

# 模糊时空数据 在数据库间的转换方法

柏禄一 著



科学出版社

# 模糊时空数据在数据库间的 转换方法

柏禄一 著

科学出版社

北京

## 内 容 简 介

本书的主体内容直接来源于作者近年来在相关领域取得的一系列研究成果。本书作为数据库领域的学术专著，目的是通过系统地介绍当前模糊时空数据建模、转换及查询技术在理论研究和应用方面的成果，一方面为数据库研究人员提供国际前沿信息，另一方面为信息领域从事时空应用的专业人员提供技术帮助。

本书可作为高等院校计算机科学与技术、智能科学与技术、信息系统等专业的研究生和高年级本科生教材，也可作为计算机及相关专业科技工作者的参考书。

### 图书在版编目(CIP)数据

模糊时空数据在数据库间的转换方法 / 柏禄一著. —北京: 科学出版社, 2018.9

ISBN 978-7-03-058791-6

I. ①模… II. ①柏… III. ①计算机网络-可扩充语言-程序设计  
IV. ①TP312

中国版本图书馆 CIP 数据核字 (2018) 第 209262 号

责任编辑: 闫悦 王迎春 / 责任校对: 王萌萌  
责任印制: 师艳茹 / 封面设计: 迷底书装

科学出版社出版

北京东黄城根北街16号

邮政编码: 100717

<http://www.sciencep.com>

保定市中华美凯印刷有限公司印刷

科学出版社发行 各地新华书店经销

\*

2018年9月第一版 开本: 720×1000 1/16

2018年9月第一次印刷 印张: 14 1/4 插页: 2

字数: 270 000

定价: 89.00元

(如有印装质量问题, 我社负责调换)

# 前 言

随着时空应用在生产、生活中的推广，一系列关于时空数据的研究应运而生。由于时空数据的时间和空间属性在实际应用中通常是模糊的，关于模糊时空数据的研究，尤其是在数据存储、数据共享及数据查询等方面，已经获得了相关领域学者的密切关注。随着对模糊时空数据在时间和空间信息表示和推理应用方面需求的不断增长，人们对模糊时空数据的管理越来越重视。

互联网技术的高速发展促进了 XML 技术的发展与应用。XML 将用户界面和结构化数据分隔开来，这种数据与显示的分离有助于数据集成与交换，因此，XML 作为数据集成与交换的媒介也逐渐被接受和认可。在互联网中，XML 被认为是下一代数据表示、交换以及查询的标准语言，具有较强的自描述性以及良好的扩展性。因此，基于 XML 的模糊时空数据的表示将成为一个发展趋势。

由于模糊时空数据自有的特点，对保存在 XML 文档中的模糊时空数据进行操作时，在 XML 文档中涉及这些操作的节点和边可能会产生不一致性，也就是违反了预先定义的空间和时间约束条件，这就可能导致数据不一致问题。在 XML 文档中，虽然这些不一致问题已经被广泛研究，但是这些研究只考虑到一般数据，对模糊时空数据一致性问题的研究还比较少。虽然已有很多研究工作致力于时态数据、空间数据或者模糊数据的一致性问题的研究，但有关模糊时空数据的一致性问题的研究还存在尚未解决的问题。因此，对这些节点和边的一致性问题修复迫在眉睫。本书提出了检查和修复模糊时空 XML 文档中不一致性问题的方法。首先，提出适用于一致性约束的模糊时空 XML 数据模型，并在数据模型的基础上分析模糊时空 XML 文档要保持一致性的条件以及可能出现的非一致状态。这些非一致状态包括在一些入射边的时间标签上出现不连续、重叠或者包含环的情况。其次，本书提出基于一致性状态的算法来检查和修复模糊时空 XML 文档。最后，通过实验测试与分析，证明本书提出的方法的正确性和有效性，并与其他有关方法的比较，从另一方面证明本书所提出的方法在修复模糊时空 XML 文档的非一致状态方面具有较好的性能。

