第10卷第1期　　软　件　学　报　　Vol.10, No.1
1999年1月　　JOURNAL OF SOFTWARE　　Jan. 1999

线性序约束关系的无损连接分解*

范志新　施伯乐

(复旦大学计算机科学系　上海　200433)

E-mail: zxfan@ms.fudan.sh.cn

摘要　对线性序约束数据库中的数据依赖和无损连接分解进行了研究.介绍了线性序约束数据库的一些基本概念,提出了线性序约束关系中多区域存在依赖和分组多区域存在依赖的概念,证明了变量集满足变量独立和多区域存在依赖是进行模式无损连接分解的充要条件.探讨了变量集在坐标线性变换下保持良性分解的情形.

关键词　线性序约束数据库,无损连接分解,多区域存在依赖,变量独立,分组多区域存在依赖.

中图法分类号　TP311

约束数据库是Kanellakis,Kuper和Revesz于1990年提出来的.[1]它综合了关系数据库、约束逻辑程序设计、空间数据库、辅助决策系统等多个领域的研究成果.其核心思想是将约束表达式作为数据来处理,为模式与约束、数据查询相结合的一致表达提供了一种有效方法.例如,把关系数据库中的每个元组看成一个等词约束合取式,如,关系$R(A,B)$中的元组(a,b)看作$((A=a)\wedge(B=b))$;在空间数据库中,把平面矩形内部区域描述为$((a_1<x<a_2)\wedge(b_1<y<b_2))$.约束数据库模型支持对空间、时间等密序量的表达和操作,也支持对不完全信息的处理.[2~4]

目前,对约束数据库中的密序约束类研究较多[2~7],在线性序约束类方面主要集中于查询的可表达性[6,8],对模式设计的研究很少.文献[5]提出密序约束保持无损连接分解的充要条件是在其规范表示的各个分块里能够进行无损连接分解,这只是一个计算问题,并没有给出约束模式进行无损连接分解所应具备的性质;文献[9]给出了密序约束中的函数依赖概念,并把它和变量独立合起来作为保持无损连接的一个充分条件,但这个函数依赖与变量独立在概念上是有交叉的,对无损连接分解来说不是必要条件.

本文提出了多区域存在依赖的概念,结合Chomicki等人提出的变量独立概念[2],得到线性序约束下保持无损连接分解的充要条件,讨论了坐标线性变换下线性序约束的转换,还给出了分组多区域存在依赖等概念.第1节介绍了线性序约束数据库的一些基本概念;第2节提出了多区域存在依赖、分组多区域存在依赖等概念,得出模式无损连接分解的充要条件;第3节探讨了坐标线性变换下线性序约束的一些问题.

1　线性序约束数据库的基本概念

按照Kanellakis等人的观点[1,3,4],约束关系可以用无量词的一阶逻辑公式表示.下面,我们先给出线性序约束的一些基本概念.

定义1.1. 线性序约束C^{LO}是指形如$\sum_{i=1}^{n}a_ix_i\theta c$的约束,其中$x_i$为实数变量,$a_i$和$c$为实常数,$\theta\in\{<,\leqslant,=,\neq,\geqslant,>\}$.

线性序约束的无量词一阶逻辑公式是一个析取范式,其中每个析取项是由线性序约束组成的合取式.令

---

* 本文研究得到国家自然科学基金资助.作者范志新,1969年生,博士,主要研究领域为数据库,知识库.施伯乐,1935年生,教授,博士生导师,主要研究领域为数据库,知识库.

