

基于本体论的应急系统知识库建模的研究

摘 要

近年来频频爆发的恐怖活动、地震、洪水、火灾、流行性传染病、煤炭安全生产以及稳定输送电能中的重大事故等突发事件对人民群众生命财产和社会安全造成极大危害。突发事件妥善处理需要分布在不同地域、不同领域的相关部门的密切配合，因此为不同领域的决策者提供对某领域知识公共而一致的理解是目前应急决策支持系统需解决的主要问题。此外决策者在面临海量数据、信息的情况下，如何满足其对信息查准、查全的要求，实现以更低的成本、更准确的数据和更迅捷的速度做出大量精确、及时的决策是应急决策支持系统面临的又一个重要问题。

本体 (Ontology) 这一概念最早来源于哲学，用于描述事物的本质。知识工程学者借用这个术语及其基本思想，来建立本体知识库。基于本体论的知识库系统建模，是目前知识工程界研究的热点。本文把本体建模引入到应急决策领域，意义在于，根据本体建模和相关知识表示标准及交换协议，建立的知识库，便于知识交流、转换和共享。本体形式化地表达了领域知识的基本概念、属性、处理方法和内在关系。这种表述可被重用和共享，便于知识的查找。基于本体建立的知识库系统可靠性高，因为可以通过更可靠的软件对形式化表达的正确性进行检查。通过本体可以规范决策任务的形式化表达，有助于任

务的解析，便于明确定义决策过程中决策者和系统的角色及应急决策过程的任务。使用本体进行建模，可以使其系统结构、文档和编码都更清晰，便于系统的维护。

本文首先对决策支持系统进行分析研究，对国防动员应急决策问题进行描述，设计了国防动员应急决策支持系统总体框架。其次，对现有知识表示方法及表示观进行综述，对国内外现有本体库进行研究，在此基础上提出适合应急决策领域本体知识的图示化建模方法DKOED(Domain Knowledge Ontology for Emergency Decision)。DKOED方法建立在本体论的基本理论之上，具有朴素的知识建模思想，是一种将语义网络的知识表示与面向对象逻辑建模、面向数据的语义建模有机地结合的方法，能够有效地对领域结构、领域中的概念、概念的性质、概念间的关系等静态知识进行建模和推理。并应用该建模方法建立应急决策领域本体模型。本文作者对本体描述语言及本体编辑工具进行分析比较，选用本体编辑器OilEd和OWL语言生成该领域本体模型的OWL+RDF+XML的形式化描述文档。

最后，在构建领域知识本体模型的基础上，提出应急决策领域知识查询方法，对查询信息源数据的规范、查询系统的框架及查询模块进行了设计，并构造了查询语句，与传统基于关键词匹配的搜索引擎进行了对比分析。分析结果表明基于本体的知识查询可以解决传统搜索引擎难以解决的问题，如同义词、下位词以及对象之间的语义关系等，更好地满足决策者对信息查全和查准的需求。

关键词：本体，应急系统，领域本体，知识库建模，知识查询

A STUDY OF ONTOLOGY-BASED KNOWLEDGE BASE MODELING FOR EMERGENCY SYSTEM

ABSTRACT

Recently, the abrupt affairs such as terrible activity, earthquake, flood, fire, epidemic, the safety of coal production and power transmission, etc. break out frequently. All these endanger the people's life property and social safety largely. The appropriate processing of abrupt affairs need cooperation of the related section in different regions, different domains, so it is the primary problem to provide the common and coherent comprehension of certain domain knowledge for decision makers. Moreover, other important question that emergency decision support system will face to is how to search the information correctly and roundly, and how to make abundant precise and timely decision with exact data and fast speed at a less cost, when decision-makers confronting with massive data and information.

The term Ontology is derived from philosophy, and is defined as "Theory of existence". In computer science, the researchers use the basic idea of Ontology to build ontological knowledge base. The modeling of Ontology-based knowledge base system is one of the research hot spots in knowledge engineering. The significance of our introduction of ontology modeling method into emergency

decision is that the knowledge base, which is established according to ontology modeling method and related knowledge representation standard and exchange protocol, would benefit for the exchanging, transferring and sharing of knowledge. Ontology, in its formalization way, represents basic concepts, attributes, processing methods and inner relationships of knowledge in decision areas. This representation can be reusable and shareable, and make it convenient for the knowledge to be searched. Ontology-based knowledge base systems are of high reliability, because the correctness of formalization can be checked through even more reliable software. Representation of the formalization of tasks can be standardized by ontology, which is helpful to the analysis of tasks, the explicit definition of the roles of decision makers and systems in decision making process and the tasks of decision making process. Ontology-based system modeling can make the structure, document and coding of the system even more distinct, thus making it easy to be maintained.

After analyzing all kinds of decision support systems, the paper recounts the problem of defence mobilization emergency decision, and designs a general framework of emergency decision support system for national defence mobilization; then existing knowledge representation methods and scientific views of knowledge representation are summarized, we researched international ontologies, and put forward a new diagrammatic method for domain ontology modeling named Domain Knowledge Ontology for Emergency Decision (DKOED). Based on the foundational theory of Ontology, DKOED integrates the methods of semantic network, object-oriented modeling and semantic data modeling. DKOED is capable of modeling and reasoning static

domain knowledge, including domain structures, concepts and properties of the concepts, and the relations among concepts, etc.. With this modeling method emergency decision domain ontology model was constructed. I analyze and compare kinds of ontology description languages and ontology editing tools, select OilEd to build the formal description document of the domain ontology model which is based on OWL+RDF+XML.

Finally, after constructing the domain knowledge ontology model, we propose the method to query the emergency decision domain knowledge, and design the specification of the query information source data, the framework of the query system and the query model, and construct the query sentence, then compare with the search engine based on traditional key words matching. The results indicate that the knowledge query based on ontology can resolve many problems that traditional search engine fails to resolve, such as synonym, subword and semantic relation between objects etc., and it can satisfy the need of decision maker to search information more comprehensively and precisely.

KEY WORDS: ontology, emergency system, domain ontology, knowledge base modeling, knowledge querying

第一章 前言

1.1 研究背景

知识是“人们在社会实践中积累起来的经验”（辞海）。知识处理是在人类知识与计算机技术相结合的基础上，开展对知识的结构、分类、获取、存取、预测、归纳、传输、转换、表示、管理、利用（包括匹配、搜索和推理等）、知识的扩展及学习机制等问题的研究。信息时代必然要经历信息处理到知识处理的转换过程。信息处理，其产品形式是软件；知识处理，其产品形式是知识服务。知识处理是计算机系统发展的必然趋势^[1]。知识处理阶段离不开知识库系统，它是知识处理的基础。知识库系统为知识处理提供基础服务功能，包括知识的表示、存储、组织管理、维护、更新和查询，及知识推理和问题求解等。

本体论(Ontology)这一概念最早来源于哲学，用于描述事物的本质。知识工程学者借用这个术语及其基本思想，目的是为了了解决知识共享问题。人们发现，知识难以共享常常是因为大家对同一件事却用了不同的术语来表达。于是人们提出，如果能找出事物的本质，并以此统一知识的组织和知识的表达，使之成为大家普遍接受的规范，就有可能解决知识共享的问题^[2]。

早期人工智能领域的逻辑学的研究者已经开始研究信息系统中支持自动推理机制的知识模拟或者知识表达。当这些技术逐渐发展成为一种标准，其中蕴涵的知识表达的理论也成为人们关注的焦点。这些理论，即知识库(Knowledge-bases)主要指与规则联系的术语的集合，限制不合理的解释并支持从基本概念和事实中推导出新的知识。正是这种知识库上的自动推理技术被认为是人工智能的一种典型的应用。在这种情况下，一个新的研究领域：知识工程(Knowledge Engineering)应运而生。

第一章 前言

1.1 研究背景

知识是“人们在社会实践中积累起来的经验”（辞海）。知识处理是在人类知识与计算机技术相结合的基础上，开展对知识的结构、分类、获取、存取、预测、归纳、传输、转换、表示、管理、利用（包括匹配、搜索和推理等）、知识的扩展及学习机制等问题的研究。信息时代必然要经历信息处理到知识处理的转换过程。信息处理，其产品形式是软件；知识处理，其产品形式是知识服务。知识处理是计算机系统发展的必然趋势^[1]。知识处理阶段离不开知识库系统，它是知识处理的基础。知识库系统为知识处理提供基础服务功能，包括知识的表示、存储、组织管理、维护、更新和查询，及知识推理和问题求解等。

本体论(Ontology)这一概念最早来源于哲学，用于描述事物的本质。知识工程学者借用这个术语及其基本思想，目的是为了了解决知识共享问题。人们发现，知识难以共享常常是因为大家对同一件事却用了不同的术语来表达。于是人们提出，如果能找出事物的本质，并以此统一知识的组织和知识的表达，使之成为大家普遍接受的规范，就有可能解决知识共享的问题^[2]。

早期人工智能领域的逻辑学的研究者已经开始研究信息系统中支持自动推理机制的知识模拟或者知识表达。当这些技术逐渐发展成为一种标准，其中蕴涵的知识表达的理论也成为人们关注的焦点。这些理论，即知识库(Knowledge-bases)主要指与规则联系的术语的集合，限制不合理的解释并支持从基本概念和事实中推导出新的知识。正是这种知识库上的自动推理技术被认为是人工智能的一种典型的应用。在这种情况下，一个新的研究领域：知识工程(Knowledge Engineering)应运而生。

近几年来, 研究人员把哲学上本体论 (Ontology) 概念引入到知识工程中, 基于 Ontology 对知识库进行建模。Ontology 建模方法属于基于知识的系统 (Knowledge-based System ,KBS) 的开发, 但一般开发知识库的方法不能完全适用于本体的建模。原因是通常开发知识库系统时, 知识工程师很难定义系统在实际领域中具体、完整的工作方式, 所以一般采用原型方法。而基于 Ontology 建模, 是为了人类、计算机对知识的共享、重用和互操作, 相对稳定、独立于具体应用。本体建模的起点是详细说明模型中涵盖的概念、实例、关系和公理等实体, 至少是初步认定这些实体的绝大多数词汇^[3]。针对某个领域抽象或概括出一组概念及概念间的关系, 构造出该领域的本体, 会使计算机对该领域的知识处理大为方便, 也便于人们在此基础上开发应用系统。本体论正成为知识库系统建模、知识获取及表示、数据库框架集成等研究领域共同关心的一个核心。

从本体研究的前景来看, 在 1998 年 6 月召开了关于本体的第一届国际会议——“信息系统中的形式化本体国际会议”。这些都标志着这一研究领域正在走向繁荣, 该领域中的研究者们对“本体”的研究主题已经有了广泛的共识, 本体的研究将成为今后一个更加活跃的研究领域。

1.2 国内外本体研究动态

目前, 在计算机科学界, 本体论方法的研究、开发和应用正在广泛开展。在很多著名的知识系统中, 如美国 D. Lenat 教授领导研制的大型常识知识库系统 CYC, Princeton 大学 Berkeley 分校研制的语言知识库 WordNet, 国内主要有中科院计算所曹存根研究员领导的建设国家公共知识基础设施 CNKI 工程, 将提供大规模的、包含多个领域的知识共享本体库; 由中科院计算机语言信息中心语言知识研究室主任董振东建设的知网工程等等, 都在一定程度上应用了本体论的方法。

1.2.1 CYC 简介

CYC 在知识表示方面与其他系统相比最大的优势是它采用了形式化

近几年来, 研究人员把哲学上本体论 (Ontology) 概念引入到知识工程中, 基于 Ontology 对知识库进行建模。Ontology 建模方法属于基于知识的系统 (Knowledge-based System ,KBS) 的开发, 但一般开发知识库的方法不能完全适用于本体的建模。原因是通常开发知识库系统时, 知识工程师很难定义系统在实际领域中具体、完整的工作方式, 所以一般采用原型方法。而基于 Ontology 建模, 是为了人类、计算机对知识的共享、重用和互操作, 相对稳定、独立于具体应用。本体建模的起点是详细说明模型中涵盖的概念、实例、关系和公理等实体, 至少是初步认定这些实体的绝大多数词汇^[3]。针对某个领域抽象或概括出一组概念及概念间的关系, 构造出该领域的本体, 会使计算机对该领域的知识处理大为方便, 也便于人们在此基础上开发应用系统。本体论正成为知识库系统建模、知识获取及表示、数据库框架集成等研究领域共同关心的一个核心。

从本体研究的前景来看, 在 1998 年 6 月召开了关于本体的第一届国际会议——“信息系统中的形式化本体国际会议”。这些都标志着这一研究领域正在走向繁荣, 该领域中的研究者们对“本体”的研究主题已经有了广泛的共识, 本体的研究将成为今后一个更加活跃的研究领域。

1.2 国内外本体研究动态

目前, 在计算机科学界, 本体论方法的研究、开发和应用正在广泛开展。在很多著名的知识系统中, 如美国 D. Lenat 教授领导研制的大型常识知识库系统 CYC, Princeton 大学 Berkeley 分校研制的语言知识库 WordNet, 国内主要有中科院计算所曹存根研究员领导的建设国家公共知识基础设施 CNKI 工程, 将提供大规模的、包含多个领域的知识共享本体库; 由中科院计算机语言信息中心语言知识研究室主任董振东建设的知网工程等等, 都在一定程度上应用了本体论的方法。

1.2.1 CYC 简介

CYC 在知识表示方面与其他系统相比最大的优势是它采用了形式化

近几年来, 研究人员把哲学上本体论 (Ontology) 概念引入到知识工程中, 基于 Ontology 对知识库进行建模。Ontology 建模方法属于基于知识的系统 (Knowledge-based System ,KBS) 的开发, 但一般开发知识库的方法不能完全适用于本体的建模。原因是通常开发知识库系统时, 知识工程师很难定义系统在实际领域中具体、完整的工作方式, 所以一般采用原型方法。而基于 Ontology 建模, 是为了人类、计算机对知识的共享、重用和互操作, 相对稳定、独立于具体应用。本体建模的起点是详细说明模型中涵盖的概念、实例、关系和公理等实体, 至少是初步认定这些实体的绝大多数词汇^[3]。针对某个领域抽象或概括出一组概念及概念间的关系, 构造出该领域的本体, 会使计算机对该领域的知识处理大为方便, 也便于人们在此基础上开发应用系统。本体论正成为知识库系统建模、知识获取及表示、数据库框架集成等研究领域共同关心的一个核心。

从本体研究的前景来看, 在 1998 年 6 月召开了关于本体的第一届国际会议——“信息系统中的形式化本体国际会议”。这些都标志着这一研究领域正在走向繁荣, 该领域中的研究者们对“本体”的研究主题已经有了广泛的共识, 本体的研究将成为今后一个更加活跃的研究领域。

1.2 国内外本体研究动态

目前, 在计算机科学界, 本体论方法的研究、开发和应用正在广泛开展。在很多著名的知识系统中, 如美国 D. Lenat 教授领导研制的大型常识知识库系统 CYC, Princeton 大学 Berkeley 分校研制的语言知识库 WordNet, 国内主要有中科院计算所曹存根研究员领导的建设国家公共知识基础设施 CNKI 工程, 将提供大规模的、包含多个领域的知识共享本体库; 由中科院计算机语言信息中心语言知识研究室主任董振东建设的知网工程等等, 都在一定程度上应用了本体论的方法。

1.2.1 CYC 简介

CYC 在知识表示方面与其他系统相比最大的优势是它采用了形式化

语言 Cycl, 将概念与陈述之间的推理性连接以机器可理解的方式编码。知识库的内容由大量的概念和关系的分类以及它们之间相互关系的丰富的形式化表示组成^[4]。知识库的布局如图 1-1 所示:

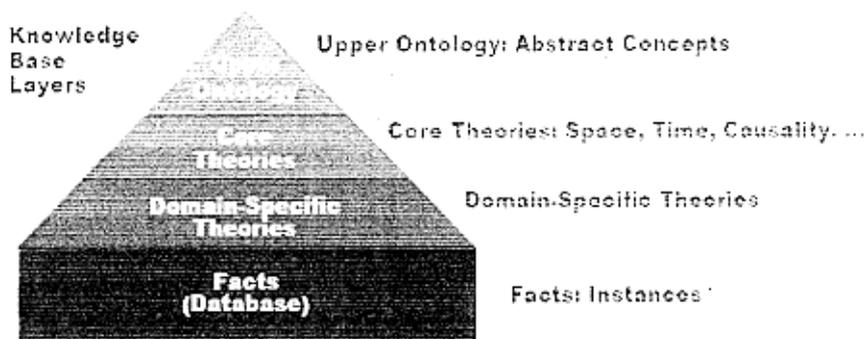


图 1-1 CYC 知识库布局

Figure 1-1 the arrangement of knowledge base in CYC

最上层是很小的抽象泛化层, 最低层是很大的现实世界事实层。Upper Ontology (上层本体) 描述非常泛化的概念之间的普遍关系。如 thing (事物) 是 CYC 中最泛化的概念, 任何事都是 thing (事物) 的一个实例。Core Theories (核心理论) 描述关于时间、空间和因果关系的普遍事实, 这些理论是所有常识推理的基础。例如, 对所有的事件 a 和 b, a 引起 b, 暗含 a 在 b 之前发生。Domain-Specific Theories (特定领域理论) 是比核心理论更加特殊化的一些理论。这些理论主要应用于特定的领域, 如军事活动、疾病传播、财政、化学等。例如, 对任何哺乳动物 m 和任何炭疽热细菌 a, m 接触到 a 就会造成 m 感染 a。Facts (Datebase) 最下面一层所包含的内容也称为底层事实, 它是对一些特定个体的描述。例如, 约翰感染了炭疽热, 这是对一个个体的特定的描述。

Cycl^[5] 是 CYC 系统的描述语言, 它是一种体系庞大而非常灵活的知识描述语言。该语言在一阶谓词演算的基础上, 扩充了等价推理、缺省推理等功能, 而且具备一些二阶谓词演算的能力。Cycl 中的基本表达式类型

包括：常量、函数、术语、谓词、量词等。常量代表个体、集合或集合的集合，如#GeorgeWBush, #Sudan, and #0-TheDigit 分别表示一个特定的个体，一个国家和一个数字。函数是带有论域并要有返回结果的，如#PresidentFn 需有一个国家作为其论域，并且返回一个总统的名字作为其结果。用于表示关系的常用术语包括#isa、#genls、#typeGenls、#disjointWith，其中#isa 关系表明某物是一个集合中的一部分；#genls 表明一个集合是另一个集合的子集；#typeGenls 表明两个集合的集合之间的子集关系；#disjointWith 表明两个集合的交集为空的情况，也就是说没有一个个体同时属于它所联系的两个集合。此外关系术语间的连接词包括：#genlPreds 和#genlInverse，关系术语连接词顾名思义是指连接关系术语的术语，#genlPreds 是指如果某事在第一个术语下为真的话，那么它在第二个术语下也为真。#genlInverse 类似于#genlPreds，但结论中的前件与后件中的位置要交换一下。Cycl 中的谓词，主要用来描述句法和语义的环境，以便写出很好形式化的句子。这类谓词主要包含：#sarity 和 #argxIsa。#sarity 指出一个谓词必须包含的论域的个数，也就是句法约束；#argxIsa 强加了语义约束，它限制术语的意义，指出谓词合法的论域类型。此外 Cycl 中还包括逻辑连接词：如#or、#and、#not 和#imply，以及量词如#forall(全称量词)和#thereExists(存在量词)等。

1. 2. 2 NKI 简介

NKI 是 1995 年曹存根研究员提出的一个在国际上首创的概念，全称是国家知识基础设施 (National Knowledge Infrastructure)。NKI 的目标是建立一个大型的可共享的知识群体。与以前人们通常说的数字化图书馆不同，国家知识基础设施的建设不是将各学科的书本拷贝进计算机，然后进行简单的主题分类处理，让读者自行查找和阅读有关的电子书本。在国家知识基础设施中，要对各学科知识 (如地理、军事、医学、历史、生物等) 进行深层次的概念分析和知识分析，研制一个可共享、可操作的庞大的专业知识群^[6]。

包括：常量、函数、术语、谓词、量词等。常量代表个体、集合或集合的集合，如#\$GeorgeWBush, \$\$Sudan, and \$\$0-TheDigit 分别表示一个特定的个体，一个国家和一个数字。函数是带有论域并要有返回结果的，如#\$PresidentFn 需有一个国家作为其论域，并且返回一个总统的名字作为其结果。用于表示关系的常用术语包括#\$isa、#\$genls、#\$typeGenls、#\$disjointWith, 其中#\$isa 关系表明某物是一个集合中的一部分；#\$genls 表明一个集合是另一个集合的子集；#\$typeGenls 表明两个集合的集合之间的子集关系；#\$disjointWith 表明两个集合的交集为空的情况，也就是说没有一个个体同时属于它所联系的两个集合。此外关系术语间的连接词包括：#\$genlPreds 和#\$genlInverse, 关系术语连接词顾名思义是指连接关系术语的术语，#\$genlPreds 是指如果某事在第一个术语下为真的话，那么它在第二个术语下也为真。#\$genlInverse 类似于#\$genlPreds, 但结论中的前件与后件中的位置要交换一下。Cycl 中的谓词，主要用来描述句法和语义的环境，以便写出很好形式化的句子。这类谓词主要包含：#\$arity 和 #\$argxIsa。#\$arity 指出一个谓词必须包含的论域的个数，也就是句法约束；#\$argxIsa 强加了语义约束，它限制术语的意义，指出谓词合法的论域类型。此外 Cycl 中还包括逻辑连接词：如#\$or、#\$and、#\$not 和#\$imply, 以及量词如#\$forall(全称量词)和#\$thereExists(存在量词)等。

1. 2. 2 NKI 简介

NKI 是 1995 年曹存根研究员提出的一个在国际上首创的概念，全称是国家知识基础设施 (National Knowledge Infrastructure)。NKI 的目标是建立一个大型的可共享的知识群体。与以前人们通常说的数字化图书馆不同，国家知识基础设施的建设不是将各学科的书本拷贝进计算机，然后进行简单的主题分类处理，让读者自行查找和阅读有关的电子书本。在国家知识基础设施中，要对各学科知识（如地理、军事、医学、历史、生物等）进行深层次的概念分析和知识分析，研制一个可共享、可操作的庞大的专业知识群^[6]。

在 NKI 中本体的描述采用框架形式, 将本体知识中的概念作为框架名, 其属性作为槽, 将属性值的类型、属性值的取值范围、属性值的数量以及其他的特征定义为侧面。

1. 2. 3 知网简介

知网(英文名称为 HowNet)是一个以汉语和英语的词语所代表的概念为描述对象, 以揭示概念与概念之间以及概念所具有的属性之间的关系为基本内容的常识知识库。知网建设的方法主要分为两个阶段, 分别为义原的提取和义原的考核与确定。知网系统包括下列数据文件和程序: 知网管理系统、知网管理工具和知网说明文件三部分, 其中知网说明文件中包括了动态角色与属性、词类表、同义、反义以及对义组的形成、事件关系和角色转换、标识符号及其说明等文件。而知识词典是知网系统的基础文件。在这个文件中每一个词语的概念及其描述形成一个记录。每一种语言的每一个记录都主要包含 4 项内容。其中每一项都由两部分组成, 中间以“=”分隔。每一个“=”的左侧是数据的域名, 右侧是数据的值^[7]。它们排列如下:

W_X= 词语

E_X= 词语例子

G_X= 词语词性

DEF= 概念定义

1. 3 选题的作用与意义

本课题研究的主题是“基于本体论的应急系统知识库建模的研究”。拟通过本体建模方法, 建立应急决策领域的本体模型, 在国防动员应急决策支持系统中利用该模型建立知识库。并在此基础上研究基于本体的信息检索, 实现决策者对信息查全率和查准率的要求。其意义有:

(1) 基于 Ontology 建立的知识库便于在不同系统之间的知识共享、知识的维护、知识的获取和互操作。

在 NKI 中本体的描述采用框架形式, 将本体知识中的概念作为框架名, 其属性作为槽, 将属性值的类型、属性值的取值范围、属性值的数量以及其他的特征定义为侧面。

1. 2. 3 知网简介

知网(英文名称为 HowNet)是一个以汉语和英语的词语所代表的概念为描述对象, 以揭示概念与概念之间以及概念所具有的属性之间的关系为基本内容的常识知识库。知网建设的方法主要分为两个阶段, 分别为义原的提取和义原的考核与确定。知网系统包括下列数据文件和程序: 知网管理系统、知网管理工具和知网说明文件三部分, 其中知网说明文件中包括了动态角色与属性、词类表、同义、反义以及对义组的形成、事件关系和角色转换、标识符号及其说明等文件。而知识词典是知网系统的基础文件。在这个文件中每一个词语的概念及其描述形成一个记录。每一种语言的每一个记录都主要包含 4 项内容。其中每一项都由两部分组成, 中间以“=”分隔。每一个“=”的左侧是数据的域名, 右侧是数据的值^[7]。它们排列如下:

W_X= 词语

E_X= 词语例子

G_X= 词语词性

DEF= 概念定义

1. 3 选题的作用与意义

本课题研究的主题是“基于本体论的应急系统知识库建模的研究”。拟通过本体建模方法, 建立应急决策领域的本体模型, 在国防动员应急决策支持系统中利用该模型建立知识库。并在此基础上研究基于本体的信息检索, 实现决策者对信息查全率和查准率的要求。其意义有:

(1) 基于 Ontology 建立的知识库便于在不同系统之间的知识共享、知识的维护、知识的获取和互操作。

在 NKI 中本体的描述采用框架形式, 将本体知识中的概念作为框架名, 其属性作为槽, 将属性值的类型、属性值的取值范围、属性值的数量以及其他的特征定义为侧面。

1. 2. 3 知网简介

知网(英文名称为 HowNet)是一个以汉语和英语的词语所代表的概念为描述对象, 以揭示概念与概念之间以及概念所具有的属性之间的关系为基本内容的常识知识库。知网建设的方法主要分为两个阶段, 分别为义原的提取和义原的考核与确定。知网系统包括下列数据文件和程序: 知网管理系统、知网管理工具和知网说明文件三部分, 其中知网说明文件中包括了动态角色与属性、词类表、同义、反义以及对义组的形成、事件关系和角色转换、标识符号及其说明等文件。而知识词典是知网系统的基础文件。在这个文件中每一个词语的概念及其描述形成一个记录。每一种语言的每一个记录都主要包含 4 项内容。其中每一项都由两部分组成, 中间以“=”分隔。每一个“=”的左侧是数据的域名, 右侧是数据的值^[7]。它们排列如下:

W_X= 词语

E_X= 词语例子

G_X= 词语词性

DEF= 概念定义

1. 3 选题的作用与意义

本课题研究的主题是“基于本体论的应急系统知识库建模的研究”。拟通过本体建模方法, 建立应急决策领域的本体模型, 在国防动员应急决策支持系统中利用该模型建立知识库。并在此基础上研究基于本体的信息检索, 实现决策者对信息查全率和查准率的要求。其意义有:

(1) 基于 Ontology 建立的知识库便于在不同系统之间的知识共享、知识的维护、知识的获取和互操作。

(2)由于基于 Ontology 的知识库建模主要研究领域知识中基本概念及其关系,所以,知识库相对稳定,也便于在分布式网络环境下对知识的存取。

(3)基于 Ontology 的知识库建模不仅是应急决策支持系统研究中的前沿,也是知识工程中的前沿课题,所以本课题的研究和探讨,不仅为知识工程中提供了例证,也为决策支持系统研究提供了有力的例证。

1.4 研究局限性

对于此题目的研究,必须面对一些哲学界和认知学界有争议的问题。本体工程是知识工程的发展和完善,那么它必然后要涉及计算机科学、语言学、认知学、逻辑学以及哲学这几门学科。当我们构造一个本体时,不可避免地会遇到上述各学科研究领域中的一些问题,有的问题甚至是在本学科中现在也没有一个明确的解决办法。

整体上来看,知识工程中对 Ontology 的研究还刚刚起步。尽管声称很多本体已被应用,但文献上并没有见到较详细的报道。就是 CYC,也未见到有关应用的情况。真正重用的实例不多,尽管构造本体的目的在于重用。从这方面来说,我们的选题具有挑战性,但是很有意义的。

1.5 文章的组织结构

全文共分为六章。

第一章介绍了课题的研究背景并概述相关领域的研究工作,提出本文的主要研究内容及选题的意义。

第二章主要对什么是知识及知识的表示方法和知识表示观进行了深入的分析 and 总结,提出了国防动员应急决策支持系统中知识的表示方法。

第三章对本体的有关理论进行了详细介绍。包括本体的定义、本体的构成、本体描述语言和编辑工具、本体功能、本体工程构建原则和方法以及现有本体的分类等内容。

第四章首先介绍了国防动员应急决策支持系统总体框架设计。其次,

(2)由于基于 Ontology 的知识库建模主要研究领域知识中基本概念及其关系,所以,知识库相对稳定,也便于在分布式网络环境下对知识的存取。

(3)基于 Ontology 的知识库建模不仅是应急决策支持系统研究中的前沿,也是知识工程中的前沿课题,所以本课题的研究和探讨,不仅为知识工程中提供了例证,也为决策支持系统研究提供了有力的例证。

