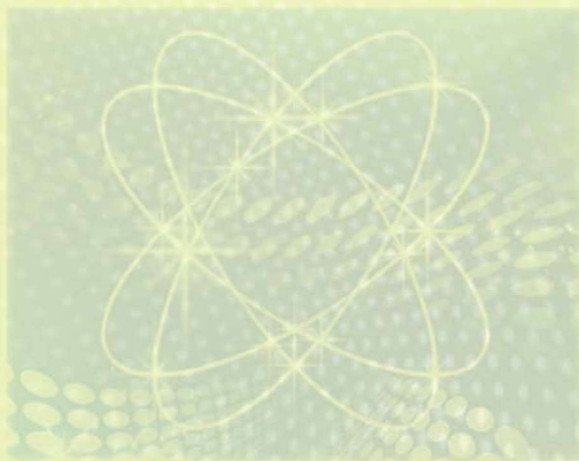


基于语义的 Web 服务集成 研究及其应用

蔡海尼 著



重庆大学出版社

基于语义的 Web 服务 集成研究及其应用

蔡海尼 著

重庆大学出版社

内容提要

本书在语义 Web 服务发现、选择和组合等领域的学术成果基础上,从基于领域本体的服务描述模型、服务匹配与发现、服务选择与组合、数据服务、服务集成框架等方面进行研究,提出了一个完整的语义 Web 服务集成方案,并基于上述方案构建了勘察设计企业协同设计平台原型。

本书可供从事服务计算研究与应用的科研技术人员、高校教师、研究生及高年级本科生阅读参考。

图书在版编目(CIP)数据

基于语义的 Web 服务集成研究及其应用/蔡海尼著. —重庆:
重庆大学出版社,2015.3

ISBN 978-7-5624-8879-8

I. ①基… II. ①蔡… III. ①语义网络—网络服务—
研究 IV. ①TP18

中国版本图书馆 CIP 数据核字(2015)第 036916 号

基于语义的 Web 服务集成研究及其应用

蔡海尼 著

策划编辑:杨粮菊

责任编辑:文 鹏 姜 凤 版式设计:杨粮菊

责任校对:贾 梅 责任印制:赵 晟

*

重庆大学出版社出版发行

出版人:邓晓益

社址:重庆市沙坪坝区大学城西路 21 号

邮编:401331

电话:(023) 88617190 88617185(中小学)

传真:(023) 88617186 88617166

网址:<http://www.cqup.com.cn>

邮箱:fxk@cqup.com.cn(营销中心)

全国新华书店经销

POD:重庆新生代彩印技术有限公司

*

开本:787×1092 1/16 印张:10 字数:124 千

2015 年 3 月第 1 版 2015 年 3 月第 1 次印刷

ISBN 978-7-5624-8879-8 定价:48.00 元

本书如有印刷、装订等质量问题,本社负责调换

版权所有,请勿擅自翻印和用本书

制作各类出版物及配套用书,违者必究

Content Abstract

Web Service Integration is the key technique for constructing the enterprise information framework based on Service Oriented Architecture (SOA). Through service integration, the services are treated as basic components to support rapid, low-cost application development in distributed or even heterogeneous environment. The aim of Service Integration is to reuse both the enterprise's internal software resources and the external services to improve the efficiency of software construction and to meet the ever-changing needs of customers.

From the technology prospective, Service Integration is a business process of service discovery, service selection, service composition, and service execution. The existing Web Service technologies cannot provide strong support for efficient service discovery, service selection and dynamic service composition. It is well appreciated that Semantic Web service provides a new solution for the aforementioned problem. With premium capability in terms of semantic description and logic reasoning, Semantic Web Service can complete the process of service integrating in an automatic, accurate, and efficient way. Semantic Web service can also create business processes and execute them dynamically. Though extensive research has been performed and abundant findings has been achieved in the field of Semantic Web Service, there has been no complete semantic-based Web service integration solution so far.