本文通讯联系人:范志新,上海200433,复旦大学计算机科学系

本文1997-11-11收到原稿,1998-02-25收到修改稿

$c_x=\sum_{i=1}^{n}a_ix_i$，由于 $c_x\leqslant c$ 等价于 $((c_x<c)\vee(c_x=c))$，$c_x\neq c$ 等价于 $((c_x<c)\vee(c_x>c))$，$c_x\geqslant c$ 等价于 $((c_x=c)\vee(c_x>c))$，$((c_x>c_1)\wedge(c_x<c_2))$ 等价于 $c_1<c_x<c_2$，这样可使每个合取式只包含 $c_1<c_x<c_2$ 和 $c_x=c$ 两种形式（对于 $c_x>c_1$ 或 $c_x<c_2$，只存在其中之一的合取式，可令 $c_1=-\infty$ 和 $c_2=+\infty$）。我们假定经过上述处理得到的范式中不含多余的析取项和合取项。

定义 1.2. 变量集 $X=\{x_1,...,x_n\}$ 上的 n 元线性序约束元组是一个约束合取式 $\psi=\bigwedge_{i=1}^{l}\varphi_i$，其中 φ_i 是线性序约束。

定义 1.3. 变量集 $X=\{x_1,...,x_n\}$ 上的 n 元线性序约束关系 R 是 X 上约束元组的集合，其对应的逻辑表达式为 $\Phi_R=\bigvee_{k=1}^{m}\psi_k$。设 D 为实数域，$\mathrm{unr}(R)=\{(x_1,...,x_n)\mid(x_1,...,x_n)\in D^n$ 且使得 $\Phi_R(x_1,...,x_n)$ 为真$\}$。这里的 unr(R)取自文献[1]中的“unrestricted relation of R”。

定义 1.4. 设 R 和 S 是变量集 X 上 n 元线性序约束关系，若 $\mathrm{unr}(R)=\mathrm{unr}(S)$，则称 R 与 S 等价。

定义 1.5. 设 R 是定义在变量集 X 上的线性序约束关系，则称 $R(X)$ 是 R 的线性序约束关系模式，R 是 $R(X)$ 的一个实例。

定义 1.6. 线性序约束关系数据库是由有限个线性序约束关系组成的集合。

一个 n 元线性序约束元组表示 n 维空间上的一个凸集。[10]如果其中所有常数都是有限的，则其所表达的凸集在 Euclid 空间中就是有界的。可以定义基于代数的线性序约束查询语言。[5,7]这里给出两个将要用到的操作。

定义 1.7. 设 R 是变量集 X 上的 n 元线性序约束关系，$Y\subseteq X$，$|Y|=m$。于是，R 在 Y 上的投影 $\mathrm{proj}(R,Y)$ 是 Y 上的一个 m 元线性序约束关系，满足 $\mathrm{unr}(\mathrm{proj}(R,Y))=\{s[Y]\mid s\in\mathrm{unr}(R)\}$。

定义 1.8. 设 R 是变量集 X 上的 n 元线性序约束关系，S 是变量集 Y 上的 m 元线性序约束关系，D 为实数域，$X\cup Y=U$，$|X\cap Y|=p$。于是，R 和 S 的自然连接 $\mathrm{nj}(R,S)$ 是 U 上的 $(n+m-p)$ 元线性序约束关系，满足 $\mathrm{unr}(\mathrm{nj}(R,S))=\{u\mid u\in D^{n+m-p},u[X]\in\mathrm{unr}(R)$ 且 $u[Y]\in\mathrm{unr}(S)\}$。

2　无损连接分解和数据依赖

线性序约束数据库也存在数据依赖，它们的存在导致数据冗余并造成操作异常，而设计良好的模式可以部分地避免这些异常。

定义 2.1. 设 $R(X)$ 是线性序约束关系模式，则 $R(X)$ 上的数据依赖 DEP 是一个映射，其定义域是 $R(X)$ 的所有实例的集合，值域为真或假。

定义 2.2. 设 X,Y,Z 是变量集，$R(X)$，$R_1(Y)$，$R_2(Z)$ 是线性序约束关系模式。若 $Y,Z\subseteq X$ 且 $Y\cup Z=X$，则称 $\langle R_1(Y),R_2(Z)\rangle$ 是 $R(X)$ 的一个分解。考虑到分解是有意义的，通常 $Y-Z\neq\varnothing$，$Z-Y\neq\varnothing$。

