

基于本体的海上交通知识库的构建

蔡永嘉

(大连海事大学信息科学技术学院 大连 116026)

摘 要 随着我国航运事业的迅速发展, 船只的数量与日俱增, 海上的交通也变得越来越拥挤。而我们知道海上的交通具有空间性、时间性、多维性等特点, 如何对海事交通知识进行准确的表达决定了上层应用系统的成败。建立针对海事群体的海事交通知识库, 利用本体通过类、属性、实例、分类法等表达结构知识、支持知识自动分类和实例识别的特点, 对海事交通领域涉及的诸多概念及其关系进行定义, 借助 OWL 等工具对其形式化表达, 并在海事交通知识查询系统中进行简单应用。试验证明, 该海事交通知识库能弥补非海事用户对海上交通空间知识的欠缺, 增强海事交通知识查询服务的通用性和改善用户体验。

关键词 海事交通; 本体; 知识库

中图法分类号 TP393 文献标识码 A

Construct knowledge base for maritime traffic based on ontology

Cai YongJia

(Department of Information Science and Technology, Dalian Maritime University, Dalian, 116026, China)

Abstract With the rapid development of China's marine industry, the number of ships are increasing steadily, and maritime traffic is becoming more and more crowded. While we know maritime traffic has relations with space and time, and its information is multidimensional. How to express maritime traffic knowledge determines the success of the application system of an upper layer. With the Advantages of ontology to express knowledge, to classify and to identify instances, A maritime traffic knowledge based was established for related users. In the ontology, the concepts and their relationships involved in marine traffic were defined and formalized with OWL. The knowledge based was used in a marine traffic knowledge query system for example. Experiments show that the marine traffic knowledge base was able to compensate for the lack of special spatial knowledge non-marine users, enhanced usability for marine traffic query service and improved the user experience.

Key words maritime traffic, ontology, knowledge base

1 引言

不管是陆上的交通还是海上的交通都促进了人类文明的发展, 人类的生产、生活离不开交通。而且 Web 上数据信息量的日益增加, 产生了海事案例信息的重用、共享困难等现象, 而本体能够存

储语义信息, 支持知识的共享和重用。各个部门为海上道路规划的规划、路径分析、信息发布与查询等提供了工具和平台。然而, 非专业的普通用户无法直接利用这些专业工具和平台进行信息查询、分析。除了培训和自我学习外, 构建具备领域知识并具有逻辑推理能力的系统来跨越非专业的用户和专家之间的知识鸿沟是当今专业系统走向大众

化的发展方向^[1]。知识库存储与领域有关的问题求解知识、说明性知识、控制策略知识以及专家经验等。通过这写专业知识的有效表达,使用户能理解知识库中的知识,并利用其作出分析和决策,获得求解问题的解答,并向用户提供清晰的解释。目前,在农业、地理、教育、电子商务、数字博物馆、临床医学等领域纷纷建立知识库,利用本体解决知识的语义层次共享、重用及知识推理等问题,取得了良好应用效果本文在交通领域利用本体和规则方法建立知识库,将交通领域各种实体和关系抽象表达为机器可以识别和共享的语义知识,并运用知识推理器,为非专家用户填补专业知识的空缺,实现了知识和语义层次上的智能化推理和查询^[2]。

2 海上交通知识领域表示

想要建立海上交通知识库的一个核心问题是其知识表示方法,通过知识表示方法既能表达专家知识,又能定义语义解释并进行正确推理^[3]。所谓知识领域表示是为述世界所作的一组约定,是知识的符号化、形式化或模型化。主要的知识表示方法包括谓词逻辑表示法、产生式表示法、框架表示法、语义网络表示法、面向对象表示法等。本文使用本体和规则相结合的方法描述海上交通知识。本体作为显式的概念化规范,具有可共享性,常用于描述共同认可的结构化知识,但存在知识表示不够丰富、推理不够实效等问题。规则侧重于陈述性知识演绎,通过逻辑程序设计实现规则知识系统。规则的参与,能够弥补本体表达能力的不足,提高实际应用的运行效率。

2.1 本体的组成

构造本体的目的是为了捕获海上交通领域的基础设施、行政区划、地标、气象等知识,通过对涉及的各种实体及其关系进行抽象,并以形式化语言进行表达,提供语义层次上的知识,使机器能够识别和推理,并为规则的应用提供语义支持。

根据知识表达的需要以及建模工具的表达能力,使用五元组表达该领域本体: $T\text{-Ontology} = (TC, TP, TI, TR, TA)$ ^[4]。TC(概念)指该领域的具有相同属性的实体或者现象的集合,如表1所示。TP(属性)用来描述概念或者实体的某个侧面,如航线的区域、出发时间、平均时速等。A(实例)是现实世界中各实体在概念空间中的对应物,是概念的特化,如表1中的“大连站”、大连湾区、塔南路、焦作市、河南理工大学等。TR(关系)主要包括基本

