

本体驱动的动态虚拟社区知识通信

王莉, 余雪丽

(太原理工大学计算机与软件学院, 太原 030024)

摘要: 知识通信是实现语义 Web 的有效途径之一。该文给出知识通信的定义和特点, 构建了基于动态虚拟社区的 P2P 知识通信架构; 为实现高效知识路由, 建立了基于本体的任务元知识库、任务实例库、规则库、本体知识库, 设计了具有发现、协调和定位服务资源等功能的路由服务 Agent, 建立了本体驱动、分散控制的快速、准确、高效的知识路由机制。

关键词: 知识通信; 动态虚拟社区; 本体; 知识路由

Ontology-driven Knowledge Communication of Dynamic Virtual Communities

WANG LI, YU Xue-li

(College of Computer and Software, Taiyuan University of Technology, Taiyuan 030024)

【Abstract】 Knowledge communication is one of the valid approach that realizes the semantic Web. A new definition of knowledge communication is put forward and a kind of P2P knowledge communication structure consisting of dynamic virtual communities based on ontology is established. In order to improve routing quality, task-meta-knowledge-base, task-case-base, rule base, ontology knowledge base are designed based on ontology. Routing service agents are created who can accomplish discovery, mediating and locating service resources. All of these constitute the decentralization controlling knowledge routing system.

【Key words】 knowledge communication; dynamic virtual community; ontology; knowledge routing

1 概述

广泛的网络应用要求信息资源之间能共享、重用、融合、协同集成和互通信, 这需要对网络信息进行规范化、结构化、语义化处理, 语义网(semantic Web)^[1]应势而生。知识是信息经过加工整理、解释、挑选和改造而形成的有结构的系统^[2], 它是实现语义网的有效途径。

目前, 语义网上知识研究的目的通常为: (1)通过对结构化、半结构化或非结构化的 Web 资源进行规范化处理并描述, 提供互联网上资源定位、采集和分类的智能处理方法。(2)由于 Web 的快速发展, 各种旧的新的 Web 知识库系统大量出现、并存, 使得分布式异构知识库之间的互联及通信、融合等成为急待解决的问题。(3)各种资源之间需要网络协同推理以更智能地为人类服务。在这些研究领域, 知识通信机制的建设都起着举足轻重的作用。

当前与知识通信相关联的研究主要有: (1)解决分布式异构知识库之间、规范化或非规范化的文件系统、数据库资源之间的互理解、互通信及更高层的知识融合等问题; (2)利用人工智能、通信工程、数学等方法, 借鉴知识理念解决现存的通信问题或构造新的通信原理。由此出现了 2 个有关知识通信的研究领域。在多 Agent 系统、语义 Web、知识网络、分布式知识库系统等研究领域主要研究第 1 种通信问题, 相关研究见文献[3~6], 在通信领域主要解决第 2 种问题, 相关研究见文献[2,7,8]。目前, 第 2 种研究主要针对通信本质, 研究通信的新型数学模型。

本文从使用技术和解决高层通信 2 方面出发, 提出一种与以上 2 个领域不同的新的知识通信概念——利用知识工具构建应用层面向知识的新通信体系, 建立各种信息、资源、

服务系统间的语义互操作, 实现全面、快速、准确的网络资源路由定位和资源动态复合。它具有以下特点:

(1)支持分散控制环境下, 异构知识库之间、知识库与非知识库、非知识库之间的不同应用的语义集成、语义互操作和协同推理。

(2)存在一种语义层的新型路由机制, 以增加知识、服务、资源发现的查全、查准、查快率, 从而快速准确地传递通信内容。

(3)知识通信机制是建立在应用层之上的高层通信, 应该是和现有网络通信技术相兼容、可嵌入、轻量级、可扩展的。

(4)在应用层上, 存在一种基于知识的通信协议, 使得通信内容传递更准确。

(5)知识通信机制具有一定智能, 能随场景不同而灵活调整通信方式, 以使通信协议最大程度减负。

2 基于动态虚拟社区的知识通信机制

语义 Web 的目标就是使机器利用 Web 数据智能协作。对 Web 资源进行规范、约束, 使其知识化, 使各种资源之间能互理解、互通信、资源共享和重用, 这成为知识通信的首要任务。另外, 资源知识化使得网络通信效能应用层得到很大提高, 优化的知识通信架构、知识路由设计是知识通信的鲜明特征。

基金项目: 国家自然科学基金资助项目(60472093); 山西省自然科学基金资助项目(2006021015); 山西省回国留学基金资助项目(2003-36); 太原理工大学青年基金资助项目

