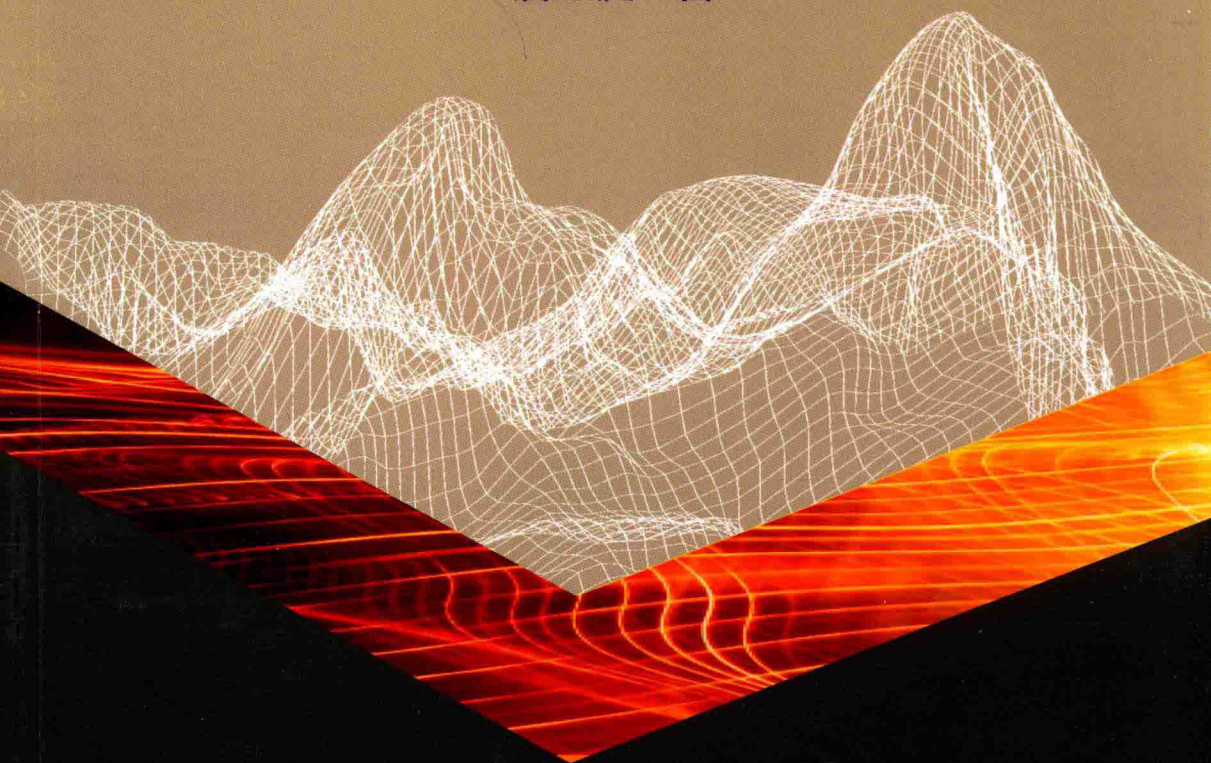


煤矿本体建模及应用

潘理虎 著



科学出版社

煤矿本体建模及应用

潘理虎 著

科学出版社

北京

内 容 简 介

近年来,煤矿安全生产技术得到快速发展,我国煤矿安全生产状况有明显改善。但由于煤矿生产环境的复杂性,煤矿事故仍时有发生,危害严重且影响社会稳定。

本书以提高煤矿安全生产监管技术水平为目的,在分析本体相关理论及煤矿领域知识的基础上,结合人工智能理论与方法,构建了煤矿领域本体模型,并将之应用于煤矿安全生产管理系统。主要阐述了基于 Jena 的本体模型推理本体半自动化构建、更新及不确定性推理等方法与技术,并分别介绍了煤矿本体在煤矿安全监测、事故逃生应急疏散及井下安全生产时空信息管理方面的应用案例。对本体理论在煤矿领域中的应用具有一定的理论与实践意义。

本书可作为高等院校煤矿安全、计算机应用等专业领域高年级本科生和研究生学习参考,亦可供煤矿安全相关从业人员及从事本体研究工作的专业人员参考。

图书在版编目(CIP)数据

煤矿本体建模及应用/潘理虎著. —北京:科学出版社,2018.11

ISBN 978-7-03-058279-9

I. ①煤… II. ①潘… III. ①煤矿-矿山安全 IV. ①TD7

中国版本图书馆CIP数据核字(2018)第158193号

责任编辑:常晓敏/责任校对:赵丽杰

责任印制:吕春珉/封面设计:耕者设计工作室

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>

北京中科印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2018年11月第一版 开本:B5(720×1000)