语义关系和空间关系。

序号	类别	概念	实例
1	路线	航线的路线	航线 1
2	码头	停靠的站点	大连站
3	航线	行驶的路线	大连到青岛
4	海上区域	各个行驶区域	大连湾区
5	标志物	标识的地方	灯塔
6	天气	行驶的天气	大风
7	行驶状态	航行的情况	空

表1 主要概念

(1)基本语义关系。基本语义关系主要包括属性关系和类属关系以及实例关系。属性关系指某个实例作为其他实例属性的关系,如“大连站”是某航线的停靠站点。类属关系表征概念之间的上下位关系,如“大连站”是“停靠站”的子类。实例关系指某个实例是某概念的具体化,如“大连站”是“停靠站”的实例。

(2)空间关系。依据空间对象实体的几何形态和知识表示的需要,本文将上述的实例抽象为点、线、面这3种类型^[5]。由于空间关系随着对象语境的不同,其语义信息会有所差别,因此,分别定义了海上交通领域中点-线、点-面、线-面、点-点、面-面、线-线等之间的空间关系。

点与线空间关系分为点在线上,点在线周边一定阈值范围内以及点远离线。点与面之间的关系分为点在面内(包括边线)。点在面外部一定阈值范围内,以及点远离面。线与面的关系分为线通过某个面域且在该区域的有断点或出入口,线与面没有交点,两者相交但不连通(如某航线经过某海区但没有停靠的站点)。点与点间的关系包括重叠、两者距离在一定阈值内以及相互远离。面与面用5种关系表达:重合;前者位于后者的面域内;前者包含后者;不相互包含但有共同面域或最近点在一定阈值范围内;相离且最近点在一定阈值之外。线与线之间的关系有3种:相交且连通;不相交;相交但不连通。

上述关系中,阈值的设定可以根据空间尺度的上下文背景进行动态设定^[6]。

TA(公理)通常是一阶谓词逻辑的表达式,代表永真断言,比如实例“大连站”属于概念“站”范畴。可以使用公理来约束信息、证明正确性或者推导新信息。在本体中加入公理则意味着可以表达更为丰富的概念间的关系,描述领域内所有值为真的事实。

2.2 本体描述

本文以 OWL 作为本体描述语言。OWL (OWL Web Ontology Language) 是 W3C 推荐的语义网本体描述语言的标准以 RDF 和 RDFS 为基础^[7], 增强了领域知识的表示与推理能力。通过布尔算子, OWL 可以递归构建复杂的类, 其语言元素还考虑了存在值、任意值和数量值等概念。同时, OWL 能描述属性的传递性、对称性、函数性等特性, 并通过公理声明某两个类等价或不交, 某两个属性等价或互逆, 某两个个体相同或不同。

下面以空间关系描述为例, 说明 OWL 表达的本体。下图 1 描述了大连站和大连湾的几何形态以及他们之间的相邻关系。

```
<rdf:RDF
.....
<owl:Ontology rdf:about="交通知识">
</owl:Ontology>
<owl:Class rdf:ID="GeoGeometry"><!--空间几何形态-->
<owl:Class rdf:about="#Polygon"><!--多边形类型-->
<rdf:subClassOf rdf:resource="#GeoGeometry"/>
</owl:Class>
<owl:Class rdf:ID="Polyline"><!--线类型-->
<rdf:subClassOf rdf:resource="#GeoGeometry"/>
</owl:Class>
<owl:Class rdf:ID="Station"><!--站点-->
<rdf:subClassOf rdf:resource="#Point"/>
</owl:Class>
<owl:Class rdf:ID="Street"><!--街道-->
<rdf:subClassOf rdf:resource="#Polyline"/>
</owl:Class>
<owl:SymmetricProperty rdf:ID="Adjacent-To"><!--具有对称特性的相邻关系-->
<rdf:range rdf:resource="#GeoDescription"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl:ObjectProperty"/>
<owl:inverseOf rdf:resource="#Adjacent-To"/>
<rdf:domain rdf:resource="#GeoDescription"/>
</owl:SymmetricProperty>
<Station rdf:ID="JiaoZuo_Station"><!--焦作火车站与站前街相邻-->
<Adjacent-To>
<Street rdf:ID="ZhanQian_Street"/>
</Adjacent-To>
</Station>
</rdf:RDF>
```