此外，由于关系数据库具有可靠性、成熟性和独立性的特点，目前大量数据都存储在传统的关系数据库中。模糊时空信息在 Web 中的共享性较差，而 XML 可以自由实现信息表现和信息交互，并且在数据库支撑方面颇具优势，对关系数据库系统更专一，实现模糊时空数据在 XML 与关系数据库之间的转换显得尤为

重要。因此，本书进行了基于 XML 的模糊时空数据建模、基于关系数据库的模糊时空数据建模、模糊时空数据从 XML 到关系数据库的转换方法以及从关系数据库到 XML 的转换方法的研究。为完成转换工作，本书分别使用 XML 和关系数据库设计了一种能捕捉模糊时空语义特点的模型；为了更好地共享以关系格式存储的模糊时空数据，并且不受平台约束，提出一种时间边的方法实现模糊时空 XML 数据与关系数据库之间的转换。这种方法的独特之处在于进行模糊时空数据转换时不需要模式信息；而且，对于 XML 文档中的模糊时空数据的时间、空间、模糊特点均在考虑之中。这种转换方法将会解决模糊时空数据在关系数据库与 XML 之间不能自由进行信息共享、信息传递的问题。接下来，针对所提出的方法还进行了实验，实验结果证明了该方法的性能优势，这种转换方法为模糊时空数据在 XML 与关系数据库之间的互操作性提供了重要参考。

与此同时，具有很多优秀特性的面向对象数据库的应用也越来越广泛，面向对象数据库把面向对象的方法和数据库技术结合起来，使数据库系统的分析与设计最大程度上与人们对客观世界的认识相一致。面向对象数据库系统是为了满足新的数据库应用需要而产生的新一代数据库系统。考虑到处理模糊时空数据的需求、XML 和面向对象数据库在处理模糊时空数据时的优势，以及 XML 和面向对象数据库在互联网应用方面的差异和联系，本书进行了关于建立模糊时空数据模型、实现模糊时空数据在 XML 和面向对象数据库之间的转换方法的研究。首先根据模糊时空数据的特性获取其语义，如模糊时间属性和模糊空间属性，提出基于 XML 和基于面向对象数据库的模糊时空数据模型。其次，对于模糊时空数据在 XML 中的表现形式，由于 XML 文档只能表示具体的数据，在实用性和代表性方面存在不足，而 XML Schema 作为 XML 文档定义语言，一方面可用于定义 XML 文档的结构，另一方面还可以验证 XML 文档是否满足指定的格式，因此，使用 XML 表示模糊时空数据更具有通用性。对于模糊时空数据在面向对象数据库中的表示，可以选择 UML 类图，将类、属性和方法都直观地表示出来。在模糊时空数据从 XML 转换到面向对象数据库时，针对 XML 的树状结构和 UML 基于类定义的结构，研究根节点、非叶子节点和叶子节点间的映射规则以及根节点与根节点之间的关系的映射规则。在模糊时空数据从面向对象数据库转换到 XML 时，本书提出一个类及其属性和方法的映射规则以及类与类之间关系的映射规则。这些映射规则对模糊时空数据在 XML 和面向对象数据库之间的互操作性有着重要的价值。

最后，本书提出了关于模糊时空数据的查询技术。由于模糊时空数据中复杂的模糊时间和模糊空间属性与传统关系型数据库中的二维表有很大差异，关于模糊时空数据查询应用的研究受到了很大的限制。XML 作为新一代互联网信息交换的标准语言，它的树型结构可以很好地解决这一限制问题，在处理模糊时空数据