作者简介: 王莉(1971-), 女, 副教授、博士研究生, 主研方向: Web 智能, 知识通信, 软件体系结构; 余雪丽, 教授、博士生导师

收稿日期: 2006-11-30 E-mail: l_lwang@126.com

图 1 是本文建立的知识通信层次结构概念图。主要有 3 层：资源层(resource layer)，知识通信中间件层(knowledge communication middleware)，需求层(requirement layer)。知识通信中间件包括 3 个主要内容：(1)用抽取(extract)、精化(refine)和封装(wrap)的方法对用户需求、知识交互等通信内容进行管理；(2)知识管理、发现(knowledge management, discovery)和知识路由机制(knowledge routing)；(3)知识通信协议(knowledge communication protocol)。

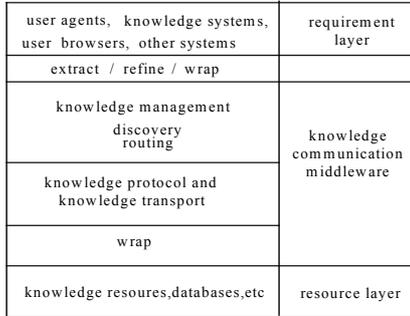


图 1 知识通信层次结构概念

图 2 是本文建立的基于动态虚拟社区的知识通信架构。

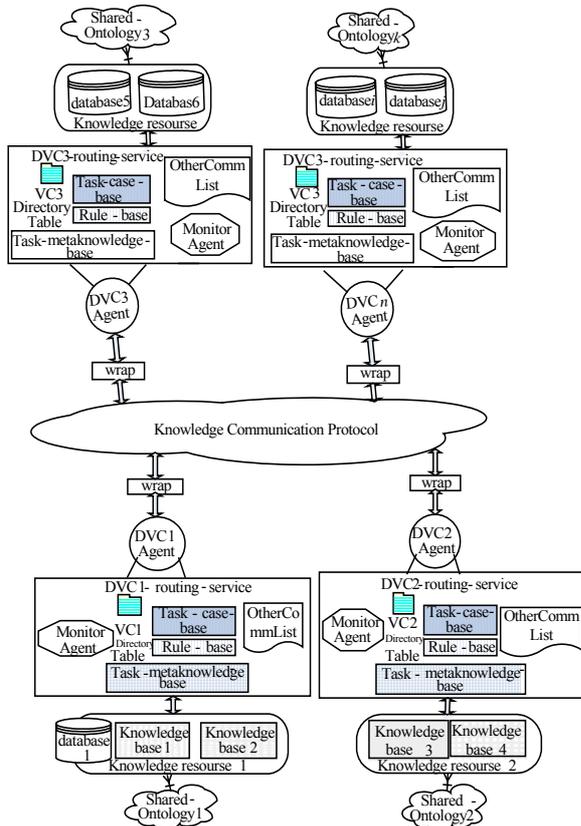


图 2 知识通信体系结构

以资源节点间领域距离、主题距离、是否共享某一本体及网络距离等多维因素为衡量指标，对 Web 上的知识库、数据库以及其他各种资源进行资源重组，建立动态的虚拟社区(dynamic virtual community)。每个虚拟社区都有一个动态虚拟社区路由服务 Agent(dynamic virtual community routing service Agent, DVC Agent)负责维护和管理本社区的任务元知识库(task meta-knowledge base)、任务实例库(task case base)、规则库(rule base)、本体知识库(shared ontology)、动态虚拟社区内的知识路由表(DVC directory table)，同时有监控

Agent(monitor Agent)对社区资源节点运行状态进行监控，辅助 DVC Agent 提供高效路由服务。虚拟社区之间是对等关系，形成一种 P2P 的通信模式，因此，每个 DVC Agent 还必须和其他 DVC Agent 进行协商交互，协作实现基于知识的高性能资源定位和服务。由此，DVC Agent 在对本社区积极维护的同时，还记载着其他动态虚拟社区的 profile 列表(othercommList)。

这里借用了网格中虚拟社区的概念^[9]，但又对它进行了扩展，因为知识通信架构建立在动态、开放、异构、自治的语义 Web 上，语义 Web 就像一个人类社会，里面充斥着各种角色的资源和服务，有一定的组织和群体，有真实的物理意义的群体，也有虚拟的精神领域的群体。这些群体一般有相同或相近的兴趣或背景，所以，知识通信体系中的虚拟社区是主题敏感的社区(topic-sensitive communities)，它的建立和动态演化分别基于以下 2 种因素：

