

- 东北财经大学青年科研人才培养项目（DUF2014Q38）成果
- 中国科学院网络化控制系统重点实验室2014年开放课题（2014NC001）成果
- 东北财经大学管理科学与工程学院青年基金培育和高新技术开发项目成果

知识协同工作流 建模、服务规划与服务组合

Modeling, Service Planning and Service Composition
in Knowledge-intensive Collaborative Workflows

高 明◎著



中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

东北财经大学青年科研人才培育项目(DUFE2014Q38)成果

中国科学院网络化控制系统重点实验室 2014 年开放课题(2014NC001)成果

东北财经大学管理科学与工程学院青年基金培育和高新技术开发项目成果

知识协同 workflow 建模、服务规划与 服务组合

高 明 著

中 国 铁 道 出 版 社

2 0 1 5 年 · 北 京

内 容 简 介

本书从知识协同和业务流程管理结合的视角研究了互联网和知识经济背景下的以知识交互为主的人机协同和服务组合中的关键问题和技术,包括:协作过程中角色、任务及资源的动态分配协调、知识资源建模和工作流过程模型的结合,知识资源和知识协同的生命周期管理;在知识分享和动态任务创建时,实现上下文感知的自适应学习和知识推送;针对工作流中随需变化的结构和模式,结合人工智能规划算法,自动生成满足用户需求和业务目标的组合服务,设计和改进相关算法用以选择功能相同但服务质量不同的候选服务对象。

本书可供高等院校相关专业师生学习使用,也可供科研单位的科研人员学习和参考。

图书在版编目(CIP)数据

知识协同 workflow 建模、服务规划与服务组合/高明

著. —北京:中国铁道出版社,2015.3

东北财经大学青年科研人才培育项目(DUFE2014Q38)

成果 中国科学院网络化控制系统重点实验室 2014 年开放课题(2014NC001)成果 东北财经大学管理科学与工程
学院青年基金培育和高新技术开发项目成果

ISBN 978-7-113-19970-8

I. ①知… II. ①高… III. ①知识管理—研究 IV.

①G302

中国版本图书馆 CIP 数据核字(2015)第 029913 号

书 名:知识协同 workflow 建模、服务规划与服务组合
作 者:高 明 著

责任编辑:金 锋 编辑部电话:010-51873125 电子信箱:jinfeng88428@163.com

编辑助理:亢丽君

封面设计:时代澄宇

责任校对:龚长江

责任印制:李 佳

出版发行:中国铁道出版社(100054,北京市西城区右安门西街8号)

印 刷:北京京华虎彩印刷有限公司

版 次:2015年3月第1版 2015年3月第1次印刷

开 本:787mm×1092mm 1/16 印张:12 字数:302千

书 号:ISBN 978-7-113-19970-8

定 价:39.00元

版权所有 侵权必究

凡购买铁道版图书,如有印制质量问题,请与本社读者服务部联系调换。电话:(010)51873174(发行部)

打击盗版举报电话:市电(010)51873659,路电(021)73659,传真(010)63549480

前 言

知识经济背景下的企业创新活动表现为复杂流程下面向内容、知识资源的多角色、跨组织的知识协同。面向知识协同的跨组织多实体工作流是传统工作流管理系统与知识管理系统的有机结合,不仅支撑和管理着企业、组织的各个实体之间的复杂业务协作关系,还能够管理和控制各类知识资源在企业、组织的各个实体间产生、流动和更新的知识协同流程和知识生命周期的全过程。

多数工作流模型和管理系统用来支持成熟的产品制造过程或固定的业务目标的实现路径管理,注重流程活动间的串行、并行等过程结构的自动化、半自动化控制。在面向知识和协同的应用,传统的工作流模型难于应付复杂的跨组织、多角色、知识密集型任务的建模和管理。要解决的主要问题包括协作过程中角色、任务及资源的动态分配协调、知识资源建模和工作流过程模型的结合,以及知识资源和知识协同的生命周期管理等。

早期的 E-Learning 系统基于传统的教育观点,注重课程资源的建设和学习内容,在计算机和互联网上的多媒体呈现方式,面向学习内容、学生培养和考评,忽视了学习活动的过程管理和实时监控,对学习活动中多实体协作和知识协同等还不够重视。工作流技术被引入 E-Learning,利用企业、组织已有的 IT 基础设施,将学习流程整合到企业的工作流程中。产品化的工作流管理软件和传统的工作流建模技术,流程结构固定,形式化描述精确,易于实现和管理,但不可避免的带来了流程“僵化”,难以灵活变更的缺点,在知识分享和动态任务创建时,很难实现上下文感知的自适应学习和知识推送。

解决此类问题应在工作流建模阶段把学习资源、学习过程的建模和管理,学习过程和学习资源的协作、协同结合在一起,引入新技术和新模型,在协作的学习流程活动中实现对知识资源的重用和积累,实现满足企业、组织成员的个性化知识需求的、情景感知的知识过滤、推送或推荐机制,以支持创造性的知识协同学习。

利用语义网和本体技术对知识资源进行语义建模和标注,构造统一的、抽象的、高层的知识本体库,借助语义网查询和推理技术,可以更好地依据动态学习环境中的上下文信息在本体知识库中提取面向用户个性化需求的知识资源推送给学习者。