方面相对容易。然而, 尽管模糊集合理论为研究模糊时空数据的查询和模糊 XML 数据的查询提供了重要的理论基础, 但是在模糊时空 XML 数据的定量查询领域所提出的相关理论还远远不够。因此, 本书提出一个名为 FSTTwigFast 的算法, 用于定量地匹配模糊时空 XML 小枝。本书通过在一般 XML 数据中加入模糊时间和模糊空间的属性、构建模糊三角模型来表示模糊时空 XML 数据。在表示模糊时空 XML 数据的基础上, 扩展了 Dewey 编码, 用于标记模糊时空 XML 数据, 以便进行特殊处理, 同时可以很好地确定 XML 文档中的模糊时空 XML 节点间的结构关系; 在处理时空 XML 文档中节点的模糊度上, 本书扩展了一般模糊 XML 数据的模糊度处理方法, 使其适用于处理时空 XML 节点的模糊度问题; 在技术细节上, 本书采用流来存储与叶子节点相匹配的 XML 文档中的节点, 存储在流中的节点通过过滤来删除不匹配节点。在过滤之后, 对叶子节点进行排序, 并为每一个匹配的叶子节点建立从叶子节点到根节点的输出链表, 形成最终的匹配结果。在提出算法之后, 进行了三组实验, 分别测试了不同查询小枝和不同时空数据查询的算法执行效率, 并将 FSTTwigFast 算法与 TwigFast 算法的查询效率进行比较。最后, 建立模糊时空 XML 数据查询系统, 展示所建立的系统在实际应用中的适用情况。

本书主要提到模糊时空数据基于 XML、关系数据库、面向对象数据库的建模方法, 以及模糊时空数据在 XML 与关系数据库和面向对象数据库之间的转换方法, 还研究了其在 XML 中的不一致性问题与查询问题。本书的研究工作将提高对模糊时空数据的管理能力, 丰富模糊时空数据的相关技术和方法, 进一步促进模糊时空数据应用技术的发展, 在时空应用研究领域具有重要的参考价值。

本书得到了国家自然科学基金(61402087)、河北省自然科学基金(F2015501049), 以及中央高校基本科研业务专项基金(N172304026)的资助。

本书同时得到了作者所在教师研究团队刘杰民、徐长明、朱琳、于长永、方淼、牛学芬、赵媛的大力支持, 也得到了作者所在学生研究团队贾志义、李银、邵珠磊、黎楠、栾静、白会磊、刘力双、王双迪、李俊腾、郝雪松的通力协作。黎楠最后对全书进行了校对和审稿。

特别感谢东北大学对本书写作工作的鼓励和支持。同时, 本书的出版得到了科学出版社的大力支持和帮助, 在此表示诚挚的感谢。

由于作者水平有限, 书中难免存在不妥之处, 敬请读者批评指正。

柏禄一

2018年1月于东北大学

# 目 录

前言

## 第一部分 基础篇

第 1 章 绪论	3
第 2 章 预备知识	7
2.1 时空数据与时空数据的模糊性	7
2.1.1 时空数据	7
2.1.2 数据的模糊性	8
2.1.3 时空数据的模糊性	9
2.2 模糊集、隶属函数和可能性分布	10
2.3 XML 的基础知识	12
2.3.1 XML 介绍	12
2.3.2 XML 的逻辑结构	13
2.3.3 文档类型声明	14
2.3.4 XML 的特点	15
2.3.5 XML 文档中的模糊性	16
2.3.6 XML 文档解析	18
2.3.7 XML 数据库查询语言	19
2.4 关系数据库的基础知识	19
2.4.1 关系数据库的特点	19
2.4.2 模糊时空数据模型	20
2.5 面向对象数据库的基础知识	22
2.6 Twig 查询	24
2.7 Dewey 编码	25
2.8 本章小结	25

