近似空间关系代数 ASRA 及应用

王生生　刘大有　胡　鹤　王新颖
(吉林大学计算机科学与技术学院,符号计算与知识工程教育部重点实验室,长春 130012)

摘　要　粗定位模型是一种基于粗集的近似区域表示模型,基于定性空间推理理论对其进行了代数形式化.通过空间关系矩阵和249种基本空间关系构造了近似空间关系代数ASRA;讨论了ASRA的公理和基本性质,研究了ASRA和RCC5关系映射中存在的不确定性;把ASRA应用于GIS,提出了基于ASRA的空间关系判定算法ASRA-RCC.与同类算法相比,ASRA-RCC能够同时支持确定和近似区域,并且具有较高的效率.

关键词　地理信息系统(420·3040)　定性空间推理　粗定位　近似空间关系代数

中图法分类号:TP181　P208　　**文献标识码**:A　　**文章编号**:1006-8961(2003)08-0946-05

Approximate Spatial Relation Algebra ASRA and Application

WANG Sheng-sheng, LIU Da-you, HU he, WANG Xin-ying
(*College of Computer Science and Technology, Key Laboratory of Symbolic Computing and Knowledge Engineering of Ministry of Education, Jilin University, Changchun* 130012)

Abstract　Uncertainty spatial information representation is a critical problem in qualitative spatial reasoning (QSR) and other spatial information process domains. Bittner put forward the rough location method which is better than other uncertainty spatial theories in special circumstance. But his theory wasn't formalized by relation algebra, thus can not be directly applied to spatial database or other spatial information applications. Based on rough set theory, we put forward a relation algebraic formalization for the rough location model and applied it to geographical informational system. Extending the RCC theory of QSR, basic spatial relation is defined by spatial relation matrix. Based on 249 basic relations and the operations on them the approximate spatial relation algebra ASRA is formed. After discussed its character and axiom, we studied the uncertainty of the maping from ASRA to RCC5. Applied it to GIS, a spatial relation judgement algorithm is given. Comparing with similar algorithms, it supports both crisp and approximate regions and is more efficient.

Keywords　GIS, Qualitative spatial reasoning, Rough location, Approximate spatial relation algebra

0　引　言

定性空间推理(Qualitative Spatial Reasoning 简称QSR)是研究空间信息表示和推理的重要手段,是GIS、机器人导航等技术的理论基础.一般把空间对象所占据的区域作为空间对象的本体.对于确定区域(边界是明晰的)的表示和推理,有比较完整的理论体系,其中具有代表性的是Cohn等提出的区域连接演算RCC(Region Connection Calculus)理论.

现实世界的空间信息都具有不确定性,空间对象的不确定性导致了其所对应的空间区域没有明确的边界,称这种区域为近似区域.近似区域的表示和推理是近年来QSR的研究热点,但目前还没有形成完整的理论[1].一种表示近似区域的模型是"蛋黄"模型(图1),它用最小扩展和最大扩展两个确定边界来刻画近似区域,该模型给出了252种JEPD(互不相交且无遗漏)关系,这些关系可聚类为40个集合[2]. 9交集扩展模型是另外一种近似区域模型,它给出了44种JEPD关系,这些关系可聚为18类[3].此外,还有ε带方法,误差带方法等其他方法.

基金项目:国家自然科学基金项目(69883003);国家863高技术项目(863-306-QN2000-1,2001AA110463)
收稿日期:2002-11-05;改回日期:2003-05-27

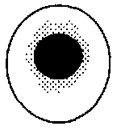

图 1　蛋黄模型

最近，粗集理论作为表示不确定性的另一种手段，已逐渐被用于解决空间不确定性问题. 瑞典的 Ahlqvist 等将粗集理论应用到 GIS 数据的分类研究中，对瑞典 Stockholm 的植被类型利用粗集分类计算法进行了重新分类实验，即利用二维扩展误差矩阵进行空间数据的多次分类，并对不同分类结果合并[4]. 粗集理论也被应用到 QSR 中，1999 年 Bittner 结合定性空间表示和粗集理论提出了空间对象的粗定位(rough location)模型. 基于粗集概念给出了不确定空间及时态对象的表示及拓扑关系的描述，定义了不确定空间对象的基本运算($\wedge_{max}$，$\wedge_{min}$，$\vee_{max}$，$\vee_{min}$)[5,6]. 和同类模型相比，粗定位模型的优势在于具有传递性，能形成等价类，可以构造非单调逻辑. 但基于粗集的不确定性空间理论还很不完善，在形式化表示、公理系统、实际应用等方面有待深入研究. 为此，在 Bittner 工作的基础上，借助空间关系矩阵对粗定位模型进行了代数形式化，建立了近似区域空间关系的代数系统 ASRA，并把 ASRA 应用于 GIS.