2018年11月第一次印刷 印张:14

字数:271 000

定价:99.00元

(如有印装质量问题,我社负责调换〈中科〉)

销售部电话 010-62136230 编辑部电话 010-62138978-8022

版权所有,侵权必究

举报电话:010-64030229; 010-64034315; 13501151303

前 言

安全生产是煤矿企业的生命线，也关系到国家社会经济的稳定发展。随着计算机技术在煤矿领域的广泛应用，该领域的安全生产和事故预防信息化技术得到快速发展，进而促进了煤矿安全生产水平的提高。本书从提高煤矿安全监测预警水平出发，将本体作为一种知识表达方式被引入煤矿安全监测监控领域，为煤矿安全生产提供统一的领域知识体系，实现信息的逻辑推理，并通过推理机制发现井下生产信息的隐含知识，从而提高煤矿井下安全事故预警的准确度，为煤矿井下生产活动的顺利进行提供保障。

本体用来表示相关领域的知识，使计算机能够像人类一样识别信息。20世纪90年代，人们将本体概念引入知识工程、信息技术、人工智能及企业管理中，使得本体理论在这些领域得到迅速发展。学者们创建了许多可广泛使用的本体，其中主要有美国普林斯顿大学研制的万维网，南加州大学的 GUM、SENSUS，CYKOAP 集团的 SAY，加州大学伯克利分校的 FRAMENET，以及得克萨斯大学的知网等。目前，本体仍是知识工程领域中一个重要的研究课题。国内外的学者为此做了很多研究，并且已经取得了一些成果。但针对煤矿安全生产领域的本体模型研究较少，且没有实现井下情境信息的智能推理与检索。由于煤矿安全生产环境及其过程的复杂性，在煤矿本体研究中还有许多问题需要结合实际应用予以解决。

本书共分为9章。第1章分析了煤矿领域本体模型的研究目的及意义，介绍了本体的国内外研究现状；第2章概述了本体的相关技术及其理论基础，指出了后续使用方法及技术的优势；第3章设计了煤矿领域本体的构建流程，实现了本体模型的构建，并给出了本体半自动化构建及更新的方法；第4章构建了用于本体推理的规则库，并引入了本体不确定性推理方法；第5章设计实现了一个基于本体推理的煤矿安全监测系统；第6章论述了基于煤矿本体构建煤矿安全知识图谱的方法与流程，以量化的方式对煤矿预警问题进行分析探究；第7章通过本体模型和煤矿相结合，搭建出了煤矿事故逃生系统的仿真本体模型；第8章在本体理论与面向对象思想的指导下，构建了一个基于本体的煤矿井下安全生产时空数据模型，设计并开发出一套煤矿井下安全生产时空信息原型系统；第9章总结本书主要的研究工作及成果。

本书得到“十二五”山西省科技重大专项项目“基于人-机-环联动煤矿井下生产全过程监测预警及重大事故救援指挥集成系统研究”和山西省-中国科学院科技合作重大项目“基于地理特征数据的煤矿安全生产关键技术研究”的支持，在

此向山西省科技厅表示感谢。

本书从规划、编写到出版离不开团队的支持和帮助。感谢太原科技大学郭勇义教授、陈立潮教授、张英俊教授，以及谢斌红、赵淑芳、赵红燕、翁自觉、谢建林、李川田等各位老师的关心和大力支持。感谢刘婷、李婉婉、张朝伟、药慧婷、张佳宇、芦飞平、李云凯、郭华等同学的辛勤工作。

感谢科学出版社各位编辑同志为本书顺利出版付出的努力。

由于自身能力及编写时间的限制，本书仍有许多不足之处，欢迎读者批评指正。

目 录

第 1 章 绪论	1
1.1 背景与意义	1
1.2 本体理论研究	2
1.3 本体应用研究	4
第 2 章 本体理论基础	6
2.1 本体定义	6
2.2 本体组成	7
2.3 本体分类	7
2.4 本体功能	9
2.5 本体评价	10
2.6 本体描述语言	12
2.6.1 基于人工智能的本体描述语言	12
2.6.2 基于 Web 的本体描述语言	13
2.7 本体构建	16
2.7.1 本体建模工具	16
2.7.2 本体构建原则	18
2.7.3 本体建模方法	19
2.8 本体推理	23
2.8.1 推理方法及其分类	23
2.8.2 常见推理机介绍	25
2.8.3 Jena 推理机	26
2.8.4 基于本体的推理技术	32
2.9 本章小结	34
第 3 章 煤矿领域本体构建和更新	35
3.1 本体构建方法及工具选择	35
3.2 煤矿领域本体构建步骤	37
3.3 煤矿领域本体实现	39
3.3.1 煤矿瓦斯监控系统本体模型	39
3.3.2 煤矿采煤工作面本体模型	41
3.3.3 煤矿掘进工作面本体模型	46
3.3.4 煤矿通风与运输系统本体模型	51