1.4 研究局限性

对于此题目的研究,必须面对一些哲学界和认知学界有争议的问题。本体工程是知识工程的发展和完善,那么它必然后要涉及计算机科学、语言学、认知学、逻辑学以及哲学这几门学科。当我们构造一个本体时,不可避免地会遇到上述各学科研究领域中的一些问题,有的问题甚至是在本学科中现在也没有一个明确的解决办法。

整体上来看,知识工程中对 Ontology 的研究还刚刚起步。尽管声称很多本体已被应用,但文献上并没有见到较详细的报道。就是 CYC,也未见到有关应用的情况。真正重用的实例不多,尽管构造本体的目的在于重用。从这方面来说,我们的选题具有挑战性,但是很有意义的。

1.5 文章的组织结构

全文共分为六章。

第一章介绍了课题的研究背景并概述相关领域的研究工作,提出本文的主要研究内容及选题的意义。

第二章主要对什么是知识及知识的表示方法和知识表示观进行了深入的分析 and 总结,提出了国防动员应急决策支持系统中知识的表示方法。

第三章对本体的有关理论进行了详细介绍。包括本体的定义、本体的构成、本体描述语言和编辑工具、本体功能、本体工程构建原则和方法以及现有本体的分类等内容。

第四章首先介绍了国防动员应急决策支持系统总体框架设计。其次,

(2)由于基于 Ontology 的知识库建模主要研究领域知识中基本概念及其关系,所以,知识库相对稳定,也便于在分布式网络环境下对知识的存取。

(3)基于 Ontology 的知识库建模不仅是应急决策支持系统研究中的前沿,也是知识工程中的前沿课题,所以本课题的研究和探讨,不仅为知识工程中提供了例证,也为决策支持系统研究提供了有力的例证。

1.4 研究局限性

对于此题目的研究,必须面对一些哲学界和认知学界有争议的问题。本体工程是知识工程的发展和完善,那么它必然后要涉及计算机科学、语言学、认知学、逻辑学以及哲学这几门学科。当我们构造一个本体时,不可避免地会遇到上述各学科研究领域中的一些问题,有的问题甚至是在本学科中现在也没有一个明确的解决办法。

整体上来看,知识工程中对 Ontology 的研究还刚刚起步。尽管声称很多本体已被应用,但文献上并没有见到较详细的报道。就是 CYC,也未见到有关应用的情况。真正重用的实例不多,尽管构造本体的目的在于重用。从这方面来说,我们的选题具有挑战性,但是很有意义的。

1.5 文章的组织结构

全文共分为六章。

第一章介绍了课题的研究背景并概述相关领域的研究工作,提出本文的主要研究内容及选题的意义。

第二章主要对什么是知识及知识的表示方法和知识表示观进行了深入的分析 and 总结,提出了国防动员应急决策支持系统中知识的表示方法。

第三章对本体的有关理论进行了详细介绍。包括本体的定义、本体的构成、本体描述语言和编辑工具、本体功能、本体工程构建原则和方法以及现有本体的分类等内容。

第四章首先介绍了国防动员应急决策支持系统总体框架设计。其次,

在深入分析了领域知识概念建模的本体论基础上,提出了一种适合应急领域本体建模的方法DKOED,对DKOED方法构建领域本体的基本策略和步骤,基本建模元素及其图示化建模符号进行了详尽的阐释。并结合国防医疗卫生动员应急决策实例进一步解释DKOED方法建模过程,利用本体编辑器Oiled生成示例本体,给出了基于XML+RDF+OWL的形式化描述。

第五章对本体在信息检索中的应用进行了深入研究,提出了应急决策领域知识查询方法。对查询信息源数据的规范、查询系统的框架及查询模块进行了设计,并构造了查询语句,提高检索的查全率和查准率,与传统基于关键词匹配的搜索引擎进行了对比分析。

第六章总结了本文研究工作的成果并对未来的工作进行了简单的讨论。

第二章 知识及知识表示

建立知识库涉及到如下几个方面的问题：一是，关于知识的表示。二是，关于知识库的结构。三是，关于知识库的管理（包括知识的获取）。由于知识库的结构是由知识的表示方法决定的，因此，只要根据具体的问题或用户需求和系统的特点，选择好知识的表示方法，那么知识库的结构也就决定了。为了能准确的选择知识的表示方法建立有效的知识库，就需要对知识进行分析。本章在对知识的定义及其层次，知识表示方法及知识表示观进行分析的基础上，提出了国防动员应急决策支持系统中知识的表示方法。

2. 1 知识的定义及其层次

2. 1. 1 知识的定义

什么是知识至今还没有一个统一定义。在《辞海》中定义为：人们在社会实践中积累起来的经验。Fischler 把知识定义为：“知识是称为模型的存储信息，被人们用于解释、预测、并对外部世界做出适当响应。”Pasaye 和 Chignell 认为：“知识是允许人们进行决策的有关外部的信息。”国际经济合作与发展组织把知识看作是：know---what, know---why, know---how, know---who。后来有人增加了 know---when, know---where, know---how many 三类。那么，什么是知识呢？我们知道，大量的数据不经过加工处理是没有多少价值的，经过分析处理的数据形成了信息。信息的作用有时间和范围限制。为了使信息在较长时间内有效，必须进行一系列内部处理，这个过程叫做综合。综合后的信息组成知识。一般来说，所谓知识，就是人们在改造客观世界的实践中所获得的基本概念、认识、经验和规律，它是人类进行智能活动的基础。从计算机科学的观点来看，知识是信息综合

第二章 知识及知识表示

建立知识库涉及到如下几个方面的问题：一是，关于知识的表示。二是，关于知识库的结构。三是，关于知识库的管理（包括知识的获取）。由于知识库的结构是由知识的表示方法决定的，因此，只要根据具体的问题或用户需求和系统的特点，选择好知识的表示方法，那么知识库的结构也就决定了。为了能准确的选择知识的表示方法建立有效的知识库，就需要对知识进行分析。本章在对知识的定义及其层次，知识表示方法及知识表示观进行分析的基础上，提出了国防动员应急决策支持系统中知识的表示方法。

2. 1 知识的定义及其层次

2. 1. 1 知识的定义

什么是知识至今还没有一个统一定义。在《辞海》中定义为：人们在社会实践中积累起来的经验。Fischler 把知识定义为：“知识是称为模型的存储信息，被人们用于解释、预测、并对外部世界做出适当响应。”Pasaye 和 Chignell 认为：“知识是允许人们进行决策的有关外部的信息。”国际经济合作与发展组织把知识看作是：know---what, know---why, know---how, know---who。后来有人增加了 know---when, know---where, know---how many 三类。那么，什么是知识呢？我们知道，大量的数据不经过加工处理是没有多少价值的，经过分析处理的数据形成了信息。信息的作用有时间和范围限制。为了使信息在较长时间内有效，必须进行一系列内部处理，这个过程叫做综合。综合后的信息组成知识。一般来说，所谓知识，就是人们在改造客观世界的实践中所获得的基本概念、认识、经验和规律，它是人类进行智能活动的基础。从计算机科学的观点来看，知识是信息综合

处理的结果，在综合过程中，信息通过相互比较，结合成有意义的链接。数据、信息和知识具有层次关系⁶。如图 2-1 所示。

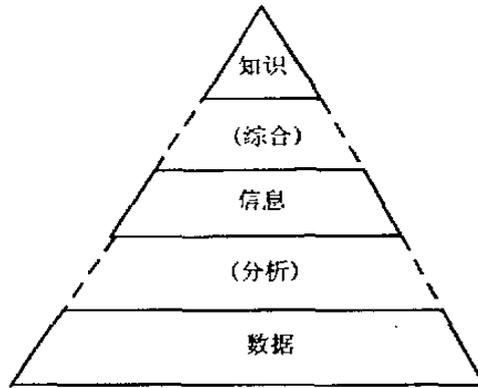


图 2-1 数据、信息和知识的层次关系

Figure 2-1 the relation of data, information and knowledge

2. 1. 2 知识的层次

人类的知识是极其丰富和庞杂的，如何规范地描述或表达这些知识，一直是人们所探讨的问题。通常人们习惯于用自然语言来表达知识，如果以自然语言表示知识直接引入到计算机，当然是最理想的了。但是，自然语言有二义性，语法和语义也难于有完善的描述。再之，计算机技术目前还不能很好的对自然语言进行有效的处理。所以，人们在不断地探讨计算机对知识的处理。

一般地，知识的表达模式为： $K=F+R+C$ 。

其中：K 表示知识项 (Knowledge items)。

F 表示事实 (Facts)：指人们对客观世界和世界的状态、属性和特征的描述，以及对事物之间关系的描述。

R 表示规则 (Rules)：指能表达在前提与结论之间因果关系的一种形式。

C 表示概念 (Concepts)：指事实 (术语) 的含义、规则的语义说明等。

为了把这些知识 (事实、规则和概念) 明白无误地用计算机所能接受的形式表示出来，必须建立一组约定的、利于把知识编码成一种适当的数

处理的结果，在综合过程中，信息通过相互比较，结合成有意义的链接。数据、信息和知识具有层次关系^[6]。如图 2-1 所示。

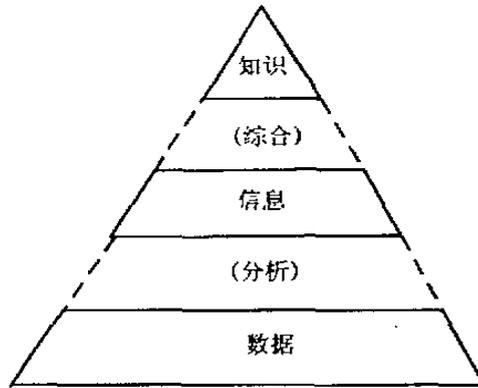


图 2-1 数据、信息和知识的层次关系

Figure 2-1 the relation of data, information and knowledge

2. 1. 2 知识的层次

人类的知识是极其丰富和庞杂的，如何规范地描述或表达这些知识，一直是人们所探讨的问题。通常人们习惯于用自然语言来表达知识，如果以自然语言表示知识直接引入到计算机，当然是最理想的了。但是，自然语言有二义性，语法和语义也难于有完善的描述。再之，计算机技术目前还不能很好的对自然语言进行有效的处理。所以，人们在不断地探讨计算机对知识的处理。

一般地，知识的表达模式为： $K=F+R+C$ 。

其中：K 表示知识项 (Knowledge items)。

F 表示事实 (Facts)：指人们对客观世界和世界的状态、属性和特征的描述，以及对事物之间关系的描述。

R 表示规则 (Rules)：指能表达在前提与结论之间因果关系的一种形式。

C 表示概念 (Concepts)：指事实 (术语) 的含义、规则的语义说明等。

为了把这些知识 (事实、规则和概念) 明白无误地用计算机所能接受的形式表示出来，必须建立一组约定的、利于把知识编码成一种适当的数

据结构，在计算机中存储起来。一旦计算机以适当的方式使用这些知识，就会产生智能行为⁹⁾。这就是知识表示要研究的问题。

从上述知识的表示模式可以看出，知识是有层次的，即，事实---概念---规则---启发式知识四个层次。其中启发性知识是关于规则的知识，是事实、概念和规则的综合。在问题求解中，利用启发性知识可以得到求解问题的捷径。对任何领域知识或学科知识的表示，都要进行知识的层次性分析，这样便于选取适当的知识表示方法。概括起来，在选择知识表示方法时，应该考虑以下几点^[10]：

(1) 所选择的知识表示方式对于特定领域知识和专家知识的特点及结构能充分地、适当地、正确地表达出来。这当然包括知识的层次性分析。具有表示某个专门领域所需知识的能力，并保证知识库中的知识是相容的。

(2) 要考虑其模块结构能否满足不断完善的需要。由于领域知识、专家知识/经验是不断完善的，其数据库也是在不断扩充和完善的。因此，选择知识表示方式时，要考虑其模块结构，以利于新的专家知识的获取和知识库的扩充、完善。

(3) 要考虑知识的表示是否条理清楚。也就是说知识表示是否简单、有效、清晰，是否便于对问题求解策略的推理和对知识库的搜索。

(4) 要考虑知识表示的自然性。即考虑知识表示的形式结构与人类对该领域的专门知识形式应该相吻合，以便容易被该领域的专家所理解，并有利于在构造和完善专家知识系统进程中，计算机软件设计人员与学科领域专家的合作。

(5) 透明性。在选择知识表示模式时，要考虑知识表示形式所表示知识的进程应该易于被人们所接受，即知识操作和控制进程易于被跟踪。

知识有多种表示方法。常用的表示方法有：谓词逻辑表示法、产生式表示法、语义网络表示法、框架表示法和面向对象的表示法等。最近人们又提出了具有语义表达能力的本体知识表示法，这也是本论文主要研究的表示方法。

2. 2 知识表示的研究

一般, 心理学家认为, 记忆、学习、问题求解、规划等是关于智能行为研究的基本课题, AI 将它们共同问题归结为表示、推理与搜索的研究。而现在, 人们则进一步认识到, 无论是推理还是搜索均在某种程度上依赖于表示^[9]。

知识表示是人工智能(AI)研究中基本问题之一, 对 AI 来说, 知识表示的研究分成两个层次: (1) 知识表示方法的研究, (2) 表示观的研究。基于“在计算机上再现智能行为”的考虑, AI 已提出了一些有价值的表示技术, 或称为表示方法 (Representation Scheme)。

表示是使用人造的体系(典型的例子是数学)对自然界事物的运动规律进行概括与抽象的模型, 而这个模型可以预言自然界这种运动的所有情况。一旦这样的表示被寻找到, 人们就认为这是对这类运动规律的更深刻的认识, 与自然的表示方法相比较, 这种表示应具有抽象性、深刻性与简洁性。对这种抽象意义上的表示, “自然界中的物体是否可被感知, 对表示没有什么直接的关系”, 换句话说, 表示与自然现象之间的形态上可以没有任何相同之处, 它是自然现象在人为体系规定下(公理)的一种解释。R. Davis 称这种意义下的表示为“替代(代理 Surrogate)”。

知识表示是概括智能行为的模型, 其特点是: (1) 智能行为所特有的灵活性问题(“常识问题”)不能概括为一类简洁的理论, 它是大量小理论的集合; (2) AI 的任务受到计算装置的约束。这就导致所采用的“表示”必须同时满足“刻画智能现象”与“计算装置可接受”这两个有时是矛盾的条件。正是对这两个条件的不同侧重导致了对“表示”的不同认识, 并由此产生 AI 研究的不同方法论。

由于对表示的作用及对实现“常识”推理的不同考虑, 导致了对表示的更深入的讨论, 这些讨论可归纳为三种“表示观”: (1) 认识论表示观; (2) 本体论表示观; (3) 知识工程表示观。“表示观”就是对智能行为的认识以及在研究中所采用的方法论。

2. 2. 1 表示方法

在 AI 中经常使用的表示方法, 几乎都是来源于研究者对智能行为在微观与宏观不同科学层次的观察与分析而抽象出的模型。根据这些表示方法的原理可以将它们分成三类:

(1) 局部表示类, 如逻辑、产生式系统、语义网络、框架、脚本、过程等。

(2) 分布表示类, 如基因、联接机制等。

(3) 直接表示类, 如各种图形、图像、声音及人造环境等。

由此, 一种知识表示方法的体系树可以被总结为图 2-2 的形式^[9]。

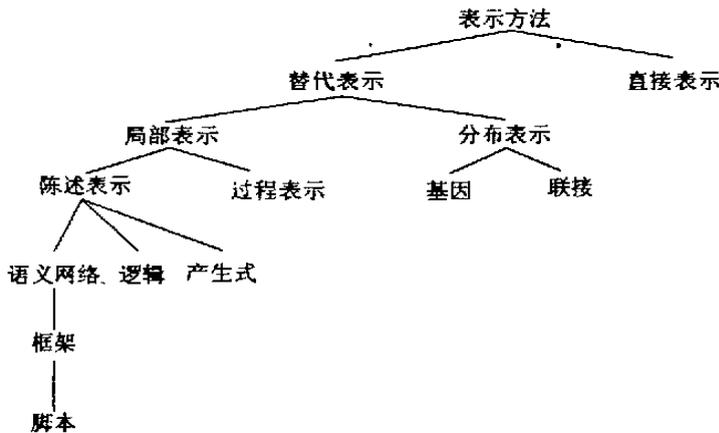


图 2-2 知识表示体系树

Figure 2-2 the architecture of knowledge representation

在这个知识表示体系树中, 局部表示是 AI 研究最充分也是正统 AI 最经常使用的表示方法。分布表示方法应该说是局部表示方法在智能行为描述上不够充分而作的补充。

而直接表示, 正在引起越来越多 AI 研究者的注意, 采用与自然世界一致的表示方法早在 60 年代初就已被提出, 现在, 这类表示方法称为“直接表示或拟真表示 (Direct or Analogical)”, 如地图、图形甚至音乐等。基于这类表示的系统是以对实体的拟真描述直接或间接参与推理为特点。

如果我们的讨论是考虑以计算机为载体对知识的编码, 则直接表示类不是一个可以完全独立于局部与分布表示类的方法。主要原因是考虑到“任何表示方法必须可以被计算机所接受”这个先决条件, 因此直接表示

类所采用的方法需要借助局部或分布表示的形式。对计算机而言，相对于局部与分布表示，直接表示可以视为外部表示，与其他内部表示相比较，它强调“表示与被表示实体间具有结构相似性”。

总之，尽管 AI 发展的历史已经提出了大量得到实践考验的有益表示方法，但是，表示问题并未真正解决。这几年，研究者正在试图对表示问题的本质进行更深入的讨论，由此形成了各种阐述研究者对智能行为研究的方法论及其本质认识的表示观。

2. 2. 2 表示观

目前，根据对“什么是表示”这一基本问题的不同理解和所采用的方法论，AI 学界对知识表示的观点主要可归纳为认识论 (Epistemology)、本体论 (Ontology)、及知识工程 (Knowledge Engineering) 三类。

认识论表示观假设：表示是对自然世界的描述，表示自身不显示任何智能行为，其唯一的作用就是携带知识，表示研究与“启发式”研究无关。

本体论表示观假设：表示是对自然世界的一种近似，它规定了看待自然世界的方式，即一个约定的集合。表示只是描述了在这个世界中，观察者当前所关心的那部分，其它部分则被忽略。

知识工程表示观：表示是对自然世界描述的计算机模型，它应该满足计算机这一实体的具体限制，因此，表示可以理解为一类数据结构及在其上的一组操作^[9]。

不同的表示观规定了智能模拟研究的不同侧重。例如，知识工程表示观强调自然世界在计算机内部某类数据结构的映象形式及对存储的内容所采用的处理方法，因此，研究知识的存储结构与对其有效地使用（推理与搜索）成为这种表示观研究的主要任务，这种表示观侧重“计算机可接受”这个条件。对认识论的表示观而言，表示是一种携带知识的理论，问题求解的有效性不在其考虑之列，它强调对自然现象（具体地说就是对常识知识）抽象与简洁的刻画。本体论的表示观则认为任何表示均是不完全的知识理论，而对其使用的有效性（计算困难程度）则是先决条件，因此，本体论的表示观强调一种聚焦的功能，“启发式”成为表示研究的一部分。

类所采用的方法需要借助局部或分布表示的形式。对计算机而言，相对于局部与分布表示，直接表示可以视为外部表示，与其他内部表示相比较，它强调“表示与被表示实体间具有结构相似性”。

总之，尽管 AI 发展的历史已经提出了大量得到实践考验的有益表示方法，但是，表示问题并未真正解决。这几年，研究者正在试图对表示问题的本质进行更深入的讨论，由此形成了各种阐述研究者对智能行为研究的方法论及其本质认识的表示观。

2. 2. 2 表示观

目前，根据对“什么是表示”这一基本问题的不同理解和所采用的方法论，AI 学界对知识表示的观点主要可归纳为认识论 (Epistemology)、本体论 (Ontology)、及知识工程 (Knowledge Engineering) 三类。

认识论表示观假设：表示是对自然世界的描述，表示自身不显示任何智能行为，其唯一的作用就是携带知识，表示研究与“启发式”研究无关。

本体论表示观假设：表示是对自然世界的一种近似，它规定了看待自然世界的方式，即一个约定的集合。表示只是描述了在这个世界中，观察者当前所关心的那部分，其它部分则被忽略。

知识工程表示观：表示是对自然世界描述的计算机模型，它应该满足计算机这一实体的具体限制，因此，表示可以理解为一类数据结构及在其上的一组操作^[9]。

不同的表示观规定了智能模拟研究的不同侧重。例如，知识工程表示观强调自然世界在计算机内部某类数据结构的映象形式及对存储的内容所采用的处理方法，因此，研究知识的存储结构与对其有效地使用（推理与搜索）成为这种表示观研究的主要任务，这种表示观侧重“计算机可接受”这个条件。对认识论的表示观而言，表示是一种携带知识的理论，问题求解的有效性不在其考虑之列，它强调对自然现象（具体地说就是对常识知识）抽象与简洁的刻画。本体论的表示观则认为任何表示均是不完全的知识理论，而对其使用的有效性（计算困难程度）则是先决条件，因此，本体论的表示观强调一种聚焦的功能，“启发式”成为表示研究的一部分。

这些表示观是从不同角度及不同描述层次解释表示的内涵而产生的不同的结论。但是,本体论表示观就不能因为其强调表示的不完善及可计算而否定它的知识携带作用,它与认识论表示的区别仅仅在于这种作用是否是唯一的。另外,由于本体论表示观承认表示与“启发式”研究之间的关系,因此,它与知识工程表示观必然紧密相关。

一般地说,认识论表示观强调知识的某种存在性研究,本体论表示观则更多考虑知识的构造性研究,而知识工程表示观则以知识系统的可实现性作为重点。显然,对任何一门学科,存在性、构造性及可实现性均是重要的,简单地否定某种表示观是不合适的,甚至是错误的。

2.3 国防动员应急决策支持系统中知识的表示

随着 TCP/IP 协议的流行和基于该协议的各种通信机制的开发,尤其是 CORBA 和 DCOM 等分布式对象技术的开发,异构硬件平台上应用软件之间的通信已基本消除障碍,然而信息和服务交互的最大障碍—通信内容的语义失配问题却始终未能有效解决。语义失配问题的根源是本体论失配—人们观察和处理事务时采用的术语概念和关系处理方法不一致。由于本体论隐含于应用程序的程序中,使得服务请求、返回结果和交换的信息易于被误解^[11]。要解决所涉及的各成员在通信过程中产生的语义冲突,只能通过定义局部的共享本体论来规范成员间的通信,以消除各成员可能对通信内容产生的误解。我们认为,清晰地表示通信内容的本体论是克服语义失配问题的唯一出路。

应急决策领域本体是一个特定研究论域的概念模型的形式化表达机制。它获取了领域内一致的知识,它不是为某个个体私有的,而是可以被一个群体所接受和共享的。该领域本体为需要共享应急决策领域的信息的决策者定义一个公共的词汇集。它包括应急领域概念的定义和概念之间的关系,这些都是机器可以理解的。

本体是论域中的一个概念(类,有时称为概念)、性质——描述每个概念不同的特征和概念的属性(人工智能中称为槽,有时候也称为角色或

者性质), 以及槽约束(刻面, 有时候也称为角色的约束)的一个形式化的显式描述。一个本体与类的实例的集合构成了知识库^[12]。

类是大多数本体的焦点。类描述了领域的概念。例如, 灾害的一个类表达了所有的灾害。特定的灾害是这个类的实例。一个类可以有子类, 子类表达了比超类更为明确的概念。

槽描述了类和实例的特征。例如, 灾害有一定的时空特征, 它有发生灾害的地点和时间, 地点可以用经纬度表示, 也可以用与之对应的地点名称表示, 它们之间可能是一个地理编码与经纬度的函数关系。因此, 可用时间、地点这样的槽描述灾害的实例。

为此, 我们提出将知识表示方法中的框架与逻辑表示法相结合, 作为本体知识的一种描述方法, 去定义国防动员应急决策支持系统中通信内容的共享本体, 并给出基于 OWL+RDF+XML 的形式化描述。

我们把应急决策领域中所涉及的概念进行分析, 用知识的框架表示来描述概念, 如对“灾害”这个基本概念的定义如下:

框架名: <灾害>

地点: 单位(经度, 纬度)

时间: 单位(年, 月, 日, 时, 分)

灾害等级: 单位(级别)

用描述逻辑来表示概念之间的关系, 以便进行知识推理。描述逻辑的基本组成部分是概念、角色和实例, 简单的概念和角色可以通过复合方式表示复杂的概念和角色。描述逻辑特别适用于表达结构化和半结构化数据之间具有层次、多重继承、聚合及其组合的情况, 被广泛用于知识建模、软件工程以及基于万维网的信息系统等领域。

描述逻辑的推理功能集中在以下两方面:

(1) 归约: 即判断一个概念是否为另一个概念的子集, 主要用于概念的自动分类。

(2) 相容: 即判断一个概念与已有的概念集(本体)是否相容, 主要用于进行概念集合的一致性检测。

根据不同的应用环境, 描述逻辑有许多变种, OWL 就是基于其中一类 SHIQ 的描述逻辑。因此, 基于 OWL 的本体建模, 可以通过转换为 SHIQ 表

达，并采用相应的推理机，实现概念的一致性检查。

下面给出基于 OWL+RDF+XML 的应急决策领域本体中对“自然灾害”的形式化描述：

```
<?xml version="1.0" encoding="UTF-8" ?>
<rdf:RDF xmlns=http://www.cyc.com/2003/04/01/cyc#
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
<owl:Ontology about="自然灾害">
</owl:Ontology>
<owl:FunctionalProperty rdf:ID="guid">
  <rdfs:label xml:lang="en">guid</rdfs:label>
  <rdfs:comment>Permanent Global Unique ID for the associated
concept.
</rdfs:comment>
  <rdfs:range
rdf:resource="http://www.w3.org/2000/01/rdf-schema#Literal" />
</owl:FunctionalProperty>
- <owl:Class rdf:ID="自然灾害">
<rdfs:label xml:lang="ch">自然灾害</rdfs:label>
<rdfs:comment> ] 旱、涝、虫、雹、瘟疫等造成的祸害
</rdfs:comment>
  <guid>99900001</guid>
<rdf:type rdf:resource="概念" />
<rdfs:subClassOf rdf:resource="灾害" />
<owl:disjointWith rdf:resource="人工灾害" />
</owl:Class>
```

通过上述文档将概念的定义及概念之间的关系进行了描述，在推理机中存放有领域理论，也就是相应的规则，由此在系统中各方交互时可以对信息进行本体论约定，并根据现有知识进行推理，以实现决策支持。

第三章 本体理论基础

本体工程是知识工程的发展和完善。知识工程发展至今,不论在关键理论、方法与技术方面,还是在具体领域应用方面,都取得了长足的进步,诸如专家系统的研制成功,知识系统的建立。与此同时,知识工程也遇到难以逾越的两大障碍:(1)如何实现知识重用。在知识工程中知识的获取极其困难,被认为是一个瓶颈问题,其代价相当昂贵。如果新的知识系统不能有效地利用现有的知识系统,不断开始“重新设计”,必将造成人力、物力的大量浪费。因此,当务之急是实现知识重用。(2)如何实现知识共享。仅有知识重用是不够的,因为它并没有涉及人机交互。在智能系统中,要实现人机交互就必须使人和机器的交流建立在对所交流领域共识的基础上。鉴于以上两个问题,人工智能专家们已经注意到:知识工程要想在目前的基础上有突破性的进展,必须向本体工程发展。因此“本体”作为本体工程的核心近年来被广泛地研究^[13]。

本体在知识库系统开发中较多应用于开发领域模型,它不仅包括了域中的知识,同时也提供了对域的正确理解。直观地讲,本体提供了建模所需的基本词汇并说明了它们之间的关系。建立大型知识库的第一步,就是设计相应的本体,这对于整个知识库的组织至关重要。本体已经在很多知识库系统中得到广泛的应用,影响较大的有美国国防高级研究计划署(DARPA)资助的高性能知识库 HPKB (High Performance Knowledge Bases)、欧洲数国联合开发的 IBROW 等。

另外,本体在自然语言处理领域中也广泛地应用。其主要原因是在自然语言处理的应用中要设法理解、表示和操纵文本的意思,而本体作为一部语义词典是必不可少的。这样,本体用以在语言独立的形式下表示文本的意思并消解歧义。在该领域中,本体主要应用于机器翻译、智能问答、智能检索等方面,应用实例有 WordNet^[14]、Mirokosmos^[15]、SENSUS^[16]等项目。

因此,在20世纪90年代初期,本体成为包括知识工程、自然语言处理和知识表示在内的诸多人工智能研究的热门课题。然而,目前有关本体论的研究还在雏形阶段,还没有形成一套完整的理论体系和构造方法,研究主要集中于:本体的理论基础、本体构造方法及技术、基于本体的知识推理、本体的具体应用等方面。