## 第二部分 模糊时空 XML 文档中的不一致性

第 3 章 适于一致性约束的模糊时空 XML 数据模型	33
3.1 适于一致性约束的模糊时空 XML 文档中的节点的分类	33

3.2	模糊时间标签	34
3.3	模糊空间标签	36
3.4	属性节点	37
3.5	适于一致性约束的模糊时空 XML 文档	38
3.6	讨论	40
3.7	本章小结	41
<b>第 4 章</b>	<b>模糊时空 XML 数据的不一致修复</b>	<b>42</b>
4.1	模糊时空 XML 文档的一致性条件	42
4.2	模糊时空 XML 文档中的不一致性的检查	43
4.3	模糊时空 XML 文档中的不一致性的修复方法	47
4.3.1	不一致类型 I 和类型 II 的修复方法	47
4.3.2	不一致类型 III 的修复方法	48
4.4	实验评估	49
4.4.1	实验环境	49
4.4.2	检查和修复不一致性算法的性能测试	52
4.4.3	内存消耗测试	54
4.4.4	不一致性修复方法的实验对比	54
4.4.5	DOM 和 DOM4J 的性能比较	60
4.5	本章小结	61
<b>参考文献</b>		<b>62</b>
<b>第三部分 模糊时空数据在 XML 与关系数据库之间的转换</b>		
<b>第 5 章</b>	<b>模糊时空数据从 XML 到关系数据库的转换</b>	<b>69</b>
5.1	模糊时空 XML 数据模型	69
5.2	XML 到关系模糊时空数据库的转换	75
5.3	评测	79
5.3.1	在气象中的应用	79
5.3.2	实验	80
5.4	本章小结	84
<b>第 6 章</b>	<b>模糊时空数据从关系数据库到 XML 的转换</b>	<b>85</b>
6.1	模糊时空数据关系模型	85
6.2	模糊时空数据关系模型到 XML 的转换	88
6.3	在气象中的应用	96
6.4	本章小结	100
<b>参考文献</b>		<b>101</b>



## 第四部分 模糊时空数据在 XML 与面向对象数据库之间的转换

第 7 章 模糊时空数据模型 .....	107
7.1 模糊时空数据 .....	107
7.2 模糊时空 XML 数据模型 .....	112
7.3 模糊时空面向对象数据模型 .....	118
7.4 本章小结 .....	123
第 8 章 模糊时空数据从面向对象数据库到 XML 的转换 .....	124
8.1 模糊时空数据从面向对象数据库到平面式 XML 的转换 .....	125
8.2 模糊时空数据从面向对象数据库到嵌套式 XML 的转换 .....	135
8.3 应用与验证 .....	144
8.3.1 建模热带气旋 Sandy .....	144
8.3.2 Sandy 的 FSODM Schema 到 XML Schema 的转换 .....	147
8.4 本章小结 .....	154
第 9 章 模糊时空数据从 XML 到面向对象数据库的转换 .....	155
9.1 模糊时空数据从平面式 XML 到面向对象数据库的转换 .....	155
9.2 模糊时空数据从嵌套式 XML 到面向对象数据库的转换 .....	160
9.3 应用与验证 (Sandy 的 XML Schema 到 FSODM Schema 的转换) .....	165
9.4 本章小结 .....	170
参考文献 .....	172

## 第五部分 模糊时空 XML 数据查询

第 10 章 模糊时空 XML 数据查询方法 .....	179
10.1 Dewey 编码扩展 .....	179
10.2 模糊度计算 .....	183
10.3 模糊时空属性匹配 .....	185
10.4 Twig 查询算法 .....	188
10.4.1 数据结构 .....	188
10.4.2 FSTTwigFast 算法 .....	189
10.5 算法分析 .....	193
10.6 实例分析 .....	194
10.7 实验评估 .....	196
10.7.1 实验环境 .....	196
10.7.2 查询效率比较 .....	199
10.7.3 不同类型时空数据查询效率比较 .....	201

---

10.7.4	FSTTwigFast 算法与 TwigStack 算法查询效率比较 .....	202
10.8	本章小结 .....	204
<b>第 11 章</b>	<b>模糊时空 XML 数据查询系统</b> .....	<b>205</b>
11.1	查询系统的设计 .....	205
11.2	查询系统的应用 .....	208
11.3	本章小结 .....	215
参考文献	.....	216
彩图		

# 第一部分 基础篇

本部分主要介绍本书的研究背景及与研究内容相关的概念知识。第 1 章介绍并分析有关模糊时空数据的研究情况, 提出对模糊时空数据的 XML 建模、关系数据库建模和面向对象数据库建模的需求, 以及对 XML 与关系数据库和面向对象数据库之间的相互转换进行研究的必要性和可操作性, 同时概括分析模糊时空数据在 XML 中的不一致性问题及修复方案, 还说明在模糊时空 XML 数据查询方面的发展及研究的必要性。第 2 章主要涉及模糊时空数据的相关知识, 并根据后面章节研究的需要进行有侧重点的讲解。