3.4	基于 FCA 的本体半自动化构建	57
3.4.1	FCA 相关理论	58
3.4.2	概念格构造算法	59
3.4.3	FCA 在领域本体构建中的应用	60
3.4.4	基于 FCA 的动态本体构建方法	61
3.4.5	煤矿动态本体构建的实现	63
3.5	基于语义相似度的本体概念更新	68
3.5.1	WordNet 简介	69
3.5.2	传统的相似度算法	70
3.5.3	基于 WordNet 的改进相似度计算	72
3.5.4	基于改进相似度算法的本体更新方法	74
3.5.5	实验及结果分析	76
3.6	本章小结	79
第 4 章	基于 Jena 的本体模型推理研究	80
4.1	基于 Jena 推理规则的创建	80
4.1.1	Jena 推理规则的定义	80
4.1.2	煤矿领域本体推理规则的创建	82
4.2	基于证据理论的本体不确定性推理	96
4.2.1	基于本体推理的煤矿安全监测	96
4.2.2	改进的 D-S 证据理论	97
4.2.3	面向井下环境评估的本体不确定性推理应用	99
4.2.4	实验及结果分析	101
4.3	本章小结	102
第 5 章	基于本体推理的煤矿安全监测系统	103
5.1	系统分析	103
5.1.1	可行性分析	103
5.1.2	需求分析	105
5.1.3	角色分析	107
5.2	系统设计	109
5.2.1	系统模块概述	110
5.2.2	详细设计	112
5.3	系统实现	116
5.3.1	注册登录界面	116
5.3.2	系统主界面	117
5.4	本章小结	122

第 6 章 基于本体的煤矿安全知识图谱研究	123
6.1 知识图谱	123
6.1.1 知识图谱的定义	123
6.1.2 知识图谱的应用	124
6.2 知识图谱的构建与设计	124
6.2.1 整体架构设计	124
6.2.2 知识抽取流程设计	127
6.2.3 本体设计	128
6.3 基于 Jena 的知识抽取	128
6.3.1 本体存储	128
6.3.2 RDF 生成	128
6.4 基于 Neo4j 的知识存储	128
6.4.1 Neo4j 介绍	129
6.4.2 知识图谱存储	129
6.5 知识图谱绘制的方法、步骤和工具	129
6.5.1 文献计量方法	129
6.5.2 统计分析方法	131
6.5.3 数据挖掘方法	131
6.5.4 知识图谱的绘制步骤	132
6.5.5 知识图谱的绘制工具	136
6.6 煤矿预警知识图谱	137
6.7 本章小结	142
第 7 章 基于本体的煤矿事故逃生应急疏散仿真系统	143
7.1 研究背景	143
7.2 建模技术简介	144
7.3 突发火灾的群体活动分析	146
7.3.1 群体逃生行为产生的主要动力性因素	146
7.3.2 火灾逃生中的主体分类	147
7.3.3 内部模型结构设计	147
7.3.4 逃生行为生成机制	148
7.3.5 判断功能与逃生模型	149
7.3.6 学习功能	150
7.3.7 主体之间的影响与策略	150
7.4 系统设计	151
7.4.1 系统总体结构	152
7.4.2 系统设计流程	153

7.4.3	本体到类图的转换	153
7.4.4	时序模块	154
7.5	主体行为设计	155
7.5.1	MineAgent 的逃生	155
7.5.2	MineAgent1 的逃生	156
7.5.3	MineAgent2 的逃生	156
7.5.4	MineAgent3 的逃生	157
7.6	基于本体概念逃生的案例演示	158
7.6.1	仿真界面	158
7.6.2	基于火灾位置的逃生仿真	159
7.6.3	基于主体地理认知差异性的仿真	160
7.6.4	基于主体特征差异性的仿真	160
7.6.5	融合传感器数据的仿真	161
7.7	本章小结	163
第 8 章	基于本体的煤矿井下安全生产时空数据模型	164
8.1	研究背景及内容	164
8.2	基础理论及相关技术	166
8.2.1	时空本体与时空对象	166
8.2.2	时空数据模型基本理论	168
8.2.3	Web 应用开发相关技术	171
8.3	煤矿井下安全生产时空数据模型	174
8.3.1	时空对象模型	175
8.3.2	面向对象的时空数据模型	177
8.4	煤矿井下安全生产时空信息系统	185
8.4.1	系统概述	185
8.4.2	系统架构设计	186
8.4.3	功能模块设计	192
8.4.4	数据库系统设计	193
8.4.5	系统实现及功能模块展示	197
8.4.6	实验与验证	204
8.5	本章小结	206
第 9 章	结束语	207
	参考文献	210