1　近似空间关系代数

1.1　确定区域空间关系

在不考虑边界的情况下，确定区域 x 把空间划分为内部和外部 2 个部分，分别用 x_I 和 x_O 表示，而边界划归到内部. 区域 x 和 y 的空间关系 R 可以通过 x_I，x_O 同 y_I，y_O 是否相交来确定

$$xRy=\begin{bmatrix} x_I\cap y_I & x_I\cap y_O \\ x_O\cap y_I & x_O\cap y_O \end{bmatrix}=\begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}\quad(1)$$

矩阵中 $r_{ij}(1\le i,j\le 2)$ 的取值是 $\{0,1\}$，0 表示不相交，1 表示相交.

约定：空间对象是非空的；空间对象是有限的，而嵌入空间(空间对象所在的空间)是无限的. 由此可以得到确定区域空间关系的 2 条公理：

公理 1　两个区域的外部必然相交；

公理 2　一个区域的每个部分(内部和外部)必然和另一个区域的某个部分相交.

这两条公理可以形式化表达为

$$\begin{cases} r_{22}=1 \\ r_{11}+r_{12}\neq 0 \\ r_{11}+r_{22}\neq 0 \end{cases}\quad(2)$$

符合式(2)的矩阵(式(1))共有 5 种，对应于基本 ***RCC5*** 关系[5](图 2). 这 5 种关系构成集合 ***RCC5***. ***RCC5*** 的所有子集构成集合 ***RCC5****，这是 ***RCC5*** 能描述的所有空间关系. 目前大多数 GIS，包括 MapInfo 等大型商业化 GIS 平台，所能描述的空间关系都是 ***RCC5**** 的子集.

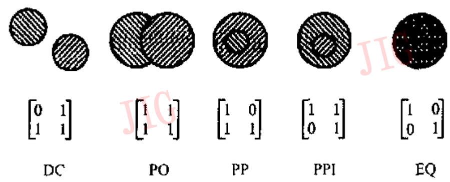

图 2　确定区域空间关系

1.2　基于粗集的不确定性空间表示

设 G 是对空间 S 的一个划分，$G=\{g_0,g_1,\cdots,g_n\}$，满足 $g_i\cap g_j=\varnothing(0\le i,j\le n,i\neq j)$而且 $S=g_0\cup\cdots\cup g_n$，g_i 称为划分单元. 规定：划分单元数目 n 是有限的；g_0 为唯一的一个无限大的划分单元.

Bittner 提出的粗定位概念是：G 是空间划分，x 是近似区域，$\Omega=\{\mathrm{fo},\mathrm{po},\mathrm{no}\}$，$x$ 用映射 $\alpha: x\rightarrow\Omega^G$ 表示. 对于 G 中的每个划分单元 g_i，$\alpha(g_i)$是 fo，表示 x 完全覆盖 g_i；$\alpha(g_i)$是 po，表示 x 部分覆盖 g_i；$\alpha(g_i)$是 no，表示 x 和 g_i 不相交[5,6]. 图 3 是近似区域粗定位表示的一个例子.

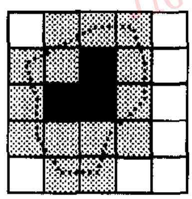

图 3　粗定位区域

1.3　近似空间关系代数(ASRA)

设空间划分 $G=\{g_0,g_1,\cdots,g_n\}$，其中 g_0 是无限的. 近似区域 x 用 3 元组来表示

$$x=(x_1,x_2,x_3),x_i\subseteq N,N=\{0,1,2,\cdots,n\}\quad(3)$$

其中，x_1，x_2，x_3 分别表示 fo，po，no 部分的划分单元编号集合，满足 $x_i\cap x_j=\varnothing(1\le i,j\le 3,i\neq j)$，

$x_1\cup x_2\cup x_3=N$. 根据公理1和公理2,显然有 $0\in x_3$ 而且 $x_1\cup x_2\neq\varnothing$.