定义 2.3. 设 $R(X)$ 是线性序约束关系模式，DEP 是 $R(X)$ 上的数据依赖，$\langle R_1(Y),R_2(Z)\rangle$ 是 $R(X)$ 的一个分解，R 是 $R(X)$ 的任一实例，$R_1=\mathrm{proj}(R,Y)$ 且 $R_2=\mathrm{proj}(R,Z)$。若 $\mathrm{unr}(\mathrm{nj}(R_1,R_2))=\mathrm{unr}(R)$ 成立，则称 $R(X)$ 的分解 $\langle R_1(Y),R_2(Z)\rangle$ 是无损连接的。

定理 2.1. 设 $R(X)$ 是线性序约束关系模式，$\langle R_1(Y),R_2(Z)\rangle$ 是 $R(X)$ 的一个分解，R 是 $R(X)$ 的任一实例，则 $\mathrm{unr}(R)\subseteq\mathrm{unr}(\mathrm{nj}(\mathrm{proj}(R,Y),\mathrm{proj}(R,Z)))$。特别地，对 R 中任一约束元组 t，也满足 $\mathrm{unr}(t)\subseteq\mathrm{unr}(\mathrm{nj}(\mathrm{proj}(t,Y),\mathrm{proj}(t,Z)))$。

例 2.1：R 是 $\{w,x,y,z\}$ 上的线性序约束关系，

t_1:　$0<w<6\wedge-1<x<4\wedge1<y-3z<2$，

t_2:　$w=8\wedge2<x<6\wedge0<y-3z<5$，

t_3:　$w=8\wedge7<x<8\wedge0<y-3z<5$

可分解成

$R_1(w,x)$：

t_{11}：　$0<w<6 \wedge -1<x<4$，

t_{12}：　$w=8 \wedge 2<x<6$，

t_{13}：　$w=8 \wedge 7<x<8$，

$R_2(w,y,z)$：

t_{21}：　$0<w<6 \wedge 1<y-3z<2$，

t_{22}：　$w=8 \wedge 0<y-3z<5$.

定义 2.4. 设 R 是 X 上的线性序约束关系，$Y,Z\subseteq X$，t 是 R 中的约束元组. 若满足

$$\mathrm{unr}(\mathrm{proj}(t,Y\cup Z))=\mathrm{unr}(\mathrm{nj}(\mathrm{proj}(t,Y),\mathrm{proj}(t,Z)))，$$

则称 Y 和 Z 在 t 中(语义)变量独立. 否则，称 Y 和 Z 在 t 中相关.

定义 2.5. 设 R 是线性序约束关系. 若存在与 R 等价的一个线性序约束关系 R'，Y 和 Z 在 R' 的每个约束元组中都是变量独立的，则称 Y 和 Z 在 R 中变量独立. 否则，称 Y 和 Z 在 R 中相关.

定义 2.6. 设 $R(X)$ 是线性序约束关系模式，$Y,Z\subset X$. 若 Y 和 Z 在 $R(X)$ 的任一实例 R 中都是变量独立的，则称 Y 和 Z 在 $R(X)$ 中变量独立.

当把约束关系 R 中的每个元组看作只含一个元组的关系时，变量独立可以保证对这些相应的关系模式分解都是无损的，而对一般的 R 来说，不能保证一定能够进行无损连接分解.

例 2.2：R 是 $X=\{w,x,y\}$ 上的线性序约束关系，

t_1：　$0<w<2 \wedge 0<x<1 \wedge 2<y<3$，

t_2：　$0<w<2 \wedge 0<x<1 \wedge 3<y<5$，

t_3：　$0<w<2 \wedge 2<x<3 \wedge 2<y<3$.