面向内容协作、知识协同的复杂应用中,企业、组织的某些具体业务会随着用户需求和业务目标而变化,企业、组织内外的协作关系存在着高度的变化性。对于此类业务协作,流程设计者无法预先对某个流程的所有环节精确地给定一个固定结构的静态流程模板;而且随着用户需求和业务目标的变化,原有的静态流程模板也可能无法适应新的环境条件。

面向工作流的自动服务规划,能够针对工作流中随需变化的结构和模式,结合人工智能规划技术和算法,自动生成满足用户需求和业务目标的服务节点执行序列,按需构造目标驱动的子流程实例;传统工作流的静态流程模板作为上层的、宏观的控制机制来统一协调高层视角中固定不变的协作活动。

规划图、HTN 及其扩展或组合在自动服务规划领域应用的最为广泛。但是上述方法侧重于研究服务组合问题的规划求解实现和相关规划算法的优化,没有考虑到在服务规划中是引入用户需求和业务目标的高层的、具有指导意义的抽象模型,还是基于较低层、细节的任务层面。如何在宏观意义上形成满足用户需求和业务目标的组合服务仅仅依靠人工智能领域的规划器是不够的,还需要针对服务组合的特点和相关应用领域做一些工作。

跨组织多实体协同的工作流中服务节点的组合问题不仅涉及流程服务节点间的逻辑顺序关系(即结构特征),也涉及服务节点存在的众多候选服务的选择和优化问题。工作流被建模和验证后,运行阶段被实例化,工作流引擎需要根据业务约束和资源约束对各个服务节点的具体服务对象(任务执行者)进行绑定,如果存在满足功能需求的候选服务对象(任务执行者)时,如何选择功能相同但服务质量(Quality of Service, QoS)不同的候选服务对象(任务执行者),是服务组合选择和优化的关键问题。

本书主要内容的逻辑关系如图 0.1 所示。第一章总结和梳理了知识协同工作流建模、服务规划和服务组合方面的基础理论和技术,第二章到第五章以此为基础,重点阐述围绕这一领域所做的主要研究工作。

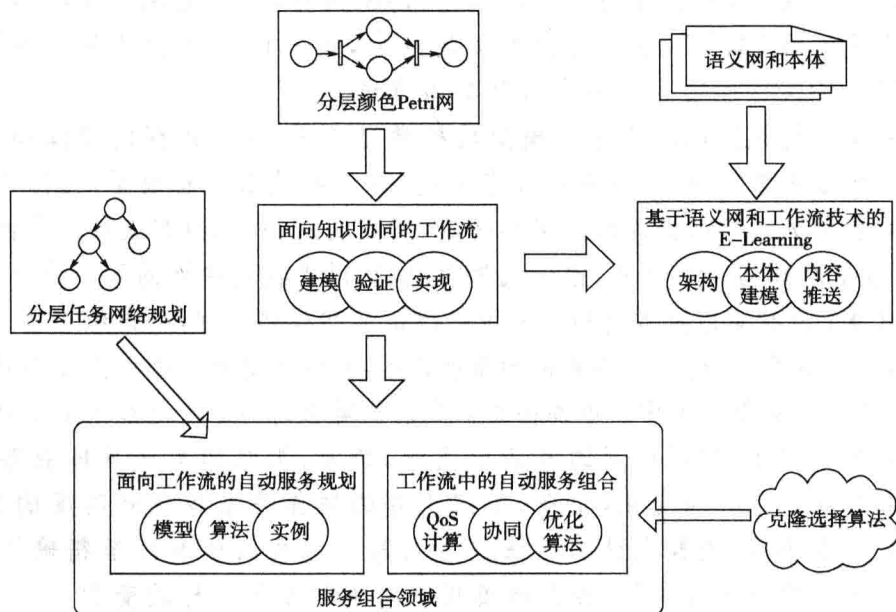


图 0.1 主要内容之间的逻辑关系

本书由东北财经大学青年科研人才培育项目(DUFE2014Q38)、中国科学院网络化控制系统重点实验室2014年开放课题(2014NC001)、东北财经大学管理科学与工程学院青年基金培育和高新技术开发项目资助出版。在本书的编写和出版过程中,得到了我的博士生导师姜继忱教授及东北财经大学管理科学与工程学院院长唐加福教授的大力支持、指导和帮助,在此向两位教授表示衷心的感谢。另外,我的家人、同事及朋友也给予我支持与鼓励,在此也一并表示感谢。