# 第1章 绪 论

空间数据库和时态数据库在 20 世纪 90 年代之前是相互独立的、没有交集的研究领域,随着时态数据库和空间数据库的发展,科研人员发现了这两种数据库的关系,并逐渐将二者结合起来进行研究,一个新的研究领域——时空数据库由此诞生。时空数据库用来存储和处理包含时态和空间信息的数据,是一种复杂的系统,成为数据库研究领域的一个重要分支。经过近些年的研究,在时空数据建模、时空信息索引与查询等方面取得了很多重要的成果。同时,时空数据库在很多领域获得了广泛的应用,这得益于其可以为时空对象提供空间和时间信息的管理。由于在实际应用中存在大量的时空数据在时空数据模型的建立方面受到了更多研究者的关注。为了使建立的模型准确有效,当前主要有两种比较成熟的方法。一种是基于场的模型,这种模型认为现实世界有很多随空间变化而变化的属性,这些属性可以用连续的函数表示。例如,在一个地图中使用压强或者温度的等高线来表示一系列常量值的点。另一种是面向对象建模方法,这种方法认为可以把现实世界完全区分为可定义的可分离对象,如在气象图中,温度区域和多雨或多雾区域都用它们在数据库中唯一的属性进行区分。

众所周知,在时空应用中,信息绝大多数都是模糊的。所以很多研究者开始将模糊性与时空数据结合起来,寻求建立模糊时空数据模型的方法。有两种建模方法使用得比较多,第一种是使用模糊集理论对时空数据对象和它们的属性建模,第二种是使用面向对象建模的方法将空间和时间上的属性与一个模型框架相结合。

使用模糊集理论的建模方法倾向于关系数据库的应用。由于传统的关系数据库具有可靠性、成熟性和独立性的特点,目前大量数据都存储在传统的关系数据库中。由于关系数据库有利于存取数据,并且对关系数据的高效访问已经发展了数十年;数据库技术发展迅猛、数据模型的丰富多样以及数据库新技术层出不穷,使得数据库的应用领域广泛深入;同时模糊值也已经被构造成模型并且可用于处理关系数据库中的模糊信息,模糊数据与关系数据库的结合可以有效地解决模糊数据的存储问题。

应用面向对象建模的方法倾向于使用面向对象数据库。与传统的数据库模型相比较,面向对象数据库模型提供了有力的面向对象建模的能力,并且因此被用

来表示和处理在时空应用中的复杂的数据。同时,面向对象数据库模型能够处理对象的结构和属性之间的关系。但是,面向对象数据库的功能是存储数据,像电子商务和医疗健康服务这些快速发展的基于 Web 的应用,在生活中越来越普遍,为了满足基于 Web 应用的发展需求,需要一种新的技术将存储在面向对象数据库中的模糊时空数据运用到基于 Web 的应用中。

XML (extensible markup language) 是一种可扩展标记语言,是标准通用标记语言的子集。具有简单性、可读性和可移植性,逐渐被作为数据交换和数据融合的媒介。近年来,随着互联网的迅速发展和广泛普及,XML 应用越来越广泛。作为一种用于标记电子文件使其具有结构性的标记语言,XML 扮演了越来越重要的角色。与此同时,XML 树型结构以及半结构化等特征,更加有利于数据管理。因此,XML 是研究模糊时空数据模型的一个很好的工具,用于解决一些传统数据库中难以解决的问题,如在时态处理、空间信息处理等方面均有所建树。在时态处理方面,相对于在传统数据库的基础上对模糊时空数据进行复杂的时空扩展,XML 数据树可以添加带标签的边来标识模糊时空数据的时间属性、空间属性以及模糊属性,使其更加简单易操作。在处理空间信息方面,前人提出使用可能性理论和相似性关系定义新的标签来表达模糊性;并提出将时间段标记成一条边,作为时间维度,然后通过增加这条被标记的边来处理时间上的信息。同理,管理空间上的信息也可用类似的方法。这种方法支持对不同的模糊时空数据之间的关系确定。在模糊时空数据的查询方面,相比于对 SQL 进行时空扩展的 STSQL (spatio temporal SQL),使用 XML 对模糊时空数据建模,可以利用 XQuery (W3C 所指定的一套标准查询语言) 直接进行查询,更加简单易操作,前人还应用丰富的数据类型定义对象之间的聚合关系和继承关系。因此,XML 是研究模糊时空数据的可靠工具。此外,诸如 XML DTD (document type declaration) 和 XML Schema 等模式描述语言的发展,也为有效地表示、推理、查询 XML 数据提供了新的契机。从本质上来说,XML Schema 更加适合对复杂的模糊时空数据进行结构化处理。