(1)静态因素。主要包括网络距离、领域距离、主题距离、是否共享某一本体等。它是指各种信息系统或知识库地域上非常接近并且领域相近，具有共享本体。在知识通信架构中，通常基于静态因素初始化虚拟社区并建立社区的静态目录表。

(2)动态因素。根据网络环境的动态性、开放性，虚拟社区不应该是一成不变的，它是可以不断动态演化的。随着知识通信交流过程的进行，由于领域和主题概念相近性、关联性的不断发生，一些新的资源会不断加入到相应的虚拟社区，同时每个虚拟社区的资源目录表也会增加新的资源信息。

在知识通信环境中，帮助用户快速准确定位所需资源是知识路由的一个主要内容。需求信息的发布者可以是用户 Agent、信息系统或各种浏览器，需求信息在进行知识通信之前，应进行加工处理。这个过程一般分为抽取、精化和封装。对于不同需求对象，建立不同的抽取策略。用户 Agent 要建立相应的用户模型；对于非规范化的信息系统要进行本体工程建设，建立它的知识概念模型；对于浏览器信息要进行过滤、本体模型建设。

需求者所提出的需求首先会被传送到本社区 DVC Agent，因为在同一社区内，知识资源所基于的本体相同或是建立在共享本体之上，所以它可以采用本社区特定表示的知识通信，直至本区域不能满足需求时，才根据本社区所记载的其他社区的 profile 列表，把任务提交到可能满足需求的别的社区路由服务处。但是在提交前要进行封装，提交任务的消息格式如下：

{task, comm, onto}，即{任务信息，虚拟社区信息，所使用的本体信息}

任务信息={请求者，任务内容}

请求者={URI, user's model}

任务内容={任务的输入，输出，任务约束}

这种优先考虑需求者自身所处虚拟社区的通信模式的好处是：因为同一社区的成员是基于同一共享本体，领域相近，所需内容经常是本社区所具有的，所以提高了资源的查准率、查精率、查快率；而且，在同一社区内通信，共享本体保证了通信信息的准确性；先由本虚拟社区解决问题，分散控制网络资源，提高了网络性能，避免了集中控制所造成的网络拥塞。

各 DVC Agent 都有相应的监测 Agent 辅助，它们监测所辖区域中资源的空闲、忙、退出状态、服务能力的变化、死

亡、新增等信息，动态更新路由表中相应资源的静态信息和动态信息。同时，当有某个正被使用的节点发生变化时，它会把这个变化反映给 DVC Agent，进行资源的重新选择或任务完成方案的重设计。

3 知识路由

在分布、动态、开放的 WWW 上基于知识建立动态虚拟社区，其主要目的就是进行准确、快速的知识路由。知识通信过程中，资源需求者将任务请求发送给 DVC Agent。DVC Agent 依照一定策略从资源路由表中查找符合需求的最佳服务，并把信息返回给需求者。

DVC Agent 主要完成 3 个功能：发现(discovery)，协调(mediating)和定位(locating)服务资源。DVC Agent 在完成这些功能的过程中需要借助于任务元知识库、任务案例库、规则库、知识路由表中的知识。

任务元知识库中存储了为完成任务所需的服务和服务协作关系知识及协作模型发生变化时的修补策略，因为一种任务的解决方案并非唯一，所以任务元知识库中不仅对任务的多种解决方法进行了描述，而且还对资源复合方式、资源在解决问题时的必然性、可能性、时序性等进行了记录，用动作或资源代替命题，采用时态逻辑的方法，进行任务知识表示和推理。例如，“水洗衣服”任务中，要顺序经过 4 个作业：盛水，加入洗涤剂，洗衣服，晾干。这 4 个环节中，加入洗涤剂、洗衣服 2 个环节分别有 2 种可选作业：加入洗衣粉，加入肥皂粉；手洗，机洗。在这 2 个环节中作业选择都是可以选做其一或都做，但不能都不做。相应工作流程见图 3。

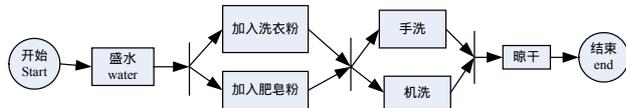


图 3 “水洗衣服”任务控制流程

其中，A 代表盛水；B 代表加入洗衣粉；C 代表加入肥皂粉；D 代表手洗；E 代表机洗；F 代表晾干，其相应的时态逻辑命题形式为

$$\square A \cdot \triangleright (B \ C) \cdot \triangleright (D \ E) \cdot \triangleright \square F$$