限于作者的理论水平,书中难免有不妥之处,敬请广大读者批评指正。

编者

2015.1

目 录

1 知识协同 workflow 建模、服务规划与服务组合的基础理论和技术	1
1.1 面向知识协同的工作流	1
1.2 workflow 建模和验证技术	3
1.3 面向服务架构、企业服务总线 and 面向资源架构	8
1.4 语义网和 workflow 技术在 E-Learning 中的应用	11
1.5 服务规划和服务组合的界定	14
1.6 自动服务规划	16
1.7 自动服务组合	17
小结	22
2 面向知识协同的工作流建模、验证和实现	23
2.1 基于分层颜色 Petri 网的工作流形式化描述和建模	24
2.2 基于分层颜色 Petri 网的工作流模型验证	34
2.3 基于 PVM 的工作流引擎设计和实现	43
2.4 基于服务总线的工作流服务交互	49
2.5 内容协同和管理	52
小结	59
3 基于语义网和 workflow 技术的 E-Learning 系统开发实证	61
3.1 语义网和本体建模中的关键技术	61
3.2 系统总体架构	68
3.3 本体驱动的个性化领域内容推荐	70
3.4 学习资源和学习流程的本体建模和系统实现	73
3.5 应用服务验证	80
小结	90
4 面向 workflow 的自动服务规划	91
4.1 规划图、HTN 和 SHOP2	91
4.2 面向目标驱动、内容协作的 GCCHTN 规划方法	96
4.3 规划知识库参考模型	97

4.4	规划领域知识库	102
4.5	规划通用方法库	106
4.6	规划领域方法库	111
4.7	规划问题求解	113
4.8	实例验证	117
	小结	122
5	workflow 中的自动服务组合	124
5.1	自动服务组合的模糊 QoS 属性集	124
5.2	自动服务组合的形式化模型和 QoS 计算	127
5.3	自动服务组合优化算法的仿真实验设计	140
5.4	标准遗传算法的改进和仿真实验	142
5.5	克隆选择算法、免疫记忆克隆规划算法的改进和仿真实验	150
5.6	并行短期记忆协同克隆选择算法	167
	小结	175
	参考文献	177

知识协同 workflow 建模、服务规划与服务组合的基础理论和技术

1.1 面向知识协同的工作流

1.1.1 工作流和工作流管理系统

工作流(Workflow)是企业计算的主要特征,它将一项工作划分为多个离散的步骤以及触发步骤转移的事件,描述系统如何在合适的时间、合适的地点以合适的资源做合适的事情。

工作流管理系统(Workflow Management Systems, WfMS)是工作流管理联盟(Workflow Management Coalition, WfMC)和目标管理小组(Object Management Group, OMG)制订的工作流系统软件的管理和开发标准,定义了工作流建模阶段和工作流执行阶段的组件和模型。

工作流管理系统在建模阶段用资源模型和过程模型来描述业务流程的两个不同方面。资源模型描述了参与过程执行的资源(人力资源和技术资源)及其属性,而过程模型则描述了过程中的活动以及活动之间的依赖关系。

工作流管理系统在执行阶段定义了工作流实例监控和工作流执行组件与用户、外部应用的交互接口,后又定义了工作流流程定义的交换接口,用于不同工作流管理系统之间交换流程定义信息。

李红臣等人对工作流模型及其形式化描述进行了综述和比较,指出 WfMC 的工作流参考模型(Workflow Reference Model)和工作流管理功能规范(Workflow Management Facility Specification)只是定义了工作流管理系统的一般性框架和所提供的服务,没有规定工作流建模的形式化模型和验证方法。与工作流相关的形式化建模和验证方法是个开放的研究课题,在学术界得到了一定的重视。

Stohr 等人从管理角度对工作流系统的应用模式进行了分类,从宏观上将其划分为面向组织的工作流、作为控制和执行测试载体的工作流和作为知识共享和学习平台的工作流;同时对商用工作流管理系统进行了总结分类,将其归纳为协作式工作流(collaborative workflow)、行政管理式工作流(administrative workflow)、特定工作流(ad-hoc workflow)和产生式工作流(production workflow)等 4 种类型。而本书主要探讨和研究面向知识协同的协作式工作流。

1.1.2 知识协同和协同效应

互联网背景下的信息服务是服务提供者通过研究、处理信息为客户提供其所需要的信息内容的过程,信息服务系统在本质上是一种协同软件,涉及服务对象协同、服务信息协同、服务流程协同和服务计算协同等。

服务对象的协同性是互联网信息服务系统的基本特征,包括对象协同感知、角色协作关系、社会网络等技术手段。通过研究企业、组织中的角色协同感知和角色特征信息,在协作中融入角色属性和流程上下文信息,能在个体协同的基础上提供更高层次的基于角色信息的服务;研究角色间的协作关系和社会网络,可以挖掘社会信息和发现社区,为个体提供扩展的信息服务。服务信息的协同则包括协同创作、协同过滤和协同检索等,用于信息分享、信息收藏、信息推送和信息推荐等信息协同活动。

随着知识经济时代的到来,知识在现代社会中的价值创造作用越来越大,企业、组织越来越重视打造基于知识创新的核心竞争力,重新调整发展战略和运营方式。知识经济背景下的企业创新活动表现为复杂流程下面向内容、知识资源的多角色、跨组织的知识协同。知识协同是不同时间和空间的多个协同实体参与的动态协作过程,协同的对象是组织内外的知识资源,以知识创新为目标。

国内外学者从不同的角度对知识协同的概念和内涵进行了研究。李霞等人认为知识协同以知识创新为目标,是由多个拥有知识资源的行为主体(组织、团队、个人)协同参与的知识活动过程,是组织优化整合知识资源的管理模式和战略手段。吴绍波等人认为知识协同过程包括知识共享、知识转移、组织学习、知识创造。