本章将着重讨论什么是本体,本体的建模,本体描述语言和编辑工具、本体功能、本体工程构建原则和方法以及现有本体的分类等内容。

3. 1 本体(Ontology)的定义

Ontology 是近年信息科学界最热门的词汇之一,国内一般将其译为“本体”。Ontology 最早是一个哲学上的概念,从哲学的范畴来说,Ontology 是客观存在的一个系统的解释或说明,关心的是客观现实的抽象本质。而人工智能领域的知识共享和重用又发展了本体的概念。在人工智能领域,解决问题的方法描述了推理行为,本体则描述知识系统的静态领域知识。它为人们及广泛异构的应用系统提供共同的领域知识理解。然而信息科学界对 Ontology 的理解也是逐步发展才走向成熟的。1991年 Neches 等人最早给出 Ontology 在信息科学中的定义:“给出构成相关领域词汇的基础术语和关系,以及利用这些术语和关系构成的规定这些词汇外延规则的定义”。后来在信息系统、知识系统等领域,随着越来越多的人研究 Ontology,产生了不同的定义。1993年 Gruber 定义 Ontology 为“概念模型的明确的规范说明”。1997年 Borst 进一步完善为“共享概念模型的形式化规范说明”。Studer 等人对上述两个定义进行了深入研究,认为 Ontology 是共享概念模型的明确的形式化规范说明,这也是目前对 Ontology 概念的统一看法^[17]。

Studer 等人的 Ontology 定义包含四层含义^[18]:概念化(Conceptualization)、明确(Explicit)、形式化(Formal)和共享(Share)。

(1) 概念模型:是指抽象出客观世界中一些现象(Phenomenon)的相关概念而得到的模型,其表示的含义独立于具体的环境状态;

因此,在20世纪90年代初期,本体成为包括知识工程、自然语言处理和知识表示在内的诸多人工智能研究的热门课题。然而,目前有关本体论的研究还在雏形阶段,还没有形成一套完整的理论体系和构造方法,研究主要集中于:本体的理论基础、本体构造方法及技术、基于本体的知识推理、本体的具体应用等方面。

本章将着重讨论什么是本体,本体的建模,本体描述语言和编辑工具、本体功能、本体工程构建原则和方法以及现有本体的分类等内容。

3. 1 本体(Ontology)的定义

Ontology 是近年信息科学界最热门的词汇之一,国内一般将其译为“本体”。Ontology 最早是一个哲学上的概念,从哲学的范畴来说,Ontology 是客观存在的一个系统的解释或说明,关心的是客观现实的抽象本质。而人工智能领域的知识共享和重用又发展了本体的概念。在人工智能领域,解决问题的方法描述了推理行为,本体则描述知识系统的静态领域知识。它为人们及广泛异构的应用系统提供共同的领域知识理解。然而信息科学界对 Ontology 的理解也是逐步发展才走向成熟的。1991年 Neches 等人最早给出 Ontology 在信息科学中的定义:“给出构成相关领域词汇的基础术语和关系,以及利用这些术语和关系构成的规定这些词汇外延规则的定义”。后来在信息系统、知识系统等领域,随着越来越多的人研究 Ontology,产生了不同的定义。1993年 Gruber 定义 Ontology 为“概念模型的明确的规范说明”。1997年 Borst 进一步完善为“共享概念模型的形式化规范说明”。Studer 等人对上述两个定义进行了深入研究,认为 Ontology 是共享概念模型的明确的形式化规范说明,这也是目前对 Ontology 概念的统一看法^[17]。

Studer 等人的 Ontology 定义包含四层含义^[18]:概念化(Conceptualization)、明确(Explicit)、形式化(Formal)和共享(Share)。

(1) 概念模型:是指抽象出客观世界中一些现象(Phenomenon)的相关概念而得到的模型,其表示的含义独立于具体的环境状态;

(2) 明确：是指所使用的概念及使用这些概念的约束都有明确的定义；

(3) 形式化：是指 Ontology 是计算机可读的，也就是计算机可处理的；

(4) 共享：是指 Ontology 中体现的是共同认可的概念集，它所针对的是团体而非个体；

Ontology 的目标是捕获相关领域的知识，提供对该领域知识的共同理解，确定该领域内共同认可的词汇，并从不同层次的形式化模式上给出这些词汇（术语）和词汇之间相互关系的明确定义。

尽管定义有很多不同的方式，但是从内涵上来看，不同研究者对于 Ontology 的认识是统一的，都把它当作是领域（领域的范围可以是特定应用中，也可以是更广的范围）内部不同主体（人、机器、软件系统等）之间进行交流（对话、互操作、共享等）的一种语义基础，即由 Ontology 提供一种共识。而且 Ontology 提供的这种共识更主要的是为机器服务，机器并不能像人类一样理解自然语言中表达的语义，目前的计算机也只能把文本看成字符串进行处理。因此，在计算机领域讨论 Ontology，就要讨论如何表达共识，也就是概念的形式化问题。

在上述这些对本体定义的不同观点中，所用到的“概念化”一词并没有给出明确的解释。Guarino^[19]对上述定义中的“概念化”给出了一种比较合理的解释，同时对概念化和本体的关系作了进一步的阐释。在此，我们概要地说明 Guarino 对“概念化”的解释。

定义 3.1 领域空间 (domain space)：领域空间定义为 $\langle D, W \rangle$ ，其中 D 表示领域， W 表示领域内事件最大状态的集合（也被称为可能世界）。

定义 3.2 概念上的关系 (conceptual relation)： $\langle D, W \rangle$ 上 n 元概念上的关系定义为 $\rho^n: W \rightarrow 2^D^n$ ，表示 W 到领域 D 上所有 n 元（普通）关系的集合的全函数。

对于概念上的关系 ρ ，集合 $E_\rho = \{\rho(w) | w \in W\}$ 包含 ρ 可接受的所有外延 (admittable extensions)。

定义 3.3 概念化 (conceptualization)：概念化定义为 $C = \langle D, W, \mathcal{R} \rangle$ ，

其中 $\langle D, W \rangle$ 为领域空间, \mathcal{R} 为 $\langle D, W \rangle$ 上概念关系 (conceptual relation) 的集合。

定义 3.4 目标的结构 (intended structure): $\forall w \in W, S_w, c$ 表示 w 关于 C 的目标的结构, $S_w, c = \langle D, R_w, c \rangle$, 其中 $R_w, c = \{\rho(w) \mid \rho \in \mathcal{R}\}$, 表示 \mathcal{R} 中概念上的关系关于 w 的外延的集合。

符号 S_c 表示概念化 C 的所有目标世界结构, $S_c = \{S_w, c \mid w \in W\}$ 。

假定逻辑语言 L 具有词汇表 V , 词汇表 V 由常量符号集合和谓词符号集合构成。Gruber 对语言模型的定义有别于标准定义, 将逻辑语言 L 的模型定义为结构 $\langle S, I \rangle$, 其中 $S = \langle D, R \rangle$ 表示一个世界结构, $I: V \rightarrow D \cup R$ 表示一个解释函数, 将 V 中的常量符号映射为 D 中元素, 将 V 中的谓词符号映射为 R 中的元素。

$\langle C, \mathcal{I} \rangle$ 表示 L 的内涵解释模型, 其中 $C = \langle D, W, \mathcal{R} \rangle$, $\mathcal{I}: V \rightarrow D \cup \mathcal{R}$ 表示一个解释函数, 将 V 中的常量符号映射为 D 中的元素, 将 V 中的谓词符号映射为 \mathcal{R} 中的元素。这种内涵解释被称为逻辑语言 L 的本体承诺 (ontological commitment)。如果 $K = \langle C, \mathcal{I} \rangle$ 是逻辑语言 L 的本体承诺, 称逻辑语言 L 通过本体承诺 K 承诺于概念化 C , 称 C 是 K 的基本概念化。

给定逻辑语言 L 和其词汇表 V , $K = \langle C, \mathcal{I} \rangle$ 是逻辑语言的本体承诺, 模型 $\langle S, I \rangle$ 与 K 兼容需要满足以下条件:

$$S \in S_c;$$

对每一个常量 c , $I(c) = \mathcal{I}(c)$;

存在一个可能世界 w , 对于每个谓词符号 ρ , I 将谓词映射为 $\mathcal{I}(\rho)$

允许的外延, 也就是存在一个概念上的关系 ρ , $\mathcal{I}(\rho) = \rho \wedge \rho(w) = I(\rho)$ 。

L 的模型 $M(L)$ 与 K 兼容的所有模型被称为 L 关于 K 的内涵模型, 记作 $I_k(L)$ 。

给定逻辑语言 L 和其本体承诺 $K = \langle C, \mathcal{I} \rangle$; L 的本体是指按照使本体的

模型集合最逼近 L 关于 K 的内涵模型集合的方式设计的公理集合。如果存在本体承诺 $K=\langle C, 3 \rangle$ 使本体 O 包含 L 关于 K 的内涵模型, 那么称语言 L 的本体 O 相似于概念化 C。

如果本体 O 的设计目的是为了描述概念化 C 的特征, 同时本体 O 相似于概念化 C, 那么称本体承诺于 C。如果逻辑语言 L 承诺于某个概念化 C, 以至本体 O 承诺于概念化 C, 那么逻辑语言 L 承诺于本体 O。

图 3-1 表示语言 L、本体 O 与概念化 C 之间关系的示意图。本体 O 是用于解释形式化词汇内涵意义的逻辑理论, 使用这种词汇表的逻辑语言 L 的内涵模型受本体承诺 K 的约束。本体通过接近这些内涵模型非直接地反映这些本体承诺, 本体 O 是语言相关的, 而概念化 C 是语言无关的。

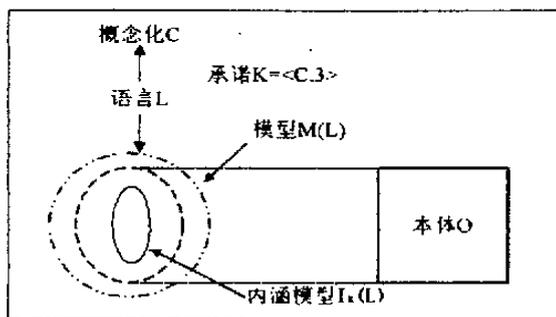


图 3-1 概念化、语言和本体关系图

Figure 3-1 the relations of conceptualization, language and ontology

3. 2 本体 (Ontology) 的建模元语

Ontology 是一种组织知识的艺术。为研究如何用 Ontology 来组织知识, Perez^[30]等人采用了分类法, 并归纳出 5 个基本建模元语: 类 (Classes) 或概念 (Concepts)、关系 (Relations)、函数 (Functions)、公理 (Axioms) 和实例 (Instances)。类或概念表示对象的集合; 关系表示领域中概念之间的交互作用, 形式上定义为 n 维笛卡尔乘积的子集: $R: C_1 \times C_2 \times \dots \times C_n$ 。如: 子类关系 (subclass-of); 函数是一类特殊的关系, 该关系的前 n-1 个元素可以唯一决定第 n 个元素。形式化的定义如下: $F: C_1 \times C_2 \times \dots \times C_{n-1} \rightarrow C_n$ 。例如 Mother-of 关系就是一个函数, 其中 Mother-of (x, y) 表示 y 是 x 的母亲, 显然 x 可以唯一确定他的母亲 y;

模型集合最逼近 L 关于 K 的内涵模型集合的方式设计的公理集合。如果存在本体承诺 $K=\langle C, 3 \rangle$ 使本体 O 包含 L 关于 K 的内涵模型，那么称语言 L 的本体 O 相似于概念化 C。

如果本体 O 的设计目的是为了描述概念化 C 的特征，同时本体 O 相似于概念化 C，那么称本体承诺于 C。如果逻辑语言 L 承诺于某个概念化 C，以至本体 O 承诺于概念化 C，那么逻辑语言 L 承诺于本体 O。

图 3-1 表示语言 L、本体 O 与概念化 C 之间关系的示意图。本体 O 是用于解释形式化词汇内涵意义的逻辑理论，使用这种词汇表的逻辑语言 L 的内涵模型受本体承诺 K 的约束。本体通过接近这些内涵模型非直接地反映这些本体承诺，本体 O 是语言相关的，而概念化 C 是语言无关的。

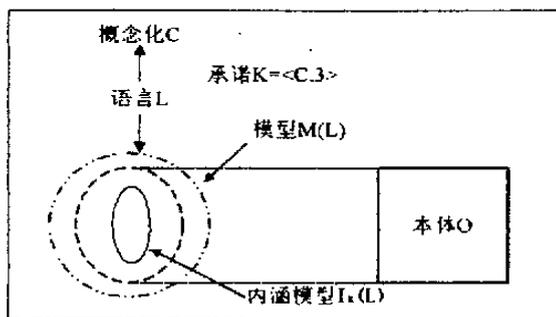


图 3-1 概念化、语言和本体关系图

Figure 3-1 the relations of conceptualization, language and ontology

3. 2 本体 (Ontology) 的建模元语

Ontology 是一种组织知识的艺术。为研究如何用 Ontology 来组织知识, Perez^[30]等人采用了分类法, 并归纳出 5 个基本建模元语: 类 (Classes) 或概念 (Concepts)、关系 (Relations)、函数 (Functions)、公理 (Axioms) 和实例 (Instances)。类或概念表示对象的集合; 关系表示领域中概念之间的交互作用, 形式上定义为 n 维笛卡尔乘积的子集: $R: C_1 \times C_2 \times \dots \times C_n$ 。如: 子类关系 (subclass-of); 函数是一类特殊的关系, 该关系的前 n-1 个元素可以唯一决定第 n 个元素。形式化的定义如下: $F: C_1 \times C_2 \times \dots \times C_{n-1} \rightarrow C_n$ 。例如 Mother-of 关系就是一个函数, 其中 Mother-of (x, y) 表示 y 是 x 的母亲, 显然 x 可以唯一确定他的母亲 y;

公理代表永真断言；实例代表元素，从语义上讲它表示的就是对象。

关系在 Ontology 中非常重要，从语义上讲，基本的关系共有 4 种：part-of、kind-of、instance-of 和 attribute-of。part-of 表达概念之间部分与整体的关系；kind-of 表达概念之间的继承关系，类似于面向对象中的父子类之间的关系，给出两个概念 C 和 D，记 $C' = \{x \mid x \text{ 是 } C \text{ 的实例}\}$ ， $D' = \{x \mid x \text{ 是 } D \text{ 的实例}\}$ ，如果对任意的 x 属于 D' ， x 都属于 C' ，则称 C 为 D 的父概念，D 为 C 的子概念；instance-of 表达概念的实例与概念之间的关系，类似于面向对象中的对象和类之间的关系；attribute-of 表达某个概念是另一个概念的属性。在实际建模过程中，概念之间的关系不限于上面列出的 4 种基本关系，可以根据领域的具体情况定义相应的关系。Ontology 正是通过这些建模元语，来组织现实世界的知识。图 3-2 用一个简单的有向图来表示一个本体。

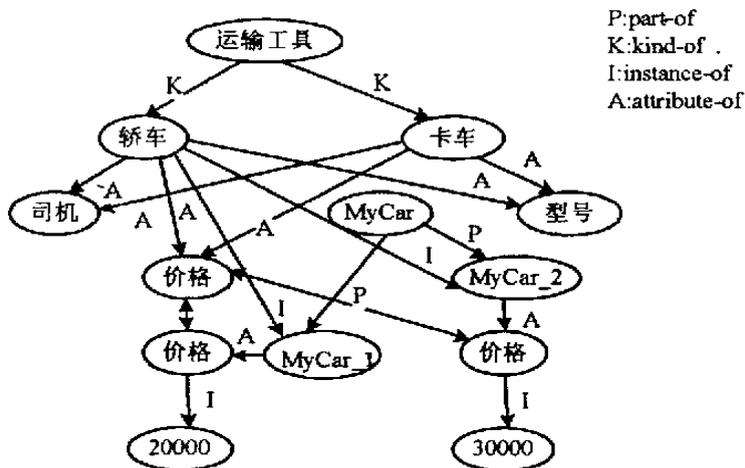


图 3-2 本体例子 (双箭头连接的是同一个概念) [12]

Figure 3-2 the example ontology(the same concepts are connected by double arrowhead)

从上面的分析可以看出，本体与面向对象的理论非常类似，但这两者本质是不同的。本体是某领域静态概念模型的描述，用公认的术语集和这些术语之间的关系来反映该领域内的知识和知识结构，不包括动态的行为。面向对象是一种软件开发方法，或者说它是一种新的软件开发规范，其主要思想是使用对象、类、继承、封装、消息和多态等基本概念来构造系统。在面向对象的理论中，对象或类不仅包括描述对象静态特征的属性，

也包括描述对象动态特征（行为）的方法。对象的属性和方法结合在一起形成一个独立的系统单位，尽可能对外隐藏对象的内部细节，提供统一的接口。为了提高面向对象领域中构件和构件库的重用能力，可以以面向对象领域为研究对象，在其上建立本体。

3. 3 本体描述语言及编辑工具

3. 3. 1 本体描述语言

为了描述并表示本体，近几年来出现了多种本体描述语言。在目前 XML 已成为数据交换格式描述标准的环境下，这些语言都以 XML 为基础来构建。下面将对几种主要的本体语言进行简要介绍^[21]。

1、XOL (XML-based Ontology Exchange Language)

这是由美国的一个生物信息学组织设计的 Ontology 语言，其目的是在针对不同领域设计的软件系统之间交换它们各自的 Ontology 定义。研究人员在仔细研究了生物信息学领域专家的信息表现需求后创建了这一语言。他们以 Ontolingua 和 OML 为基础，把 OKBC-Lite (Open Knowledge Based Connectivity 协议的一个子集) 的强大表现能力与 OML 语法结合，在 XML 基础上制定了 XOL。目前还没有应用工具支持 XOL 的 Ontology 开发，但是由于 XOL 基于 XML 语法，所以我们可以使用 XML 工具来编辑 XOL 文件。

2、SHOE (Simple HTML Ontology Extension)

由美国马里兰大学开发的一种 Ontology 语言，它是对 HTML 的简单扩展，可以用来向 HTML 文档或是其他 Web 文档中插入机器可读的语义知识。随后，马里兰大学对 SHOE 进行了一些改进，使它可以用 XML 语法来表示。SHOE 使得 Web 代理能够从 Web 页面或文档中收集有意义的信息，并能够提高查询和知识获取能力。这一过程包括三个步骤：首先，用 SHOE 定义一个 Ontology；其次，用定义的 Ontology 信息对 HTML 页面进行注释；最后，用一个 Web 代理搜索带有用 SHOE 标识的页面，获取其中的语义并存放在知识库中，以备日后的语义查询所需。

也包括描述对象动态特征(行为)的方法。对象的属性和方法结合在一起形成一个独立的系统单位,尽可能对外隐藏对象的内部细节,提供统一的接口。为了提高面向对象领域中构件和构件库的重用能力,可以以面向对象领域为研究对象,在其上建立本体。

3. 3 本体描述语言及编辑工具

3. 3. 1 本体描述语言

为了描述并表示本体,近几年来出现了多种本体描述语言。在目前 XML 已成为数据交换格式描述标准的环境下,这些语言都以 XML 为基础来构建。下面将对几种主要的本体语言进行简要介绍^[21]。

1、XOL (XML-based Ontology Exchange Language)

这是由美国的一个生物信息学组织设计的 Ontology 语言,其目的是在针对不同领域设计的软件系统之间交换它们各自的 Ontology 定义。研究人员在仔细研究了生物信息学领域专家的信息表现需求后创建了这一语言。他们以 Ontolingua 和 OML 为基础,把 OKBC-Lite (Open Knowledge Based Connectivity 协议的一个子集)的强大表现能力与 OML 语法结合,在 XML 基础上制定了 XOL。目前还没有应用工具支持 XOL 的 Ontology 开发,但是由于 XOL 基于 XML 语法,所以我们可以使用 XML 工具来编辑 XOL 文件。

2、SHOE (Simple HTML Ontology Extension)

由美国马里兰大学开发的一种 Ontology 语言,它是对 HTML 的简单扩展,可以用来向 HTML 文档或是其他 Web 文档中插入机器可读的语义知识。随后,马里兰大学对 SHOE 进行了一些改进,使它可以用 XML 语法来表示。SHOE 使得 Web 代理能够从 Web 页面或文档中收集有意义的信息,并能够提高查询和知识获取能力。这一过程包括三个步骤:首先,用 SHOE 定义一个 Ontology;其次,用定义的 Ontology 信息对 HTML 页面进行注释;最后,用一个 Web 代理搜索带有用 SHOE 标识的页面,获取其中的语义并存放在知识库中,以备日后的语义查询所需。

也包括描述对象动态特征（行为）的方法。对象的属性和方法结合在一起形成一个独立的系统单位，尽可能对外隐藏对象的内部细节，提供统一的接口。为了提高面向对象领域中构件和构件库的重用能力，可以以面向对象领域为研究对象，在其上建立本体。

3. 3 本体描述语言及编辑工具

3. 3. 1 本体描述语言

为了描述并表示本体，近几年来出现了多种本体描述语言。在目前 XML 已成为数据交换格式描述标准的环境下，这些语言都以 XML 为基础来构建。下面将对几种主要的本体语言进行简要介绍^[21]。

1、XOL (XML-based Ontology Exchange Language)

这是由美国的一个生物信息学组织设计的 Ontology 语言，其目的是在针对不同领域设计的软件系统之间交换它们各自的 Ontology 定义。研究人员在仔细研究了生物信息学领域专家的信息表现需求后创建了这一语言。他们以 Ontolingua 和 OML 为基础，把 OKBC-Lite (Open Knowledge Based Connectivity 协议的一个子集) 的强大表现能力与 OML 语法结合，在 XML 基础上制定了 XOL。目前还没有应用工具支持 XOL 的 Ontology 开发，但是由于 XOL 基于 XML 语法，所以我们可以使用 XML 工具来编辑 XOL 文件。

2、SHOE (Simple HTML Ontology Extension)

由美国马里兰大学开发的一种 Ontology 语言，它是对 HTML 的简单扩展，可以用来向 HTML 文档或是其他 Web 文档中插入机器可读的语义知识。随后，马里兰大学对 SHOE 进行了一些改进，使它可以用 XML 语法来表示。SHOE 使得 Web 代理能够从 Web 页面或文档中收集有意义的信息，并能够提高查询和知识获取能力。这一过程包括三个步骤：首先，用 SHOE 定义一个 Ontology；其次，用定义的 Ontology 信息对 HTML 页面进行注释；最后，用一个 Web 代理搜索带有用 SHOE 标识的页面，获取其中的语义并存放在知识库中，以备日后的语义查询所需。

3、OML (Ontology Markup Language)

OML 是华盛顿大学以 SHOE 为部分基础开发的一种 Ontology 语言。实际上,也可以把它看作是用 XML 表示的 SHOE,所以 OML 和 SHOE 具有很多相似的特征。OML 包括四个不同的层次:OML 核心层描述了这门语言的逻辑,并包含在其他的几个层中;简单 OML 层直接映射到 RDF/RDFS;简化 OML 层包括了概念图的内容;标准 OML 是 OML 中最富有表现能力的部分与 XOL 一样,目前也没有支持 OML 的编辑工具,但是由于它同样以 XML 为基础,所以可用 XML 工具来处理 OML 文档。

4、RDF/RDFS (Resource Description Framework and RDF Schema)

RDF 是由万维网联盟开发的一项描述 Web 资源的规范,其目的是在 XML 的基础上提供一种标准的、可交互的数据语义描述方式。并非局限于 Web 环境,它也可被用来对服务、过程和商业模式进行精确的描述。RDF 数据模型采用了语义网络 (Semantic networks) 的知识表示机制。它由三个部分组成:资源 (resource)、属性 (property) 和声明 (statement)。资源是 RDF 表达式的描述对象,并用 URI 来唯一标识;属性则定义了所描述资源的一些特点或是用来描述资源的关系;声明则在一个特定的范围内为所描述资源的某一属性指定一个值。

RDF 数据模型并没有提供定义属性、资源之间关系的机制,这一角色由 RDFS 来担当。RDFS 提供了一种类似于框架表示法 (frame-based approach) 的机制来定义知识模型。

5、OIL (Ontology Interchange Language)

OIL 是在项目 OntoKnowledge 中开发出来的。这一项目的目的是在 Web 资源间提供语义层次的交互性。它的语法和语义建立在已有规范 (OKBC、XOL、RDF/RDFS) 的基础上,并提供在框架表示法中通用的建模原语 (概念、关系、函数、公理和实例) 来创建 Ontology。不仅如此,它还具有正式语义和一阶谓词逻辑中的推理支持能力。

OIL 建立在 RDF/RDFS 之上,分为以下层次:核心 OIL 层,其中包括 OIL 原语并建立了到 RDF/RDFS 原语的直接映射;标准 OIL 层含有完整的 OIL 模型,具有比 RDF/RDFS 中更多的原语;实例 OIL 层中加入了概念实例和原有模型的角色;密集 OIL (Heavy OIL) 层作为将来 OIL 扩展之用。

OIL 的编辑工具有 OilEd 、 Protege2000 和 WebODE 等, OIL 语法不仅可以用 XML, 还可以用 ASCII 来表现。

6、DAML+OIL (DARPA Agent Markup Language+OIL)

DAML+OIL 由来自美国和欧盟的联合委员会开发。DAML 是美国 DARPA (国防高级研究署) 的一个项目, 其目标是以 XML 为基础创建一种支持语义交互的语言, 它与 OIL 具有相同的目的。

DAML+OIL 在 RDF/RDFS 基础上建立, 实质上就是在后者的基础上加入了一个含有属性和类的词汇库; 从它的名字上可以看出来, 它与 OIL 有着紧密的联系。它取代了最初的规范 (DAML-ONT), 并向其中加入了很多 OIL 语言的内容。OilEd 、 OntoEdit、 Protege2000 以及 WebODE 等工具都可以用来编辑 DAML+OIL Ontology。

7、OWL (Ontology Web Language)

OWL 由万维网联盟的 Web Ontology 工作组设计, 是 DAML+OIL 的修订本。它的语法与 DAML+OIL 的非常相似, 因此可以很容易地被转换为后者。万维网联盟 Web Ontology 工作组最初于 2002 年 3 月发布了 OWL 规范工作草案。2003 年 2 月, 万维网联盟又发布了新的 OWL 规范工作草案。OWL 改进了基于 RDF/RDFS 的 DAML+OIL 的一些不足, 借助于万维网联盟这样的权威组织支持, 有望成为将来的 Web Ontology 语言标准。

上述几种 Ontology 语言的关系如图 3-3 所示。

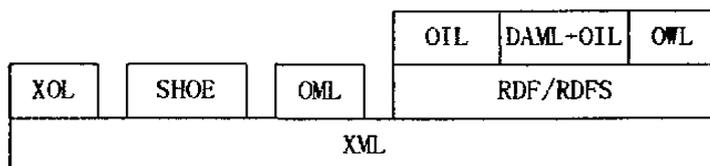


图 3-3 几种 Ontology 语言的关系图

Figure 3-3 the relations of kinds of ontology languages

本论文中应急决策领域本体的描述语言采用了 XML+RDF+OWL 相结合的形式, 而 OWL 语言将来极有可能成为 Ontology 语言的国际标准, 因此我们有必要对 OWL 语言做一简要的介绍, 以获得对 Ontology 语言更直观的认识。

3. 3. 2 本体描述语言 OWL

网络本体语言 OWL 是由国际万维网联盟 W3C 发布的共享本体的标记语言，用来描述万维网文档和应用中的类以及类之间的关系。它的作用是，通过定义类以及类的属性来形式化地描述领域，并通过 OWL 的形式化语义对类进行某种程度上的逻辑推理。

OWL 是针对各方面需求设计而成的。比如它既要保持对 DAML-ONT/OIL/RDFS 的兼容性，又要保证更强大的语义表达能力，还要保证描述逻辑的可判定推理等。诸多的功能需求使得开发人员对之难以做到完美的折衷。于是，开发人员针对各类特征的需求定制了相应风格的变体。主要包括：OWL DL、OWL Lite 和 OWL Full 三种形式。OWL DL 将保证可判定推理、表达能力强作为首要目标，它忽略了对 RDFS 的兼容性，主要针对概念、性质、个体之间关系的描述，以保证强大的语义表达能力。OWL Lite 关注的重点是一个更简洁的本体语言，它是 OWL DL 的一个子集。OWL Lite 降低了 OWL DL 中的公理约束，保证一个迅速高效的推理过程。它和一些已实现的逻辑推理系统（如 FaCT, RACER）非常相似。OWL Full 将保持对 RDF 的向上兼容性作为首要目标，它将 RDF (S) 扩展为一个完备的本体语言，克服了 OWL DL 与 RDFS 在语义上的种种冲突。但是，OWL Full 消除了基数限制中对可传递性质的约束，因此不能保证可判定推理^[23]。总体上说，从 OWL Lite、OWL DL、到 OWL Full 是一个逐步扩展的过程，用户可以根据需求灵活选用。

