步工作的重点所在。同时,我们将进一步改进基于 MDHDM 的设计工具,使其更好地支持 Web 应用程序的设计。

参考文献

- 1 Miller J, Mukerji J. MDA Guide Version 1.0. 1. Document number omg/2003-06-01. Retrieved from: <http://www.omg.org/mda>, 2003
- 2 Frankel D. Model Driven Architecture. Applying MDA to Enterprise Computing. John Wiley & Sons, 2003
- 3 Retschitzegger W, Schwinger W. Towards Modeling of Data Web Applications -A Requirement's Perspective. In: Proc. of the America's Conf. on Information Systems, 2000. 149~155
- 4 de Janeiro R, Plata L. A Software Architecture for Structuring Complex Web Applications. In: The 11th Intl. World Wide Web Conf. Hawaii, USA, 2002
- 5 Schwabe D, Rossi G. An Object Oriented Approach to Web-Based Application Design. Theory and Practice of Object Systems (TAPOS), Special Issue on the Internet, 1998, 4(4): 207~225
- 6 Koch N, Baumeister H, Mandel L. Extending UML to Model Navigation and Presentation in Web Applications. In Modeling Web Applications, Workshop of the UML'2000. Ed. Geri Winters and Jason Winters, York, England, Oct. 2000

(上接第152页)

```

Next k;
For l=1 to q
    If  $SD-r(t_i, u_{l左}) < \epsilon$  or  $SD-r(t_i, u_{l右}) < \epsilon$  Then
 $G_{DMA} = G_{DMA} \cup u_i$ ;
    Next l;
Next i;
 $R_{DA} = R_{DA} - G_{DA}$ ;  $R_{MB} = R_{MB} - G_{MB}$ ;  $R_{DMA} = R_{DMA} - G_{DMA}$ ;
{DA 表出发的处理}  $G_{MB} = \phi$ ;  $G_{DMA} = \phi$ ;
For j=1 to m
    For k=1 to p
        If  $SD-r(w_j, v_{k左}) < \epsilon$  or  $SD-r(w_j, v_{k右}) < \epsilon$ 
Then  $G_{MB} = G_{MB} \cup v_k$ ;
        Next k;
        For l=1 to q
            If  $SD-r(w_j, u_{l左}) < \epsilon$  or  $SD-r(w_j, u_{l右}) < \epsilon$ 
Then  $G_{DMA} = G_{DMA} \cup u_i$ ;
            Next l;
        Next j;
 $R_{MB} = R_{MB} - G_{MB}$ ;  $R_{DMA} = R_{DMA} - G_{DMA}$ ;
{MB 表出发的处理}  $G_{DMA} = \phi$ ;
For k=1 to p
    For l=1 to q
        If  $(SD-r(v_{k左}, u_{l左}) < \epsilon$  or  $SD-r(v_{k右}, u_{l右}) < \epsilon$  or
 $SD-r(v_{k左}, u_{l右}) < \epsilon$  or  $SD-r(v_{k右}, u_{l左}) < \epsilon)$ 
Then  $G_{DMA} = G_{DMA} \cup u_i$ ;

```

```

      Next  $l$ ;
      Next  $k$ ;
       $R_{DMA} = R_{DMA} - G_{DMA}$ ;
END.

```

min-out 算法的时间复杂度为 $O(mn^2)$, 其中, m 为关系中属性个数, n 为关系中元组的数目。无论是表内冗余还是表间冗余, 消除冗余算法的复杂度较高, 需要系统大量的时间和空间, 因此要尽量减少使用的频率, 在确定会产生冗余的情况下才进行冗余信息处理。

参考文献

- 1 刘永山,杨楠. 模糊空值环境下的关系模型与关系操作. 计算机工程与科学, 1997, 19(1): 73~77
- 2 郝忠孝. 空值环境下数据库导论. 北京: 机械工业出版社, 1996. 1~76
- 3 何新贵. 模糊知识处理的理论与技术. 北京: 国防工业出版社, 1998. 14~86, 321~343