The solution of Web Service Integration mainly includes two aspects: data integration and application integration. An integrated framework of semantic Web service is built and applied in the field of survey and design.

前言

面向服务的计算(Service Oriented Computing, SOC)和面向服务的体系架构(Service Oriented Architecture, SOA)代表了分布式计算和软件开发的最新发展方向,其基本思想是把服务作为基础构件,支持快速、低成本的分布式甚至异构环境的应用组合开发,解决不同系统/平台间的应用整合问题。Web 服务集成是构建基于 SOA 的企业信息框架的关键技术,其目标是实现企业现有内部软件资源和外部服务的重用,并在此基础上组合成具有一定功能的软件系统,以提高软件构造的效率,满足不断变化的用户需求。

目前的 Web 服务技术在服务发现效能、动态服务选择与组合方面存在局限性。语义 Web 服务为上述问题提供了新的解决途径,利用其良好的语义描述能力和逻辑推理能力,能自动、准确、高效地进行服务集成,动态生成和执行业务流程。目前学术界的研究成果主

要集中在语义 Web 服务的发现、选择和组合方面,尚未提出一个完整的基于语义的 Web 服务集成方案。

本书从基于语义的服务描述模型、服务匹配与发现、服务选择与组合、数据服务、服务集成框架等方面进行研究,并构建了勘察设计企业协同设计平台原型。

首先,针对现有 Web 服务描述语言 WSDL (Web Service Description Language) 不能提供语义信息的局限性,建立了基于语义的 Web 服务描述模型。该模型将描述 Web 服务语义的主流规范 OWL-S (Web Ontology Language for Services) 和现有的 WSDL, UDDI 等成熟技术充分结合,用 WSDL 描述服务名称、服务描述, URI 地址信息等具体信息,用 OWL-S 描述服务功能等抽象信息,同时支持基于语法和语义的匹配与发现;并在功能语义的基础上增加服务的 QoS (Quality of Service) 本体描述,为服务选择和组合提供支持。

其次,针对基于 UDDI (Universal Description, Discovery and Integration) 的传统 Web 服务发现机制只能提供关键字查找,不能进行语义匹配的问题,提出了一种基于语义的 Web 服务分层匹配模型。该模型从三个层次各层采用相应的算法,逐步缩减匹配问题的规模,改进了

服务匹配的效率和精准度。在此基础上,对 UDDI 进行语义扩展,构建支持语义的服务发现框架,并针对实际 SOA 系统中服务访问规模大、频率高的特点,进行了性能优化。同时,提出了一种基于迭代学习的任务分解算法,优化服务的选择与组合过程;并针对服务组合过程中运行环境的异构性和资源负载的动态变化,设计了一种面向任务的自适应服务组合机制。

最后,从数据集成和应用集成两个方面进一步分析 Web 服务集成中的关键问题,构建了基于语义的 Web 服务集成框架,并应用于勘察设计企业协同设计平台。

本书由国家自然科学基金面上项目“基于异构服务网络分析的 Web 服务推荐研究(NO. 61379158)”、国家“十一五”科技支撑计划重点项目“基于 SOA 技术的勘察设计企业应用软件架构系统(NO. 2007BAF23B0302)”、重庆市科委自然科学基金“SOA 中基于语义的自适应服务组合机制研究及其应用(NO. CSTC2010BB2244)”等项目资助。此外,本书的出版还得到了重庆大学软件学院的大力支持,在此一并表示衷心的感谢。

限于作者的学识水平,书中疏漏和错误之处在所难免,恳请读者批评指正。

著者

2014 年 10 月

目 录

第 1 章 绪 论	1
1.1 研究背景	1
1.2 国内外研究现状	3
1.3 主要研究内容	8
1.4 本书结构	10
第 2 章 基于语义的 Web 服务描述	12
2.1 Web 服务语义描述	13
2.2 基于语义的 Web 服务描述模型	18
2.3 语义 Web 服务描述模型构建	23
本章小结	31
第 3 章 基于语义的 Web 服务发现	32
3.1 单概念语义匹配算法	33
3.2 概念集合语义匹配算法	44
3.3 基于描述特征的服务聚类	50