OWL 对于客观世界的描述主要从概念和属性两个方面进行，与其相应的描述手段是面向对象域 (Object Domain) 的方式和面向数据类型域 (Datatype Domain) 的方式。面向对象域的描述方式采用 RDFS 和 OWL 自身的语法进行，用于描述概念间分类化、层次化的继承关系以及相互间的关联关系；在进行面向数据类型域的描述时，OWL 支持 XML Schema 的所有数据类型进行概念属性的定义与表达。因此，OWL 通过对概念、概念属性及其相互间关系的描述，构成概念的复杂关系网络^[23]。

OWL 中的概念由类 (Class) 来表示，它可以是名字 (如 URI) 或表达式，而且提供大量的构造子来建立表达式，OWL 强大的表达能力正是由它

所支持的概念构造子、性质构造子，以及各种公理 (Axioms) 所决定的。

● OWL Lite 语言构造子

<p>RDFSchema特征:</p> <ul style="list-style-type: none"> <u>Class(Thing, Nothing)</u> <u>rdfs:subClassOf</u> <u>rdf:Property</u> <u>rdfs:subPropertyOf</u> <u>rdfs:domain</u> <u>rdfs:range</u> <u>Individual</u> 	<p>(In)Equality:</p> <ul style="list-style-type: none"> <u>equivalentClass</u> <u>equivalentProperty</u> <u>sameAs</u> <u>differentFrom</u> <u>AllDifferent</u> <u>distinctMembers</u> 	<p>性质特征:</p> <ul style="list-style-type: none"> <u>ObjectProperty</u> <u>DatatypeProperty</u> <u>inverseOf</u> <u>TransitiveProperty</u> <u>SymmetricProperty</u> <u>FunctionalProperty</u> <u>InverseFunctionalProperty</u>
<p>性质限制:</p> <ul style="list-style-type: none"> <u>Restriction</u> <u>onProperty</u> <u>allValuesFrom</u> <u>someValuesFrom</u> 	<p>基数限制:</p> <ul style="list-style-type: none"> <u>minCardinality(only 0 or 1)</u> <u>maxCardinality(only 0 or 1)</u> <u>cardinality(only 0 or 1)</u> 	<p>头信息:</p> <ul style="list-style-type: none"> <u>Ontology</u> <u>imports</u>
<p>类相似:</p> <ul style="list-style-type: none"> <u>intersectionOf</u> 	<p>版本信息:</p> <ul style="list-style-type: none"> <u>versionInfo</u> <u>priorVersion</u> <u>backwardCompatibleWith</u> <u>incompatibleWith</u> <u>DeprecatedClass</u> <u>DeprecatedProperty</u> 	<p>评注性质:</p> <ul style="list-style-type: none"> <u>rdfs:label</u> <u>rdfs:comment</u> <u>rdfs:seeAlso</u> <u>rdfs:isDefinedBy</u> <u>AnnotationProperty</u> <u>OntologyProperty</u>

● OWL DL 和 OWL Full 中用来扩展 OWL Lite 的构造子

<p>类公理:</p> <ul style="list-style-type: none"> • <u>one of dataRange</u> • <u>disjointWith</u> • <u>equivalentClass</u> (应用于类表达式) • <u>rdfs:subClassOf</u> (应用于类表达式) 	<p>类表达式的布尔组合:</p> <ul style="list-style-type: none"> • <u>unionOf</u> • <u>complementOf</u> • <u>intersectionOf</u>
<p>任意基数:</p> <ul style="list-style-type: none"> • <u>minCardinality</u> • <u>maxCardinality</u> • <u>cardinality</u> 	<p>填充信息:</p> <ul style="list-style-type: none"> • <u>hasValue</u>

具体地说，OWL 通过 subClassOf、subPropertyOf 形成概念及其关系的分类化、层次化结构，通过 sameClassAs、samePropertyAs、inverseOf、equivalentTo 等形成概念间的同义、反义等语义关系，通过 intersectionOf、unionOf、complementOf、oneOf 等构建概念间的逻辑组合关系，通过 domain、range、toClass、hasValue、cardinality、min Cardinality、max Cardinality 等对关系约束进行描述，通过 disjointWith、uniqueProperty、unambiguousProperty、transitiveProperty 等实现对概念及其关系的公理定义。

3. 3. 3 本体编辑工具

在本体表示中,有许多已经研发出来的本体编辑器,其中最具代表性的就是 Racer 和 OilEd。其中 Racer 是基于 OWL 的,而 OilEd 则是基于 DAML+OIL 的,其现有版本也提供了对 OWL 的支持。在本节内容中,我们将分别简要地介绍一下这两个本体编辑器。

1. Racer

Racer^[21]是一种基于 OWL 的推理器,具有高度优化的推理系统。它可以被看作是语义网的一种核心推理 Agent。Racer 现在支持多种有关在 OWL 语言中定义的 Ontology 的推理服务,这些服务可以通过基于网络的 API 供给其它 Agent 使用。Racer 现在正被多种客户使用,这些客户包括 Ontology 编辑, Ontology 发展和可视化工具,以及基于网络的用来探索和分析 OWL Ontology 的原型等。

Racer 通过类似 HTTP 和 TCP 这样的协议访问(Racer 这种程序后来也被称为 Racer 服务器),无论从本地文件还是远程网络服务器,Racer 都可以读取 OWL 知识库。Racer 同时也是一个 HTTP 客户机。反过来,其它需要推理服务的客户程序也可以通过基于 TCP 的协议来访问 Racer。如今在大多数 Racer 的应用中,RICE 是最具有代表性的。与 OilEd 相比,RICE 是一种使用 TCP 协议来得到广泛的询问回答功能的,而 OilEd 则可以被看作是一种使用 DIG 协议和 Racer 服务器通讯的客户。

其中,DIG 协议是一种基于 XML 和 HTTP 的标准,它是用来将客户程序连接到表达逻辑推理引擎的,DIG 允许知识库的分配,并使客户能够提出标准表达逻辑询问,作为一种标准和一种不常使用的协议命令,与系统相关的申明和询问的所有不同形式不可能被 DIG 所完全包括,更不用说包含指令的长期询问,因此,为了接受指令和询问,Racer 也提供了一种基于 TCP 的接口,对于交互式的使用 Racer,Racer 支持的语言并不是基于 XML 或者 RDF 的,这样做的好处就是用户们可以同时输入询问,这些询问可以被直接发送到 Racer 服务器,但是,Racer 的 TCP 接口也可以被 JAVA 或者 C++程序方便的访问,这两种语言都有相应的 API。

在表达逻辑的术语中,Racer 包含有具有一个 T-BOX 和一个 A-BOX 的

知识库：一个具有特定名字的对象；另外，OWL 也允许存在概念中的个体。例如，表达所有人都来源于一个叫亚当的人这样一个事实要求指定一个存在于理念(T-BOX)中的个体，而一个在概念中提到的个体的表达性则存在于 A-BOX 中。另外，在检验合适的 SHIQ 时，Ontology 可以从一个 OWL 文件中产生，Racer 可以直接读取 OWL 文件并将它们表示为逻辑知识库。下面简单的介绍一下 T-BOX 支持的请求。

(1) 概念的一致性：是否一个概念描述的对象集合是空的；

(2) 概念包含：是否两个概念描述的对象集合之间存在子集关系；

(3) 找到所有在一个 T-BOX 中提到的不一致的概念。注意不一致的概念可能是建模错误的结果；

(4) 确定一个概念的父亲和孩子：一个概念的父亲是一个 T-BOX 中包含这个概念的最明确的概念，一个概念的孩子是 T-BOX 中提到的这个概念包含的最概括的概念名。

A-BOX 支持的请求如下：

(1) 检查一个 A-BOX 的一致性：是否给定的限制太强了，它们之间是否相互矛盾；

(2) 实例检测：是否一个个体代表的对象是一个特定请求概念描述的对象集合中的一员，这个个体被称为这个请求概念的一个实例。

(3) 实例取回：从一个 A-BOX 中找到所有个体，以便这些个体所代表的对象可以被证明是被一个请求概念描述的一个对象集合中的一员。

(4) 一个个体的直接类型的计算：从一个 T-BOX 中找到最详细的概念名称使得一个给定的个体是这个概念的一个实例。

(5) 计算与一个个体有关的任务的填补者。

除此之外，Racer 所具有的系统特色还有如下几方面：

(1) 优化：多种为 Racer 发展、实现和研究的优化技术，它们与 T-BOX、A-BOX 和具体值有关，Racer 设计的目的之一是自动选择适应于当前输入的先进的优化技术。

(2) 持续性：与数据库系统相似，对于请求回答，复杂的数据结构被 Racer 计算和内部使用。T-BOX 和 A-BOX 的请求回答处理的内部结构被保存到磁盘，它们可以被快速访问，如果 Racer 服务器重新启动，还可以被

重用。

(3) 多用户支持：线程安全，锁定，装载平衡：在一个分布式系统环境下，同一时间可能有多个 Agent 连接到一个服务器，请求必须被同步，所以像数据库一样，如线程安全，锁定，装载平衡等问题必须被解决，例如，如果多个 Racer 服务器被启动，请求可以被自动指定到空闲 Racer 服务器上，Racer 代理解决了这些问题，Racer 代理是 Racer 系统的一部分。

2. OilEd

OilEd^[25]是一个使用比较简单的本体编辑器，具有可视化的图形用户界面。它是基于 DAML+OIL 设计的一个本体编辑器。现在的 OilEd 版本也提供对 OWL 的支持。OilEd 允许用户创建和编辑任何本体，这些本体描述将自动通过 OilEd 转化为用 DAML+OIL 或 OWL 编写的本体文件。开发 OilEd 的最初目的是为了提供一个比较简单的免费本体编辑器供使用者构造简单的本体。

OilEd 并不是一个完全的、功能强大的本体开发环境。它并不支持大规模的本体的开发、移植和集成，以及版本的升级等其它的和构建本体有关的功能。但是，本体开发人员可以使用 OilEd 创建简单的基本的用 DAML+OIL 或 OWL 编写的本体模型。

与 Racer 不同，在 OilEd 中，是通过集成 FaCT 推理机实现推理过程的。FaCT 推理机是由英国曼彻斯特大学研制的。FaCT 推理机是基于 SHIQ 逻辑的。在对本体描述进行检验时，通过 OilEd 提供的连接功能，可以将其连接到 OilEd 所集成的推理机 FaCT 之上，由 FaCT 推理机可以完成对本体描述文档的一致性检测和包含等逻辑关系的检验等逻辑推理功能。

一般说来，OilEd 具有如下的特点：

(1) 良好的用户界面，可供对于本体描述语法不太熟悉的本体设计者比较容易的构建本体，减轻本体设计者的负担。

(2) 在构建本体的过程中，可以自动生成由统一的 DAML+OIL 或 OWL 等多种本体建模语言描述的本体文件，这大大简化了对于本体描述文件的统一处理的复杂度。

(3) 具有一定的推理功能，能够对所要描述的本体进行一致性检测

和包含等逻辑关系的检验，极大的提高了构建本体的准确性，降低了构建本体的出错率。

3. 4 本体的功能

一般来说，本体的功能主要是为了实现“重用”。另外还有一个重要的功能就是信息集成。基于上述两个主要的功能，我们把本体的使用功能粗略的划分为以下三类：

1. 信息交换：

本体的核心概念是知识共享。通过减少概念和术语上的歧义，本体描述为某一组织或是工作小组提供了一个统一框架或是规范模型，使得来自不同背景，持不同观点和目的的人员之间的理解和交流成为可能，并保持语义上的一致性。

2. 互操作：

绝大多数的应用程序使用本体实现不同系统之间的互操作，即不同系统或是工具之间的数据传输。这种应用还可以细分为两类：轻量级(Lightweight)和重量级(Heavyweight)。基于本体的轻量级应用是一种全面向语法形式化和转换的本体使用。它只能保证人们使用相同的词汇，但不能保证他们对相同的词汇有相同的解释或是理解。因此这种应用中本体不包含有语义信息，或者说是一种较低层次的本体的应用，比如 XML 语言中的 DTD。而基于本体的重量级应用是一种对知识表示语言中的建构和约束作普通的语义解释的使用方式，目的是支持本体的不同使用者之间的进行语义层面的信息共享和互操作。本体共享的数据交换的关键是本体的知识表示语言必须保证对本体一致无歧义的解释。而这类本体应用中的本体还支持另外一些功能服务，比如支持推理的查询，更新和一致性检查功能。这些是由本体知识表示语言所采用的推理机制所决定的。

3. 系统工程：

本体还可以应用到软件系统的设计和实现上。在软件开发的规格说明中，本体论通过对需要解决的问题和任务的理解描述，可以帮助我们在需求分析、信息获取中提高明确性，减小分析代价。同时，本体可以作为需

和包含等逻辑关系的检验，极大的提高了构建本体的准确性，降低了构建本体的出错率。

3. 4 本体的功能

一般来说，本体的功能主要是为了实现“重用”。另外还有一个重要的功能就是信息集成。基于上述两个主要的功能，我们把本体的使用功能粗略的划分为以下三类：

1. 信息交换：

本体的核心概念是知识共享。通过减少概念和术语上的歧义，本体描述为某一组织或是工作小组提供了一个统一框架或是规范模型，使得来自不同背景，持不同观点和目的的人员之间的理解和交流成为可能，并保持语义上的一致性。

2. 互操作：

绝大多数的应用程序使用本体实现不同系统之间的互操作，即不同系统或是工具之间的数据传输。这种应用还可以细分为两类：轻量级(Lightweight)和重量级(Heavyweight)。基于本体的轻量级应用是一种全面面向语法形式化和转换的本体使用。它只能保证人们使用相同的词汇，但不能保证他们对相同的词汇有相同的解释或是理解。因此这种应用中本体不包含有语义信息，或者说是一种较低层次的本体的应用，比如 XML 语言中的 DTD。而基于本体的重量级应用是一种对知识表示语言中的建构和约束作普通的语义解释的使用方式，目的是支持本体的不同使用者之间的进行语义层面的信息共享和互操作。本体共享的数据交换的关键是本体的知识表示语言必须保证对本体一致无歧义的解释。而这类本体应用中的本体还支持另外一些功能服务，比如支持推理的查询，更新和一致性检查功能。这些是由本体知识表示语言所采用的推理机制所决定的。

3. 系统工程：

本体还可以应用到软件系统的设计和实现上。在软件开发的规格说明中，本体论通过对需要解决的问题和任务的理解描述，可以帮助我们在需求分析、信息获取中提高明确性，减小分析代价。同时，本体可以作为需

求分析基础上软件设计时的基础,以自动或是半自动的方式检查它们之间的一致性,从而提高软件系统的可靠性。本体还可以通过对系统内部各个功能模块和它们之间的联系详细描述达到软件的重用性。

3. 5 本体 (Ontology) 工程

Ontology 作为通信、互操作和系统工程的基础,必须经过精心的设计。实际上,Ontology 的创建过程是一个非常费时费力的过程,需要一套完善的工程化的系统方法来支持,特定的专用 Ontology 还需要专家进行参与。通用的大规模 Ontology 很少,大多 Ontology 只是针对某个具体领域或应用而创建的。在实际应用中,不同 Ontology 之间常常需要进行映射、扩充与合并处理,以及根据特定的需要由一个大的 Ontology 提取满足要求的小 Ontology 等操作,此外,当现实的知识体系发生变化时,先前创建的 Ontology 也必须做出相应的演化以保持 Ontology 与现实的一致性,这都是 Ontology 工程所需研究的问题。

3. 5. 1 构造本体 (Ontology) 的原则

如何构造 Ontology 是 Ontology 工程中的最基本问题。Gruber 提出了指导 Ontology 构造的五个原则^[26]:

1、清晰 (Clarity): Ontology 必须有效地说明所定义术语的意思,定义应该是客观的,与背景独立的,当定义可以用逻辑公理表达时,它应该是形式化的,定义应该尽可能地完整,所有定义应该用自然语言加以说明。

2、一致性 (Coherence): Ontology 应该是一致的,也就是说,它应该支持与其定义相一致的推理,它所定义的公理以及用自然语言进行说明的文档都应该具有一致性。

3、可扩展性 (Extendibility): Ontology 应该为可预料到的任务提供概念基础,它应该可以支持在已有的概念基础上定义新的术语,以满足特殊的需求,而无需修改已有的概念定义。

求分析基础上软件设计时的基础,以自动或是半自动的方式检查它们之间的一致性,从而提高软件系统的可靠性。本体还可以通过对系统内部各个功能模块和它们之间的联系详细描述达到软件的重用性。

3. 5 本体 (Ontology) 工程

Ontology 作为通信、互操作和系统工程的基础,必须经过精心的设计。实际上,Ontology 的创建过程是一个非常费时费力的过程,需要一套完善的工程化的系统方法来支持,特定的专用 Ontology 还需要专家进行参与。通用的大规模 Ontology 很少,大多 Ontology 只是针对某个具体领域或应用而创建的。在实际应用中,不同 Ontology 之间常常需要进行映射、扩充与合并处理,以及根据特定的需要由一个大的 Ontology 提取满足要求的小 Ontology 等操作,此外,当现实的知识体系发生变化时,先前创建的 Ontology 也必须做出相应的演化以保持 Ontology 与现实的一致性,这都是 Ontology 工程所需研究的问题。

3. 5. 1 构造本体 (Ontology) 的原则

如何构造 Ontology 是 Ontology 工程中的最基本问题。Gruber 提出了指导 Ontology 构造的五个原则^[26]:

1、清晰 (Clarity): Ontology 必须有效地说明所定义术语的意思,定义应该是客观的,与背景独立的,当定义可以用逻辑公理表达时,它应该是形式化的,定义应该尽可能地完整,所有定义应该用自然语言加以说明。

2、一致性 (Coherence): Ontology 应该是一致的,也就是说,它应该支持与其定义相一致的推理,它所定义的公理以及用自然语言进行说明的文档都应该具有一致性。

3、可扩展性 (Extendibility): Ontology 应该为可预料到的任务提供概念基础,它应该可以支持在已有的概念基础上定义新的术语,以满足特殊的需求,而无需修改已有的概念定义。

求分析基础上软件设计时的基础,以自动或是半自动的方式检查它们之间的一致性,从而提高软件系统的可靠性。本体还可以通过对系统内部各个功能模块和它们之间的联系详细描述达到软件的重用性。

3. 5 本体 (Ontology) 工程

Ontology 作为通信、互操作和系统工程的基础,必须经过精心的设计。实际上,Ontology 的创建过程是一个非常费时费力的过程,需要一套完善的工程化的系统方法来支持,特定的专用 Ontology 还需要专家进行参与。通用的大规模 Ontology 很少,大多 Ontology 只是针对某个具体领域或应用而创建的。在实际应用中,不同 Ontology 之间常常需要进行映射、扩充与合并处理,以及根据特定的需要由一个大的 Ontology 提取满足要求的小 Ontology 等操作,此外,当现实的知识体系发生变化时,先前创建的 Ontology 也必须做出相应的演化以保持 Ontology 与现实的一致性,这都是 Ontology 工程所需研究的问题。

3. 5. 1 构造本体 (Ontology) 的原则

如何构造 Ontology 是 Ontology 工程中的最基本问题。Gruber 提出了指导 Ontology 构造的五个原则^[26]:

1、清晰 (Clarity): Ontology 必须有效地说明所定义术语的意思,定义应该是客观的,与背景独立的,当定义可以用逻辑公理表达时,它应该是形式化的,定义应该尽可能地完整,所有定义应该用自然语言加以说明。

2、一致性 (Coherence): Ontology 应该是一致的,也就是说,它应该支持与其定义相一致的推理,它所定义的公理以及用自然语言进行说明的文档都应该具有一致性。

3、可扩展性 (Extendibility): Ontology 应该为可预料到的任务提供概念基础,它应该可以支持在已有的概念基础上定义新的术语,以满足特殊的需求,而无需修改已有的概念定义。

4、 编码偏好程度最小 (Minimal encoding bias) : 概念的描述不应该依赖于某一种特殊的符号层表示方法, 因为实际系统可能采用不同的知识表示方法。

5、 约定最小 (Minimal ontological commitment) : Ontology 约定应该最小, 只要能够满足特定的知识共享需求即可, 这可以通过定义约束最弱的公理以及只定义通信所需的词汇来保证。

3. 5. 2 本体 (Ontology) 建模的主要方法

目前, 建立本体大多采用手工方式, 远远没有成为一种工程性的活动。在建立各自的本体时, 都有自己的原则、标准和定义, 缺乏公认的建模方法, 影响了本体的重用、共享和互操作。但是, 研究人员在不断地探索本体的开发方法。目前知识工程界比较成型的建模方法主要有^[26]:

1、 Mike Uschold & King 的“骨架”法

建立在企业本体基础之上, 是相关商业企业间术语和定义的集合, 该方法只提供开发本体的指导方针。目前企业本体在爱丁堡大学人工智能研究所及它的合作伙伴—IBM, Lloyd’ s Register , Logica UK Limited and Unilever。“骨架”法流程见图 3-4。

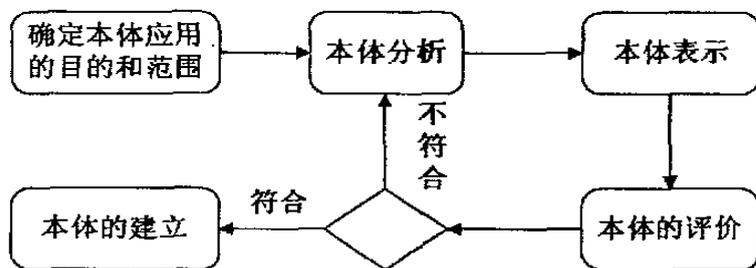


图 3-4 “骨架”法流程图

Figure 3-4 the flow chart of “framework” method

(1) 确定本体应用的目的和范围: 根据所研究的领域或任务, 建立相应的领域本体或过程本体, 领域越大, 所建本体越大, 因此需限制研究的范围。

(2) 本体分析: 定义本体所有术语的意义及其之间的关系, 该步骤需领域专家的参与, 对该领域越了解, 所建本体就越完善。

(3) 本体表示: 一般用语义模型表示本体。

(4) 本体评价:建立本体的评价标准是清晰性、一致性、完善性、可扩展性。清晰性就是本体中的术语应被无歧义的定义;一致性指的是术语之间关系逻辑上应一致;完整性,本体中的概念及关系应是完整的,应包括该领域内所有概念,但很难达到,需不断完善;可扩展性,本体应用能够扩展,在该领域不断发展时能加入新的概念。

(5) 本体的建立:对所有本体按以上标准进行检验,符合要求的以文件的形式存放,否则转(2)。

2、Gruninger & Fox 的“评价法”(又称 TOVE)

这个方法用于构造多伦多虚拟企业本体工程,由多伦多大学企业集成实验室研制,使用一阶逻辑进行集成。TOVE 本体包括企业设计本体、工程本体、计划本体和服务本体。TOVE 流程见图 3-5。

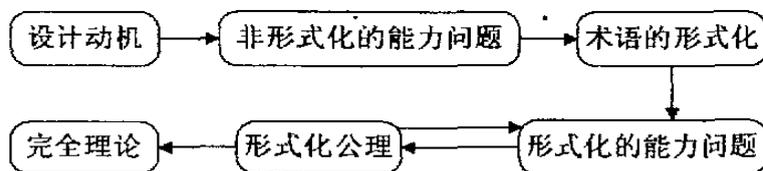


图 3-5 TOVE 流程图

Figure 3-5 the flow chart of TOVE

(1) 定义直接可能的应用和所有解决方案,提供潜在的非形式化的对象和关系的语义表示。

(2) 能力问题作为约束条件,包括能解决什么问题和如何解决,这里的问题用术语表示,答案用公理和形式化定义回答。由于是在没有形式化 Ontology 之前进行的,所以叫非形式化的能力问题。

(3) 术语的规范化:从非形式化能力问题中提取非形式化的术语,然后用 Ontology 形式化语言进行定义。

(4) 形式化的能力问题:一旦能力问题脱离了非形式化,Ontology 术语已定义,则能力问题自然形式化了。

(5) 形式化公理:术语定义所遵循的公理用一阶谓词逻辑表示,包括定义的语义或解释。

(6) 说明问题的解决方案必须是完全的。

3、Berneras et al. 方法(KACTUS 工程)

这种方法开发本体由应用开发控制。所以每一个应用都有相应的表示

该应用所需的 Ontology。这些本体既能重用其它的 Ontology，也能被后继应用集成，应用于电子网络的开发。具体的开发过程如下：

(1) 应用的说明。提供应用的上下文和应用模型所需的组件。

(2) 相关本体论范畴的初步设计。搜索已存在的 Ontologies，进行提炼、扩充。

(3) Ontology 的构造。最小关联原则用来确保模型既相互依赖，又尽可能一致，以至得到最大同构。

KACTUS 是欧洲的 ESPRIT 项目，支持 EXPRESS 和 Ontolingua，目的是关于技术系统生命周期过程中的知识重用，用 CML 语言描述，CML 是 KADS 工程语言，非形式化的，不能被程序执行。

4、Methontology 方法

这种方法由马德里大学工艺分校开发人工智能图书馆使用。它分为三个阶段：

(1) 管理阶段。这一阶段的系统规划包括任务的进展情况、需要的资源、如何保证质量等问题。

(2) 开发阶段。规范说明→概念化→形式化→执行→维护。

(3) 维护阶段。包括知识获取、系统集成、评价、文档说明、配置管理。

用这种方法开发的 Ontology 有 (Onto)²Agent：基于 Ontology 的 WWW 代理，关于 Ontology 使用参考 Ontology 作为知识源进行一定约束条件的重新知识获取；Chemical OntoAgent：基于 Ontology 的 WWW 化学教育代理，允许学生学习化学，自测该领域的技巧；Ontogeneration：使用域本体(化学家)和语言本体来产生西班牙文本描述，来作为对学生关于化学领域问题的查询的回答。

5、SENSUS 的描述

这个 Ontology 用于自然语言程序，由 ISI(信息科学研究所)自然语言组企图为机器翻译提供广泛的概念结构，共有 50 000 多个概念(电子类知识)，为了能在 Sensus 基础上构造特定领域的 Ontology，必须把不相关的术语从 Sensus 中剪除掉。具体过程如下：

(1) 定义“叶子”术语(和 Sensus 无关)；

(2) 把叶子术语手工地和 Sensus 术语相连;

(3) 找出叶子节点到 Sensus 根的“路”;

(4) 增加和域相关并且没有出现的概念;

(5) 用启发式思维找出全部的特定的域的术语:对于某些有两条以上路经过的节点必是一棵子树的父节点,那么这棵子树上的所有节点都和该域相关,是要增加的术语。对于高层节点通常有多条路经过,则很难判断。现已存在使用 Sensus 构造的用于军事领域的 Ontology,包括武器、原油、飞机等。该方法使用的术语如下:

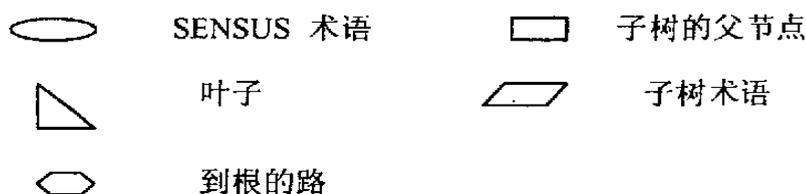


图 3-6 Sensus 术语图

Figure 3-6 the terms of Sensus

上述方法都是从各自的工作实践出发,总结出的建立本体的轮廓,还都不够成熟。表现在:一,缺少本体建模标准、指导原则的可操作性的方法;二,对特定领域具有良好应用的建模方法,缺少集成研究;三,开发者大多直接使用自己熟悉的形式化语言对本体进行编码,影响了本体的共享和重用。相对而言, Mcthontology 方法还是比较成熟的。不管何种方法,都应本着两个基本原则:提供共享和翻译工具。

一般地,把建立 Ontology 模型的过程分为两个阶段:第一阶段是用自然语言和图表来描述领域模型,形成 Ontology 原型;第二阶段是使用知识表示语言把 Ontology 模型进行编码,形成便于人们交流、无歧义的可被软件解释的 Ontology。

本文在对上述本体建模方法分析研究的基础上,提出了基于本体论的应急系统知识库建模的方法。该方法有以下两个优点:首先,在本体建模的第一阶段采用了 DKOED(Domain Knowledge Ontology for Emergency Decision) 图示化建模方法,它建立在对面向对象逻辑建模方法 UML 扩展的基础上,采用 UML 类图构建的图示化领域本体具有图示化建模直观、便捷的特点,不仅为本体建模提供了标准,而且有利于开发人员和分析人员

的沟通，大大提高了该方法的可操作性。其次，在本体建模的第二阶段，采用了 OWL 本体描述语言对建立的本体模型进行编码。OWL 是由万维网联盟的 Web Ontology 工作组设计，将来极有可能成为 Ontology 语言的国际标准，因此用 OWL 描述的本体文档有利于本体模型的共享和重用。