显然，$\{w,x\}$ 和 $\{w,y\}$ 在 R 中变量独立，但 R 在 $\{w,x,y\}$ 上的分解 $\langle\{w,x\},\{w,y\}\rangle$ 不是无损的. R 在 $\{x,y\}$ 上的投影是一个矩形区域的内部. 假定还存在元组

t_4：　$0<w<2 \wedge 2<x<3 \wedge 3<y<5$，

则 $R(X)$ 的分解 $\langle\{w,x\},\{w,y\}\rangle$ 就是无损的.

为此，我们引入多区域存在依赖的概念.

定义 2.7. 设 $R(X)$ 是线性序约束关系模式，$Y,Z\subseteq X$，$Y\cup Z=X$，$Y\cap Z=W$. 若对 $R(X)$ 的任一实例 R，均满足

$$\mathrm{unr}(\bigcup_{t\in R}\mathrm{nj}(\mathrm{proj}(t,Y),\mathrm{proj}(t,Z)))=\mathrm{unr}(\mathrm{nj}(\mathrm{proj}(R,Y),\mathrm{proj}(R,Z)))，$$

则称 $R(X)$ 满足多区域存在依赖 MRED(multi-region existence dependencies)：$W\overset{m}{\rightarrow}Y|Z$ 或 $W\overset{m}{\rightarrow}(Y-W)|(Z-W)$. 若 $Z=W$，则称该 MRED 是平凡的.

注意，多区域存在依赖与关系数据库里的多值依赖概念相似：多值依赖 $A\rightarrow\rightarrow B|C$ 是指对相同 A 值来说，B 上投影与 C 上投影的自然连接是 $B\cup C$ 上投影的迭加；而多区域存在依赖 $A\overset{m}{\rightarrow}B|C$ 是指对 A 的相同区域而言，B 上投影与 C 上投影的自然连接是 $B\cup C$ 上投影的迭加. 但二者是有差别的：在多值依赖中，由 B 和 C 的投影经自然连接可完全恢复 $B\cup C$；而在多区域存在依赖中，由 B 和 C 的投影经自然连接不一定能恢复 $B\cup C$.

例 2.3：R 是 $\{w,u,v\}$ 上的线性序约束关系，

t_1：　$2<w<3 \wedge 0<u<1 \wedge 2<v<3 \wedge 1<v-u<2$，

t_2：　$2<w<3 \wedge 5.5<u+v<6.5 \wedge 4.5<v-u<5.5$，

t_3：　$2<w<3 \wedge 2<u<3 \wedge 2<v<3 \wedge -1<v-u<0$，

t_4：　$2<w<3 \wedge 2<u<3 \wedge 5<v<6$.

显然，R 满足多区域存在依赖 $w\overset{m}{\rightarrow}u|v$，但 $\mathrm{unr}(\mathrm{nj}(\mathrm{proj}(t,wu),\mathrm{proj}(t,wv)))$ 中的点并非都在 R 中，比如点 $(2.7,0.4,2.5)$ 就不在 R 中.

引理 2.1. 设 $R(X)$ 是线性序约束关系模式，$Y,Z\subseteq X$，$Y\cup Z=X$，$Y\cap Z=W$，R 是 $R(X)$ 的任一实例，则下式成立：

$$\mathrm{unr}(\bigcup_{t\in R}\mathrm{nj}(\mathrm{proj}(t,Y),\mathrm{proj}(t,Z)))\subseteq\mathrm{unr}(\mathrm{nj}(\mathrm{proj}(R,Y),\mathrm{proj}(R,Z))).$$

引理 2.2. 设 $R(X)$ 是线性序约束关系模式，R 是 $R(X)$ 的任一实例，则下列结论成立：

$$\mathrm{unr}(R)=\mathrm{unr}(\bigcup_{t\in R}t)=\bigcup_{t\in R}\mathrm{unr}(t).$$