第1章 绪论

煤炭产业是我国经济发展的支柱产业之一，随着经济的发展，我国对煤炭资源的需求逐渐增加。然而，与国外煤炭开采技术先进的国家相比，国内煤炭开采过程中发生事故数量和危害程度仍然偏高。大量研究成果证实，对煤矿安全本质认识不够全面，管理不够科学，煤矿安全监测技术不够先进，是导致事故发生的主要原因。本体可以提供对相关领域知识的共同认可与理解，实现知识的重用和共享。对煤矿生产过程进行本体建模，可对煤矿生产要素及其关系提供统一的形式化表示方法，形成本体知识和推理规则。基于本体知识进行的逻辑推理，可对煤矿安全监控系统产生的数据进行量化分析，进而发现井下潜在的危险信息，提高煤矿灾害预警的效率，保障煤矿生产的顺利进行。

1.1 背景与意义

我国煤炭资源丰富，开采规模大。作为煤炭生产和消费大国，煤炭工业对我国地方经济的发展发挥巨大作用。但煤矿生产面临井下作业人员集中、生产系统工艺复杂、技术装备水平低等诸多问题，导致矿难事故频频发生（李婧，2015）。为降低煤矿事故死亡率，同时实现资源的永续利用，减轻环境压力，国家致力于开发使用新能源，但煤炭产业的基础性地位难以改变。为保证煤炭行业的可持续发展，发挥其在经济发展中的促进和支撑作用，必须严格控制矿井事故的发生，加强对煤矿安全管理工作的的高度重视和监督。

近年来，国家大力推进煤矿安全生产技术，调整能源结构，加强监管力度，提高安全生产能力，煤矿安全生产形势明显好转（孟现飞，2014）。但与世界主要产煤国家相比，我国煤矿百万吨死亡率仍高于其他国家几十倍，甚至上百倍（靳运章，2016）。根据中华人民共和国国家安全生产监督管理总局公布的数据统计，2000~2017年，全国各省由于煤矿事故致死的人数达到59 977人，年平均死亡人数为3332人，煤矿安全现状不容乐观。矿井事故的发生不仅造成经济上的巨大损失，还威胁到煤矿职工的生命安全（丁振，2016）。因此，构建一个完整、通用的煤矿安全生产模型，对降低事故发生率、保障人员及设备的安全，促进煤炭产业的稳步发展具有重大意义。

计算机、网络和传感通信技术的迅猛发展及物联网的出现，为煤矿的安全监测和预警救援带来了新思路。针对当下中国煤矿安全信息化存在的大量问题，已有许多专家、学者应用计算机技术来解决该问题，如从20世纪90年代初开始，

加拿大国际镍公司研究自动采矿技术,并将在 2050 年实现某矿山的无人采矿,所有设备将通过卫星操纵,以此实现机械自动化采矿。1993 年,芬兰开始了智能矿山的计划,研究自动采矿技术,同时设计了采矿实时过程控制、高速通信网络、资源实时管理、新机械应用、自动采矿与设备遥控等 23 个专题。1999 年,美国对地下煤矿的自动定位与导航技术进行研究,并获得了商业化研究成果。智慧矿山是煤矿企业信息化发展的终极目标,也是煤矿安全生产领域的重点研究方向之一,已吸引了国内外众多研究机构参与到矿山的智慧化建设中,并且取得了显著成果。目前,很多国家都已制定了矿山信息化发展的长远计划,目的是实现煤矿生产的自动化和智能化。中国对煤矿生产自动化、智能化的研究也非常活跃,在计算机应用于煤矿安全生产的建设研究中也取得了喜人成果,一批煤矿安全建设优秀项目与成果脱颖而出,如开滦集团的企业信息化与电子矿图系统、潞安集团的建下压煤开采设计系统、神华集团神东公司的综合自动化采煤系统等。继“数字地球”概念出现后,吴立新等围绕数字矿山的相关问题展开了一系列的讨论与研究,包括中国矿业大学、煤炭科学研究总院等高校和研究单位都积极地进行了相关研究,继而涌现出了一大批关于智慧矿山建设的优秀项目与成果,其中山东招金集团研发的三维地测采煤生产辅助决策系统为煤矿安全开采提供了保障。