第四章 应急决策支持系统设计及领域本体建模

目前,国内外对本体知识库的研究都集中在常识知识库的建立上,本文所作的应急决策支持系统本体知识库是对建立一个领域本体库的尝试。所谓领域本体(不引起歧义时文中亦简称本体)指:在一个特定的领域中可重用,提供该特定领域的概念定义和概念之间的关系,提供该领域中发生的活动以及该领域的主要理论和基本原理等的一种本体,一个领域内的概念通常具有内涵定义精确,外延易于确定的特点^[2]。

4. 1 应急决策支持系统设计

4. 1. 1 问题描述

应急决策是国家有关部门在突发事件发生前及发生后采取紧急措施,为满足突发事件需求,对人力、物力、财力及其他各种资源进行统一组织、指挥、协调、控制的一系列活动。目的是将突发事件扼制在萌芽阶段,将事件所造成的灾害降到最低。衡量一个应急决策成功与否的标准是在保证国家安定团结的基础上,将事件对人民群众造成的损失降低的程度。

近年来频频爆发的战争、地震、火灾、流行性传染病、煤炭安全生产以及稳定输送电能中的重大事故等突发事件对人民群众生命财产和社会安全造成极大危害。尤其自“911”事件、2003年爆发伊拉克战争、SARS蔓延和2004年初禽流感之后,国家与人民生命安全问题凸现出来,我国格外重视危机管理、突发事件的预警与应急决策系统的建设,应运而生了各类应急决策支持系统^[27-29]。这些系统集成广义模型于模型库中,定量分析辅助决策,并利用不确定性推理、机器学习、神经网络、数据挖掘和 Agents 等技术,以专家系统为定性分析辅助决策。

在当今世界人们面临海量数据、信息的情况下,分布在不同地域、不同领域的决策者要能以更低的成本、更加迅捷的速度、更加准确的数据作出大量及时、精确的决策。故决策系统知识的准确性与系统反应的快速时

第四章 应急决策支持系统设计及领域本体建模

目前,国内外对本体知识库的研究都集中在常识知识库的建立上,本文所作的应急决策支持系统本体知识库是对建立一个领域本体库的尝试。所谓领域本体(不引起歧义时文中亦简称本体)指:在一个特定的领域中可重用,提供该特定领域的概念定义和概念之间的关系,提供该领域中发生的活动以及该领域的主要理论和基本原理等的一种本体,一个领域内的概念通常具有内涵定义精确,外延易于确定的特点^[2]。

4. 1 应急决策支持系统设计

4. 1. 1 问题描述

应急决策是国家有关部门在突发事件发生前及发生后采取紧急措施,为满足突发事件需求,对人力、物力、财力及其他各种资源进行统一组织、指挥、协调、控制的一系列活动。目的是将突发事件扼制在萌芽阶段,将事件所造成的灾害降到最低。衡量一个应急决策成功与否的标准是在保证国家安定团结的基础上,将事件对人民群众造成的损失降低的程度。

近年来频频爆发的战争、地震、火灾、流行性传染病、煤炭安全生产以及稳定输送电能中的重大事故等突发事件对人民群众生命财产和社会安全造成极大危害。尤其自“911”事件、2003年爆发伊拉克战争、SARS蔓延和2004年初禽流感之后,国家与人民生命安全问题凸现出来,我国格外重视危机管理、突发事件的预警与应急决策系统的建设,应运而生了各类应急决策支持系统^[27-29]。这些系统集成广义模型于模型库中,定量分析辅助决策,并利用不确定性推理、机器学习、神经网络、数据挖掘和 Agents 等技术,以专家系统为定性分析辅助决策。

在当今世界人们面临海量数据、信息的情况下,分布在不同地域、不同领域的决策者要能以更低的成本、更加迅捷的速度、更加准确的数据作出大量及时、精确的决策。故决策系统知识的准确性与系统反应的快速时

效性是衡量分布式智能决策支持系统,特别是应急决策系统的两个重要指标。

解决知识表示准确性,提高搜索查准率应从根本上对非结构化或半非结构化信息使用结构化知识表示,特别是在决策信息的利用上,对具体实体的物理活动抽象为信息结构模型描述,并在数学抽取的模型感知层上进行识别、理解、决策处理。而本体就是知识结构性的基本描述。本体提供领域的公共词汇表,以不同的形式化级别定义术语的涵义和术语间的关系。故以分类学方式组织、并包含典型建模原语的本体能够提供对某领域公共而一致的理解,符合应急系统知识表示的要求,能实现异构系统、跨研究群体的知识复用与共享;也有助于在构造新的知识系统时,从已有本体库获取知识,提高开发速度和准确性。

本课题以国防动员项目为背景着重考虑分布在不同地域、不同领域的决策者在面临海量数据、信息的情况下,为满足对信息查准、查全的需求,在决策支持系统中采用本体(Ontology)表示领域知识,实现知识共享重用。由于基于本体的搜索方法是基于知识的,语义上匹配的,在查准率和查全率上有更好的保证,从而实现以更低的成本、更准确的数据和更迅捷的速度做出大量精确、及时的决策。

4. 1. 2 国防动员应急决策支持系统总体框架设计

国防动员是一项系统工程,涉及到国民经济的方方面面,它由许多具体部门和行业动员组成,其中应急决策支持系统是国防动员系统的重要组成部分。突发事件妥善处理需要分布在不同地域、不同领域的相关部门的密切配合,因此,这也是一种分布式应急决策问题。

该系统以数据仓库为基础,以决策支持(DSS)为研究手段,以地理信息系统为表现方法,以事件救援为主线,以协调优化各类资源加强快速反应能力为目标,以城市为单位,建立分布式网络知识库,协调政府各职能部门有效地进行数据采集、管理、分析和共享,通过对各类突发事件异常情况的研究,利用案例匹配和专家推理机制两种思路进行科学的归纳、总结、决策。从技术上保证对突发事件的快速反应和正确应对。

效性是衡量分布式智能决策支持系统,特别是应急决策系统的两个重要指标。

解决知识表示准确性,提高搜索查准率应从根本上对非结构化或半非结构化信息使用结构化知识表示,特别是在决策信息的利用上,对具体实体的物理活动抽象为信息结构模型描述,并在数学抽取的模型感知层上进行识别、理解、决策处理。而本体就是知识结构性的基本描述。本体提供领域的公共词汇表,以不同的形式化级别定义术语的涵义和术语间的关系。故以分类学方式组织、并包含典型建模原语的本体能够提供对某领域公共而一致的理解,符合应急系统知识表示的要求,能实现异构系统、跨研究群体的知识复用与共享;也有助于在构造新的知识系统时,从已有本体库获取知识,提高开发速度和准确性。

本课题以国防动员项目为背景着重考虑分布在不同地域、不同领域的决策者在面临海量数据、信息的情况下,为满足对信息查准、查全的需求,在决策支持系统中采用本体(Ontology)表示领域知识,实现知识共享重用。由于基于本体的搜索方法是基于知识的,语义上匹配的,在查准率和查全率上有更好的保证,从而实现以更低的成本、更准确的数据和更迅捷的速度做出大量精确、及时的决策。

4. 1. 2 国防动员应急决策支持系统总体框架设计

国防动员是一项系统工程,涉及到国民经济的方方面面,它由许多具体部门和行业动员组成,其中应急决策支持系统是国防动员系统的重要组成部分。突发事件妥善处理需要分布在不同地域、不同领域的相关部门的密切配合,因此,这也是一种分布式应急决策问题。

该系统以数据仓库为基础,以决策支持(DSS)为研究手段,以地理信息系统为表现方法,以事件救援为主线,以协调优化各类资源加强快速反应能力为目标,以城市为单位,建立分布式网络知识库,协调政府各职能部门有效地进行数据采集、管理、分析和共享,通过对各类突发事件异常情况的研究,利用案例匹配和专家推理机制两种思路进行科学的归纳、总结、决策。从技术上保证对突发事件的快速反应和正确应对。

应急决策支持系统以快速正确应对各种突发事件为核心,因此数据尤其是空间数据是本系统的基础,将各类空间数据及其属性数据与其执行紧急任务时所需的各资源的信息相链接,通过各政府职能部门的通力协作构建一个完整的数据协作体,是本系统实施的基础。

在此基础上,利用组件技术,采用完全面向对象的设计,实现图数一体化;形成系统数据的整体流转,实现数据高共享,保证数据一致性;并从数据设计做起,实现各领域各级决策者与数据及各子系统间的无缝连接;在统一数据及组件框架的基础上,构筑各类突发事件应急决策支持系统。系统采用分布式三层系统构架,具有高度的适应性和可扩展性,即可以在某一专题领域内建立专业的应急决策支持系统,也可以在统一协作数据的基础上构成一个相对完善的综合应急决策支持系统。其系统框图如下所示:

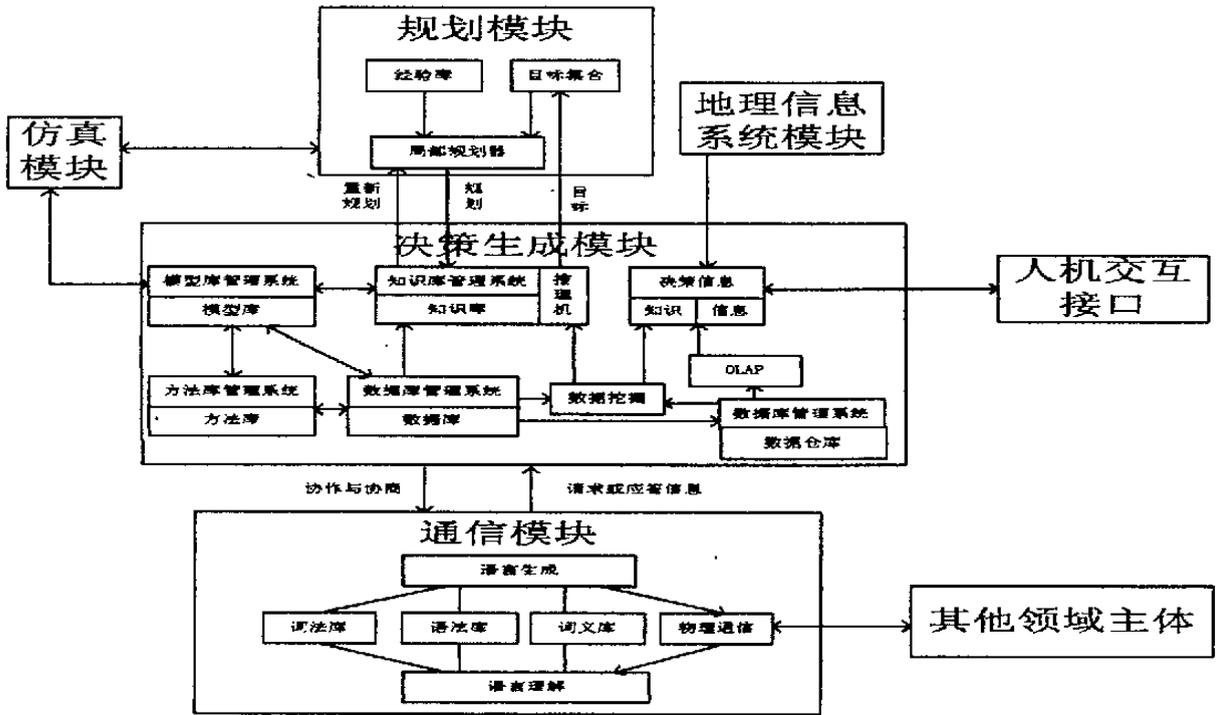


图 4-1 国防动员应急决策支持系统总体框架

Figure 4-1 the overall structure of emergency decision support system

国防动员应急决策支持系统总体框架由以下几大模块组成,分别是决策生成模块、仿真模块、规划模块、地理信息系统(GIS)模块、通信模

块及人机交互接口。其中决策生成模块是系统的核心组成部分，它由知识库、数据库、模型库和方法库构成，并采用了数据挖掘和数据仓库技术，在知识库中集成了推理机。

知识库管理部分，集中管理决策问题领域的知识（规则和事实），包括知识的获取、表达、管理等功能。推理机用来识别并解答用户提出的问题，分为确定性推理和不确定性推理两大类。

数据库用来存储大量与决策问题领域有关的数据，一般组织成易于进行大量数据操作的形式，典型的数据组织模型有网络模型、层次模型、关系模型等形式。数据库由数据库管理系统来管理和维护。数据库管理系统必须具有数据库建立、删除、修改、维护、数据存储、检索、排序、索引、统计等功能。

模型库系统由模型库和模型库管理系统所组成。模型是对客观事物的一种抽象描述，人们通过对模型的认识来增加对复杂问题的理解和处理。用模型来辅助决策已经是人们的共识。模型库是将众多的模型按一定的结构形式组织起来，通过模型库管理系统对各个模型进行有效的管理和使用。模型库像数据库一样，是一个共享资源。模型库中的模型可以重复使用，即可以被不同系统所调用，避免了冗余。通过模型库可以将多个模型组合起来构成更大的模型。模型库的模型除了数学模型以外，还有数据处理模型、图形和图像模型、报表模型、智能模型等，这是组合多个模型所需要的。多种类型的模型自然扩充了辅助决策的能力。

方法库系统由方法库和方法库管理系统组成。方法库系统是决策支持系统的一个重要组成部分。方法库由基本方法和标准算法组成，它为模型提供基本模块和程序。一个模型有多个不同的方法，例如：运输问题模型可以有：表上作业法、标号法、图上作业法等三种不同的方法。尽管三者的效果相同，但具体求解方法是不同的，三者的程序也不相同。多种方法组成一个模型，例如，预测模型可以由相关分析方法和线性回归方法二者共同组成。相关分析法从大量试验数据中选择那些相关性显著的因子，舍弃那些相关性不显著的因子（由相关系数进行筛选）。线性回归方法是对那些显著因子的变化数据，利用最小二乘法建立线性回归方程。之后，再利用该回归方程来预测今后的情况。综上所述，模型和方法是不同的，模

型接近实际问题，方法接近基础和求解算法，模型是由方法组成的。

数据仓库是面向主题的、集成的、稳定的、不同时间的数据集，用于支持经营管理中决策的制定过程。传统数据库用于事务处理，也叫操作型处理，是指对数据库联机进行日常操作，即对一个或一组记录的查询和修改，主要为企业特定的应用服务。用户关心的是响应时间、数据的安全性和完整性。数据仓库用于决策支持，也称分析型处理，用于决策分析，它是建立决策支持系统的基础。

联机分析处理（OLAP）是一种软件技术，它使分析人员能够迅速、一致、交互地从各个方面观察信息，以达到深入理解数据的目的。这些信息是从原始数据转换过来的，按照用户的理解，它反映了企业真实的方方面面。

人机交互接口是决策支持系统中的组成部分之一，是决策支持系统与用户之间的交互界面。用户通过“人机交互接口”控制实际决策支持系统的运行，决策支持系统既需要用户输入必要的信息（用于控制）和数据（用于计算），同时要向用户显示运行的情况以及最后的结果。作为决策支持系统的人机交互系统，既包括一般意义下的人机交互接口，还包括将模型库部件、数据库部件组合成系统的系统集成功能。目前人机交互方式有多种形式，一般有菜单、填表、命令语言、屏幕显示、窗口、报表输出等。

国防动员应急决策支持系统总体框架的设计中还考虑了一些辅助系统的设计，如：规划模块，主要通过以往经验和对突发事件处理的最终目标建立预警机制，制定相应的应急预案。仿真模块，通过对一些突发事件的仿真模拟，对应急预案进行完善，对相关部门人员进行培训，以增强应对突发事件的能力。地理信息系统（GIS）模块，为突发事件的准确定位，以及应急物资的调配提供支持。目前，面向主体（Agent）的软件开发已成为软件开发的一种重要方法，我们设计了用于与其他相关领域开发的软件 Agent 进行协调通信的通信模块。

4. 2 应急决策领域本体建模

本论文主要讨论应急决策领域知识的本体表示问题，针对系统总体框

型接近实际问题，方法接近基础和求解算法，模型是由方法组成的。

数据仓库是面向主题的、集成的、稳定的、不同时间的数据集，用于支持经营管理中决策的制定过程。传统数据库用于事务处理，也叫操作型处理，是指对数据库联机进行日常操作，即对一个或一组记录的查询和修改，主要为企业特定的应用服务。用户关心的是响应时间、数据的安全性和完整性。数据仓库用于决策支持，也称分析型处理，用于决策分析，它是建立决策支持系统的基础。

联机分析处理（OLAP）是一种软件技术，它使分析人员能够迅速、一致、交互地从各个方面观察信息，以达到深入理解数据的目的。这些信息是从原始数据转换过来的，按照用户的理解，它反映了企业真实的方方面面。

人机交互接口是决策支持系统中的组成部分之一，是决策支持系统与用户之间的交互界面。用户通过“人机交互接口”控制实际决策支持系统的运行，决策支持系统既需要用户输入必要的信息（用于控制）和数据（用于计算），同时要向用户显示运行的情况以及最后的结果。作为决策支持系统的人机交互系统，既包括一般意义下的人机交互接口，还包括将模型库部件、数据库部件组合成系统的系统集成功能。目前人机交互方式有多种形式，一般有菜单、填表、命令语言、屏幕显示、窗口、报表输出等。

国防动员应急决策支持系统总体框架的设计中还考虑了一些辅助系统的设计，如：规划模块，主要通过以往经验和对突发事件处理的最终目标建立预警机制，制定相应的应急预案。仿真模块，通过对一些突发事件的仿真模拟，对应急预案进行完善，对相关部门人员进行培训，以增强应对突发事件的能力。地理信息系统（GIS）模块，为突发事件的准确定位，以及应急物资的调配提供支持。目前，面向主体（Agent）的软件开发已成为软件开发的一种重要方法，我们设计了用于与其他相关领域开发的软件 Agent 进行协调通信的通信模块。

4. 2 应急决策领域本体建模

本论文主要讨论应急决策领域知识的本体表示问题，针对系统总体框

架中决策生成模块中的知识库进行基于本体的建模,以便为不同领域的决策者提供对应急决策领域知识公共而一致的理解,满足其在面临海量数据、信息的情况下,对信息查准、查全的要求。

从哲学领域借鉴过来的本体论,如何科学地指导领域知识的建模与开发是一个亟待解决的问题。为此,首先必须明确本体论对现实世界进行抽象、建模的基本观点和方法。然后研究如何将这些基本的观点和方法应用于特定领域知识建模,包括模型的建立、模型的描述、模型的检验、模型的应用方法等。

4. 2. 1 领域知识建模的本体论基础

Mario Bunge^[30]关于本体论的理论被广泛地作为检验系统分析和设计方法的基础理论。本文采用 Bunge 提出的本体论框架作为领域知识建模的本体论基础。Bunge 系统地阐述了一系列用于表示现实世界现象的高层和抽象的构造符。Bunge 的工作兼容并收了历史上和同时代的大量研究成果,使用数学方法严格地阐述了他的方法体系。下面首先介绍 Bunge 本体论模型与静态知识建模相关的主要概念和观点。

事物 (Thing)

基本假定 4.1: 世界是由具有性质的事物构成的。

性质 (Properties)

事物都具有性质。具体事物的性质被称为实质性质 (substantial) 或性质。概念上的事物的性质被称为形式化性质、属性或断言。

性质可以依赖某个或多个事物,分为内在性质 (intrinsic) 和相互性质 (mutual): 内在性质仅依赖于某个具体事物。相互性质依赖两个或多个具体事物。

属性 (Attributes)

事物性质的存在不依赖于人类的感知与否。属性是根据人类的感知指派给事物的特征。

基本假定 4.2: 一般地,每个性质都能够被表示为命题函数 A 。命题函数 A 定义为: $T_1 \times \dots \times T_n \times V_1 \times \dots \times V_m \rightarrow$ 关于 A 的陈述,其中,

$T_k(k=1, \dots, n)$ 表示事物的集合, $V_j(j=1, \dots, m)$ 表示值的集合; 并且每个特定的性质都能够被表示为如下形式的属性函数, $A(t_1, \dots, t_n, v_1, \dots, v_m)$, 其中 $t_i \in T_i, v_j \in V_j$ 。

类 (Class) 和种类 (Kind)

定义 4.1: 性质的范围是具有这些性质的事物的集合。也就是说, 如果 Θ 是所有事物的集合, P 是所有性质的集合, 范围函数 S 定义为 $S: P \rightarrow 2^\Theta$ 。

定义 4.2: 事物的子集被称为类, 当且仅当子集满足以下条件: 存在一个性质, 事物的子集是这个性质的范围。也就是说, 事物集合 Θ 的子集 X 被称为类, 当且仅当 $\exists p \in P, X = S(p) \in 2^\Theta$ 。

定义 4.3: 事物集合 X 是另一个事物集合 Y 的子类, 当且仅当 X 是 Y 的真子集。相反地, Y 是 X 的超类。

推论 4.1: $S(p1) = X, S(p2) = Y, p1, p2 \in P$, 当且仅当 $S(p1) \subset S(p2)$ 时, X 是 Y 的子类。

定义 4.4: R 是某些性质的集合, R -kind 为 R 中所有性质的范围的交集, 称为 R 种类。

当一个种类包含的所有事物的性质组合遵守“规律”时, 这个种类被称为自然种类 (natural kind)。

功能模式 (Functional Schema), 状态 (State) 和规律 (Law)

定义 4.5: T 是具有相同性质集合的事物集合。功能模式 $X_m = \langle M, F \rangle$, M 是非空的集合, F 是在 M 上定义的函数的有限序列, $F = \langle F_1, \dots, F_n \rangle$; 其中, $F_i: M \rightarrow V_i, V_i$ 是一个值域, 每个 (属性) 函数均表示事物的一个性质。

定义 4.6: 假定事物 X 是由功能模式 $\langle M, F \rangle$ 描述的, 函数 $F_i: M \rightarrow V_i$,

用于表示事物的第 i 个状态函数。 $F(t) = \langle F_1(t), \dots, F_n(t) \rangle$ 在确定点 t 的值的集合表示 X 在 t 的状态, $t \in M$ 。

定义 4.7: 任何关于事物功能模式的构成部分的可能取值、或者关于事物功能模式的构成部分结合 (combinations) 的约束, 被称为规律。

原理 4.1: 没有两个事物具有完全相同的特定性质集合。

原理 4.2: 一个事物如果被命名, 只要事物的自然种类不发生变化, 事物将保持它的名称不变, 此即为标称不变性 (Nominal invariance)。

交互作用 (Interaction)

定义 4.8: X 作用于事物 Y , 当且仅当 Y 遍历 M 的给定子集时, X 存在时 Y 的状态有别于 X 不存在时。如果事物 X 和事物 Y 中的一个事物作用于另外一个事物, 那么事物 X 和 Y 之间存在交互作用:

组合 (Composition)

原理 4.3: 两个事物能够组合 (associate) 形成另外一个事物。

定义 4.9: 当且仅当组合事物的某个性质是其构成部分的性质时, 称该性质为继承的 (inherited), 否则称该性质为突现的 (emergent)

原理 4.4: 每个组合事物都具有突现的性质。

4. 2. 2 应急决策领域本体图示化建模方法 DKOED

以领域知识建模的本体论基础理论为指导, 作者提出了一种适用于应急决策领域本体建模方法 (Domain Knowledge Ontology for Emergency Decision, DKOED)。DKOED 方法将语义网络的知识表示方法与面向对象逻辑建模、面向数据的语义建模方法有机地进行了结合, 能够有效地对领域结构、领域中的概念、概念的性质、概念间的关系等静态知识进行建模。

本节首先对 DKOED 方法采用的面向本体的分析方法、建立领域本体的基本步骤进行了阐述。其次, 给出了领域本体的基本建模元素、基本建模元素的图示化建模符号。

4. 2. 2. 1 DKOED 的分析方法和主要步骤

用于表示事物的第 i 个状态函数。 $F(t) = \langle F_1(t), \dots, F_n(t) \rangle$ 在确定点 t 的值的集合表示 X 在 t 的状态, $t \in M$ 。

定义 4.7: 任何关于事物功能模式的构成部分的可能取值、或者关于事物功能模式的构成部分结合 (combinations) 的约束, 被称为规律。

原理 4.1: 没有两个事物具有完全相同的特定性质集合。

原理 4.2: 一个事物如果被命名, 只要事物的自然种类不发生变化, 事物将保持它的名称不变, 此即为标称不变性 (Nominal invariance)。

交互作用 (Interaction)

定义 4.8: X 作用于事物 Y , 当且仅当 Y 遍历 M 的给定子集时, X 存在时 Y 的状态有别于 X 不存在时。如果事物 X 和事物 Y 中的一个事物作用于另外一个事物, 那么事物 X 和 Y 之间存在交互作用:

组合 (Composition)

原理 4.3: 两个事物能够组合 (associate) 形成另外一个事物。

定义 4.9: 当且仅当组合事物的某个性质是其构成部分的性质时, 称该性质为继承的 (inherited), 否则称该性质为突现的 (emergent)

原理 4.4: 每个组合事物都具有突现的性质。

4. 2. 2 应急决策领域本体图示化建模方法 DKOED

以领域知识建模的本体论基础理论为指导, 作者提出了一种适用于应急决策领域本体建模方法 (Domain Knowledge Ontology for Emergency Decision, DKOED)。DKOED 方法将语义网络的知识表示方法与面向对象逻辑建模、面向数据的语义建模方法有机地进行了结合, 能够有效地对领域结构、领域中的概念、概念的性质、概念间的关系等静态知识进行建模。

本节首先对 DKOED 方法采用的面向本体的分析方法、建立领域本体的基本步骤进行了阐述。其次, 给出了领域本体的基本建模元素、基本建模元素的图示化建模符号。

4. 2. 2. 1 DKOED 的分析方法和主要步骤

DKOED 方法对领域本体的基本观点为：

1. 某个领域的本体就是关于该领域的一个公认的概念集, 其中的概念含有公认的语义, 这些语义通过概念之间的各种关系来体现。本体通过它的概念集及其所处的上下文来刻画概念的内涵。

2. 本体的基本组成元素是概念, 概念元素是具有结构的, 本体是一组结构化成分的组合。

3. 概念之间存在多种类型的关系, 关系在组成本体的各元素之间起着至关重要的作用。

4. 本体根据抽象层次的不同, 也存在继承关系。

从总体上来讲, 面向本体的分析方法是面向对象方法的延伸。本体论的哲学定义告诉我们, 本体论研究实体的存在性和实体存在的本质。面向本体的分析方法将面向对象方法向更深入、更本质的方向推进。本体中概念的参照物是所有具有与概念同样性质的事物的类。从这种意义上说, 概念表示事物的特定性质, 概念元素与面向对象中的类具有对应关系, 概念的实例与面向对象中的对象具有对应关系。

目前, 领域本体建模方法 (如 Ontolingua、Flogic、KL-ONE 等) 的研究集中于如何建立领域静态知识模型, 因此, DKOED 方法主要用于领域静态知识模型的建立。

以 Bunge 本体论框架为基础, 作者给出一种 DKOED 方法中领域本体的形式化定义。

定义 4.10: 领域本体

$\text{Domain_Ontology} ::= \langle \text{meta_info}, \text{CONCEPTS}, \text{RELATIONS}, \text{RULES}, \text{AXIOMS}, \text{INDIVIDUALS} \rangle$; 其中:

$\text{meta_info} ::= \langle \text{ONTO}_{\text{title}}, \text{ONTO}_{\text{relation}}, \text{ONTO}_{\text{creator}}, \text{ONTO}_{\text{subject}}, \text{ONTO}_{\text{discription}}, \text{ONTO}_{\text{publisher}}, \text{ONTO}_{\text{date}}, \text{ONTO}_{\text{format}}, \text{ONTO}_{\text{identifier}}, \text{ONTO}_{\text{soucre}}, \text{ONTO}_{\text{language}}, \text{ONTO}_{\text{right}} \rangle$;

$\text{CONCEPTS} ::= \{\text{concept}\}$;

$\text{RELATIONS} ::= \{r(c_1, c_2) \mid c_1, c_2 \in \text{CONCEPTS}, r \text{ 为关系的名称}\}$;

$\text{RULES} ::= \{\text{RULE}\}$;

$\text{AXIOMS} ::= \{\text{AXIOM}\}$;

INDIVIDUALS ::= {individual | $\sigma(\text{individual}) \in \text{CONCEPTS}$ }。

meta_info 定义本体的基本元信息，包括本体的名称、继承其它本体的集合、创建者、主题、对本体内容的描述、发布者、日期、格式、标识、资料的来源、语言、版权等元信息。

CONCEPTS 为领域内以术语形式出现的相应概念集合，概念具有刻画事物性质的属性。

RELATIONS 为领域内概念之间的二元关系集合。一般地，本体中概念之间至少存在以下三类二元关系：继承关系、聚合关系、关联关系，概念和概念中属性之间的关系也是一类重要的关系，称为属性关系。