证明:由 unr 的定义知,$unr(R)=unr(\bigcup_{t\in R}t)$. 于是,对任一 $a\in unr(R)$,有 $a\in\bigcup_{t\in R}t$,从而 a 至少属于 R 中某个 t 的 $unr(t)$,因此,$unr(R)\in\bigcup_{t\in R}unr(t)$. 由 $unr(R)$的定义知,对于任一满足 t 的 $b\in unr(t)$,b 都在 $unr(R)$中,因此,$\bigcup_{t\in R}unr(t)\in unr(R)$. 所以,$unr(R)=\bigcup_{t\in R}unr(t)$. 引理结论成立. □

定理 2.2. 设 $R(X)$是线性序约束关系模式,$Y,Z\subset X,Y\cup Z=X,Y\cap Z=W,Y\neq W,Z\neq W$. $R(X)$的一个分解 $\langle R_1(Y),R_2(Z)\rangle$是无损的,当且仅当 Y 和 Z 在 $R(X)$中变量独立且满足多区域存在依赖 $W\overset{m}{\rightarrow}Y\ Z$.

证明:设 R 是 $R(X)$的任一实例.

先证充分性. 由 Y 和 Z 在 $R(X)$中变量独立和引理 2.2.知:

$$unr(R)=unr(\bigcup_{t\in R}proj(t,X))=unr(\bigcup_{t\in R}nj(proj(t,Y),proj(t,Z))),$$

再由 Y 和 Z 在 $R(X)$中满足多区域存在依赖 $W\overset{m}{\rightarrow}Y|Z$,可知:

$$unr(\bigcup_{t\in R}nj(proj(t,Y),proj(t,Z)))=unr(nj(proj(R,Y),proj(R,Z))),$$

因此,$unr(R)=unr(nj(proj(R,Y),proj(R,Z)))$成立. 从而,分解$\langle R_1(Y),R_2(Z)\rangle$是无损的.

其次证明必要性. 根据定理 2.1,$unr(t)\subseteq unr(nj(proj(t,Y),proj(t,Z)))$成立,于是得 $unr(R)=\bigcup_{t\in R}unr(t)\subseteq\bigcup_{t\in R}unr(nj(proj(t,Y),proj(t,Z)))$. 根据引理 2.1,$unr(\bigcup_{t\in R}nj(proj(t,Y),proj(t,Z)))\subseteq unr(nj(proj(R,Y),proj(R,Z)))$成立,因为$\bigcup_{t\in R}unr(nj(proj(t,Y),proj(t,Z)))=unr(\bigcup_{t\in R}nj(proj(t,Y),proj(t,Z)))$,所以有 $unr(R)\subseteq unr(\bigcup_{t\in R}nj(proj(t,Y),proj(t,Z)))\subseteq unr(nj(proj(R,Y),proj(R,Z)))$. 由假设 $R(X)$的分解$\langle R_1(Y),R_2(Z)\rangle$是无损的,即满足 $unr(R)=unr(nj(proj(R,Y),proj(R,Z)))$. 进而,

$$unr(R)=unr(\bigcup_{t\in R}nj(proj(t,Y),proj(t,Z))),\tag{1}$$

$$unr(\bigcup_{t\in R}nj(proj(t,Y),proj(t,Z)))=unr(nj(proj(R,Y),proj(R,Z)))\tag{2}$$

均满足. 由式(2)得,$R(X)$满足多区域存在依赖 $W\overset{m}{\rightarrow}Y|Z$. 将 R 中的每个约束元组 t 变为 $nj(proj(t,Y),proj(t,Z))$,由式(1)知,$unr(nj(proj(t,Y),proj(t,Z)))\subseteq unr(R)$,从而变化后的约束关系 R' 与 R 等价,且每一个约束元组均满足 $unr(proj(t,Y\cup Z))=unr(nj(proj(t,Y),proj(t,Z)))$. 于是,$Y$ 和 Z 在 R 中变量独立. 由 R 的任意性知,Y 和 Z 在 $R(X)$中变量独立.