知识协同是知识管理成功的关键因素之一,协同化的知识管理成为新一轮的研究和实践的热点,关注的重点在于企业、组织如何重构、整合和优化内外部知识资源,使不同时段、来源于不同实体的知识有效的集成、互补及共享,进而实现有效的知识重用和快速知识创新。

近年来,开始有学者关注知识协同中存在着的协同效应。陈昆玉等人认为企业知识协同通过整合组织内外部的知识资源,使组织学习、利用和创造知识的整体效益能够大于各独立组成部分总和的效应。杨利军研究了供应链中的知识协同,提出基于供应链各节点的企业在知识获取、知识流动、知识共享等方面的协作中存在既竞争又合作的特殊关系。李春娟对国内知识协同进行了较为全面和系统的研究,提出知识协同是一项强调动态性的过程,知识协同可以实现“ $1+1>2$ ”的协同效应。

具体的,康峰等人针对多指标决策方法解决团队伙伴选择不理想的问题,根据知识协同团队的特点建立了伙伴选择的指标体系,并通过计算和谐系数与不和谐系数构建和谐矩阵与不和谐矩阵,确定关键核选出占优伙伴,该方法能够很好地处理知识协同团队伙伴的整体和谐性问题。

而冯博等人在知识创新团队的伙伴选择问题中着重考虑了伙伴之间的协同关系与协同效应,并建立了团队伙伴选择的数学模型,该方法具有一定的可行性和实用价值。

1.1.3 知识协同、知识管理系统和工作流管理系统的融合

Gartner 在 2009 年业务流程管理(Business Process Management, BPM)的调研报告中

指出, BPM 已经逐步被纳入企业的核心战略目标, 面临着前所未有的发展机遇。但随着知识经济环境下的企业和组织的知识协同和创新活动日益增多, 对传统的工作流管理理论、方法和工具提出了严峻的挑战。基于工作流技术的业务流程管理与企业内容管理、知识管理系统的融合已经成为一种趋势。

知识管理系统是知识协同的必要技术辅助手段和支撑平台, 协同性对传统的基于内容的知识管理系统提出了更高的要求。晏浩构建的协同知识管理模型(CKMM), 考虑了企业知识流动的特点, 在模型中融入了协同工作环境、知识发现与过滤技术、人际互动和企业文化。杨善林等人认为协同交互式知识管理系统是在现有其他信息系统基础上集成、建立起来的协同交互式知识管理系统。

面向知识协同的跨组织多实体工作流是传统工作流管理系统与知识管理系统的有机结合, 不仅支撑和管理着企业、组织的各个实体之间的复杂业务协作关系, 还能够管理和控制各类知识资源在企业、组织的各个实体间产生、流动和更新的知识协同流程和知识生命周期的全过程。

在技术层面, 构建面向知识协同的多角色、多实体和跨组织的协作平台, 涉及软件工程和计算机科学的多个领域和关键技术, 诸如面向服务架构(Service-Oriented Architecture, SOA)、工作流技术、多 Agent 建模技术、语义 Web 和本体、人工智能相关方法与技术等。

1.2 工作流建模和验证技术

1.2.1 工作流的过程模型

在工业界, BPMN(Business Process Management Notation)是业界普遍认可的一种概念级工作流建模标记语言, 由 OMG 制订和维护, 目前最新版为 BPMN2(Business Process Management Notation 2)。与 BPEL(Business Process Execution Language, 又称 WS-BPEL, 前身为 BPEL4WS)不同的是, 它不是一种执行语言, 而是用于不同角色、领域和知识背景的非专业人员和流程设计者间共享、交流有关流程设计思想和经验的通用图元模型。

由于 WfMC 的工作流参考模型和 BPMN 规范缺乏清晰、一致和可验证的数学或图论基础, 许多学者尝试用不同的理论、模型和技术对它进行形式化模型描述并加以验证。工作流的过程模型的形式化描述和相关的建模语言、方法的设计一直以来都是工作流领域的重要研究内容之一。

用于工作流形式化描述和建模的方法有很多, 诸如: 基于图论(有向图、状态图)的方法、面向对象方法、形式语言方法、状态—实体—活动模型(State-Entity-Activity-Model, SEAM)、ECA(Event-Condition-Action)规则语言、UML(Unified Modeling Language)建模语言、信息控制网(Information Control Net, ICN)以及 Petri 网等。

Sadiq 等人用有向无环图(Directed Acyclic Graph, DAG)来描述工作流过程。图中任务节点表示活动、子过程等, 协作节点表示选择或合并结构; 有向边表示节点执行和数据流动顺序。流程执行从唯一的开始节点起, 到唯一的结束节点终止。

ECA 定义的业务规则简单易懂: 当事件(event)发生时, 如果条件(condition)为真, 则执行某个动作(action), 否则执行另一个动作(alternative action)。Knolmayer 等人使用了扩

展的 ECA(Event-Condition-Action)规则语言来描述业务过程。Van 等人也基于 ECA 构造了事件驱动流程链(Event-driven Process Chains,EPCs)来对流程建模,利用精心设计的因果事件流来驱动流程引擎与外部实体的交互。

Van 等人还利用 UML 的活动图 and 状态图来描述流程中不同实体的交互行为,以及实体在整个生命周期中的状态随事件或操作的变迁。

Ellis 等人用改造于 Petri 网的信息控制网(Information Control Net,ICN)来进行 workflow 建模。ICN 中,用变迁表示活动,库所表示控制节点,用控制节点表示与(and)和或(or)的活动控制关系。