RULES 为规则的集合。任何关于概念或关系的可能取值或它们之间可能的组合关系的约束被称为规则。一般地，可以将规则定义为 $\text{RULE}: X_1 \times \dots \times X_n \times T \rightarrow S$ ，其中， $X_1, \dots, X_n \in \text{CONCEPTS} \cup \text{RELATIONS}$ ； T 表示规则被“观察”的所有条件，通常为时间实例的集合； S 表示关于 X_1, \dots, X_n 的陈述。RULE 表示在某些观察点的条件下， S 对 X_1, \dots, X_n 可能取值或组合关系的陈述成立。

AXIOMS 为公理的集合，用于定义规则或关系之间的关系和规则。 $\text{AXIOM}: X_1 \times \dots \times X_n \rightarrow S$ ，其中， $X_1, \dots, X_n \in \text{CONCEPTS} \cup \text{RELATIONS} \cup \text{RULES}$ 。AXIOMS 表示在任何条件下 S 对 X_1, \dots, X_n 可能取值或组合关系的陈述均成立。

INDIVIDUALS 为领域内概念实例的集合。其中， $\sigma: \text{INDIVIDUALS} \rightarrow \text{CONCEPTS}$ 为概念实例到所属相应概念的映射，概念的实例具有所属概念定义的属性。

● DKOED 的分析方法

人类认识世界的过程从本质上讲就是对事物进行抽象和区分的过程，典型的过程包括三个方面的内容^[31]：

1. 从现实世界中区分特定的客体及其属性。
2. 区分事物的整体及其组成部分。
3. 对不同种类的事物给出形式化表示，然后在此基础上加以区分。

DKOED 分析方法的基本概念和方法也是建立在以上三种思维组织模式之上的。作为面向对象方法的延伸，首先采用面向对象方法中自底向上

的分析方法,标识对象、通过对具体对象的抽象化来发现类、标识类的分类结构与组装结构。这些工作与面向对象分析方法中相应的工作基本保持一致。

在面向对象方法中,确定分类结构与组装结构是识别类和对象之间关系的主要工作,通过识别对象之间的消息连接体现对象之间的其它关系。在面向本体的方法中,确定概念之间的关系是一个非常重要的分析阶段,不仅要确定纵向的分类关系,而且需要识别概念之间的横向联系。因此,DKOED方法在这一阶段,采用自顶向下的方法。以“行为”为中心自顶向下的关系分析方法基本步骤如下:

步骤 1: 识别领域最顶层的行为种类 Act_1, \dots, Act_n 。

步骤 2: 识别各种行为种类涉及的概念,作为概念之间关系的集合 R 。

步骤 3: 识别各种行为种类中的所有行为的集合 $SAct_1, \dots, SAct_n$ 。

步骤 4: 识别集合 $SAct_1, \dots, SAct_n$ 中的行为涉及的概念。如果 R 中的关系不能够体现识别出的概念所涉及的某个行为 x , 那么在 R 中增加反映行为 x 的关系。

步骤 5: 识别各种行为种类之间的依赖关系 D 。

步骤 6: 识别各种行为种类内部的行为依赖关系, 合并到 D 中。

R 作为领域中概念之间关系的集合, D 作为领域中规则的集合。

● DKOED 方法的主要步骤:

步骤 1: 采用面向对象的方法,标识对象、通过对具体对象的抽象化来发现类、标识类的分类结构与组装结构。

步骤 2: 采用自顶向下的分析方法确定概念之间的横向关系。

步骤 3: 采用扩展的 UML 语言定义概念和概念之间的关系,建立领域本体的图示化模型。

4.2.2.2 DKOED 方法中领域本体的图示化建模符号

DKOED 方法采用类表示领域中涉及的概念,采用对象表示概念的实例——事实个体。DKOED 方法采用 UML 类图的标准建模符号、类图的 2 种标准约束(完全、不完全)以及使用 UML 的约束扩展机制针对属性、关系继承、不相交继承、完全-不相交继承定义的扩展约束,形成了一套标准的

领域本体建模符号。

1. 类



说明：在 DKOED 中构造领域本体时，首先要定义类，也就是通过定义类描述领域中涉及的概念。类的实例为对象，用于表示概念的实例——事实个体。DKOED 中类的图示化建模符号采用 UML 类图中的“类”表示。

规则：

- (1) 类必须具有一个名称，在同一个领域本体中类的名称是唯一的；
- (2) 类的名称应该使用能够反映类所代表的概念的名称。

2. 属性

说明：类的属性放置在类名字的下方，用来描述该类的对象所具有的特征。属性具有类型，属性的类型反映属性的种类。采用 UML 约束扩展机制表示 DKOED 方法中的属性定义。属性的语法格式为：

属性名：类型名=初值 {m, n, #, (枚举类型属性值)}。

其中，属性名：类型名=初值是 UML 定义属性的语法，{m, n, #, (枚举类型属性值)} 为扩展约束。基数约束 m 和 n 为可选项，缺省为 m=1, n=1；“#”为可选项，表示该属性是类的“限定性的性质”；(枚举类型属性值) 是可选项，为属性定义枚举类型的属性值。

规则：

- (1) 每个类可以具有多个属性，不同的类可以具有相同的属性；
- (2) 在通常的情况下，不同的类可以具有相同的属性名，不同类中的同名属性可以具有不同的类型。但是为了属性定义语义的清晰，DKOED 方法要求类型一致的属性可以出现在多个类中，类型不一致的属性必须使用不同的属性名。

(3) 本体中属性的类型为系统预定义的类型，也可以在 {性质串} 中为属性定义枚举类型的属性值。

(4) 属性定义位于类名称的下方，m, n 表示同一属性在类实例中能

够出现的次数, m, n 是正整数, $m \leq n$, $n=*$ 时表示无上限。属性定义中带有标签值“#”的属性基数约束只能为 $m=1, n=1$ 。

3. 关系

图示化建模符号	建模符号名称
$\frac{\text{roleA 关系名称 roleB}}{mA..nA \quad nB..nB}$	关系

说明: 关系用于描述类与类之间的语义连接。对象是类的实例, 类与类之间的关系代表了对象之间的关系。类与类之间存在多种连接方式, 每种连接的语义各不相同, 但是外部的表现形式相似, 所以通称为关系。DKOED 中关系的图示化建模符号采用 UML 类图中的“关联关系”表示。

任何所连接的实体超过两个的关系, 都能够转换为一组二无的多对一关系集合, 而不丢失任何信息。因此, DKOED 中的关系设计为二元关系。

规则:

(1) 关系是双向的, 可以在关系的每一个方向上为关系起一个名称。这个名称就是类在关系中所扮演角色的名称(参见角色建模符号的定义);

(2) 关系必须具有一个名称。在同一个本体中, 关系的名称唯一确定了与关系相关的角色的名称;

(3) 关系的名称要能够反映关系所代表的交互关系;

(4) 一个类可以与它自身存在关系, 这种关系被称为递归关系。

4. 角色

说明: 在 DKOED 中, 角色用于指示类在关系中的作用。同时, 角色还可用于表示类的实例在关系中的基数约束。DKOED 中的角色采用 UML 类图中“关联关系”中的“角色”表示, 基数约束采用 UML 中的多重性约束表示。

规则:

(1) 角色的名称唯一确定了角色所在的关系, 具有继承关系的类可以具有相同名称的角色, 角色的名称要能够反映相应类在关系中的地位和作用;

(2) roleA 是类 A 与类 B 之间的关系中类 A 扮演的角色, 类 A 为角色的实施者, 类 B 为 roleA 作用的接收者;

(3) roleA 和 roleB 分别是类 A 与类 B 在关系 R 中的角色, 基数约束分别为 $m_A..n_A$, $m_B..n_B$ 。角色基数约束的含义是类 A 的同一个实例至少要与类 B 的 m_B 个实例构成 m_B 个 R 的实例; 类 A 的同一个实例最多能够与类 B 的 n_B 个实例构成 n_B 个 R 的实例, 反之亦然。

基数约束允许的组合为: $0..n$ 、 $m..n$ 、 $0..*$ 、 $m..*$ 、 $1..1$, 其中 m 、 n 是正整数, $m \leq n$, $n=*$ 表示无上限, $1..1$ 是角色基数的缺省约束。

5. 继承

图示化建模符号	建模符号名称
	基本继承
	完全继承
	不相交继承
	完全-不相交继承

说明: 一个类的信息能够被另一个类继承, 继承其它类的类不仅可以具有自己的信息, 而且能够具有被继承类中的信息, 这种机制被称为继承。具有继承关系的两个类中, 被继承的类称为父类, 继承其它类信息的类称为子类。DKOED 方法采用 UML 类图中的“类”、继承关系的两种标准约束(完全、不完全)、扩展约束“不相交继承”和扩展约束“完全-不相交继承”表示类之间多种继承关系的图示化建模符号。

规则:

(1) 类的继承关系可以是多层的。也就是说, 一个子类本身还可以作为另一个类的父类, 层层继承下去。在这种情况下, 类与类之间存在祖先和后代关系。

(2) 继承同一个父类的子类之间可以存在如下三种约束:

- 完全继承: 父类的所有子类都被穷举完毕, 不可能再有其它未列出的子类存在, 这种约束称为完全继承约束;

- 不相交继承: 父类的所有子类两两之间互不相交, 这种约束称为不相交继承约束;

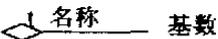
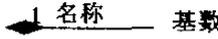
- 完全-不相交继承: 父类的所有子类同时满足完全继承和不相交继承的约束;

没有附加以上约束的继承被称为基本继承或不完全继承, 是继承的缺

省情况。

(3) 一个子类能够同时继承多个父类，被继承的多个父类之间不能存在继承关系。

6. 聚合

图示化建模符号	建模符号名称
 名称 基数	普通聚合
 名称 基数	复合聚合

说明：聚合是一类特殊的关系，用于表示不同类之间的“整体与部分”关系。DKOED 方法采用 UML 类图中的“聚合关系”、“复合聚合关系”表示类之间聚合关系的图示化建模符号。

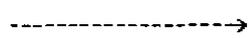
规则：

(1) 可以为部分方的类定义在聚合关系中的基数约束，部分的基数约束为 $m..n$ ，其中 m 、 n 为正整数， $m \leq n$ ， $n=*$ 时表示无上限；整体方类的基数约束为 1；

(2) 如果部分类完全隶属于整体类，称这样的聚合为复合聚合。即如果整体类不存在，则部分类也没有存在的价值；

(3) 聚合关系的名称在同一个领域本体中是唯一的。

7. 关系的继承

图示化建模符号	建模符号名称
	关系继承

说明：一个关系能够被另一个关系继承。DKOED 方法采用 UML 类图中的“依赖关系”表示关系之间继承的图示化建模符号。

规则：

(1) 如果关系 $R1$ 被关系 $R2$ 继承，那么 $R1$ 所连接的类是 $R2$ 连接的类自身或祖先，或者 $R2$ 所连接的类聚合为 $R1$ 所连接的类；

(2) 如果关系 $R1$ 被 $R2$ 继承，那么 $R1$ 和 $R2$ 两个方向的上角色基数分别满足以下约束：

$R2$ 的最小值约束 $\geq R1$ 的最小值约束；

$R2$ 的最大值约束 $\leq R1$ 的最大值约束。

8. 类不相交

图示化建模符号	建模符号名称
<u>disjoint</u>	类不相交

说明：可以在继承关系以外，定义类之间的不相交关系。DKOED 方法通过在 UML 类图中使用名称为“disjoint”的“关联关系”表示类不相交关系的图示化建模符号。

规则：

(1) 可以在两个具有相同祖先，但不在该祖先继承关系的同一个层次上的类之间定义类不相交关系；

(2) 可以在不存在祖先关系，但是其中一个类的祖先与另外一个类的祖先或另外一个类本身存在聚合关系的类之间定义类不相交关系。

4.2.2.3 DKOED 中规则的表达

有关规则的表达具有不同的抽象层次。在 DKOED 方法中，规则的表达分为“本质层”和“应用层”两个层次。

DKOED 中领域本体对类（概念）及类（概念）之间各种联系的定义（包括继承、聚合、类不相交、关系、关系继承、角色等）本身就是一种对领域认识的公理化定义，同时隐含了本体创建者对领域内各种基本规则的表达。这些公理化的定义和基本规则的表达，构成了 DKOED 方法中领域本体的“本质层”规则表达。

针对“应用层”规则的表达，不同的应用可以采用 UML 语言中的 OCL 语言或其它语言在领域本体之上定义特定的应用规则，然后选择应用规则具体的实现策略。在应急决策领域本体知识库中应用层规则的表达采用产生式规则的表达方法，产生式规则可以利用数据库中的表来实现。我们将应用层规则与概念（事实）分开存放，这样做的好处是便于编辑和推理方法的灵活使用。我们对应急决策领域应用层规则进行了定义，以下列出了其中的部分规则：

IF 发生地震 THEN 成立应急决策指挥组 AND 制定应急方案

IF 发生火灾 THEN 派出消防人员 AND 派出医务人员

IF 火山爆发 AND 有人员伤亡 THEN 成立应急决策指挥组 AND 征召

烧伤科医生

IF 发生洪水 THEN 征召武警人员 AND 制定应急方案

4. 2. 3 应急决策领域本体模型的建立

本体建模的实质就是建立某一领域的信息模型。一个构造良好且能够实用的本体需要领域知识专家、人工智能专家、数据库专家、Web 技术人员甚至是语言学家的共同参与才能取得成功。作者在此构建的应急领域本体仅仅是一个示例,虽然会有不少考虑不周之处,但是对于演示本体的构造方法,并以此为基础进行基于本体的知识搜索研究却是足够的。

我们把建立应急决策领域本体模型的过程分为两个阶段:第一阶段是用 DKOED 方法建立领域本体的图示化模型;第二阶段是使用本体描述语言 OWL 对领域本体模型进行编码,形成便于人们交流、无歧义的可被软件解释的本体文档。

4. 2. 3. 1 用 DKOED 方法建立领域本体的图示化模型

本节结合一个比较完整的建模实例详细说明了 DKOED 方法的建模过程,总结了 DKOED 方法的特点。以下概要说明利用 DKOED 方法对应急领域本体建模的过程,

1、应急领域知识的获取

为了获得应急决策领域的相关知识,作者多次到山西省军区进行了调研,查阅了大量有关应急决策方面的论文,并对应急决策是国家有关部门在突发事件发生前及发生后采取紧急措施,为满足突发事件需求,对人力、物力、财力及其他各种资源进行统一组织、指挥、协调、控制的一系列活动进行分析。总之,通过对应急决策内容的分析及与领域专家的相互交流,我们以应对各种突发灾害的决策制定为例,经过标识对象、发现类、标识类的分类结构与组装结构,得到领域中的概念分类体系和概念的组合结构,如图 4-2 所示:

烧伤科医生

IF 发生洪水 THEN 征召武警人员 AND 制定应急方案

4. 2. 3 应急决策领域本体模型的建立

本体建模的实质就是建立某一领域的信息模型。一个构造良好且能够实用的本体需要领域知识专家、人工智能专家、数据库专家、Web 技术人员甚至是语言学家的共同参与才能取得成功。作者在此构建的应急领域本体仅仅是一个示例, 虽然会有不少考虑不周之处, 但是对于演示本体的构造方法, 并以此为基础进行基于本体的知识搜索研究却是足够的。

我们把建立应急决策领域本体模型的过程分为两个阶段: 第一阶段是用 DKOED 方法建立领域本体的图示化模型; 第二阶段是使用本体描述语言 OWL 对领域本体模型进行编码, 形成便于人们交流、无歧义的可被软件解释的本体文档。

4. 2. 3. 1 用 DKOED 方法建立领域本体的图示化模型

本节结合一个比较完整的建模实例详细说明了 DKOED 方法的建模过程, 总结了 DKOED 方法的特点。以下概要说明利用 DKOED 方法对应急领域本体建模的过程,

1、应急领域知识的获取

为了获得应急决策领域的相关知识, 作者多次到山西省军区进行了调研, 查阅了大量有关应急决策方面的论文, 并对应急决策是国家有关部门在突发事件发生前及发生后采取紧急措施, 为满足突发事件需求, 对人力、物力、财力及其他各种资源进行统一组织、指挥、协调、控制的一系列活动进行分析。总之, 通过对应急决策内容的分析及与领域专家的相互交流, 我们以应对各种突发灾害的决策制定为例, 经过标识对象、发现类、标识类的分类结构与组装结构, 得到领域中的概念分类体系和概念的组合格式结构, 如图 4-2 所示:

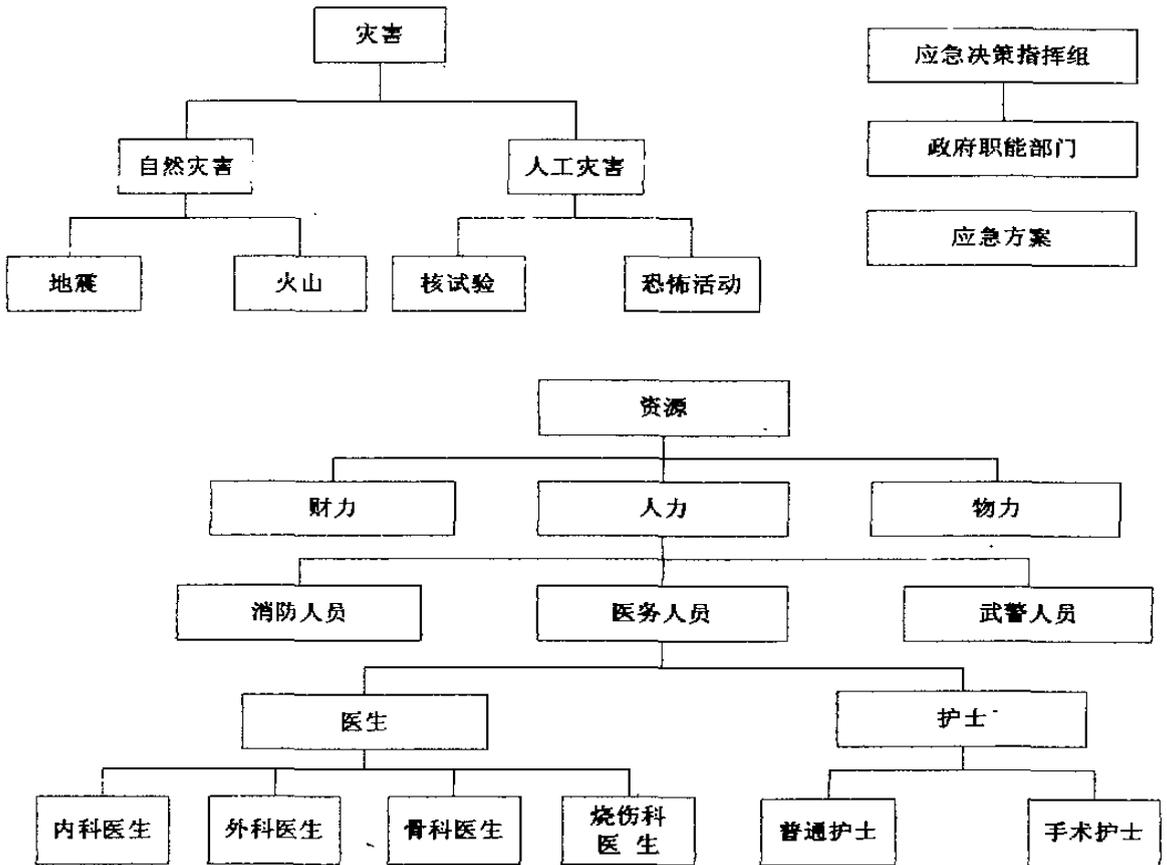


图 4-2 领域本体的概念分类体系

Figure 4-2 the concept classification of domain ontology

2、概念之间关系的分析

我们采用 4.2.2.3 节介绍的 DKOED 分析方法, 自顶向下的分析并确定概念之间的横向关系, 得到概念关系表 4-1。

表 4-1 领域实例的概念关系表

Tabal 4-1 the relations of domain concepts in the example

关系名称	概念 A	概念 A 的角色	概念 B	概念 B 的角色
因果	灾害	起因	应急决策指挥部	后果
承担与分配	地震	分配给	政府职能部门	承担
资源调度	资源	被调度	应急决策指挥部	调度
参与	医生	加入	应急方案	包含
参与	护士	加入	应急方案	包含
需求	灾害	需要	资源	分配给
制订	应急方案	归属	应急决策指挥部	编写

3、概念、关系的图示化描述

利用 DKOED 方法中提供的图示化建模符号,我们定义了如图 4-3 所示的领域本体模型。

图 4-3 中忽略了每个概念的属性定义,图中因果关系表示每个灾害必须至少成立一个应急决策指挥部,一个应急决策指挥部可以处理一到多起灾害;承担与分配关系表示地震发生时必须至少有一个政府职能部门承担灾害的处理事宜;制订关系表示每个应急决策指挥部必须编写一到多个应急方案;参与关系表示一个应急方案中可以包含有一到多名医生(护士),每个医生(护士)只能加入一个应急方案中;资源调度关系表示应急决策指挥部与资源之间存在调度与被调度的关系;承担与分配关系与因果关系存在继承关系;资源与物力、人力、财力存在复合聚合关系;人力与消防人员、医务人员、武警人员存在继承关系;医务人员与医生、护士存在不相交继承关系;医生与外科医生、内科医生、骨科医生、烧伤科医生存在继承关系;护士与普通护士、手术护士存在完全不相交继承关系;应急决策指挥部与政府职能部门之间存在普通聚合关系;灾害与自然灾害、人工灾害之间存在完全不相交继承关系;自然灾害与地震、火山之间及人工灾害与核试验、恐怖活动之间存在继承关系;护士与武警人员存在类不相交关系。

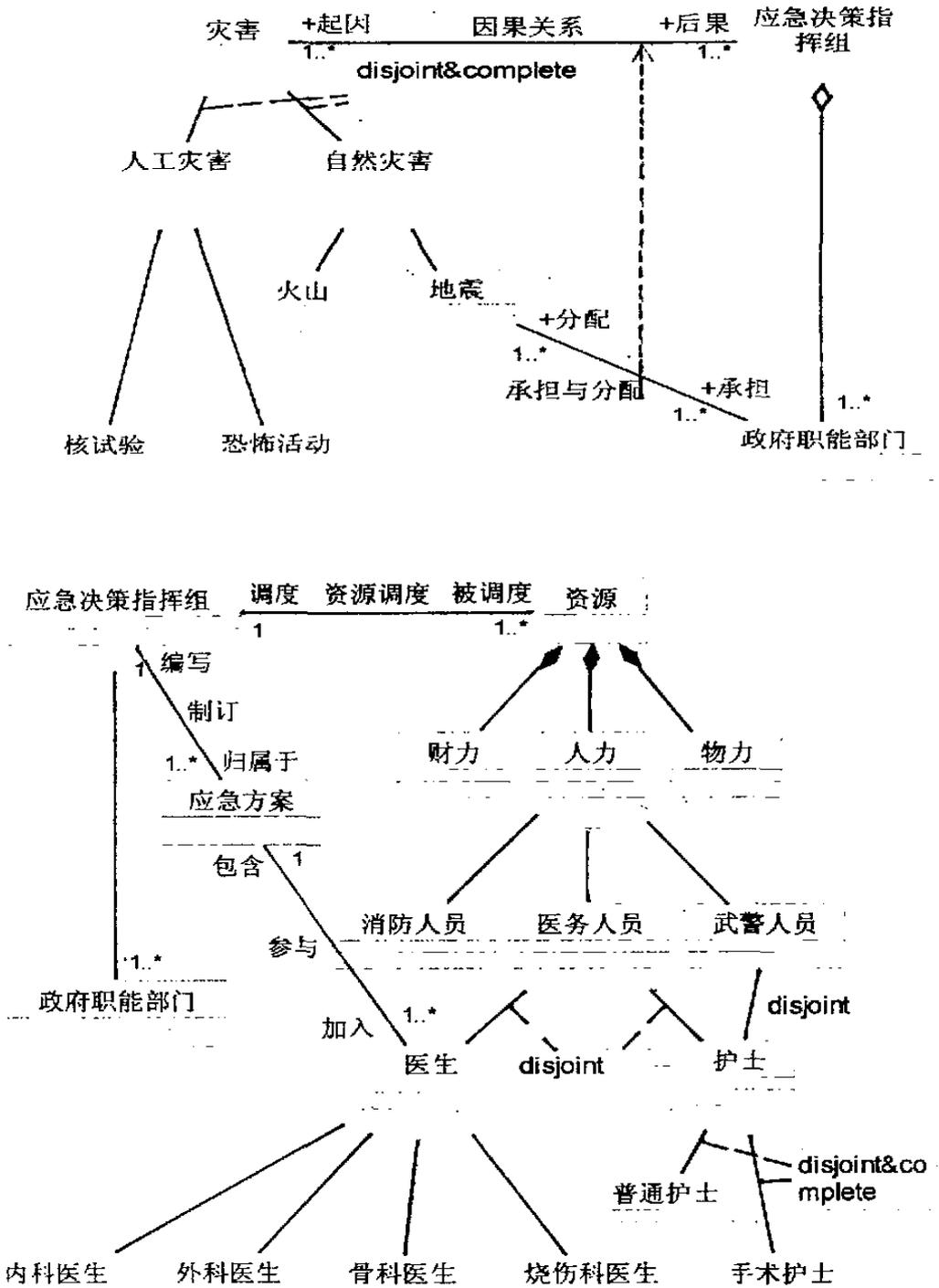


图 4-3 领域本体模型

Figure 4-3 the domain ontology model

从以上对应急决策领域本体建模的过程中,可以看到 DKORD 方法具有以下特点:

1、DKOED 方法直接建立在本体论的基本理论之上,具有朴素的知识建模思想。其知识建模的基本要素:类、属性、关系、角色、继承、聚合与本体论基本理论中的相应概念具有明确的对应关系。在本体论基本理论中,能够通过这些相应的概念构建系统的静态模型。

2、DKOED 方法的基本建模元素涵盖了概念数据建模方法 EER 的建模能力,继承了面向对象逻辑建模能力。具有图示化建模直观、便捷的特点,采用 UML 类图构建的图示化领域本体有利于开发人员和分析人员的沟通。

4. 2. 3. 2 用 OilEd 生成领域本体模型的 OWL+RDF+XML 文档

在利用 DKOED 方法建立领域本体模型的图示化表示后,我们现在采用本体编辑器 OilEd (参见 3.3.3 节)对本体模型进行形式化描述,并利用 OilEd 具有的推理功能对所要描述的本体进行一致性检测和聚合等逻辑关系的检验,最后生成基于 XML+RDF+OWL 的知识库文档。

在用 OilEd 生成本体文档之前,我们首先要在 OilEd 中对用 UML 图示化表示的领域本体中的类及类的属性,以及类与类之间的关系进行定义,图 4-4——4-7 分别显示了类、类的属性、类的实例以及类与类之间关系的公理定义。