总之,物联网与计算机技术的普及,为加强煤炭生产的安全管理,提高煤矿安全监测水平,保障煤矿工人的生命安全提供了很大的技术支持。20 世纪 60 年代,计算机科学界引入了哲学中本体的概念,为进一步提高煤矿安全管理自动化、信息化水平,同时加强煤矿安全监测和事故预警,提供了新的途径。本体作为一种抽象化的知识表示模型,可以清晰地描述煤矿领域复杂多变的信息,并且能够有效组织和管理领域知识,使知识得到更好的重用和共享。通过构建煤矿安全生产领域通用的本体模型,构建煤矿井下人-机-环之间的关系,建立煤矿井下数据信息的统一规范,提供对该领域知识的共同理解。基于本体的监控系统实时监控设备及环境的状态、人员的行为,利用推理机和本体规则对生产过程中的情境信息进行分析 and 推理,发现潜藏在生产中的不安全因素,实现对井下人员、设备及环境安全状况的监测与预警,避免由传统煤矿监控监测系统对环境变化不敏感或工作人员的不安全行为所引起的煤矿事故,对于煤矿安全生产具有较大的理论与实践意义。

1.2 本体理论研究

“本体论”是哲学研究发展得来的一个概念,主要研究客观事物存在的本质和组成。随着科学技术的快速发展,知识工程、信息检索及机器学习等领域出现了大量的本体研究。20 世纪 80 年代,人工智能领域借鉴了本体的概念,国内外研究人员对本体建模做了深入探索。美国斯坦福大学知识系统实验室(knowledge

systems laboratory, KSL)对本体建模工具和本体应用领域都进行了深层次的研究,对后期研究人员构建本体模型具有指导意义。同时,该实验室的 Gruber 在 1993 年就提出了用于知识工程领域的本体定义,为本体概念的发展指出了方向。德国卡尔斯鲁厄大学的应用情报学和规范描述方法研究所 (Institute of Applied Informatics and Formal Description Methods, AIFB) 的 Rudi Studer、Alexander Maeche 对本体理论和数学表达作了深入的研究,该研究所进行了 20 余项有关本体的研究课题,涉及知识的表示与推理、本体工程及形式概念分析等多个领域,将本体的理论研究进一步推向了侧重于知识发现的应用阶段(龚资,2007)。2001 年,万维网联盟(W3C)开始研究本体,随后将 OWL 推荐为 W3C 的正式标准,为软件代替人工处理信息及实现信息推理奠定了基础(徐立广等,2006)。

为满足应用领域的需求,在本体推理的相关研究领域,有众多组织都投入了对于本体推理技术的研究,其中包括耶鲁大学、斯坦福大学、曼彻斯特大学等高校, DARPA、ERCIM 等研究机构, IBM、HP 等跨国企业。中国在本体推理技术方面的研究也在不断深入,已取得了较突出成果的有清华大学、北京大学和相关科研院所等。这些研究组织经过长期的研究,取得了一系列具有深刻影响的研究成果。如 2002 年 Kowalsky 等提出的简单本体的推理分析思想,2004 年 Qu 提出的 Web 本体语言的逻辑推理算法,2006 年 Seiji Koide 和 Hideaki Takeda 在亚洲语义网会议上论述的 OWL Full 语言的正向推理算法,为本体推理机的应用提供了技术支撑。近年来涌现出了大量本体推理机的研究成果(Choi et al., 2013)。例如, W3C 研发了基于描述逻辑的本体推理机,该推理机能够用来测试本体;还有目前更多使用的集成在语义网开发平台的推理引擎 Jena 和德国 Karlsruhe 大学开发的支持安全规则的 KAON2。Jena 推理机是美国 HP 实验室开发的一个开放的 Java 工具包,有较为完整的对本体进行存储、解析及推理的接口,目前是最受欢迎、应用最广泛的开源软件。2010 年 10 月, Jena 被成功纳入 Apache 软件基金会,并于 2012 年成为其顶级项目,可见 Jena 受到了本体研究人员的广泛欢迎。