综上所述,定理结论成立. □

变量独立和多区域存在依赖都是线性序约束关系中的数据依赖,它们保证了模式的无损连接分解. 我们还可以定义其他的数据依赖,下面给出另一种重要的依赖型.

定义 2.8. 设 $R(X)$是线性序约束关系模式,$Y,Z\subseteq X,Y\cup Z=X,Y\cap Z=W$. 若对 $R(X)$的任一实例 R,都存在一个等价约束关系 R',使得在 R' 上存在一个划分$\{G_1,\ldots,G_p\}$,$1\leqslant i,j\leqslant p,i\neq j$,满足:

(1) $unr(\bigcup_{t\in R}nj(proj(t,Y),proj(t,Z)))=unr(nj(proj(G_i,Y),proj(G_i,Z)))$;

(2) $unr(proj(G_i,Y))\cap unr(proj(G_j,Y))=\varnothing$;

(3) $unr(proj(G_i,Z))\cap unr(proj(G_j,Z))=\varnothing$,

则称 $R(X)$满足分组多区域存在依赖 PMRED(partitioned multi-region existence dependencies):$W\overset{p}{\rightarrow}Y|Z$ 或 $W\overset{p}{\rightarrow}(Y-W)|(Z-W)$.

定理 2.3. 线性序约束关系模式 $R(X)$上的多区域存在依赖都是分组多区域存在依赖.

可见,多区域存在依赖是分组多区域存在依赖的一种特殊形式.

3 线性序约束的坐标线性变换

先来看一个例子.

例 3.1:R 是$\{x,y,z\}$上的线性序约束关系,

$$t_1:\qquad 1<x<3\wedge 3<y+z<5\wedge 1<y-z<3.$$

显然,$\{x,y\}$与$\{x,z\}$不是变量独立的. 但若令 $y'=y+z,z'=y-z$,则$\{x,y'\}$与$\{x,z'\}$却是变量独立的.

定义 3.1. 设变量集 $X=\{x_1,...,x_n\}$，$X'=\{x'_1,...,x'_n\}$，X'中的每个变量都是 X 中变量的线性组合，即 $x'_i=\sum_{k=1}^{n}a_{ik}x_k$（其中 a_{ik} 为实常数，$1\leqslant i\leqslant n$）。可以用矩阵来表示。令

$$Y=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},Y'=\begin{pmatrix}x'_1\\ \vdots\\ x'_n\end{pmatrix},A_{nn}=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&\ddots&\vdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix},$$

则有 $Y'=A_{nn}Y$. 若 $|A_{nn}|\neq 0$，则称 $x'_1,...,x'_n$ 是线性无关的；否则，称 $x'_1,...,x'_n$ 是线性相关的.

定理 3.1. 设变量集 $X=\{x_1,...,x_n\}$，$X'=\{x'_1,...,x'_n\}$，线性序约束关系模式 $R(X)$，Y,Y',A_{nn}同定义 3.1. 若存在 $|A_{nn}|\neq 0$，使得 $Y'=A_{nn}Y$，则 R 在 X' 与 X 上是一一对应的.

略证：由 $|A_{nn}|\neq 0$ 知 A_{nn}^{-1} 存在，从而 $Y=A_{nn}^{-1}Y'$. 易知，R 在 X 上的每一点均在 X' 上有一点对应，且 R 在 X' 上的每一点也均在 X 上有一点对应. 故，R 在 X' 与 X 上一一对应. □

例 3.2：在例 3.1 中，取

$$A_{33}=\begin{pmatrix}1&0&0\\ 0&1&1\\ 0&1&-1\end{pmatrix},$$

从而得到新坐标系$\{x',y',z'\}$下的 R'，

$$t_1:\quad 1<x'<3\wedge 3<y'<5\wedge 1<z'<3.$$

它在$\{x',y',z'\}$上满足变量独立和多区域存在依赖 $x'\overset{m}{\rightarrow}y'|z'$，故存在无损分解$\langle\{x',y'\},\{x',z'\}\rangle$. R 在坐标系$\{x,y\}$上的投影是一个平行四边形.