3.4 语义 Web 服务分层匹配	58
3.5 语义 Web 服务发现	68
本章小结	81
 第 4 章 基于任务分解的服务选择与组合	 82
4.1 基于任务分解的服务选择	83
4.2 基于任务分解的语义 Web 服务选择 算法	92
4.3 基于任务分解的服务组合框架 ...	98
4.4 仿真实验及分析	102
本章小结	108
 第 5 章 基于语义的 Web 服务集成	 109
5.1 基于 ESB 的企业服务集成	109
5.2 ESB 中服务发现和选择模式的优化	111
5.3 面向服务架构的数据服务	120
5.4 基于语义的 Web 服务集成框架	123
本章小结	128
 第 6 章 结论与展望	 129
6.1 结 论	129
6.2 展 望	131
参考文献	133

第 1 章 绪 论

1.1 研究背景

面向服务的计算(Service Oriented Computing, SOC)和面向服务的体系架构(Service Oriented Architecture, SOA)代表了分布式计算和软件开发的最新发展方向,其基本思想是把服务作为基本构件,支持快速、低成本和简单的分布式甚至异构环境的应用组合开发,为软件向网络服务的转换提供了理论支持。作为其实现的主流技术,Web 服务定义了应用程序如何在 Web 上实现互操作性,促使 Web 从信息共享平台演化为功能共享、资源共享的公共平台。Web 服务具有平台无关性、语言无关性、松耦合性等优势,解决了不同系统/平台间的应用整合问题,为业务流程组合提供了技术基础。

随着 Web 服务规范的完善和相关技术的发展,互联网已成为一个 Web 服务资源库,为各个领域提供了大量可用的服务。出于 Web 服务的模块性和可重用性方面的考虑,单个的 Web 服务通常只能提供有限的设计逻辑和实现逻辑,复杂业务应用需要一组 Web 服务通过编排不同的执行顺序以及服务之间的交互协作来实现。Web 服务集成的目标是实现企业现有内部软件资源和外部服务的重用,并在此基础上组合成具有一定功能的软件系统,以提高软件构造的效率,满足不断变化的用户需求。这正是 SOC 与 SOA 的核心思想。从 SOA 的角度看,服务集成是一个服务发现、选择和组合执行的业务流程。现行的 Web 服务技术在服务描述、服务发现等方面已有成熟的规范,但在服务选择、动态服务组合方面,由于涉及逻辑判断和推理,尤其是服务语义等问题,还缺乏有力的技术支撑。

语义 Web 服务为服务集成提供了新的解决途径。由于支撑 Web 服务的 WSDL(Web Service Description Language)、UDDI(Universal Description, Discovery and Integration)等传统技术本身缺乏服务内容、服务能力等方面的描述,不能很好地支持服务的自动定位、智能匹配和自动组合。语义 Web 服务技术结合 Web 服务和语义 Web,利用语义 Web 良好的语义描述能力和逻辑推理能力,以自动、准确、高效地进行服务发现、服务匹配、服务组合、服务监控和服务调用为目标,动态生成和执行业务流程,能更好地实现企业应用系统的集成。构建语义 Web 的主要方法是建立本体(Ontology),赋予 Web 服务功能性和非功能性的语义描述信息,通过过程本体(Process Ontology)和过程控制本体(Process Control Ontology)来描述服务模型。过程本体使服务能够进行规划、动态组合以及交互操作,过程控制本体使得智能代理能够监控服务请求的执行。在语义 Web 环境下,服务之间可以通过本体进行通信、资源共享和协同工作。目前,在基