虽然国内对本体理论及技术的研究相对落后于国外相关领域的研究,但是 20 世纪 90 年代初开始,致力于本体研究的科研力量层出不穷。其中,中国科学院最早于 1998 年就开始了对本体的研究。其下属的多个研究所和实验室一直都引领着国内本体的研究潮流。2000 年,中国科学院的陆汝钐院士、北京大学的金芝教授在“天马”专家系统开发环境研究中,提出了面向本体的需求分析(ontology-oriented requirements analysis, OORA)模型,该模型将本体作为增强面向对象方法表现力的关联手段,以此来实现信息系统的需求分析。中国农业科学院的李景在 2009 年进行了领域本体构建方法的研究,为本体方法在人工智能及知识工程领域的应用提供了指导。清华大学的学者研发了面向语义 Web 的本体系统支撑软件。Zhu 等在产品组装中使用了 SWRL/SQWRL 规则推理。Abdul-Ghafour 等利用本体及其推理有效地共享智能 CAD 设计中涉及的知识。Zhong 等(2012)提出通

过事件本体的推理技术来判断事件类的影响因素,对分析事件场景及灾害预警具有重要意义。Lee 等(2014)提出了基于本体推理的客户投诉智能处理方案,该解决方案为企业提供了一种基于知识的信息化方式来解决客户投诉问题。王毅等(2014)设计开发了基于本体的知识表示与推理模型,为建模知识的重用提供了有效的方法。此外,在本体构建方法的研究方面也有了一定成果,已有的本体建模方法有七步法、TOVE 法、KACTUS 法、Sensus 法等,但都缺乏通用性,没有形成一种完整统一的本体构建方法(Grandi, 2016; Chuprina, 2016)。

因此,本体相关技术的研究一直受到学术界的关注和重视。国内外对本体建模及应用的研究都比较活跃,也取得了一定成果。这些研究对煤矿领域进行本体建模、对煤矿安全本质的认识及煤矿井下完备信息的建立与检索具有很重要的意义。

1.3 本体应用研究

本体作为一种表达概念间层次结构与语义关系的模型,已被广泛应用在很多领域。特别在搜索引擎、知识工程、电子商务、信息抽取及软件复用等方面,都有很大进展(王晔等, 2014)。美国 Lenat 教授研制出了大型的常识知识库系统,能够在此基础上进行自然语言的学习、理解、问题求解与帮助等人类智能活动的研究(时卫静, 2009)。Embley 等提出了一种基于本体的文本信息抽取方法,利用本体从自然语言中进行信息的抽取与查询(安源源, 2014)。Chang 等提出了基于本体的知识管理方法,并且将该方法与图形化的模型工具相结合(段绍林, 2015)。Nitishal 等(2013)认为,通过知识共享可以实现制造系统的互动能力,并利用本体语言进行描述。除此之外,国际上许多研究机构在语义 Web 中引入本体层的研究,推出了一系列的相关标准、开发工具及解决方案,如修订 RDF 模型、开发出可扩展的 Protégé 等(毕强等, 2010)。

本体在国内的应用研究起步比较晚,但也取得了一定成果。1995 年曹存根提出了首创概念 NKI,陆汝钤院士构建了 Pangu 常识知识库,哈尔滨工程大学研究人员构建了企业本体论系统并对其作了相关研究等。近年来,针对本体建模及应用的研究逐渐兴起,如中国科学院的李景进行了以花卉学本体建模为例的本体理论及农业文献检索系统的应用研究,设计并构建了国内首个具有一定推理能力的领域本体模型,该系统能够实现智能推理与检索、概念检查与纠错等功能,具有本体检索系统的特有优势,为本体模型智能推理方面的研究提供了思路。北京林业大学的贾雪峰做了关于林业关键词表构建林业领域本体的研究,提出了基于本体中语义概念关系及其扩展机制的领域本体智能检索系统,在一定程度上对传统的检索模型进行了改进与完善。黄津津等(2012)提出了一种基于本体的 E-learning 个性化服务解决方法,并开发了相应的服务系统,为用户的学习和答疑提供服务。魏晓萍(2013)构建了肝炎病毒蛋白领域的本体模型,为生物学信息的快速检索