于是，对给定的变量集 X，希望能够找到一个 $A_{nn}(A_{nn}\neq 0)$，通过这样变换后使得 R 在获得的新变量集 X'与 X 上一一对应，而且具有某些良好的性质. 这样，第 1、2 节中给出的定义和定理在 X'上依然成立. 如果 R 在变量集 X 上不满足变量独立和多区域存在依赖，但若存在 X 的一个满秩线性变换，使得 R 在变换后的变量集 X'上满足变量独立和多区域存在依赖，那么也可以进行无损连接的分解.

4 结　论

本文研究了线性序约束数据库中关系模式的无损连接分解. 通过引入多区域存在依赖，证明了线性序约束关系进行模式无损连接分解的充要条件是变量集满足变量独立和多区域存在依赖. 我们给出了分组多区域存在依赖的定义以及它与多区域存在依赖的关系，探讨了变量集在坐标线性变换下保持等价的性质，为某些类型的线性序约束关系提供了在变换后的新变量集下进行良性分解的途径. 研究线性序约束关系中的其他数据依赖是未来的工作.

致谢　感谢王宇君博士和田增平博士提供的资料，感谢汪卫博士和陈明博士参与讨论并提出了宝贵的意见.

参考文献

1 Kanellakis P C, Kuper G, Revesz P. Constraint query languages. In: Proceedings of 9th ACM Symposium on Principles of Database Systems. 1990. 299～313

2 Chomicki J, Goldin D Q, Kuper G M. Variable independence and aggregation closure. In: Proceedings of 15th ACM Symposium on Principles of Database Systems. 1996. 40～48

3 Grumbach S, Su Jian-wen. Finitely representable databases. In: Proceedings of 13th ACM Symposium on Principles of Database Systems. 1994. 289～300

4 Kanellakis P C. Constraint programming and database languages: a tutorial. In: Proceedings of ACM 14th Symposium on Principles of Database Systems. 1995. 46～53

5 Kanellakis P C, Goldin D Q. Constraint programming and database query languages. Lecture Notes on Computer Science, Springer-Verlag, 1994, 789: 96～120

6 Grumbach S, Su Jian-wen. Dense-order constraint databases. In: Proceedings of 14th ACM Symposium on Principles of

Database Systems. 1995. 66～77

7 Goldin D Q, Kanellakis P C. Constraint query algebras. Constraints Journal, 1996,1(1):24～32

8 Paredaens J, den Bussche J V, Gucht D V. First-order queries on finite structures over the reals. In: Proceedings of 10th IEEE Symposium on Logic in Computer Science. 1995. 79～87

9 Wang Yu-jun, Shi Bo-le, Qu Yun-yao. Decomposition and lossless Join in constraint database. Lecture Notes on Computer Science, Springer-Verlag, 1997,1191:80～91

10 Preparata F P, Shamos M I. Computational Geometry——an Introduction. David Gries ed. New York: Springer-Verlag, 1985

Lossless Join Decomposition in Dense Linear Order Constraint Databases

FAN Zhi-xin　SHI Bo-le

(*Department of Computer Science　Fudan University　Shanghai*　200433)

Abstract　In this paper, the data dependencies and lossless join decomposition in dense linear order constraint databases are studied. Several concepts of the dense linear order constraint databases are introduced. The concept of the multi-region existence dependencies is defined and studied. It is proved that the lossless join decomposition of constraint relational scheme preserves if and only if they satisfies variable independence and multi-region existence dependencies. Another dependency——partitioned multi-region existence dependencies is given. The issue of variable set under linear transformation is also proposed.

Key words　Dense linear order constraint database, lossless join decomposition, multi-region existence dependencies, variable independence, partitioned multi-region existence dependencies.