近似区域 x,y 间的空间关系 R 用 3×3 矩阵表示:

$$R=\begin{bmatrix}x_1\cap y_1 & x_1\cap y_2 & x_1\cap y_3\\ x_2\cap y_1 & x_2\cap y_2 & x_2\cap y_3\\ x_3\cap y_1 & x_3\cap y_2 & x_3\cap y_3\end{bmatrix}=\begin{bmatrix}r_{11} & r_{12} & r_{13}\\ r_{21} & r_{22} & r_{23}\\ r_{31} & r_{32} & r_{33}\end{bmatrix}\quad(4)$$

其中,r_{ij} 的取值范围是{0,?,1},0表示不相交,1表示相交,?表示不确定.定义部分序关系:$0\leq ?\leq 1$,并且"+"运算定义为部分序关系取上界,"×"运算定义为部分序关系取下界.在此基础上进行空间关系矩阵之间的加乘运算.

r_{33} 固定为1,R 可以用 1×8 矩阵表示:$R=(r_{11},r_{12},r_{13},r_{21},r_{22},r_{23},r_{31},r_{32})$

由公理2可推得

$$R\times\begin{pmatrix}1&1&0&1&1&0&1&1\\1&1&1&1&1&1&0&0\end{pmatrix}^{\mathrm{T}}=(1\quad 1)\quad(5)$$

满足式(5)的空间关系共有249种($2^8-4-4+1$),都具有物理意义,称为基本空间关系.所有基本空间关系的集合记为 β,每种关系用 e_i 表示($1\leq i\leq 249$).

R 由基本空间关系构成:

$$R=\bigcup_{1\leq i\leq m}e_{r_i}\quad(6)$$

其中,$1\leq r_i\leq 249|,1\leq m\leq 249,r_i\neq r_j(1\leq i,j\leq m)$

近似区域空间关系 R 共有 $2^{|\beta|}$ 种,记所有 R 组成的集合为 I.每个空间关系是基本空间关系的集合($R\subseteq\beta$),如果所有 β 的子集组成的集合用 2^β 表示,则 $I=2^\beta$.

显然($R,\subseteq$)是一个格,如果定义⊕、⊗和－运算分别为集合的并、交和补,则(R,⊗,⊕,－,$\varnothing,\beta$)同集合代数等价也是布尔代数系统,称为近似空间关系代数,简称ASRA.基本空间关系 e_i 是ASRA的基底.

1.4 ASRA的不确定性

下面讨论ASRA映射到RCC5存在的不确定性.

设 x 为近似区域,x 所对应的确定区域用©(x)表示.

定理1 x,y 为近似区域,©(x)和©(y)的相交问题只有在 x,y 仅有po部分相交时才具有不确定性.

证明 表1列举了 x,y 各部分相交的所有情况,只有po交po具有不确定性.

根据定理1和表1,©(x)和©(y)的内部是否相交可以用 $r_{11}+r_{12}+r_{21}+?\times r_{22}$ 表示.

表1 判定©(x)和©(y)是否相交

	fo	po	no
fo	√	√	×
po	√	?	×
no	×	×	×

同理可以得到判定©$(x)_{\mathrm{I}}\cap$©$(y)_{\mathrm{O}}$ 和©$(x)_{\mathrm{O}}\cap$©$(y)_{\mathrm{I}}$ 的表达式,它们统一表达为

$$C(e)=(©(x)_{\mathrm{I}}\cap©(y)_{\mathrm{I}},©(x)_{\mathrm{I}}\cap©(y)_{\mathrm{O}},©(y)_{\mathrm{O}}\cap©(y)_{\mathrm{I}})$$
$$=e\times\begin{pmatrix}1&1&0&1&?&0&0&0\\0&1&1&0&?&1&0&0\\0&0&0&1&?&0&1&1\end{pmatrix}^{\mathrm{T}}\quad(7)$$

其中,e 是ASRA的原子,即基本空间关系.