Van 是最早在工作流建模中引入 Petri 网的学者之一。他定义了 workflow 网(Workflow Net,WF-Net),用 Petri 网的标识(Token,又称托肯、标记)表示 workflow 的状态,还利用 Petri 网理论对 WF-Net 的有界性、活性等进行了证明。

上述方法是基于不同的应用场合和设计原则提出的,对不同类型的工作流过程的描述能力存在差异,有些方法侧重于 workflow 管理系统的实现,而有些侧重于对 workflow 形式化模型和语义的分析和验证。

1.2.2 工作流的资源模型

工作流的过程模型主要是对工作流过程结构的形式化描述,而工作流的资源模型侧重于流程中的数据和执行者、角色、组织架构进行描述,而且关注如何为活动节点分配执行人员和资源,包括用户指派、资源指派和资源访问控制。WfMC 在工作流参考模型中定义的资源模型是抽象化、高层的和指导性质的,没有具体设计和实现。

Crampton 等人对工作流中的用户指派机制进行了研究,对具体任务、抽象任务和任务角色三种执行模式进行了分析。徐伟等人提出了一种面向服务的、基于服务授权历史的动态责任分离约束的工作流访问控制模型——SOWAC 模型。把服务视为流程任务的抽象执行和访问控制的基本单元,用服务的访问控制替代流程任务的访问控制。

描述资源模型的元数据模型在不同的 workflow 管理系统内通常各不相同,不利于 workflow 管理系统与其他系统的信息共享和集成应用,特别是和资源、知识密集型应用,如内容管理系统、知识管理系统集成时,建模和管理的难度大,信息冗余和不一致问题变得突出,有部分学者开始注意到这方面的问题。

陈大峰等人开发的基于指派规则和资源管理服务的资源指派系统 RAS,将资源管理功能从 workflow 机中独立出来,以方便地实现资源模型的共享,且具有良好的可扩充性。他根据项目实践,总结出设计 workflow 的资源指派,要考虑支持组件级的插接以提供可扩充性;保持资源模型和过程模型的相互独立性和抽象性;考虑执行者的能力评估、资源优化配置等属性,用于复杂应用的优化,提供更好的灵活性。

王青等人使用本体技术来解决不同构件之间的元数据映射,提出一个基于本体扩展的 workflow 构件,并实现了基于构件的本体扩展 workflow 的框架。

1.2.3 工作流管理系统的开源实现

OFBiz 较早的开源项目,构建在 J2EE 1.4 规范上,为基于 Web 的企业和电子商务应用提供了 workflow 支持,但其应用组件和 workflow 引擎相互耦合,配置、管理和定制开发较为复杂,

不易于扩展。

Activiti 和 ProcessMaker 是开源的轻量级 workflow 管理系统,可以作为独立的(Stand-alone) workflow 服务器提供一个流程交互的平台。

PVM(原 JBPM)在开源领域是较有影响力的项目之一,得到了广泛的应用。PVM 的流程建模语言,早期是基于 jPDL 的,它是 WfMC workflow 参考模型提出的 XPDL(XML Process Definition Language)语言的一种 Java 实现。PVM 的近期版本,结合了 BPMN2 的流程定义和语义规范,能够支持 BPMN2 中的各种关键概念和定义,并且易于扩展加入自定义的模型和语义。

PVM 提供了一个轻量级的嵌入式 workflow 引擎,可以选用 Hibernate 或 EJB3 作为流程引擎的数据持久化机制。它可以被架设为单独的流程服务器,也可以作为应用系统的嵌入式构件,和应用软件运行在同一 JVM 中,架构灵活,易于部署和扩展。

PVM 提供了事件驱动的扩展机制,降低了业务逻辑与流程引擎的耦合性。但其对多角色和多用户的交互、动态表单以及应用上下文感知(Context-Aware)等缺少功能支持或仅提供较为简单的机制,无法完全满足资源密集型的复杂流程的建模和运行,需要深入和大量的定制化和改进。

1.2.4 面向知识协同的 workflow 建模

多数 workflow 模型和管理系统用来支持成熟的产品制造过程或固定的业务目标的实现路径管理,注重流程活动间的串行、并行等过程结构的自动化、半自动化控制。在面向知识和协同的应用,特别在支持企业、组织知识创新活动的 workflow 中,传统的工作流模型难于应付复杂的跨组织、多角色、知识密集型任务的建模和管理。

面向知识协同的 workflow 建模要解决的主要问题包括:协作过程中角色、任务及资源的动态分配协调;知识资源建模和 workflow 过程模型的结合,知识资源和知识协同的生命周期管理;强调“人”在面向知识协同 workflow 中的主体地位,设计基于流程上下文和知识上下文的情景感知机制;提高知识聚合和知识重用的效率等。相关的研究还不是太多。

Van 等人提出了面向知识密集型任务的活动流模型(Case Handling),用有向无环图将流程中的子活动根据先后、依赖关系进行描述,对每项活动定义数据依赖、产出关系和约束,来协调知识密集型任务的流程建模和运行。Tsai 等人从宏观流程和微观流程分层设计的角度探讨了面向服务架构下的人机参与的复杂流程的设计和建模。