然而,由于模糊时空数据自有的特点,在对模糊时空 XML 文档进行操作(更新、插入或删除)时,其节点和边可能会产生不一致性。这些不一致性主要有时间不一致、时空转换不一致、时空不一致三种。由于节点间的时间信息不同步以及信息传输延迟导致的节点间在时间理解上存在差异,称为时间不一致。产生原因为两种应用均有各自的时钟标准,导致对同一时间的观测值不同,从而引发了时间不一致。由于节点间的时间不一致,同时导致了它们对同一实体的位置观测存在差异,称为时空转换不一致。时间不一致通过时空转换不一致,进而造成了不同节点对同一实体的位置观测差异,这种差异导致了时空不一致。在模糊时空 XML 文档的操作中,这些不一致状态的出现,导致不能对未

进行修复的模糊时空 XML 进行数据操作, 因此本书提出了一种不一致性修复方法。

XML 数据以纯文本的形式进行存储, 提供了一种独立于软件和硬件的数据存储方法, 可以在不兼容的系统中自由交换数据, 是各种应用程序之间进行数据传输的最常用的工具。为了更好地共享数据信息, 对模糊时空数据在 XML 和关系数据库之间的研究存在理论意义和实际价值。现有的关于数据在关系数据库与 XML 之间的转换比较成熟的方法的应用范围都是常规数据, 但是这些方法对于模糊数据来说是不完全适用的。XML 的灵活性和简单性为在关系数据库中存放模糊时空 XML 数据增加了难度。针对模糊时空数据在关系数据库与 XML 之间的转换, 还需要进行深入的研究。这种转换将会解决模糊时空数据在关系数据库与 XML 之间不能自由进行信息共享、信息传递的问题, 提高模糊时空数据在关系数据库与 XML 之间的互操作性。

同时, 模糊时空数据在 XML 和面向对象数据库之间的研究也存在很大价值。为了扩展模糊时空数据在 XML 和关系数据库之间转换的研究领域, 近些年来, 还提出了模糊时空数据在面向对象数据库与 XML 之间转换的研究方向。随着计算机技术的发展, 人们对数据的处理能力也在不断增强, 对复杂数据的应用范围也有所扩大, 通过对模糊时空数据的研究, 人们可以更加有效地使用这些模糊的、动态的、空间的数据, 如对自然灾害的预测、移动目标的监测等方面。本书提出了对模糊时空数据在面向对象数据库和 XML 之间双向转换的方法体系的研究, 将从支持模糊时空语义的数据模型入手, 提出表达能力更强的模糊时空 XML 数据模型和模糊时空面向对象数据模型, 用以处理大量存在于网络应用和面向对象数据库中的模糊时空信息。本书深入研究了模糊时空数据在 XML 和面向对象数据库中的表示方法; 同时, 基于所建立的模糊时空数据模型和表示方法, 深入研究模糊时空数据在 XML 和面向对象数据库之间的相互转换, 并使用实际时空应用中的模糊时空数据, 对所提出转换方法作进一步的评估和论证。