由于式(1)中 r_{22} 固定为1,RCC5基本关系也可以采用相交关系三元组($x_{\mathrm{I}}\cap y_{\mathrm{I}},x_{\mathrm{I}}\cap y_{\mathrm{O}},x_{\mathrm{O}}\cap y_{\mathrm{I}}$)表达,形式上与 $C(e)$ 相同.通过 $C(e)$ 可以建立ASRA基本空间关系和 $RCC5$ 之间的对应关系."?"既可以和"0"对应又可以和"1"对应.每个ASRA的原子 e 对应的 $RCC5$ 基本空间关系的集合用 $RCC5(e)$ 表示.例如,$C(e)=(?,?,?)$,$\boldsymbol{RCC5}(e)$={DC,PO,PP,PPI,EQ}

如果|$\boldsymbol{RCC5}(e)$|>1,则 e 对应的 $\boldsymbol{RCC5}$ 关系存在着不确定性,这样的 e 称为不确定基本空间关系,所有不确定基本空间关系构成 β_u;否则,e 称为确定基本空间关系,所有确定基本空间关系构成 β_c

$$\beta=\beta_c+\beta_u\quad(8)$$

$\boldsymbol{RCC5}$ 中的DC关系对应的ASRA基本空间关系有9种,PO有178种,PP有11种,PPI有11种,EQ有1种.ASRA的确定基本空间关系有210种,不确定基本空间关系有39种.

如果某个ASRA关系仅由 β_c 中的确定空间基本关系构成,则称为确定ASRA关系,用 R_c 表示;否则称为不确定ASRA关系,用 R_u 表示.所有确定ASRA关系的集合是 I_c,所有不确定ASRA关系的集合是 I_u.以 β_c 中元素为基底也可以构成一个布尔代数系统(R,⊗,⊕,－,$\varnothing,\beta_c$).ASRA关系具有如下性质:

(1) $R_c\subseteq\beta_c$

(2) $R_u\cap\beta_u\neq\varnothing$

(3) $I_c=2^{\beta_c}$

(4) $I=I_c+I_u$

设 $R\in RCC5^*$,$AC(R)$和 $AU(R)$是 R 在ASRA中的上近似和下近似关系,分别定义为

$$AC(\boldsymbol{R}) \overset{\text{def}}{\equiv} \{e \mid e \in \beta, RCC5(e) \subseteq \boldsymbol{R}\} \tag{9}$$

$$AU(\boldsymbol{R}) \overset{\text{def}}{\equiv} \{e \mid e \in \beta, RCC5(e) \otimes \boldsymbol{R} \neq \varnothing\} \tag{10}$$

如果近似区域间的 ASRA 关系是 $AC(\boldsymbol{R})$，那么它们之间的 RCC5 关系确定为 $\boldsymbol{R}$；如果近似区域间的 ASRA 关系是 $AU(\boldsymbol{R})$，那么它们之间的 RCC5 关系可能为 $\boldsymbol{R}$.

定理 2　$AC(\boldsymbol{R})$和 $AU(\boldsymbol{R})$具有如下性质：

(1) $AC(\boldsymbol{R}) \subseteq AU(\boldsymbol{R})$

(2) $AC(-\boldsymbol{R}) = -AU(\boldsymbol{R})$

(3) $AU(-\boldsymbol{R}) = -AC(\boldsymbol{R})$

(4) $AC(\boldsymbol{R} \oplus \boldsymbol{S}) = AC(\boldsymbol{R}) \oplus AC(\boldsymbol{S})$

(5) $AU(\boldsymbol{R} \oplus \boldsymbol{S}) = AU(\boldsymbol{R}) \oplus AU(\boldsymbol{S})$

(6) $AC(\boldsymbol{R} \otimes \boldsymbol{S}) = AC(\boldsymbol{R}) \otimes AC(\boldsymbol{S})$

(7) $AU(\boldsymbol{R} \otimes \boldsymbol{S}) = AU(\boldsymbol{R}) \otimes AU(\boldsymbol{S})$

(8) $AU(\boldsymbol{R}) \oplus AC(\boldsymbol{R}) = AU(\boldsymbol{R})$

(9) $AU(\boldsymbol{R}) \otimes AC(\boldsymbol{R}) = AC(\boldsymbol{R})$

(10) $AC(\boldsymbol{R})^{\circ} = AC(\boldsymbol{R}^{\circ})$

(11) $AU(\boldsymbol{R})^{\circ} = AU(\boldsymbol{R}^{\circ})$

证明：

从式(9)和式(10)可直接得到性质 1.