于语义和本体的服务发现、选择和组合等领域已形成了一些有价值的研究成果,但尚未提出一个完整的基于语义的 Web 服务集成方案。

本书研究基于 SOA 构建企业信息框架模型的关键技术,提出并实现了一种基于语义的 Web 服务集成方案,并应用于勘察设计领域。该方案以基于领域本体的服务描述、服务发现和服务组合为主要技术手段,并充分利用 WSDL,UDDI 等成熟的 Web 服务规范,封装企业现有服务,提供层次化服务构件,把人事、办公、财务、项目管理、工程设计等相关系统服务融入 SOA 构架中,实现勘察设计企业的服务集成、数据集成和应用集成。

1.2 国内外研究现状

目前关于语义 Web 服务的主要研究方向有:语义描述、自动服务发现、自动服务选择、自动服务组合和自动服务协商等^{[1][2]}。

1) 服务语义描述

Web 服务的语义描述是实现服务的自动发现、调用、组合和监控等功能的基础和前提。现今主流的语义 Web 服务描述框架是 OWL-S(Web Ontology Language for Services)。OWL-S 是 W3C 建议的基于 OWL 语言的 Web 服务上层本体,它首先用本体来表达领域内的语义信息,然后再利用本体中的概念对服务的功能信息进行描述。OWL-S 不仅可以描述服务的语义,而且能够进行适当的推理,具有支持自动服务发现、组合、调用、互操作和执行监控的能力,没有提供相应的解决方案。由于 OWL-S 与 WSDL 各具优势且有互补性,考虑到方案的实用性,用 OWL-S 的过程模型表示服务的抽象描述,用业内公认的 WSDL 规范来表示服务的具体

描述,重用 WSDL 及其支持软件,是较为理想的 Web 服务描述模型。文献^[3]给出了一种同时支持 WSDL 和 OWL-S 描述的服务发现系统匹配方案。对于两者的结合,从研究成果的角度,常见的语义描述方式是用相关本体在 Web 服务的 WSDL 描述中加入注解,即对 Web 服务进行语义标注,也可在此基础上将其转换为 OWL-S; Zhang Lei 等人提出将从 WSDL 文档中抽取出的隐含语义信息转换为 OWL 本体的格式的方案^[4]。Farrag 等提出了一种将 WSDL 转为 OWL-S 的映射算法,作为服务发现机制的组成部分,并进行了实验验证^[5]。从应用的角度,目前较为成熟的研究成果集中体现在半自动化语义标注的框架和工具。

与此同时,随着服务技术和需求的不断发展,单纯的功能性服务描述已不能满足用户从众多功能等价的服务中寻求最优服务的需求。QoS (Quality of Service) 是描述非功能属性的主要指标集,涉及 SOA 协议栈的各个层次,需要扩展相应的描述模型。OWL-S 等主流的语义描述框架均主要考虑功能方面,缺乏对 QoS 的本体描述和语义支持。在语法模型方面,常见的 QoS 指标包括性能、互操作性、可靠性、可访问性、安全等。文献^[6]给出了本体 WSMO-QoS 的描述,采用 WSMO 建模描述 Web 服务的 QoS。文献^[7]展开对获取、存储、度量 QoS 等基础支持技术的研究,提出了一个可扩展的 Web Service QoS 信息管理框架,支持 QoS 模型定义、QoS 度量方法、QoS 信息采集。文献^[8]分析 QoS 的特性,提出一种估算和预测服务组合 QoS 的方法,可以对满足任何概率分布模型的 QoS 作出较准确的估算。文献^[9]在 OWL-S 的基础上扩展,提出一个轻量级的 QoS 本体模型,并基于此研究语义 Web 服务自动组合框架。语法 QoS 模型以系统语义完全确定为前提,注重实现细节;语义 QoS 模型则主要用于定义 QoS 的指标集合,指标的测量方法,取值范围和度量单位等,注重表示概念和概念间的关系,不强调实现。在上述研究成果的基础上,建立统一的