对于 XML 查询技术的需求也日益增长, 大量的 XML 数据查询规范和查询技术在众多应用的需求下被提出来, 如 Lorel、XML-QL、XQL、Quilt、XQuery 等。这些查询语言都将路径表达式作为核心内容, 用路径表达式来描述 XML 中的元素在数据层次中的定位。目前众多 XML 查询处理方法中, 应用最广泛的是结构连接和整体 Twig 查询。结构连接就是将 Twig 模式分解为一系列二元结构关系, 结构连接算法从 XML 数据库中获得与二元结构匹配的数据, 将这些匹配的数据连接起来形成最终结果。这种方法会产生大量冗余的中间结果, 当内存无法容纳时, 频繁地换页会产生大量 I/O 操作, 严重影响系统的性能。然而, 整体匹配算法无须执行对给定的 Twig 查询进行分解处理再合并的操作, 可以最大程度地减少

不必要的中间结果,从而达到提高查询性能的目的。所以,近年来整体匹配的 Twig 查询成为一个热门研究课题。

整体 Twig 查询的研究日趋成熟,但是该查询方法是针对一般数据的。随着地理信息系统等大量基于时空信息应用的发展,时空数据应运而生。由于时空数据的时间和空间属性在实际应用中通常是模糊的,关于模糊时空数据的研究,尤其是查询模糊时空数据,已经获得了相关领域学者的广泛关注。XML 作为新一代互联网信息交换的标准语言,具有查询模糊时空数据的能力。同时,模糊时空数据中复杂的模糊时间和模糊空间属性不同于传统关系型数据库中的二维表,这点使得模糊时空数据查询应用的研究受到了很大的限制。XML 的树型结构可以很好地解决这一限制问题,同时在处理模糊数据方面也相对容易。然而,尽管关于模糊集合理论的研究为研究查询模糊时空数据和查询模糊 XML 数据提供了很好的理论基础,但是在定量查询模糊时空 XML 数据领域被提出的相关理论还是不够充足。在众多基于时空数据应用的需求下,定量查询模糊时空 XML 数据的研究迫在眉睫。

后续章节将围绕模糊时空数据在 XML 文档中的不一致性、在 XML 与关系数据库之间的转换和 XML 与面向对象数据库之间的转换,以及 XML 查询展开。



## 第2章 预备知识

本章主要介绍与模糊时空数据研究相关的基础内容，作为后面章节研究的理论基础，主要涉及时空数据的特点、模糊性的概念和时空数据中的模糊性，模糊集、隶属函数和可能性分布的介绍，XML 的基础知识、XML 中的模糊性和 XML 数据库查询语言、关系数据库与面向对象数据库的相关知识，Twig 查询以及 Dewey 编码的概念。

### 2.1 时空数据与时空数据的模糊性

#### 2.1.1 时空数据

时空数据是可以同时包含时间信息和空间信息的特殊数据，具有多源、海量、更新快速的综合特点。时空数据由于包含时间信息、空间信息和对象属性三方面的固有特征，呈现出多维、语义多变、时空信息动态化的复杂性。具体特点如下。

(1) 时空数据包含对象以及时间和空间等方面的关联关系。

(2) 时空数据具有属性随时间和空间改变以及动态变化、多维演化的特点，这些基于对象的时空变化是可度量的，其变化过程可以通过构建时空数据关联模型来描述对象的关联映射。

(3) 时空数据具有多尺度的特性，在建立时空数据关联模型时，需要考虑到关联机制的多尺度选择；针对不同尺度的时空数据关联模型，可实现对象关联关系的尺度转换与重建，进而实现时空数据关联模型的多尺度特性。

(4) 时空数据的时空变化具有随时、多尺度、多维度以及动态关联等特点，对约定的关联约束可进行面向对象分类，建立面向对象的关联约束机制，根据关联约束之间的相关性可建立面向对象的关联约束法则。

时空数据具有时间和空间两个维度上的属性，能更准确地描述对象的信息，可以实时地观察对象的阶段性行为特征，以及参考时空关联约束模型、观察和预测某特定的阶段性行为发生的可能性。可针对时空数据对象的观察与预测问题，研究空间数据对象行为的建模和规则库构建，为异常事件的数据挖掘和主动预警提供研究基础。