其中， \square 是必然算子； \triangleright 是可能算子； $\cdot \triangleright$ 代表动作执行先后顺序关系， \triangleright 代表可选。这样，就可以根据任务元知识库推理得到详细概念层的多种任务解决策略。

DVC Agent 的 3 个功能具体如下：

(1)发现，就是在得到详细、规范的服务需求信息后，在路由表中进行服务匹配。

(2)协调，主要指当不能查找到符合任务完成策略的服务信息时，可对任务完成策略进行调整，主要有 3 种协调方案：根据任务案例库进行基于案例的任务推理(reasoning)、任务转换(transforming)和任务分解(decomposing)。采用案例推理，先计算任务和案例之间的语义距离，再推理找到最为类似的服务案例。若案例推理失败，则可以转而进行其余 2 种协调策略。对于原子服务需求，即在任务元知识库中此任务无分解规则，则采用任务语义转换策略，利用推理方法，转换为别的相近或相关的服务需求。例如，小张想获得小李的电话号码，可是小李并没注册她的电话号码，但是小李的丈夫小王注册了家庭电话，在不能直接获得小李电话号码的前提下，可以通过获得小王的家庭电话得到小李的联系方式。这就是任务语义转换策略。还有一种协调方法是任务分解，它主要针对复杂任务而言，当一个任务需要多个服务的合作，而且

需要是在一定的合作方式下才能完成时，就必须采取分解协调策略。由于同一个任务的解决方案不止一个，因此必须设立合理优化的推理规则以便在最短时间内得到最可行的服务策略。在查找无结果的情况下，还可以适当调整分解策略。不管哪种协调方法，都要借助于规则库和任务元知识库。

(3)定位，在得到多个可选的服务后，根据服务的动态属性(质量属性、意图、价格策略等)，综合多属性考虑，建立一个质量最优的服务方案。

DVC Agent 的活动图见图 4 所示。

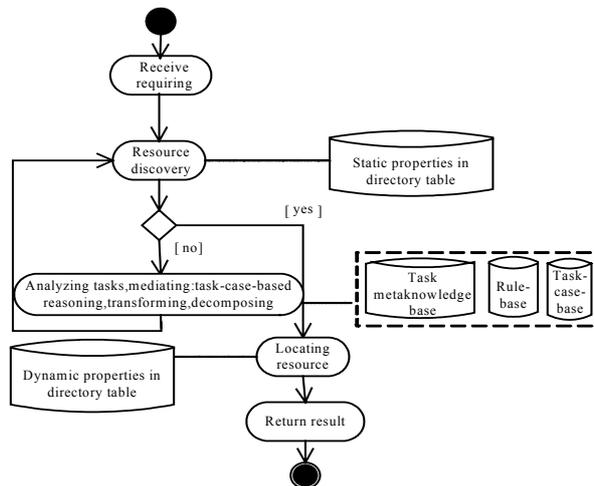


图 4 路由服务 Agent 的活动图

4 结束语

本文建立的基于本体的动态虚拟社区间的知识通信架构正在进行部分应用，笔者利用医疗卫生系统、消防系统 2 个现有系统，对他们的现有资料进行语义处理，建立虚拟社区管理，组成了一个分布式智能决策应急系统，以应对公共社会危机事件。

随着语义 Web 的迅速发展，越来越多的异构资源分布在网，建立这些资源之间的通信是使知识共享重用以及联合解决问题的关键所在。本文就是在这种广泛的需求下，利用本体技术，同时借用网格中的虚拟社区概念，建立了基于本体的动态虚拟社区的知识通信架构，这仅是项目的初始工作，下一步将优化虚拟社区构建方法，建立合理的社区间和社区内的通信模式，实现知识高性能的互联互通。

参考文献

- 1 Lee T B, Hendler J. Publishing on the Semantic Web[J]. Nature, 2001, 410(6832): 1023-1024.
- 2 陆汝钊. 研究知识科学, 发展知识工程, 推进知识产业[J]. 信息技术快报, 2003, 3(6).
- 3 Wu Zhaohui, Chen Huajun, Xu Jiefeng. Knowledge Based Grid: A Generic Grid Architecture for Semantic Web[J]. J. of Comput. Sci. & Technol., 2003, 18(4): 462-473.
- 4 Preece A, Hui K, Gray A, et al. The KRAFT Architecture for Knowledge Fusion and Transformation[J]. Knowledge-based Systems, 2000, 13(2/3).
- 5 Hui K, Preece A. An Infrastructure for Knowledge Communication in AKT Version 1.0[Z]. (2001-12). <http://www.csd.abdn.ac.uk/research/akt/aktbus/aktbus-1.0>.

(下转第 226 页)