有学者研究了流程中的上下文感知。通过在流程的特定节点定义相应的触发条件和约束,可为不同上下文环境中的流程参与者提供个性化的知识体验。

窦万春等人针对基于知识应用和交互的复杂 workflow 系统,对知识应用特性进行了分析,并构造了对应知识聚合与信息再生的过程单元,探讨了基于知识应用和交互的 workflow 系统中的控制问题,提出了一种基于区域控制和交互的控制策略。

Van 等人提出将结构弱化的 workflow 用于面向流程和协作的知识管理系统能够在动态变化的环境中提供主动的信息服务,更好地支持知识密集型活动。为此,他们设计了一个基于 Agent 的软件框架,验证结果表明这一理念是有效的。

有学者探讨了跨组织 workflow 模型的建模问题。Corrêa 等人设计和开发的 JamSession 工具,为不同组织间的工作流协作、整合和事务合作的协调执行和形式化验证提供了一种可

行的解决方案。

也有学者尝试对知识协同过程中的知识流进行建模和分析。张青等人借助 Petri 网对高校科技管理系统中的知识流进行分析,描述和刻画了知识流的特征及相互关系,分析了高校科技管理系统中知识流的动力机制。丁漪杰等人使用了颜色 Petri 网对知识流建模,指出基本 Petri 网用于复杂系统建模时,模型变得庞大,且很难描述各种知识资源,而颜色 Petri 网可以形象的描述知识的异步并发,实现对知识流的监控,能够克服基本 Petri 网的上述缺点。

1.2.5 原生的半结构化和非结构化资源存储系统

知识资源是多种格式的、多变的半结构化和非结构化的资源类型,知识资源不仅包括具体的文档、内容和多媒体片段,也可以是元数据类型的资源,知识资源的存储和管理一般都是基于操作系统的文件系统或大型关系数据库,树形结构的 XML 虽然能描述半结构化的资源,但存储性能低且描述能力仍然不够灵活。原生的支持半结构化和非结构化的资源存储机制是知识资源存储和管理所迫切需要的。

基于 JSR170 和 JSR283 规范定义的 Java Content Repository(JCR)是一个统一的资源管理和存储的抽象模型,支持原生的半结构化和非结构化的资源定义和存储。Jackrabbit 是 JCR 的一个开源的参考实现。

基于 JCR 的内容仓库是一个高级的信息管理系统,是对传统数据库的扩展,它能够提版本控制、全文检索,访问控制,内容分类、内容元数据定义、内容事件监视等服务。JCR 支持多仓库、多结构和节点的自定义内容管理模型,允许为每个仓库定义节点名称空间,防止不同来源和不同应用程序域之间的项目和节点类型的命名冲突,其模型如图 1.1 所示。

JCR 内容仓库由一个或多个 workspace 构成。每个 workspace 是树状结构,有唯一的根节点(root node),树上的元素(item)可以是节点(node)或属性(property),每个 node 可以有零到多个子节点和零到多个子属性,可表述的内容结构很灵活。以 blog 内容仓库模型为例,每个 root node 的子节点代表一个 blog 实体,blog 实体有关的数据存储在 blogEntry 节点的 property 里,blogAttachment 的 property 也可以支持存储二进制的图片文件,如图 1.2 所示。

JCR 规范在 javax.jcr 名称空间上定义了一个能与内容仓库互相访问的、独立的、标准化的和统一的 API,允许以厂商中立的方式访问任何符合规范的仓库实现,无须捆绑到任何特定的底层架构上,后端数据存储可以是文件系统、WebDAV 仓库、XML 文件或关系数据库。利用 JCR 不与特定底层架构绑定的特性,可实现异构数据平台间的资源共享。