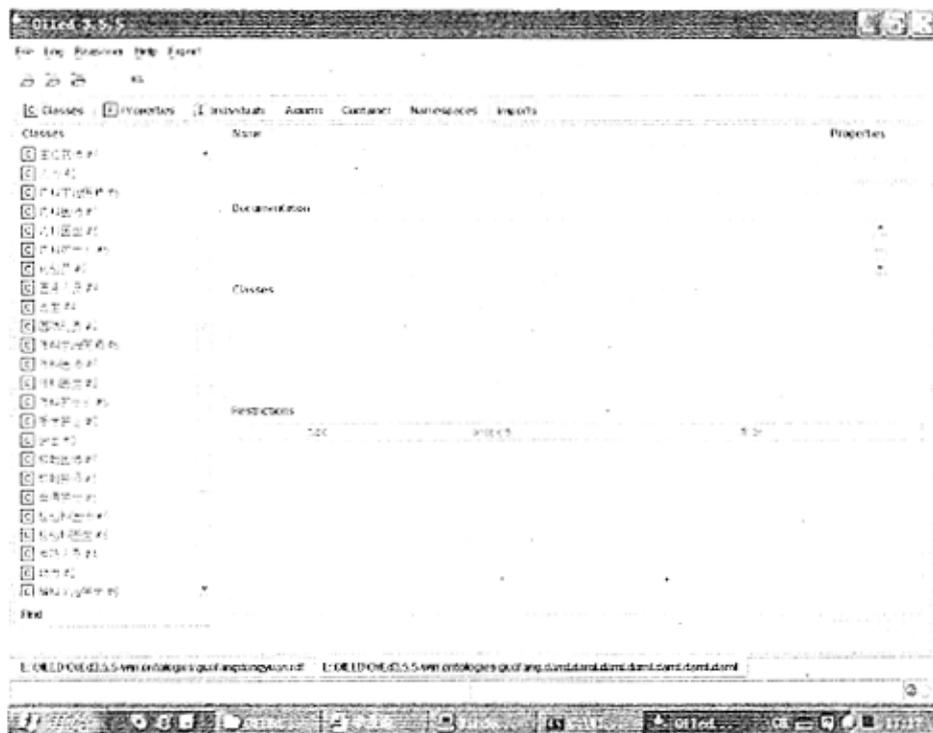


图 4-4 领域本体中类的定义

Figure 4-4 the class definitions of domain ontology

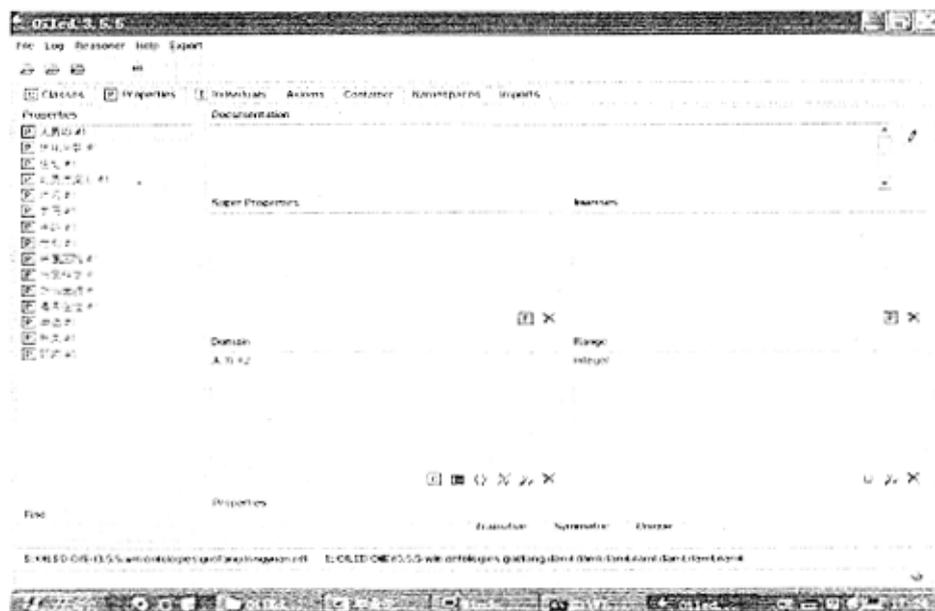


图 4-5 领域本体中类的属性定义

Figure 4-5 the class property definitions of domain ontology

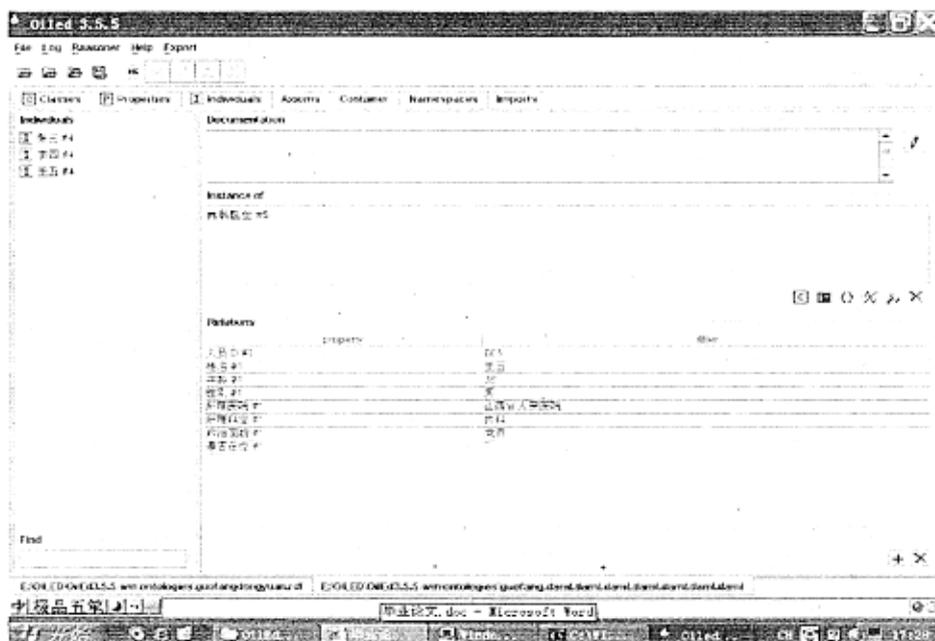


图 4-6 领域本体类的实例定义

Figure 4-6 the individual definitions of domain ontology

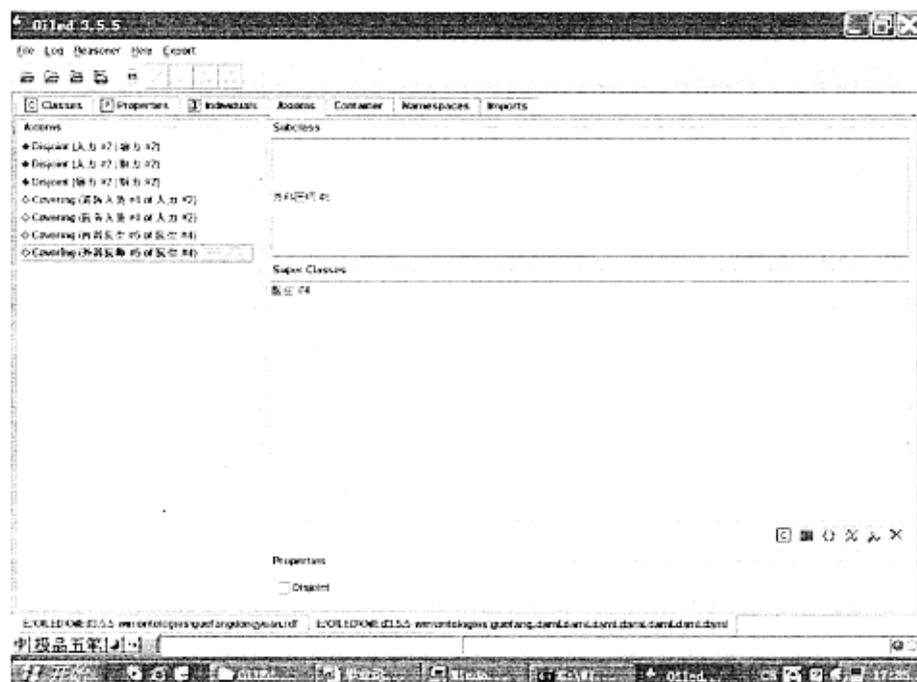


图 4-7 领域本体公理定义

Figure 4-7 the Axiom definitions of domain ontology

在进行了上面的定义之后，我们可以对本体描述进行检验。通过 OilEd 提供的连接功能，可以将其连接到 OilEd 所集成的推理机 FaCT 之上，由 FaCT 推理机完成对本体描述文档的一致性检测和包含等逻辑关系的检验。

利用 **Connect to Reasoner...** 菜单项我们将 OilEd 与推理机 FaCT 进行连接。一旦连接成功，**Classify Project** 按钮将激活，通过此分类功能我们可以完成以下任务：

- 将本体描述翻译成描述逻辑 SHIQ，并提交给推理机以便对本体概念进行分类；
- 对本体中的每一个概念（类）进行一致性检查；
- 对本体中的每一个概念（类），由分类器确定其父类。

在对概念进行分类后，编辑器的活动日志中可得到一个分类情况的记录。在此记录中我们可以看到任何不满足一致性检查的概念，并且这些概念在概念（类）的列表中将用红色突出显示。在这种情况下，我们可以对本体概念之间的关系进行修正，加入或丢弃一些包含关系，与此同时一些不满足信息也可能加入到本体模型中，经过不断的修正，最后得到满足一致性检测和包含等逻辑关系检验的正确本体模型。

在完成对本体模型的检验之后，我们可以得到本体模型中类的层次关系如图 4-8 所示：

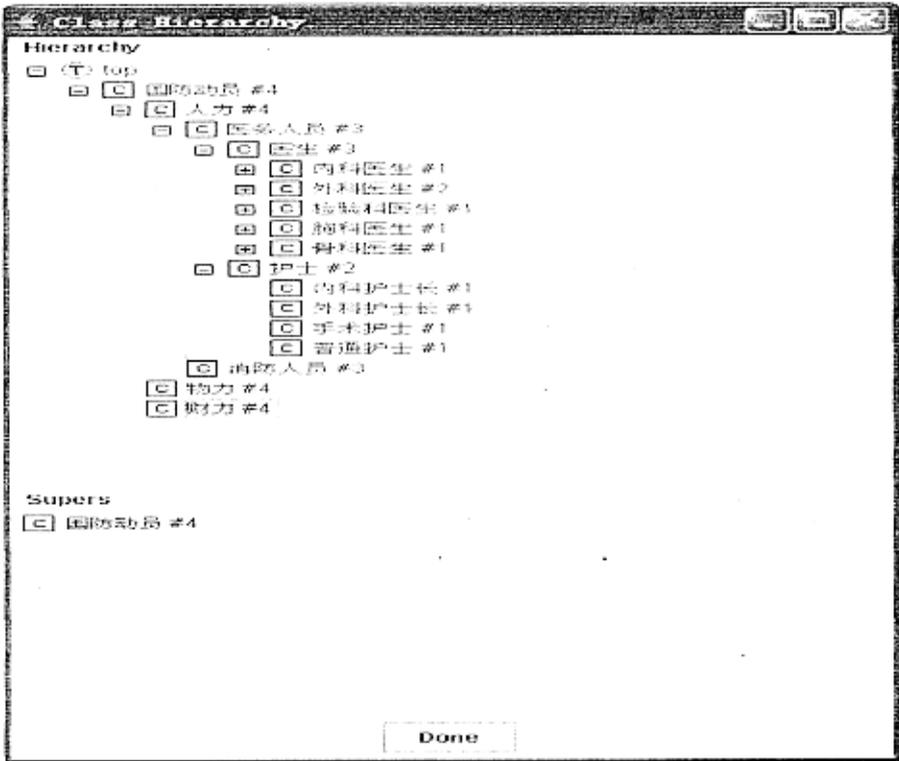


图 4-8 领域本体类的层次结构

Figure 4-8 the class hierarchy of domain ontology

最后我们利用Oiled 的输出功能菜单选择输出本体的基于OWL+RDF+XML的文档。下面列出了用Oiled 生成的国防动员应急决策领域本体中关于“医务人员”概念的部分文档。

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
- <rdf:RDF xmlns:owl="http://www.w3.org/2002/07/owl#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:xsd="http://www.w3.org/2001/XMLSchema#">
    <owl:Ontology rdf:about="" />
    <owl:Class
rdf:about="file:/F:/yzh(important)/ontology/OILED /Oiled
3.5.5-win/ontologies/guofangdongyuan.daml#%D2%BD%CE%F1%C8%CB%D
4%B1">
```

```

    <rdfs:label>医务人员</rdfs:label>
  - <rdfs:comment>
  - <![CDATA[从事医疗事务的人员]]>
    </rdfs:comment>
  - <rdfs:subClassOf>
    <owl:Class
rdf:about="file:/F:/yzh(important)/ontology/OILED /Oiled
3.5.5-win/ontologies/guofangdongyuan.daml#%C8%CB%C1%A6" />
    </rdfs:subClassOf>
    </owl:Class>
    <owl:Class
rdf:about="file:/F:/yzh/ontology/OILED/Oiled3.5.5-win/ontologi
es/guofang.daml#%C8%CB%C1%A6">
    - <owl:disjointWith>
    <owl:Class
rdf:about="file:/F:/yzh/ontology/OILED/Oiled3.5.5-win/ontologi
es/guofang.daml#%CE%EF%C1%A6" />
    </owl:disjointWith>
    </owl:Class>
    <owl:Thing
rdf:about="file:/F:/yzh/ontology/OILED/Oiled3.5.5-win/ontologi
es/guofang.daml.daml#%C0%EE%CB%C4">
    - <rdfs:comment>
    - <![CDATA[ ]]>
    </rdfs:comment>
    <rdfs:label>李四</rdfs:label>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Thing" />
    </owl:Thing>

```

在Oiled中我们也可以很方便地实现对知识库的管理，如进行知识的增加、修改、删除等操作，在进行以上操作后必须对其进行一致性的检查，在确定没有产生冲突的情况下，可以重新输出新的本体文档，实现对知识库的更新。

4. 2. 4 应急系统知识库结构设计

应急决策支持系统知识库是经过合理组织的关于应急领域的陈述型知识和过程型知识的集合, 知识通过一定的表示, 存储在知识库中。知识库与传统的数据库不同, 知识库不但包含大量的简单事实, 还包括了用于推理及问题求解的条件和规则。

通过对应急决策领域知识内容的分析, 我们知道, 应急决策领域本体知识库主要是介绍概念及概念之间关系的。这些知识大多是以陈述性知识为主。对于过程性知识, 应急事件的处理, 也是描述各个概念、任务及其之间关系的。因此, 在应急系统知识库的设计时, 我们采用以基于关系模型的知识表示为主, 另外还使用到产生式规则、语义网络方法等。对概念、任务的描述以基于关系模型的知识表示为主, 对概念之间关系的描述以语义网络为主, 对事件处理规则的描述以产生式规则为主。我们在设计和实现时, 基于关系模型存储的是事实库, 是各个概念和相关的属性, 每一个概念是一个记录, 记录中有多个属性; 基于语义网络存储的是概念之间的关系, 在该存储表示中, 概念的值及其属性都在事实库中存储, 语义网络仅代表概念相互之间的关系及是什么关系; 基于产生式规则, 存储的是条件和规则表, 代表的是根据一定的任务所进行的推理过程。

1、概念关系的数据结构设计

我们通过本体方法描述了概念之间的关系。为了系统实现, 我们把这些关系图看作有向网。网中的弧代表关系, 弧上的标识代表关系名, 结点代表概念。这样对领域本体的图示化表示就转化为有向网表示。对有向网的存储和操作也就表示了对领域本体的图形存储和操作。下面我们对其有关的概念、存储和操作方法进行论述。

定义 4.11: 图 G 由两个集合 V 和 E 组成, 记为 $G=(V, E)$, 其中 V 是顶点的有限非空集, E 是 V 中顶点偶对有限集, 这些顶点偶对称为边。一般地, $V(G)$ 和 $E(G)$ 分别表示图 G 的顶点集合和边集合。如果 $E(G)$ 为空, 表示图 G 只有顶点而没有边。

定义 4.12: 对于一个图 G , 若边集 $E(G)$ 是有向边的集合, 则图 G 为有向图。此时, 边称为弧, $E(G)$ 为弧的集合。

定义 4.13: 在有向图 G 中, 若存在弧 $\langle V_i, V_j \rangle \in E(G)$, 则 V_i 为弧尾结点, V_j 为弧头结点, V_i 和 V_j 互为邻接点, V_i 是 V_j 的入边邻接点, V_j 为 V_i 的出边邻接点, 以结点 V_i 为尾的出边的个数称为 V_i 的出度, 以 V_j 为入边的个数称为 V_j 的入度。有向图 G 中结点的度是其入度和出度的和。

对于有向图有多种存储表示, 如邻接矩阵表示、邻接表表示和十字链表表示等。对有向图的操作都是以深度优先搜索和广度优先搜索算法为基础的。例如图 4-3 表达的本体模型的一部分, 用有向图的邻接表表示如图 4-9 所示。

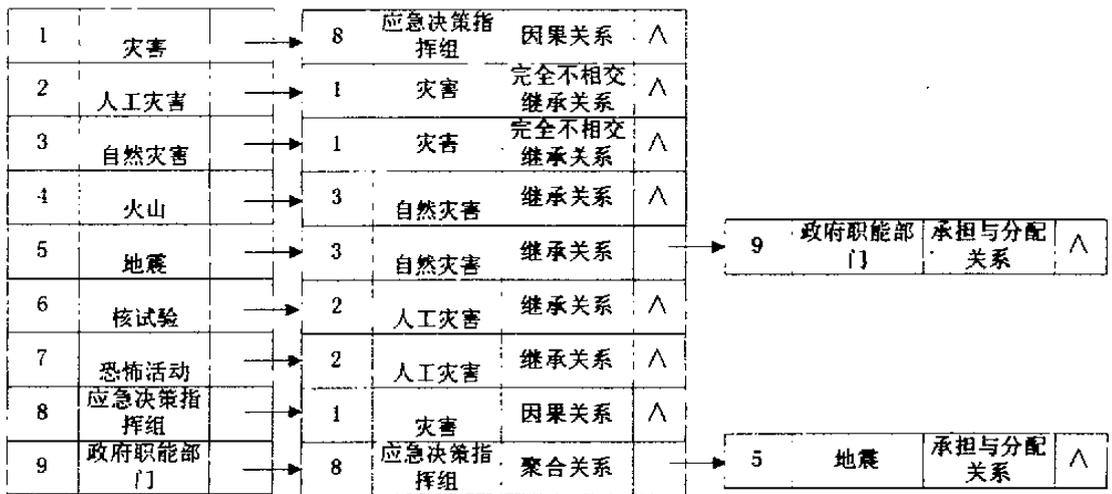


图 4-9 概念关系的一种存储结构

Figure 4-9 a sort of memory structure for concept relations

其中, 头结点表中的每一个记录有三个域: 结点编号、结点数据或指针(指向事实库)和指针域(指向以该结点为出边的另一个邻接点)。

结点有四个域: 结点编号、结点数据或指针(指向事实库)、关系类型域(表示从头结点到该结点之间是什么关系, 如头结点 3, 其邻接的结点是 1, 在结点 1 中的关系类型是“完全不相交继承”, 表示自然灾害是灾害的一个子类)、指针域(指向以相同的弧尾结点的另一个邻接点)。

有了如此存储结构就可以进行深度优先搜索和广度优先搜索, 对于概念语义模型的这样的存储结构, 在知识的搜索或查询中, 可以按照概念名称作为关键字进行搜索, 而且不必要把所有的结点都搜索一遍, 只要搜索到就可以退出。当搜索到所需要的概念时, 就可以进行编辑等操作。至于采取什么存储结构来存储有向图更方便操作, 要根据具体目的、要求等来

决定。

由于系统生成邻接表是动态的，而领域知识本体是固定的，其存储要相对稳定，所以采用邻接矩阵表示较为合适。在邻接矩阵中，边或弧都赋予权值，不同的权值代表不同的关系类型。有向图 G 的 n 阶邻接矩阵 A 定义为：

$$A[i, j] = \begin{cases} W_{ij} & \text{if } \langle V_i, V_j \rangle \in E(G) // \text{不同的 } W_{ij} \text{ 表示不同的关系类型} \\ 0 \text{ 或 } \infty & \text{其它 // 不存在直接关系} \end{cases}$$

此时，形式化描述为：`int adj_network[n][n]`, n 表示 G 的顶点数。

总之，把领域本体的图示化表示，可以转化为有向图的表示，并且存储结构中带有语义。这样便于知识的存储和查询等处理。把关系与事实分开存储，也有助于知识查询中问题与答案的自动生成，即对每一个概念，可以根据它与其它概念之间的关系，该概念的属性生成问题与答案，而只需要一个算法搜索程序。因为，只需要重复搜索有向网即可。

2、事实库的数据结构设计

上述的语义网络存储的是语义关系，结点内容要存储到事实库中。这里的事实库是指按一定方式组织起来的应急领域概念知识库，其中主要存放的是概念及其属性值。我们在设计事实库时，定义如下的数据项，这些数据项都可以用 XML 标注，使得在不需要推理的情况下可直接在 Web 上使用。事实库的数据结构如图 4-10 所示。

概念	属性1	属性2	属性n
----	-----	-----	-------	-----

图 4-10 事实库的数据结构图

Figure 4-10 the structure of concept database

3、规则库的数据结构设计

领域本体包含了与问题求解方法无关的静态知识，我们也可以从推理和问题解决的角度对领域知识进行描述，使得领域知识的描述与其使用方式紧密地结合起来。也就是说，把描述特定任务和问题解决方法的知识表达，与领域知识的表达结合起来，但要分层次进行。这样便于任务和方法的共享。

对于应急系统知识库，由于它是用于辅助决策人员制定决策的，并具

有一定的智能化。但不论是任务的完成还是问题的求解，其推理步骤都是按照一定的条件或规则来进行的。在领域知识库中存储的都是概念及概念之间的关系。所以推理是根据一定的条件、规则和事实，得出结论或结果的过程。因此，条件、规则和事实要按照一定的格式来描述。一般地，条件、规则和事实的形式化描述如下：

规则框架包括两部分：条件表和规则表。

条件表可以定义为一个多元结构，如一个四元组：

CONDITION (ID, OBJECT, VALUE, RULE_ID)

其中，ID：表示条件号，唯一标识一个条件；

OBJECT：条件的主体，如“是否发生某灾害”等；主体由属性 VALUE 来确定的；

VALUE：主体的属性，RULE_ID 为对应的规则号。

规则表可以也定义为一个多元组，如一个四元组：

RULE (RULE_ID, CONDITIONNUM, CONJ, ACTION)

其中，RULE_ID：表示规则号，唯一标识一条规则；

CONDITIONNUM：表示该规则包含的条件数；

CONJ：表示条件之间的关系，由 AND、OR 和 NOT 等组成，它是一个逻辑表达式；

ACTION：表示满足逻辑表达式后要进行的操作、活动或动作，即规则的结论部分；

关于规则库数据结构设计，我们可以根据上面描述的方法，建立条件表和规则表，形成规则库。

需要说明的是：上述结构都可以形式化表示。我们仅给出了图形化的表示，在实现时，可以很方便地用 C 或 C++ 及 PASCAL 等结构化语言表达。

第五章 基于本体的应急决策知识查询

5.1 本体在信息检索中的应用

自上世纪90年代以来,随着计算机在人类社会和生活的各个方面的广泛应用,数字信息的共享和利用已不再是研究人员的专利,大众对数字信息的需求也越来越强烈,众多的信息系统被开发出来,为用户提供其感兴趣的内容。数字信息一方面为人们的日常工作和生活带来了帮助,另一方面,大量的信息又使人们不知所措。如何组织和提供信息就成为信息系统要解决的关键问题。

目前,信息检索技术可分为3类^[32]:全文检索(Text retrieval)、数据检索(Data retrieval)和知识检索(Knowledge retrieval)。

全文检索的特点是把用户的查询请求和全文中的每一个词进行比较,不考虑查询请求与文件语义上的匹配,这种方式虽然可以保证查全率,但是查准率却大大地降低了。

数据检索的特点是查询要求和信息系统中的数据都遵循一定的格式,具有一定的结构,允许对特定的字段检索(例如:作者=“王刚”)。数据检索需要有标识字段的方法。数据检索的性能取决于所使用的标识字段的方法和用户对这种方法的理理解,因此具有很大的局限性。数据检索支持语义匹配的能力也较差。

知识检索强调的是基于知识的、语义上的匹配,因此在查准率和查全率上有更好的保证。目前知识检索是信息检索研究的重点,特别是面向Web信息的知识检索。常规的直接基于关键词的信息检索技术已不能满足用户在语义上和知识上的需求,寻找新的方法也就成为目前研究的热点。本体具有良好的概念层次结构和对逻辑推理的支持,因而在信息检索,特别是在基于知识的检索中得到了广泛的应用^[33]。基于本体的信息检索的基本

第五章 基于本体的应急决策知识查询

5.1 本体在信息检索中的应用

自上世纪90年代以来,随着计算机在人类社会和生活的各个方面的广泛应用,数字信息的共享和利用已不再是研究人员的专利,大众对数字信息的需求也越来越强烈,众多的信息系统被开发出来,为用户提供其感兴趣的内容。数字信息一方面为人们的日常工作和生活带来了帮助,另一方面,大量的信息又使人们不知所措。如何组织和提供信息就成为信息系统要解决的关键问题。

目前,信息检索技术可分为3类^[32]:全文检索(Text retrieval)、数据检索(Data retrieval)和知识检索(Knowledge retrieval)。

全文检索的特点是把用户的查询请求和全文中的每一个词进行比较,不考虑查询请求与文件语义上的匹配,这种方式虽然可以保证查全率,但是查准率却大大地降低了。

数据检索的特点是查询要求和信息系统中的数据都遵循一定的格式,具有一定的结构,允许对特定的字段检索(例如:作者=“王刚”)。数据检索需要有标识字段的方法。数据检索的性能取决于所使用的标识字段的方法和用户对这种方法的理理解,因此具有很大的局限性。数据检索支持语义匹配的能力也较差。

知识检索强调的是基于知识的、语义上的匹配,因此在查准率和查全率上有更好的保证。目前知识检索是信息检索研究的重点,特别是面向Web信息的知识检索。常规的直接基于关键词的信息检索技术已不能满足用户在语义上和知识上的需求,寻找新的方法也就成为目前研究的热点。本体具有良好的概念层次结构和对逻辑推理的支持,因而在信息检索,特别是在基于知识的检索中得到了广泛的应用^[33]。基于本体的信息检索的基本

设计思想可以总结如下:

(1) 在领域专家的帮助下,建立相关领域的本体。

(2) 收集信息源中的数据,并参照已建立的本体,把收集来的数据按规定的格式存储在元数据库(关系数据库、知识库等)中。

(3) 对用户检索界面获取的查询请求,查询转换器按照本体把查询请求转换成规定的格式,在本体的帮助下从元数据库中匹配出符合条件的数据集。

(4) 检索的结果经过定制处理后,返回给用户。

需要说明的是,如果检索系统不需要太强的推理能力,本体可用概念图的形式表示并存储,数据可以保存在一般的关系数据库中,采用图的匹配技术来完成信息检索。如果要求比较强的推理能力,一般需要用一种描述语言(如:Loom, Ontolingua, OWL等)表示本体,数据保存在知识库中,采用描述语言的逻辑推理能力来完成信息检索。由于本体具有能通过概念之间的关系来表达概念语义的能力,所以能够提高检索的查全率和查准率。

目前本体应用在信息检索中的著名项目包括 (Onto)² Agent、Ontobroker和SKC (Scalable Knowledge Composition)。这3个项目也分别代表了3个方向。(Onto)² Agent的目的是为了帮助用户检索到所需要的WWW上已有的本体,主要采用了参照本体。参照本体是以WWW上已有的本体为对象建立起来的Ontology,它保存有各类Ontology的元数据。Ontobroker面向的是WWW上的网页资源,目的是为用户检索到所需要的网页,这些网页含有用户所关心的内容^[34]。SKC是一个正在进行的项目,其目标是解决信息系统语义异构的问题,实现异构的自治系统之间的互操作。该项目希望通过在本体上建立一个代数系统,用这个代数系统来实现各本体之间的互操作,从而实现异构系统之间的互操作。

5. 2 应急决策知识查询系统设计

Internet技术的发展为分布在不同地域、不同领域的决策者提供了方便快捷的信息获取手段。然而要从庞大的分布式知识库中检索信息,特别

设计思想可以总结如下:

(1) 在领域专家的帮助下,建立相关领域的本体。

(2) 收集信息源中的数据,并参照已建立的本体,把收集来的数据按规定的格式存储在元数据库(关系数据库、知识库等)中。

(3) 对用户检索界面获取的查询请求,查询转换器按照本体把查询请求转换成规定的格式,在本体的帮助下从元数据库中匹配出符合条件的数据集。

(4) 检索的结果经过定制处理后,返回给用户。

需要说明的是,如果检索系统不需要太强的推理能力,本体可用概念图的形式表示并存储,数据可以保存在一般的关系数据库中,采用图的匹配技术来完成信息检索。如果要求比较强的推理能力,一般需要用一种描述语言(如:Loom, Ontolingua, OWL等)表示本体,数据保存在知识库中,采用描述语言的逻辑推理能力来完成信息检索。由于本体具有能通过概念之间的关系来表达概念语义的能力,所以能够提高检索的查全率和查准率。

目前本体应用在信息检索中的著名项目包括 (Onto)² Agent、Ontobroker和SKC (Scalable Knowledge Composition)。这3个项目也分别代表了3个方向。(Onto)² Agent的目的是为了帮助用户检索到所需要的WWW上已有的本体,主要采用了参照本体。参照本体是以WWW上已有的本体为对象建立起来的Ontology,它保存有各类Ontology的元数据。Ontobroker面向的是WWW上的网页资源,目的是为用户检索到所需要的网页,这些网页含有用户所关心的内容^[34]。SKC是一个正在进行的项目,其目标是解决信息系统语义异构的问题,实现异构的自治系统之间的互操作。该项目希望通过在本体上建立一个代数系统,用这个代数系统来实现各本体之间的互操作,从而实现异构系统之间的互操作。

5. 2 应急决策知识查询系统设计

Internet技术的发展为分布在不同地域、不同领域的决策者提供了方便快捷的信息获取手段。然而要从庞大的分布式知识库中检索信息,特别

是在应急决策支持系统中，需要快速准确的信息检索方法的协助。目前大多数决策支持系统在信息检索的方法上一般都存在信息导引能力差的问题，不能快速准确地为用户提供其所关心的信息。为解决这个问题，我们将本体论应用于本系统，构造基于本体的智能查询系统，以满足决策者对信息查全率和查准率的要求。

5. 2. 1 查询信息源数据的规范表示

要对具有分布性的网上信息进行查询，则必须具有一个统一和规范化的描述手段对网上信息进行一些约束。根据上节中基于本体的信息检索的基本设计思想我们可知，在建立领域本体的基础上，下一步可以按照已建立的本体对信息源数据进行规范化，并将其按规定的格式存储在元数据库中。然而如何使用本体中定义的类、属性、实例来描述广泛分布在网上的大量不规则信息仍是一个值得探讨的问题。一般的做法是通过加入标注来对网上信息加以描述，就像使用HTML、XML这样的置标语言对已有数据进行标注一样。

例如，可以对现有某所医院的Web页面使用已经定义的本体对其中的“医务人员”进行标注。标注可以手工完成，在理解了已有Web页面内容之后，参照相关本体对所关心的内容用OWL语言重新进行陈述。在HTML语言的<meta>元素中指明这是用OWL标注过的页面，而且标注生成的内容嵌入一个自定义标识<annotatedByOWL></annotatedByOWL>中。这样做的目的是方便页面内容收集工具识别出用OWL标注过的页面以及获取标注内容。

5. 2. 2 系统结构框架

在上述讨论和分析的基础上，以下给出了应急决策知识查询的系统结构图，如图5-1所示。

是在应急决策支持系统中，需要快速准确的信息检索方法的协助。目前大多数决策支持系统在信息检索的方法上一般都存在信息导引能力差的问题，不能快速准确地为用户提供其所关心的信息。为解决这个问题，我们将本体论应用于本系统，构造基于本体的智能查询系统，以满足决策者对信息查全率和查准率的要求。

5. 2. 1 查询信息源数据的规范表示

要对具有分布性的网上信息进行查询，则必须具有一个统一和规范化的描述手段对网上信息进行一些约束。根据上节中基于本体的信息检索的基本设计思想我们可知，在建立领域本体的基础上，下一步可以按照已建立的本体对信息源数据进行规范化，并将其按规定的格式存储在元数据库中。然而如何使用本体中定义的类、属性、实例来描述广泛分布在网上的大量不规则信息仍是一个值得探讨的问题。一般的做法是通过加入标注来对网上信息加以描述，就像使用HTML、XML这样的置标语言对已有数据进行标注一样。

例如，可以对现有某所医院的Web页面使用已经定义的本体对其中的“医务人员”进行标注。标注可以手工完成，在理解了已有Web页面内容之后，参照相关本体对所关心的内容用OWL语言重新进行陈述。在HTML语言的<meta>元素中指明这是用OWL标注过的页面，而且标注生成的内容嵌入一个自定义标识<annotatedByOWL></annotatedByOWL>中。这样做的目的是方便页面内容收集工具识别出用OWL标注过的页面以及获取标注内容。

5. 2. 2 系统结构框架

在上述讨论和分析的基础上，以下给出了应急决策知识查询的系统结构图，如图5-1所示。

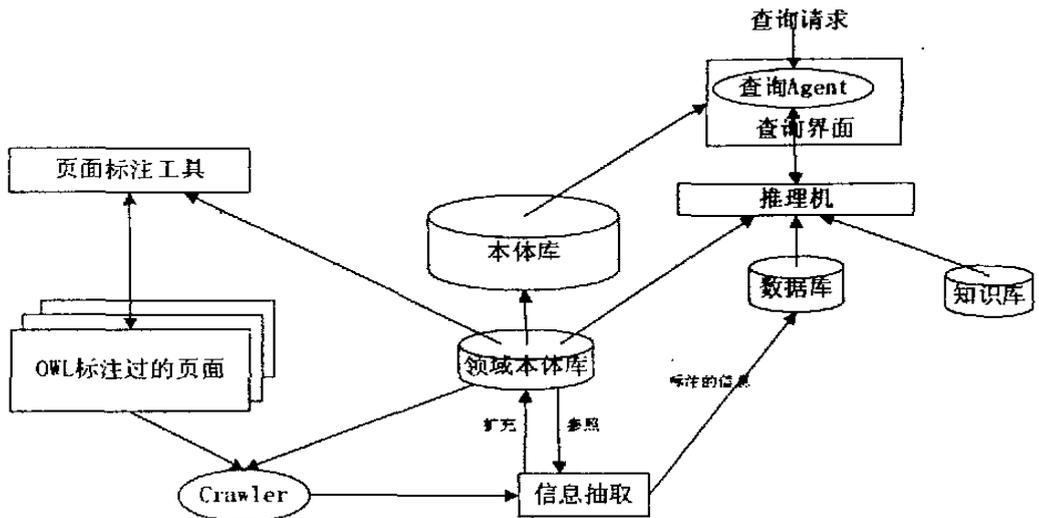


图5-1 应急决策知识查询系统结构图

Figure 5-1 the structure of emergency decision knowledge querying system

在图5-1中，关于领域本体的构建和信息源页面标注已经在本文第四章和第五章5.2.1节中做了介绍。下面将对系统的其他组成部分进行介绍。

Crawler是一个用来收集页面的“爬虫”程序。它与传统搜索引擎中所使用的页面收集器构造基本一致，只是它仅收集用OWL标注过的页面。它通过查看页面meta元素的http-equiv属性值是否为“OWL”来判断该页面是否经过OWL标注。若是，则从其中截取预定义标识<annotatedByOWL>所包含的文档，并把它传递给信息抽取模块。

信息抽取模块的功能有两个：一是参照已有的领域本体从Crawler所收集的标注文档中提取出页面中所标注的信息。这些信息都以二元谓词的形式表示：(predicate #subject #object)。这种比较规范的格式可以方便我们利用已有的成熟技术（如关系数据库）来高效地存储、查询、维护这些数量巨大且动态变化的信息。信息抽取模块的另一个功能是把从页面中发现的新的本体提取出来，经过相应处理后存入领域本体库中，对已有领域本体库进行扩充，同时也供以后的信息抽取参照之用。

本体库是各种本体的存储模块，收集不同的本体，包括一些顶层的本体和诸多领域的本体。它为查询Agent分析查询信息属于哪个领域，从而缩小查询范围，并对决策者查询信息进行规范提供帮助。

数据库用来存储从网页上所收集的大量用OWL语言标注过的信息。这些信息是推理机用来处理查询请求所需的原始信息。

知识库中存储了大量的常识性知识，它们是推理的前提。这些常识性知识由人们手工输入，并采用二元谓词的形式来表示，可以被看作是一些永真断言。为了获得更强的推理能力，知识库中需要存储大量这样的断言。

推理机接收查询Agent提交的查询语句，根据由数据库中检索得到的信息、知识库中的推理前提、以及领域本体库中相关的本体进行推理，得出满足用户需要结果。由于知识库、数据库和本体库中的数据均以谓词逻辑的形式表达，因此我们可以利用成熟的基于谓词逻辑的推理机制进行推理，这是一项相当成熟的技术。

查询模块中的查询Agent是用户进行查询的接口。在此模块中利用本体库来辅助分析查询信息属于哪个领域，并对决策者查询信息进行规范。从两个角度入手，增强了信息的导引能力。为决策者准确地查询和搜索相关决策信息提供了保证。

5. 2. 3 查询 Agent 的设计思想

首先，我们将术语的每种含义和它相应的领域建立对应关系。

定义 5.1: 函数 $term(O_i)$ 从领域 D_i 对应的本体 O_i 中求出该领域的术语集(包括同义词);

定义 5.2: 函数 $keydef(O_i, keyword)$ 从本体 O_i 中求出关键字 $keyword$ 的定义;

设 O_1, O_2, \dots, O_n 分别是领域 D_1, D_2, \dots, D_n 的本体, 术语集 $T_i = term(O_i)$, 其中 $(0 \leq i \leq n)$. $KS = \{key_1, key_2, \dots, key_n\}$ 为检索信息 Index 中所给出的关键字。

规则 5.1: 任一次检索信息中所给出的关键字若不出现在该领域的本体中, 则说明该检索与这一领域无关, 即 $KS \cap T_i = \emptyset \vdash Index \notin D_i$, 这里 $1 \leq i \leq n$. 例如, 检索的关键字 $key =$ “手术器械” 我们在消防本体知识库中没有该关键字的本体定义, 则可以断定该次检索与消防领域无关。

利用这一规则, 我们可以滤掉不相关的领域, 缩小检索范围。并由此得到所有可能与所查询信息相关的领域 DS_1, DS_2, \dots, DS_k , 其中 $KS \cap TS_j \neq \emptyset, s_1 \leq s_j \leq s_k$.

知识库中存储了大量的常识性知识，它们是推理的前提。这些常识性知识由人们手工输入，并采用二元谓词的形式来表示，可以被看作是一些永真断言。为了获得更强的推理能力，知识库中需要存储大量这样的断言。

推理机接收查询Agent提交的查询语句，根据由数据库中检索得到的信息、知识库中的推理前提、以及领域本体库中相关的本体进行推理，得出满足用户需要结果。由于知识库、数据库和本体库中的数据均以谓词逻辑的形式表达，因此我们可以利用成熟的基于谓词逻辑的推理机制进行推理，这是一项相当成熟的技术。

查询模块中的查询Agent是用户进行查询的接口。在此模块中利用本体库来辅助分析查询信息属于哪个领域，并对决策者查询信息进行规范。从两个角度入手，增强了信息的导引能力。为决策者准确地查询和搜索相关决策信息提供了保证。

5. 2. 3 查询 Agent 的设计思想

首先，我们将术语的每种含义和它相应的领域建立对应关系。

定义 5.1: 函数 $term(O_i)$ 从领域 D_i 对应的本体 O_i 中求出该领域的术语集(包括同义词);

定义 5.2: 函数 $keydef(O_i, keyword)$ 从本体 O_i 中求出关键字 $keyword$ 的定义;

设 O_1, O_2, \dots, O_n 分别是领域 D_1, D_2, \dots, D_n 的本体, 术语集 $T_i = term(O_i)$, 其中 $(0 \leq i \leq n)$. $KS = \{key_1, key_2, \dots, key_n\}$ 为检索信息 Index 中所给出的关键字。

规则 5.1: 任一次检索信息中所给出的关键字若不出现在该领域的本体中, 则说明该检索与这一领域无关, 即 $KS \cap T_i = \emptyset \vdash Index \notin D_i$, 这里 $1 \leq i \leq n$. 例如, 检索的关键字 $key =$ “手术器械” 我们在消防本体知识库中没有该关键字的本体定义, 则可以断定该次检索与消防领域无关。

利用这一规则, 我们可以滤掉不相关的领域, 缩小检索范围。并由此得到所有可能与所查询信息相关的领域 DS_1, DS_2, \dots, DS_k , 其中 $KS \cap TS_j \neq \emptyset, s_1 \leq s_j \leq s_k$.

其次，我们还可以在本体库的协助下，规范用户查询的信息以提高查询效率。这里，我们将 Stefano Borgo 等人的规范面向对象构件的思想应用于对用户查询信息的规范。查询 Agent 与用户交互，根据用户提出的关键字集 $KEY = \{key_1, key_2, \dots, key_n\}$ ，查询 O_1, O_2, \dots, O_n ，从领域 DS_1, DS_2, \dots, DS_k (其中 $\exists key (key \in KEY) \text{ and } (key \in T_d), j = 1, 2, \dots, sk$) 对应的 OS_1, OS_2, \dots, OS_k 中找出关键字 key_i 的定义，然后分别将领域以及在该领域下的定义罗列给用户。用户可以依此进行选择。这样在检索过程中系统能知道用户所关心的领域，为检索提供了更加精确的信息。

5. 2. 4 查询界面中查询语句的构造

按照 5.1 节中基于本体的信息检索的基本设计思想可知，我们要从查询界面中获取用户的查询请求，并把它按照本体转换成规定的格式，然后在本体的帮助下由推理机从元数据库中匹配出符合条件的数据集合，经过定制处理后返回给用户。

在此我们使用“应急医疗队的医生”这一查询请求来举例。在查询系统的知识库中对“应急医疗队的医生”有这样解释：如果一个医生被征募到应急方案中，那么他（她）就是应急医疗队的医生。在此用变量 ? doc 来表示待考察的对象，用 eddo 来表示应急决策领域本体的名称空间，则它的类型应该是 eddo:Doctor，具体表述为：`rdf:type(?doc, eddo:Doctor)`。

若 ? doc 被征募到应急方案中，那么就可知他（她）是应急医疗队的医生。在此，用变量 ? enlist 来表示征募。? doc 被 ? enlist 可这样表示：`eddo:enlist(?doc, ?enlist)`。

这样，我们可以在知识库的帮助下得到完整的查询语句：