由 $AC(-\boldsymbol{R}) = \{e \mid RCC5(e) \subseteq -\boldsymbol{R}\} = \{e \mid RCC5(e) \cap \boldsymbol{R} = \varnothing\} = -\{e \mid RCC5(e) \otimes \boldsymbol{R} \neq \varnothing\} = -AU(\boldsymbol{R})$可得性质 2.

对称地可以得到性质 3. $AC(\boldsymbol{R})$和 $AU(\boldsymbol{R})$间的关系如图 4 所示，阴影部分是不确定关系部分.

用集合运算可以得到性质 4～性质 7，从性质 1 可以得到性质 8，性质 9.

由 $RCC5(\boldsymbol{e})$的定义可得到

$$RCC5(\boldsymbol{e}^{\circ}) = RCC5(\boldsymbol{e})^{\circ}$$

从而推出性质 10，性质 11.

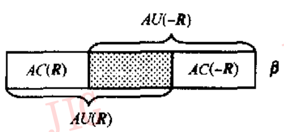

图 4　$AC(\boldsymbol{R})$和 $AU(\boldsymbol{R})$间的关系

2　ASRA 在 GIS 中的应用

GIS 中不确定信息的表示和分析目前还处于研究阶段，商业化 GIS 平台多数都不支持近似区域，并且一些系统对确定和不确定信息是分开处理的，这不符合空间数据的特点. 为此，基于 ASRA 提出了一种确定和近似区域的统一表示方法和拓扑关系判定算法，实现平台是组件式地理信息系统平台 CGIS[12].

空间表示建立的方法是：首先，对一个项目内的所有图层建立相同的空间划分 $\boldsymbol{G}$. 划分单元的形状和粒度要根据具体情况而定，可以根据四叉树空间索引矩形划分，也可以根据特定图层（如行政区划）划分；然后，把不确定性空间对象直接用 ASRA 的三元组（式(3)）表示，对于确定空间对象采用下面的算法建立 ASRA 表示.

算法 1　ASRA 空间表示建立算法：

for 每个确定空间对象 x　// 不确定空间对象直接用三元组表示

$x.f = x.p = \varnothing$；$x.o = \{0\}$；//($x.f, x.p, x.o$)是 x 的属性，对应于式(3)三元组.

for($i=1$；$i<=n$；$i++$)//$\boldsymbol{G}$ 有 $n+1$ 个划分单元 $g_0,\cdots,g_n$

{　if($g_i \cap x = g_i$) then $x.f = x.f \cup i$；//g_i 完全在 x 内部

else if ($g_i \cap x = \varnothing$) then $x.o = x.o \cup i$；//g_i 完全在 x 外部

else $x.p = x.p \cup i$；}

下面是确定和近似空间区域的拓扑关系统一判定算法

算法 2　RCC5 拓扑关系判定算法 ASRA-RCC

输入　空间对象 $\boldsymbol{x}, \boldsymbol{y}$ 和空间关系 $\boldsymbol{R} \in \boldsymbol{RCC5}^{*}$

输出　$\boldsymbol{xRy}$ 成立返回 true，不成立返回 false，不能确定返回 unknown

(1) 根据 $\boldsymbol{x}, \boldsymbol{y}$ 的粗集三元组求出 ASRA 基本空间关系矩阵 $\boldsymbol{e}$；

(2) if ($\boldsymbol{e} \in AC(\boldsymbol{R})$) return true；

if ($\boldsymbol{e} \notin AU(\boldsymbol{R})$) return false；

(3) if x 或 y 是不确定空间对象 then return unknown；

(4) for x 的每条边 d_1，y 的每条边 d_2

if d_1 和 d_2 相交（不包括重合情况）then $\boldsymbol{R}_1 = \{PO\}$；goto (6)；

(5) if x 的所有顶点都在 y 内部 then $\boldsymbol{R}_1 = \{PP\}$；

elseif y 的所有顶点都在 x 内部 then $\boldsymbol{R}_1 = \{PPI\}$；

elseif x 的所有顶点都在 y 边上 then $\boldsymbol{R}_1 = \{EQ\}$；

else $\boldsymbol{R}_1 = \{DC\}$；

(6) if ($\boldsymbol{R}_1 \in \boldsymbol{R}$) then return true；else return

false;