与查询提供了一种有效途径。王新媛用简化的本体架构构建了微博信息推荐系统，有效地提高了信息推荐的质量和精度（王新媛，2015）。

本体理论在知识工程、软件复用、信息检索等方面有了较多应用，但在煤矿行业研究较少。目前已有的研究有：郭华（2014）进行了煤矿瓦斯监控系统的本体模型的研究，该研究将本体模型引入煤矿瓦斯监控系统，通过对煤矿安全生产领域进行需求分析，建立了瓦斯监控的本体模型框架，并对该模型进行了评价，得到全面、可靠的本体模型，为解决瓦斯监控系统中海量数据、存储结构复杂导致的集成及共享问题，提供了行之有效的解决方案；药慧婷等（2016）构建了掘进工作面本体模型，为智能决策系统的实现提供了支持和保障；孟现飞构建了基于本体的煤矿事故预警知识库，为煤矿事故预防和预警提供了保障；戴露进行了煤矿地测数据的地理本体与网络服务的研究，为整个煤矿地测领域的管理、集成共享、网络服务提供可参考的意见；李靖进行了煤矿安全事件本体中案例推理的研究与应用，在事故发生后，通过案例推理，将事故处理的解决方案及时提供给决策人员，从而减少事故损失（李靖，2015）。

第 2 章 本体理论基础

本体是对共享概念模型的形式化规范说明，它通过使用一定范围内普遍认可的术语和概念，使对应领域中所涉及的知识具有唯一性，是目前广泛使用的一种知识表示方法。煤矿生产领域所涉及的学科众多、内容复杂，领域知识难以被生产人员利用。将本体方法运用到煤矿生产领域，建立领域知识本体模型，为煤矿知识共享和重用提供了一种新方法。

2.1 本体定义

17 世纪，日本学者首次将德国经济学院学者 Goclenius 提出的 Ontology 称为本体论。“本体”最初来源于哲学领域，其定义为“本体是关于事物客观存在的，是对客观存在事物的系统性解释和说明，关心的是客观现实的抽象本质”(Grandi, 2016)。随着信息技术的不断发展，本体技术逐步应用到人工智能领域，并被认为是能在一定程度上表示共享的知识和概念。本体的核心是获取现实世界中相关现象的本质。研究表明，由于人们选择不同的术语表达同一个事情，因此造成了知识难以共享的局面，为了解决知识共享中存在的问题，需获取事物的原本特质，采用统一的知识体系与结构及表达形式，使人更易接受。通过获取领域中的相关本质知识，使不同的主体之间进行交流，从而解决信息工程中出现的知识共享方面的问题。哲学上的本体概念是一个描述性的范畴，它和自然科学领域的本体概念有许多区别，不仅体现在其适用性上，而且体现在其研究的最首要目标上。科学上的本体概念是一个可以被量化的研究对象。哲学上的本体可以对某一领域的客观存在进行分类，这样的分类可用于量化的进程。科学上的本体不仅包含了某一领域知识的一些类、属性和实例，而且提供了研究这些实体和实体间关系的一个量化方法。由于本体在知识表示、信息检索系统和软件工程等方面都有应用，不同领域的学者不断加深对本体的理解，因此本体的概念经历了从开始被提出到逐渐成熟的过程 (Jiang et al., 2009)。

本体一直被计算机科学所用，但没有明确的定义，它的含义有许多种。科学领域中有关本体最经典的定义是 Gruber 提出的，即本体是“概念模型明确的规范化说明”(Gruber, 1993)。随后 Borst 对其定义加以延伸，将共享性纳入其中 (Borst, 1997)。Studer 等 (1998) 经过进一步分析和扩展，用 4 层含义概括本体，即概念模型、形式化、明确、共享性。所谓概念模型，是指将现实世界中的现象或事物抽象和简化为系统化的概念而得到的模型。形式化是指本体可以为计算机所理解

和处理。明确意味着本体所描述的概念的类型及这些概念的约束说明要有确切的定义。共享性是指本体以突出事物的共性为主, 所描述的概念是相关领域中共同认可的, 即公认的概念集。Perez 等(1999)在此基础上又归纳了本体的 5 个基本建模原语, 即类、关系、函数、公理、实例, 为本体的创建提供了特定的功能实现要素。

尽管本体的定义形式各不相同, 但就本质而言, 都是为了描述特定领域知识, 明确定义共同认可的术语集及术语间的关系, 为领域内的主体提供一种共识。本体研究世界万物存在的意义和本质。人们之所以构建本体, 是希望现有的知识库能在以后的研究领域中继续使用。为了使不同的知识系统共同享用某种相同的知识库, 需要建立本体模型。当用户查询领域知识时, 可减少响应时间, 且减少各个领域术语的混用现象, 避免了资源的过于重复和浪费。自本体被引入信息科学领域中用作知识表示的方法以来, 在语义网、信息检索、自然语言处理、知识问答系统等领域均得到了广泛使用。其应用可分为 3 类, 即人与组织之间的信息交流、系统之间的互操作、软件工程(郭晓黎, 2016)。