1.2.6 工作流的模型验证方法

工作流的模型验证是指将工作流模型转换为一个形式化验证工具的输入,然后使用该

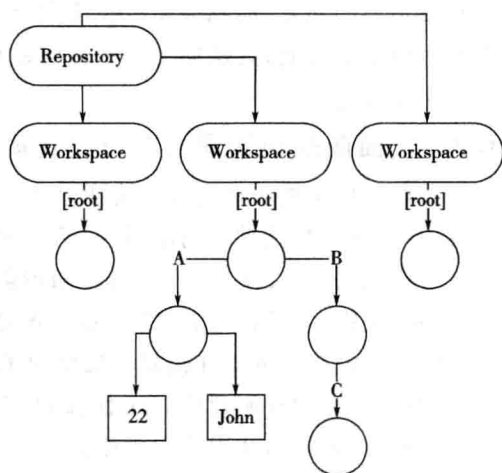


图 1.1 多工作区的 JCR 仓库模型

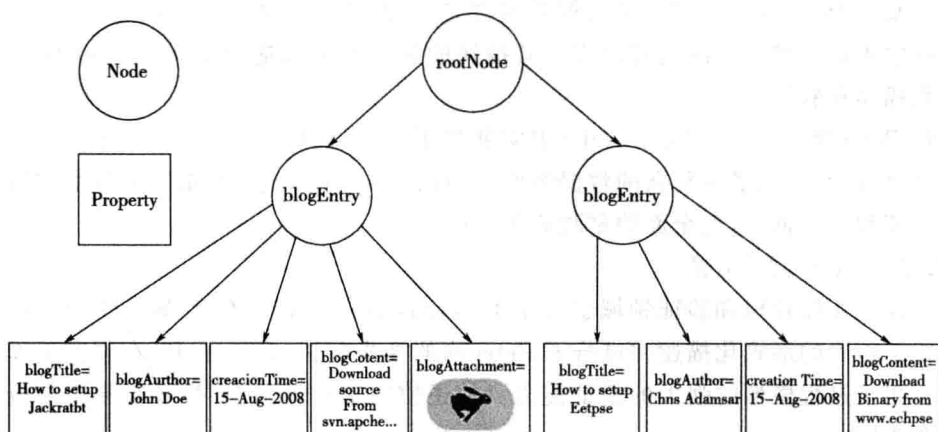


图 1.2 基于 JCR 的 blog 仓库示例

工具进行验证和分析。工作流的模型验证方法常被应用到 Web 服务组合验证领域。下面, 本书对用于工作流和 Web 服务组合的模型验证方法进行一个简要的回顾和综述。

自动机理论可用于描述服务交互过程, 配合时序逻辑来描述系统目标, 然后用自动机模型检测方法验证流程的正确性。Fu 和 Nakajima 等人把组合 Web 服务转换为 SPIN 输入语言, 使用 LTL 线性时序逻辑公式描述验证属性, 最后用 SPIN 模型检测器来验证组合服务的正确性。

进程代数是采用代数形式描述并验证系统的理论, 包括通信系统演算 (Calculus of Communicating System, CCS), 通信顺序进程 (Communication Sequential Processes, CSP) 和 π 演算等。

Antonio 等人使用 CCS 描述 WSCI, 来验证两个或更多 Web 服务是否行为兼容。Overdick 等人利用 π 演算对 BPEL 描述的 Web 服务组合流程进行建模, 借助 π 演算的操作语义和推演能力自动判断信息交互的兼容性和逻辑的正确性。龙慧云基于进程代数理论, 给出了对 BPEL4WS 建立基于传值 CCS 形式化模型的方法和具体实现算法, 并对一个实例进行建模验证其有无死锁的性质。

(1) Petri 网和颜色 Petri 网

Petri 网诞生于 1962 年, 由 C. A. Petri 提出, 用于描述系统中可能发生的各种状态变化以及变化之间的关系。Petri 网描述的系统模型基于图形, 简单直观, 同时又具有严密的数学基础和完善的性质分析方法, 以及一系列成熟的建模和验证工具, 已经成为学术界和工业界解决验证问题的一种有力的数学模型。

随着学者对 Petri 网理论的深入研究和 Petri 网在不同领域的应用, 学者对基本 Petri 网模型进行了各项扩展, 提出了多种高级 Petri 网模型, 如颜色 Petri 网、谓词/变迁网、时间 Petri 网、随机 Petri 网和模糊 Petri 网等。广泛应用在分布式软件系统、分布式数据库系统、并发并行计算、柔性制造和通信协议等领域, 用于形式化验证和性能评价等方面。

Petri 网易于建模工作流的顺序、选择和并行执行过程, 在此基础上可以验证模型的活性、死锁/活锁、有界性和公平性等性质, 被广泛应用到工作流和 Web 服务组合的建模和模型验证中。

但是基本 Petri 网不支持复杂数据类型的定义,无法在流程模型中描述用户自定义的属性及流程中不同类型数据的流转过程。其描述的流程模型仅适用于工作流和 Web 服务组合的过程描述和验证。

颜色 Petri 网(Colored Petri-Net, CPN)扩展了 Petri 网的 Token 机制,支持 Token 的颜色集定义,使网系统能描述复杂的数据类型。颜色 Petri 网不仅可以描述流程的控制流还能同时描述流程的数据流,更全面的刻画业务流程。

(2) 基于 Petri 网的方法

Van 在工作流建模和验证领域引入了 Petri 网,定义了工作流网(WF-Net)。对工作流模型进行了精确的形式化描述并进行了 Petri 网的性质分析,证明了 Petri 网在工作流中的适用性和用于过程模型的强大的验证能力。他还使用工作流网对 Web 服务组合流程进行了描述和验证。

YAWL(Yet Another Workflow Language)是 Van 在之前提出的 WF-Net 上的进一步改进。他们发现在对现实世界中的某些复杂工作流建模时,一些特殊的控制活动诸如多线索流程实例、复杂同步等被映射到高级 Petri 网是困难的。他们仍然以 Petri 网为基础,在重新定义的工作流语言中充分考虑到了这些特殊需求和性质,设计了一个通用的工作流语言 YAWL 和相应的工具。不过 YAWL 依然是面向工作流结构和过程控制的。

Chi 等人设计和实现了一个基于 Petri 网模型的图形化服务组合工作流设计工具。Yoo 等人为了对组合服务的互相依赖性、输入输出参数的语义一致性进行验证,也采用了 Petri 网模型和工具。