ASRA-RCC 首先用 ASRA 判定,如果不能判定,再进行几何计算来判定.与同类算法相比,ASRA-RCC 不仅解决了确定和近似区域之间拓扑关系的判定问题,并且能够提高确定区域拓扑关系的判定效率.几何信息的数据量大、精度高,几何计算非常耗时.ASRA-RCC 只在 ASRA 无法判定的情况下才需要几何运算,算法的执行时间取决于 ASRA 直接判定的概率.在国家基础空间数据库(包括全国行政区划图,地形图和气象、植被、土地覆盖等专题图)中的实验结果(如图 5)表明,ASRA-RCC 的执行时间是传统四叉树算法的 10%到 40%,并且数据规模越大,效果越好(图 6).算法的效率还依赖于空间划分,如以行政区划为划分单元往往能收到很好的效果,因为很多对象的位置都同行政区划相关.

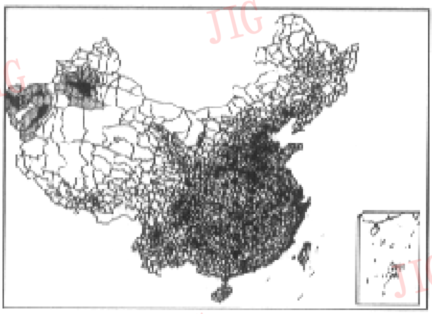

图 5 以行政区划为划分单元

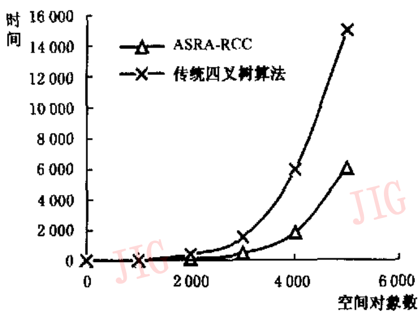

图 6 平均执行时间对比

3 结 论

本文对 Bittner 的模型进行了代数形式化,并研究了其公理、性质和不确定性分析方法,最后将其应用于 GIS 建立了一个近似和确定区域的统一模型(ASRA).同原模型相比,ASRA 表达更严密,实用性更强.ASRA 的应用可以推广到空间推理、空间数据挖掘、空间决策支持系统等很多涉及到不确定性空间信息的领域.同时对定性空间推理理论工作的实用化有一定意义.

参 考 文 献

1 Cohn A G, Hazarika S M. Qualitative spatial representation and reasoning: An overview[J]. Fundamental Informatics, 2001, 46(1-2):1~29.

2 Lehmann F, Cohn A G. The EGG/YOLK reliability heirarchy: semantic data integration using sorts with prototype[A]. In: Adam N R, Bhargava B K, Yesha Y, eds. Proceedings Conference on Information Knowledge Management[C], New York: ACM Press, 1994:272~279.

3 Clementini E, Di Felice P. Approximate topological relations [J]. International Journal of Approximate Reasoning, 1997: 100~110.

4 Ola Ahlqvist, Johannes Keukelaar, Karim Oukbir. Rough classification and accuracy assessment[J]. International Journal Geographical Information Science[J], 2000, 14(5):475~496.

5 Thomas Bittner. Approximate qualitative temporal reasoning [J]. Annals of Mathematics and Artificial Intelligence, 2002, 36(1-2):39~80.

6 Thomas Bittner, Stell J G. Vagueness and rough location[J]. GeoInformatica, 2002, 6(2):99~121.

7 王生生,刘大有. 混合维定性空间查询语言 MQS-SQL[J]. 电子学报, 2002, 30(12A):1995~1999.

王生生 1974 年生,讲师,现为吉林大学计算机学院在职博士生.主要研究方向为空间推理、时空推理、地理信息系统等.发表论文 7 篇.

刘大有 1942 年生,教授,博士生导师,现任吉林大学计算机学院(兼软件学院)院长和国务院学位委员会学科评议组成员等职务.主要研究知识工程与 ES、DAI 与智能 Agent,空间推理与 GIS,粗集、格机与数据挖掘,智能软件等.

胡 鹤 1976 年生,现为吉林大学计算机学院博士研究生.主要研究兴趣包括时空知识表示、描述逻辑和语义网等.

王新颖 1979 年生,现为吉林大学计算机学院硕士研究生.研究方向为图形学、GIS.