```
rdf:type(?doc, eddo:Doctor) ^ eddo:enlist(?doc, ?enlist)
```

上述查询语句也可以用 KIF 格式重写如下：

```
And(Type ?doc, eddo:Doctor)
```

```
(PropertyValue eddo:enlist?doc ?enlist)
```

这样做的好处显而易见：仅用 ASCII 字符表达上述复杂的谓词合式公

其次，我们还可以在本体库的协助下，规范用户查询的信息以提高查询效率。这里，我们将 Stefano Borgo 等人的规范面向对象构件的思想应用于对用户查询信息的规范。查询 Agent 与用户交互，根据用户提出的关键字集 $KEY = \{key_1, key_2, \dots, key_n\}$ ，查询 O_1, O_2, \dots, O_n ，从领域 DS_1, DS_2, \dots, DS_k （其中 $\exists key (key \in KEY) \text{ and } (key \in T_j), j = 1, 2, \dots, sk$ ）对应的 OS_1, OS_2, \dots, OS_k 中找出关键字 key_i 的定义，然后分别将领域以及在该领域下的定义罗列给用户。用户可以依此进行选择。这样在检索过程中系统能知道用户所关心的领域，为检索提供了更加精确的信息。

5. 2. 4 查询界面中查询语句的构造

按照 5.1 节中基于本体的信息检索的基本设计思想可知，我们要从查询界面中获取用户的查询请求，并把它按照本体转换成规定的格式，然后在本体的帮助下由推理机从元数据库中匹配出符合条件的数据集，经过定制处理后返回给用户。

在此我们使用“应急医疗队的医生”这一查询请求来举例。在查询系统的知识库中对“应急医疗队的医生”有这样解释：如果一个医生被征募到应急方案中，那么他（她）就是应急医疗队的医生。在此用变量 ? doc 来表示待考察的对象，用 eddo 来表示应急决策领域本体的名称空间，则它的类型应该是 eddo:Doctor，具体表述为：`rdf:type(?doc, eddo:Doctor)`。

若 ? doc 被征募到应急方案中，那么就可知他（她）是应急医疗队的医生。在此，用变量 ? enlist 来表示征募。? doc 被 ? enlist 可这样表示：`eddo:enlist(?doc, ?enlist)`。

这样，我们可以在知识库的帮助下得到完整的查询语句：

```
rdf:type(?doc, eddo:Doctor) ^ eddo:enlist(?doc, ?enlist)
```

上述查询语句也可以用 KIF 格式重写如下：

```
And(Type ?doc, eddo:Doctor)
```

```
(PropertyValue eddo:enlist?doc ?enlist)
```

这样做的好处显而易见：仅用 ASCII 字符表达上述复杂的谓词合式公

式,将有助于查询表达式的传递和处理。在把查询语句转化为一阶谓词逻辑表达式后,我们就可以利用传统的基于一阶谓词逻辑的自动推理机完成推理和判断工作。

5.3 与传统基于关键词匹配检索方法的比较

如 5.1 节所述,目前 Web 上的搜索引擎一般仅能提供基于关键词的检索。由于关键词间仅有简单的“与”“或”关系,而缺少相互之间的语义关系,再加上一词多义和下位词的原因,使得查准率和召回率均较低。不仅如此,基于关键词的检索方法只能进行简单的匹配操作来发现已有的相关信息,而不能在此基础上进行推理,进而得出更多有意义的结果,也就是说,它不能提供更高级的智能查询功能。下面我们通过一个查询实例来分析本文所设计的应急决策知识查询系统是如何解决上述问题的。

在此,仍以“应急医疗队的医生”为例进行查询,我们发现这样的查询很难由基于关键词的搜索引擎来完成。本文所设计的应急决策知识查询系统将解决如下问题:

1、解决同义词问题:在上述查询语句中,“应急医疗队”一词是指应急决策领域中“医务人员”这一对象,在该领域中“应急医疗队”是“医务人员”的一个同义词。如在所查询的页面上只有“医务人员”的字样,那么用“应急医疗队”作为关键词进行匹配的搜索方式就不能找出满意的结果。

本系统利用所建立的本体对概念的明确定义来解决这一问题。通过为用户提供所查询领域的规范化查询关键词来帮助用户明确自己所提交的“应急医疗队”指的是应急决策领域本体中“医务人员”这一概念。

2、解决下位词问题:在这个例子中,就是应该如何理解“医生”这样一个概念的问题。由应急决策领域本体中的定义可知,“内科医生”和“骨科医生”等都是“医生”的下位词(内科医生和骨科医生都是医生的子类)。这样图 5-1 中的推理模块就可以参照应急决策领域本体定义从数据库中检索符合用户要求的有关“内科医生”和“骨科医生”等的记录。

3、解决对象之间的语义关系问题:本查询系统在接收到查询语句“应

式,将有助于查询表达式的传递和处理。在把查询语句转化为一阶谓词逻辑表达式后,我们就可以利用传统的基于一阶谓词逻辑的自动推理机完成推理和判断工作。

5.3 与传统基于关键词匹配检索方法的比较

如 5.1 节所述,目前 Web 上的搜索引擎一般仅能提供基于关键词的检索。由于关键词间仅有简单的“与”“或”关系,而缺少相互之间的语义关系,再加上一词多义和下位词的原因,使得查准率和召回率均较低。不仅如此,基于关键词的检索方法只能进行简单的匹配操作来发现已有的相关信息,而不能在此基础上进行推理,进而得出更多有意义的结果,也就是说,它不能提供更高级的智能查询功能。下面我们通过一个查询实例来分析本文所设计的应急决策知识查询系统是如何解决上述问题的。

在此,仍以“应急医疗队的医生”为例进行查询,我们发现这样的查询很难由基于关键词的搜索引擎来完成。本文所设计的应急决策知识查询系统将解决如下问题:

1、解决同义词问题:在上述查询语句中,“应急医疗队”一词是指应急决策领域中“医务人员”这一对象,在该领域中“应急医疗队”是“医务人员”的一个同义词。如在所查询的页面上只有“医务人员”的字样,那么用“应急医疗队”作为关键词进行匹配的搜索方式就不能找出满意的结果。

本系统利用所建立的本体对概念的明确定义来解决这一问题。通过为用户提供所查询领域的规范化查询关键词来帮助用户明确自己所提交的“应急医疗队”指的是应急决策领域本体中“医务人员”这一概念。

2、解决下位词问题:在这个例子中,就是应该如何理解“医生”这样一个概念的问题。由应急决策领域本体中的定义可知,“内科医生”和“骨科医生”等都是“医生”的下位词(内科医生和骨科医生都是医生的子类)。这样图 5-1 中的推理模块就可以参照应急决策领域本体定义从数据库中检索符合用户要求的有关“内科医生”和“骨科医生”等的记录。

3、解决对象之间的语义关系问题:本查询系统在接收到查询语句“应

急医疗队的医生”后可以根据知识库中预先存入的作为推理前提的断言得出结果：“被征募到应急方案中的医生“(形式化的表示方法为 `enlist(eddo:Doctor, eddo: emergency scheme)`)。这一结果还可以反馈给用户,以得到其确认,若用户并不满意系统的这一解释,则可以修改自己的查询要求,并再次提交。

通过以上的分析,可以看到基于本体的知识查询可以解决传统搜索引擎难以解决的问题,如同义词、下位词以及对象之间的语义关系等,更好地满足决策者对信息查全和查准的需求。

第六章 总结与展望

本体建模的实质就是建立某一领域的信息模型。一个构造良好且能够实用的本体需要领域知识专家、人工智能专家、数据库专家、Web 技术人员甚至是语言学家的共同参与才能取得成功。本文在对知识表示及国内外有关本体库研究的基础上,提出适合应急决策领域本体建模的方法 DKOED,并利用该方法对应急决策领域本体模型进行了构造。同时对基于本体的智能查询系统进行了设计,为应急决策知识查询提供支持。

6.1 课题内容小结

本课题研究中,从始至终都体现着“全局着眼、局部实现”的思想,在内容上可以分成三大块:一是对国防动员应急决策支持系统总体框架的设计,二是应急决策领域本体建模,三是基于本体的应急决策知识查询。

在以上三部分内容中,第一部分内容首先从全局着眼,对应急决策支持系统总体框架进行了设计,在系统中采用了四库模型,并引入数据挖掘和数据仓库技术,在知识库中加入推理机,支持在知识库上的自动推理技术。同时也考虑了一些辅助系统的设计,如仿真模块、地理信息系统 GIS 的设计等。第二部分内容,也就是应急决策领域本体建模是本课题研究的重点内容。以领域知识建模的本体论基础理论为指导,作者提出了一种适用于应急决策领域本体建模方法(Domain Knowledge Ontology for Emergency Decision,DKOED)。DKOED 方法将语义网络的知识表示方法与面向对象逻辑建模、面向数据的语义建模方法有机地进行了结合,建立了应对突发灾害的应急决策示例本体模型,并利用本体编辑器 OilEd 生成该模型的基于 OWL+RDF+XML 的形式化文档,为基于本体的应用研究奠定了基础。第三部分内容介绍了应急决策领域知识查询,对查询信息源数据的规范、查询系统的框架及查询模块进行了设计,并构造了查询语句,最后与传统基于关键词匹配的搜索引擎进行了对比分析。

6.2 课题评价与展望

尽管论文中对本体建模及知识查询的分析和设计还是初步的,甚至说是尝试性的,系统的实现也是示例性的,但对于一个大型应用系统的前期开发而言,本课题能够根据所开发的课题领域做出比较系统的研究分析,并结合一个子系统的设计实例研究归纳了一些本体知识库建模的方法,这对日后进一步进行该课题的研究开发工作,还是有一定启迪作用的。比如,参照应急决策领域本体建模方法,继续对领域中相关概念(术语)、概念之间的关系和公理进行分析,不断对本体知识库内容进行扩充,以形成一个真正切实可行的为决策人员提供支持的本体知识库。同时,对应急决策支持系统的其它部分进行设计开发,以形成一个完整可用的决策支持系统。

可以看到对于此题目的研究,必须面对一些哲学界和认知学界有争议的问题。本体工程是知识工程的发展和完善,它必然要涉及计算机科学、语言学、认知学、逻辑学以及哲学这几门学科。当我们构造一个本体时,不可避免地会遇到上述各学科研究领域中的一些问题,有的问题甚至在本学科中现在还在探讨中,希望本文中本体知识表示及建模的研究对 AI 及应急决策是一种有益的探讨与尝试。

参考文献

- [1] 于晓东, 基于 Ontology 知识库系统建模与应用研究, [学位论文], 北京, 国家图书馆, 2003
- [2] 高颖, 领域本体的形式分析, [学位论文], 北京, 国家图书馆, 2003
- [3] 邓志鸿, 唐世渭, 张铭等, Ontology 研究综述, 北京大学学报(自然科学版), 2002, 38(5), 730-738
- [4] <http://www.cyc.com> 2004.5.4
- [5] CycL. <http://www.cyc.com/tech.html#cycl> 2004.6.4
- [6] 张宇翔, NKI 本体理论中一些基本关系的研究, [学位论文], 云南师范大学, 2002
- [7] <http://www.keenage.com> 2004.6.5
- [8] 曾庆田, 曹存根, 基于本体的数学知识获取与知识继承机制研究, 微电子学与计算机, 2003, 9, 19-27
- [9] 王珏, 袁小红, 关于知识表示的讨论, 计算机学报, 1995, 18(3), 212-224
- [10] 王万森, 人工智能原理及其应用, 北京, 电子工业出版社, 2000, 23-27
- [11] 钟凌燕, 高济, 一种基于 UML 的本体论建模方法, 计算机工程, 2003, 29(2), 38-39
- [12] 宋炜, 张铭, 语义网简明教程, 北京, 高等教育出版社, 2004, 196-202
- [13] B. Chandrasekaran, J. R. Josephson, V. R. Benjamins, What Are Ontologies, and Why Do We Need Them, IEEE Intelligent Systems, 1999, 14(1), 20-26
- [14] Wordnet. <http://www.cogsci.princeton.edu/~wn/>
- [15] Mikrokmos. <http://crl.nmsu.edu/Research/Projects/mikro/>

- [16] SENSUS.<http://www.isi.edu/natural-language/resources/sensus.html>
- [17] 汪鹏, Ontology 知识表示的艺术, 计算机教育, 2004, 7, 41-43
- [18] Studer R, Benjamins V R, Fensel D, Knowledge Engineering, Principles and Methods, Data and Knowledge Engineering, 1998, 25(122), 161-197
- [19] Guarino N, Semantic Matching : Formal Ontological Distinctions for Information Organization, Extraction, and Integration, In: Paziienza M T, eds. Information Extraction : A Multidisciplinary Approach to an Emerging Information Technology, Springer Verlag, 1997, 139-170
- [20] Perez A G, Benjamins V R, Overview of Knowledge Sharing and Reuse Components: Ontologies and Problem Solving Methods, In: Stockholm V R, Benjamins B, Chandrasekaran A, eds. Proceedings of the IJCAI-99 workshop on Ontologies and Problem Solving Methods (KRR5) 1999, 1-15
- [21] 熊垚飞, 基于 Ontology 的 Web 信息集成研究, [学位论文], 北京, 国家图书馆, 2003
- [22] <http://www.w3.org/TR/owl-features/2004.7>
- [23] Smith M K, Welty C, McGuinness D, OWL Web Ontology Language Guide. <http://www.w3.org/tr/2003/WD-owl-guide-20030331>
- [24] <http://www.sts.tu-harburg.de/~r.f.moeller/racer/> 2004.10
- [25] <http://oiled.man.ac.uk/2004.10>
- [26] 杨秋芬, 陈跃新, Ontology 方法学综述, 计算机应用研究, 2002, 4, 5-7
- [27] 张荣梅, 智能决策支持系统研究开发及应用, 北京, 冶金工业出版社, 2003
- [28] 施益强, 伍良, 陈玲, 海上溢油事故应急反应系统的框架研究, 灾害学, 2002, 12, 87-91
- [29] 于跃海, 何建敏, 郑瑞强等, 基于案例推理的 ICU 急诊诊断方案生成系统, 东南大学学报 (自然科学版), 2001, 3, 6-9
- [30] Wand Y, R. Weber, Mario Bunge's Ontology as a formal foundation for information systems concepts, In: P. Weingartner, G.J.W. Dorn (Eds.),

Studies on Mario Bunge's Treatise, Rödopi, Atlanta, 1990, 123-149

- [31] 徐振宇, 基于本体的 Web 数据语义信息的表示与处理方法研究, [学位论文], 北京, 国家图书馆, 2002
- [32] S.Lawren, C.L.Giles, Searching the World Wide Web, *Science*, 1998, 280 (5360), 98-100
- [33] Uschold M, Gruninger M, ONTOLOGIES、Principles、Methods and applications, *Knowledge Engineering Review*, 1996, 11(2), 93-155
- [34] Guarino N, Masolo C, Vetere G, OntoSeek : Content-Based Access to the Web, *IEEE Intelligent Systems*, 1999, 14(3), 70-80
- [35] Adam M. Gadowski, Claudio Balducelli, Sandro Bologna , etc. Integrated Parallel Bottom-up and Top-down Approach to the Development of Agent-based Intelligent DSSs for Emergency Management. TIEMS98: Disaster and Emergency Management: International Challenges for the Next Decade. The Fifth Annual Conference of The International Emergency Management Society, Washington, D.C., May 19-22, 1998
- [36] Adam M. Gadowski, Sandro Bologna, Giovanni DiCostanzo, etc. An Approach to the Intelligent Decision Advisor (IDA) for Emergency Managers, TIEMS'99, The Sixth Annual Conference of The International Emergency Management Society, Delft, Netherlands, June 8-11, 1999, 1-14

致 谢

光阴荏苒，三年的研究生生涯一晃而过。在三年的学习生活中我得到了许多老师和同学们的无私关怀和帮助，使我终身难忘。

我衷心地感谢导师谢红薇教授，三年来她给予了我无微不至的关怀和孜孜不倦的教诲！谢老师严谨的治学态度、渊博的专业知识、为人宽厚的处世态度、一丝不苟的工作作风给我留下了极为深刻的印象，使我获得了一份宝贵思想财富。

感谢余雪丽教授、王莉老师，在学习和生活中对我的无私帮助和关怀。感谢同课题组的张晓波、李娟莉、胡坤等同学对我论文研究工作的支持，跟他们一起讨论和解决问题的日子是愉快而令人难忘的。同时我也要感谢同实验室的曹禹、王泰翔、王辉、吴广平、李冰清、乔玲玲、岳一领等同学给予我的帮助。

最后，我要特别感谢我的家人以及关心我的亲人和朋友，在我多年的求学路上他们给了我无微不至的关怀，他们的勉励和支持使我得以顺利地完成论文。

作者攻读硕士学位期间所发表的论文

- 1、谢红薇，袁占花，张晓波，余雪丽. 基于本体论的应急系统知识表示的研究. 计算机工程与应用 2005 年第 13 期 刊出.
- 2、谢红薇，张晓波，袁占花，余雪丽. 基于遗传算法求解应急决策系统中的最优路径. 计算机应用 2005 年第 4 期 刊出.