Yang 等人用分层颜色 Petri 网(Hierarchical Colored Petri-Net, HCPN)来描述 Web 服务组合的过程模型,用于检测流程的行为特征如死锁、活锁等,并给出了相应的转换规则和算法。李景霞等人也用分层颜色 Petri 网来描述 Web 服务组合问题,从控制流和数据流两方面对 Web 服务组合流程建模并进行仿真和验证。Ni 等人设计了一种基于颜色 Petri 网的服务组合模型,把 OWL-S 描述的语义 Web 的组合服务转换为 CPN 模型进行可达性、有界性和语义一致性的验证。

由于 BPMN 规范缺乏清晰、一致和可验证的数学或图论基础,许多学者尝试用不同的理论、模型和技术对它进行形式化模型描述并加以验证。Raedts 等人定义了一组模型转换规则和算法,将 BPMN 表述的工作流转换为 Petri 网模型,然后转换为 mCRL2 模型进行模型验证。Dijkman 等人对 BPMN 的语义和流程模型进行分析的基础上,提出将 BPMN 映射到 Petri 网是可行的,并实现了相应的算法和工具,借此对 BPMN 的工作流模型进行静态性质分析。

CPN 也被应用到其他类似领域,如网格的任务调度。

1.3 面向服务架构、企业服务总线 and 面向资源架构

1.3.1 面向服务架构

面向服务架构(Service-oriented Architecture, SOA)是一种面向企业级服务的系统架构,系统中具体应用程序的功能是由一些松散耦合且具有统一接口定义方式的组件(服务)

组合构建起来的,并提供一个抽象的服务层(接口),对服务使用者隐藏了服务的实现细节,具有可复用,灵活和可扩展,位置透明性以及协议无关性等诸多优势。

在传统的 JavaEE 服务器中通常利用资源适配器(JCA)、Java 消息(JMS)或点对点的 Web 服务来与外部系统交互,在复杂场景中,信息集成的可管理性和性能会急剧下降。

面向服务架构和 Web Service 等新兴技术的应用简化了信息集成和互操作的难度和成本,能够实现异构系统间的自动化、半自动化协作。IBM WebSphere、Oracle(BEA) Weblogic 等信息企业界巨头均提供了商业性的信息集成套件,采用 BPEL 协议来编排和组合各种基于 Web Service 的服务流程。基于 Web Service 和企业服务总线(Enterprise Service Bus, ESB)的企业信息集成是结合了数据传输、转换、流程整合与管理等关键技术的新一代信息整合模式,优于传统的点对点整合方式。

但是以服务为单位的集成和互操作粒度较粗,异构系统间仍然缺少面向知识和资源的统一视角。

1.3.2 企业服务总线

企业信息集成的方式从传统的以应用服务器为中心的点对点交互模式,演变为基于企业服务总线(Enterprise Service Bus, ESB)的总线模式。企业服务总线构建在基于多协议、通用、事件驱动架构的资源集成基础设施之上,充分地利用了 Web 服务,提供了一些标准的功能和模式来解决一些通用的问题,同时能够提供足够的灵活性和可扩展性以满足面向服务计算环境下的企业信息集成。传统的中间件产品可以更好的专注于自己擅长的领域,在企业软件系统内以高耦合和高性能的处理方式来工作,通过企业服务总线又能高效的与其他外部系统交互。

Glassfish 和 Apache ServiceMix 实现了基于 JBI(Java Business Integration, JSR-208)规范定义的插件式组件集成架构的 ESB,对 JSR 标准的支持能够提供更开放的扩展性,具有一定的吸引力;而 MuleSource 提供了基于通用消息对象(Universal Message Objects, UMO)和流水线事件驱动模型(Staged Event Driven Architecture, SEDA)的 Mule ESB,开发效率和性能有不错的表现,拥有一定的市场份额;微软公司在 Dot Net Framework、Share Point Portal Server 和 Office System 中也添加了统一的通信平台(Windows Communication Foundation, WCF),使得基于以 Web 服务通信为主的服务组件更容易的交互和集成。

企业服务总线支持通信各方面的服务交互和管理,充当 SOA 中服务提供者和请求者之间的连接服务的中间层,因此服务彼此不直接交互。它是一个灵活的连接框架,可促进可靠而安全的系统集成,并同时减少应用程序接口的数量、大小和复杂度。在综合研究了企业服务总线的相关文献、技术和产品的基础上,给出了 ESB 的参考模型,如图 1.3 所示。

ESB 具有很高的灵活性,同时支持嵌入式的轻量级应用和单独运行的企业级应用,成本可控;支持的通信协议众多,能够支持企业的多数连接性需求,如大型主机事务(CICS),ERP,DBMS、中间件服务

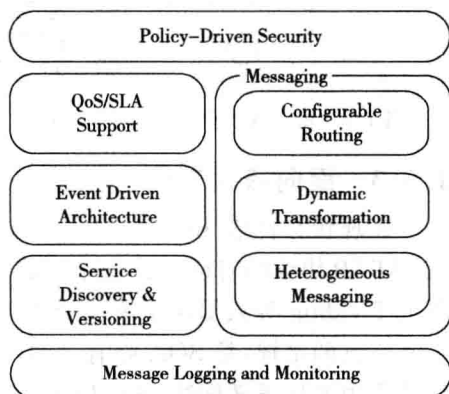


图 1.3 ESB 的参考模型