2.2 本体组成

一般而言, 研究者将本体表示为 O , 划分为 5 个基础建模元语, 即类、关系、函数、公理和实例。其中类又称为概念, 表示为 C , 指任何事物, 既可以是一般层面上的概念, 也可以是工作描述、行为、功能或任务等; 在语义层, 类指的是对象的集合, 一般定义为框架结构, 包含概念名称、概念之间的关系集合及用自然语言表述的概念。关系指的是概念在领域内的相互作用, 用 n 维笛卡儿积的子集表示为 $R: C_1 \times C_2 \times \cdots \times C_n$, 指概念类 C_1, C_2, \cdots, C_n 之间存在 n 元关系。函数指的是一种特殊的关系, 其中前 $n-1$ 个元素能够唯一的决定第 n 个元素。将其形式化地定义为 $F: C_1 \times C_2 \times \cdots \times C_{n-1} \rightarrow C_n$ 。公理指永真断言, 表示为 A 。实例指概念的基本元素, 表示为 I , 语义层面上讲, 实例就是概念的对象, 每个概念都有很多相关的实例, 但是每个实例只能属于一个概念(仇宝艳, 2009)。用数学表达式可将本体构成元素表示为: $O = \{C, R, F, A, I\}$ 。

本体概念间的基本关系从语义上分为 4 种: part-of、kind-of、instance-of、attribute-of。但在实际构建本体模型过程中, 不仅存在以上的 4 种基本关系, 也不一定要严格按照本体的 5 类元语来构建本体模型, 一般根据具体情况制定相应的关系来满足应用的需要。

2.3 本体分类

按照不同的属性, 可以对本体进行不同的分类。按照描述对象不同可划分为

4 类，如表 2.1 所示，这种方法也是学者普遍认同的分类方法（云红艳等，2015）。

表 2.1 本体按描述对象分类

本体类型	特点
领域本体 (domain ontology)	关于某特定学科领域或现实世界一部分的本体，表达的是那些适合于该领域术语的特殊含义
通用本体或顶级本体、上层本体 (general or upper ontology)	关于具有普遍意义的客观世界的常识本体，描述最普通的概念及概念之间的关系，独立于特定的问题和领域
问题求解本体、应用本体 (problem-solved ontology/application ontology)	以具体问题求解为描述对象的本体，描述的是特定问题求解的概念及概念之间的关系
表示本体、元本体 (representation ontology/meta-ontology)	关于知识表示语言的本体，在表示本体中，类、对象、关系、属性等术语都经过严谨的分析和定义

根据不同的角度，本体还可以分成许多种类型。现在比较流行的是对本体进行研究之后，提出通过对本体的具体描述程度及对某个领域的依赖程度来判别分类。还有学者根据了解本体的具体描述程度、细化程度的高低对本体进行分类，即分为参考本体（详细程度高）和共享本体（详细程度低），参考本体比较具体，共享本体则比较抽象、简单。

如图 2.1 所示，常见的依照本体对某个领域依赖程度分类的本体分层结构，包括领域本体、顶级本体、应用本体及任务本体 4 种类型。

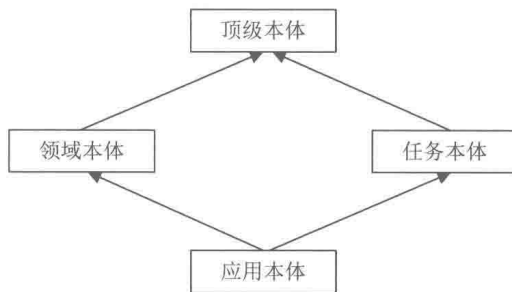


图 2.1 本体分层结构

在本体分层结构中，第一层是用来描述比较常用的概念及概念之间各个关系的相关顶级本体，用它能够描述时间、空间等各个领域，但是与具体的某个领域没有实质性的关系。第二层则是具有描述某一个相关领域信息的领域本体，及为了解决某个问题而构建的任务本体，以上两种本体可以使用上一层中的本体概念。最下面的一层则是对于某一种应用构建的应用本体，它对于相应的领域依赖程度比较大，而且是十分具体的。通过对以上两个分类方法进行细化，发现 10 种类别的本体，它们分别是语句本体、知识本体、表达本体、顶级本体、领域本体、元本体、应用本体、分别本体、方法本体、任务本体，不过以上的这些本体不够细化，它们之间存在交叉重叠。