图1 基于 OWL 的本体描述示例

Fig.1 Example for ontology description by OWL

2.3 规则的构建

虽然本体描述能通过纵向类属分类和属性、关系等语义手段进行组织, 但对于属性操作, 如属性规则知识的表达等, OWL 的表达能力远远不够, 而这正是空间关系推理所需用到的属性操作 SWRL (Semantic Web Rule Language) 正是为了弥补 OWL 语言表达不足而发展的, 它是以 OWL DL 与 OWL Lite 以及 Unary/Binary Datalog RuleML (Rule Markup Language, 规则标记语言) 为基础的规则描述语言, 利用高度抽象的语法表达 OWL 本体所叙述的知识, 实现了 Horn-like 规则与 OWL 知识库的结合^[8]。本文采用 SWRL 对空间推理规则进

行形式化表达, 通过空间关系和出行规则的形式化表达, 实现了交通知识的计算机可读和空间关系的自动处理。

3 海上交通知识的推理

本文建立的知识库充分发挥了本体和规则两者的优势: 一方面, 本体提供复杂的概念定义; 另一方面, 规则表达基于多元关系的陈述性知识。基于本体和规则知识库的推理的任务主要分两个部分, 即基于本体推理和基于规则的推理。

3.1 基于本体的推理

本体推理主要用于完成以下任务。第一, 本体满足性测试及一致性测试, 避免产生形式和语义的矛盾。第二, 包含性测试, 划分概念层。第三, 实例检测, 为个体确定相应的概念描述, 并对属于某概念的个体进行分类。第四, 辅助推理功能。如对概念名字和个体名字的检索, 与角色相关的个体对的检索及角色层次的检索等。

3.2 基于规则的推理

基于规则的推理可借助 JESS (Java Expert System Shell) 来完成。该工具利用 Lisp-Like 语法来描述规则与事实, 并且可以通过规则来推论事实。

4 应用分析

本文以海上交通出行经常用到的航行查询和航线线路查询为例, 说明海上交通知识的应用场景及模式。

基于海上参照的查找基本流程如下。判断查找对象类型。遍历海上交通知识库中的概念及实例, 若查找对象是一个实例, 则将其作为结果直接返回; 若查询到多个实例, 则进入下一步条件约束; 若查找对象是概念, 则记录该概念及其所有派生概念包含的所有实例再进行下一步约束。

航线路查询依赖于前期建好的交通知识库, 通过出行方式选择, 粗略判断出行起点和终点的类型, 然后再建立该类型地名与用户输入的出发地和目的地之间的关系, 进而找出真正的出发地和目的地。然后, 根据不同的地名类型和出行策略, 选择不同的演算方法, 求得出行的方案。

5 结束语

本文通过本体和规则方法,建立了海上交通知识库,将基于语义和空间关系的推理结合起来。本体对相关领域的概念及关系给予了充分的表达,为基于语义和知识层次的查询奠定基础,提高了查询的查准率和查全率。规则弥补了本体表达能力不足的缺陷,使用规则发现隐含丰富的语义知识,可以减少运算量,优化查询算法,提高查询效率。知识库及推理系统的建立弥补了非领域用户的知识鸿沟,为用户提供了更加方便的查询方式。

参 考 文 献

- [1]. RATCHATA PEACHAVANISH, HASSAN A KA RI-MI. Ontological Engineering for Interpreting Geospa-tial Queries [J] . Transactions in GIS , 2007, 11(1):115-130.
- [2]. Vol. N. Ontology Learning from Text: Methods, Evaluation and Applications[J]. Computational Linguistics, 2005, 32(4):págs. 569-572.
- [3]. 徐恒, 马飞. 基于 RDF 的本体知识库设计与实现[C]// The, International Conference on Computational Intelligence and Industrial Application. 2010.
- [4]. Kaneiwa K, Mizoguchi R. Ontological Knowledge Base Reasoning with Sort-Hierarchy and Rigidity.[C]// Principles of Knowledge Representation and Reasoning: Proceedings of the Ninth International Conference. 2004:278-288.
- [5]. 张元发. 基于本体的船舶领域知识获取研究[D]. 大连海事大学, 2010.
- [6]. Oliver D E, Rubin D L, Stuart J M, et al. Ontology development for a pharmacogenetics knowledge base.[J]. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing, 2015:700.
- [7]. Luciano J S, Andersson B, Batchelor C, et al. The Translational Medicine Ontology and Knowledge Base: driving personalized medicine by bridging the gap between bench and bedside[J]. Journal of Biomedical Semantics, 2011, 2 Suppl 2(2):1-21.
- [8]. Cheng G, Du Q, Ma H. The Design and Implementation of Ontology and Rules Based Knowledge Base for Transportation[C]// International Conference on Computer Science and Software Engineering. IEEE, 2008:1035-1038.