语义描述模型,作为后继服务发现、服务选择与组合的基础,是本书首先要解决的问题。

2) 服务匹配与发现

在服务发现领域,占据业界主导地位的 UDDI 只能实现语法层次的关键字精确匹配,搜索的查准率和查全率都不高。针对上述缺陷,基于语义的服务发现通过基于共享的领域本体,利用逻辑推理来实现 Web 服务的语义匹配和搜索,提高服务发现的效能,其关键技术是服务的语义匹配,即如何计算服务请求描述和已发布服务描述之间的语义相似程度。文献^[10]根据请求服务与发布服务的输入输出匹配情况将匹配度分为4个等级:精确匹配、插入匹配、包含匹配及不匹配,通过建立语义分类树来判断概念间的关系,查准率高,但匹配度不能得到细致的区分。文献^[11]提出一种基于 OWL-S 的匹配算法,利用描述逻辑推理机来判断输入输出概念之间的关系,实现服务的自动语义匹配。文献^[12]将服务匹配问题转换为获取请求的最佳覆盖 (BestCover) 问题:给定一个请求和知识库,找出最佳覆盖的服务集合,该集合中的每个服务描述都包含与请求尽可能多的公有信息和尽可能少的多余信息。该算法能有效地在服务描述和请求之间进行灵活匹配,但仍不能提供精确的等级区分。文献^[13]把服务匹配分为两个阶段:第一阶段,先根据请求服务的功能性需求(服务做什么、它的输入输出、前提及结果)返回满足基本需求的服务;第二阶段,根据非功能性需求为当前请求确定最适合的服务。

基于相似度计算的服务匹配算法根据语义距离来实现近似服务匹配。文献^[14]根据词语在 WordNet 中出现的频率来生成概念权重,改进基于距离的语义相似度算法,使相似度计算结果更加符合实际,匹配结果更加准确。文献^[15]针对传统方法计算相似度时基于同一本体和同一度量方法所带来的准确度低的问题,提出一种支持多领域本体的语义相似

度计算方法,从而提高匹配的准确率。文献^[16]则对概念相似度算法进行扩展,使用 WordNet、常识和人的直觉来计算句子之间的相似度。

综上所述,基于推理的匹配方法精度相对较低,出现匹配程度相同结果的几率大,基于相似度的匹配算法则能对 Web 服务进行更准确的区分,而分阶段的匹配思想考虑的影响因素更为全面。在匹配效率方面,基于推理的服务匹配实现比较简单,可在匹配之前对本体库进行预处理,获得线性时间复杂度;基于相似度计算的服务匹配算法本身比较复杂,实现难度大,又不能进行预处理,致使在最坏的情况下算法的复杂度很高,严重影响匹配阶段的性能。因此,如何借鉴各类方法的优点,实现高效精准的服务发现,是本书的研究重点之一。

3) 服务选择与服务组合

服务选择与服务组合针对单个 Web 服务无法满足服务请求者复杂需求的情况,从服务发现产生的若干候选服务集中分别筛选出原子服务,将这些服务按特定逻辑流程组合起来,以最优的大粒度组合服务形态达到用户需求。

随着 Web 服务应用和技术的发展,存在大量提供相同或相似功能的服务,单纯基于功能的服务选择显然不能达到最优服务组合的目的。因此,基于 QoS 的 Web 服务选择问题受到广泛关注,研究成果集中在服务或服务组合的 QoS 属性模型、QoS 计算方法和服务选择算法的构造等方面。现有的 Web 服务体系结构未涉及 QoS 描述,因此不支持基于 QoS 的 Web 服务选择。在 QoS 模型方面,文献^[17]提出一种 UDDI 扩展模型,Web 服务提供者同时提供功能信息和 QoS 信息,生成包含服务质量描述的注册文档,并将 QoS 确认成功的服务注册到 UDDI 中心。上述方案充分利用现有协议框架,易于实现,但未考虑服务请求者对服务 QoS 的评价。在计算模型和选择算法方面,文献^[18~21]对 OWL-S 进行扩展,对具

有相似功能的候选服务进行最优 QoS 筛选,具有可扩展性和可重用性。文献^[22]针对传统基于 QoS 感知的服务选择执行效率低的问题,提出一种可生成局部最优解的新算法,该算法能随着执行时间的增加不断自动优化解决方案。

在服务组合领域,业务流程驱动的 Web 服务组合方法以工作流技术为基础,为业界所推崇,如组合服务执行引擎 BPEL4WS 和 BPWS4J,其描述能力和功能非常强大,但服务协同过程复杂且需要人工参与。因此,学术界对服务的自适应和动态组合问题进行了深入研究,也称即时任务求解的 Web 服务组合方法。该方法求解用户即时提交的需求,根据用户的服务需求,动态地从服务库自动选取若干服务进行组合。上述研究通常主要包含基于人工智能理论的服务组合和基于图搜索的服务组合两类。文献^[23,24]提出基于智能 Agent 和 DAML-S 的服务自动组合方法,提供用户约束和可重用的高层过程并以此实现自动化服务组合。中国科学院的史忠植等对基于主体的语义 Web 服务自动组合进行了研究,提出基于多 Agent 的 Web 服务组合平台架构,并给出了服务的自动组合算法^[25]。总体来说,目前基于人工智能理论的服务自动组合技术使用难度较大,当规划空间规模很大时,这类方法的复杂度会急剧上升。文献^[26]中基于参数推导图自动构建 Web 服务组合流程,并自动将参数推导图转换成以 Web 服务为节点、以服务间依赖关系为边的服务组合流程图。文献^[27]则提出了一种多 QoS 属性约束下的服务组合模拟退火算法,通过对组合服务进行有向图建模,将最优服务组合问题转化为在有向图中搜索最优多约束路径问题。基于图搜索的方法不需要特定的表现形式或推理系统,易于实施,但当可用服务数量很大时时间复杂度高。因此,如何缩小搜索空间、提高服务和组合服务的重用率以实现快速组合是其要解决的关键问题^[28~30]。

服务选择与服务组合问题是 Web 服务集成的关键和难点所在,相关研究目前多处于探索阶段^[31~38],缺乏公认的标准和成熟的技术^[39]。本书根据勘察设计企业服务集成的需求,对已有研究成果^[40,41]进行整合和改进,提出一种基于任务分解的服务选择与自适应服务组合机制,有效支持企业 SOA 信息框架模型的构建。

1.3 主要研究内容

构建基于 SOA 的勘察设计企业信息框架模型,实现勘察设计企业的服务集成,关键在于服务描述、服务发现、服务选择和组合各阶段统一建模,合理选择和设计算法,最终形成满足需求的服务集成方案。本书的主要研究内容如下:

1) 建立统一的语义描述模型

建立统一的语义描述模型是首先要解决的基本问题。WSDL 是业界广泛认可的描述 Web 服务功能的协议,有成熟的技术框架作支撑,但仅从功能和语法层面描述 Web 服务,缺乏语义信息。OWL-S 是描述 Web 服务语义的主流规范,支持自动服务发现、组合、调用、互操作和执行监控,但没有提供相应的解决方案。同时随着 Web 服务应用的深入,服务数量激增,单纯功能性描述已不能适应服务优化的要求,QoS 作为描述 Web 服务非功能性属性的主要指标集,是服务选择和服务组合的重要依据。基于上述分析,本书建模的主要思想是:用 WSDL 表示服务的具体描述,用 OWL-S 表示服务的抽象描述,同时支持 WSDL 和 OWL-S 的匹配与发现,并在 OWL-S 功能语义描述的基础上增加服务 QoS 的语义描述本体,为服务选择和服务组合提供支持。