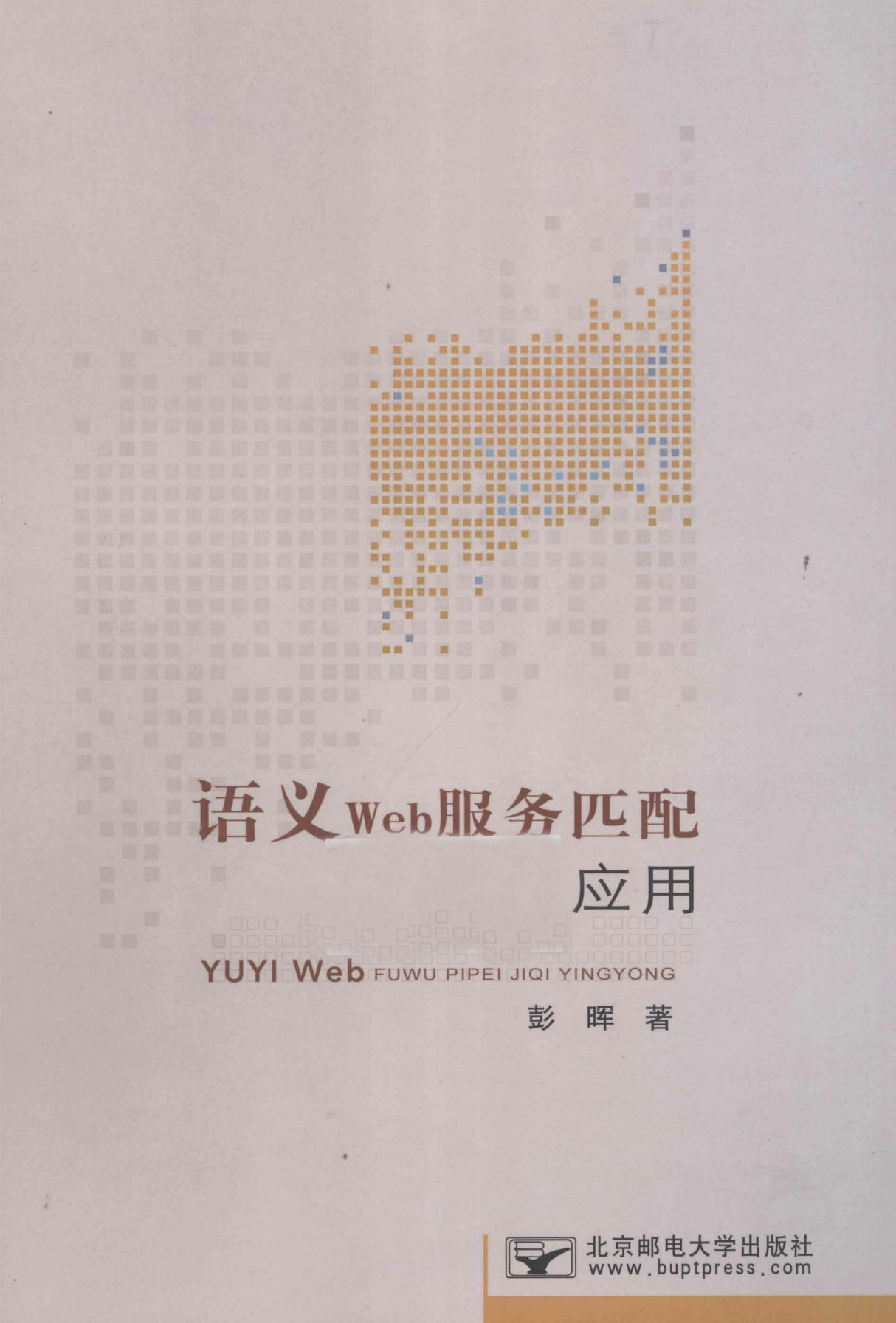




语义Web服务匹配 应用

YUYI Web FUWU PIPEI JIQI YINGYONG

彭 晖 著



北京邮电大学出版社
www.buptpress.com

语义 Web 服务匹配及其应用

彭 晖 著



·北京邮电大学出版社
[www. buptpress. com](http://www.buptpress.com)

内 容 简 介

本书系统介绍了语义 Web 服务研究的相关概念、研究现状、最新进展、语义 Web 服务匹配算法及语义 Web 服务组合的策略。在理论研究的基础上,以旅游 Web 服务和网络教学 Web 服务为例,详细介绍了语义 Web 服务匹配与组合的方法和过程,介绍了语义服务开发平台和工具。本书共 7 章,内容分别是语义 Web 服务的相关基础知识、语义 Web 服务描述、基于本体概念语义相似度的 Web 服务匹配、基于逻辑表达式的 Web 服务匹配算法、综合 Web 服务组合策略和应用实例。

本书可作为从事语义 Web、Web 服务等方面研究的研究人员以及对此感兴趣的读者的研究参考书,也可以作为高等院校计算机、自动化、电子信息、系统工程等专业的本科或研究生教材。

图书在版编目(CIP)数据

语义 Web 服务匹配及其应用 / 彭晖著. -- 北京:北京邮电大学出版社, 2015.1

ISBN 978-7-5635-4252-9

I. ①语… II. ①彭… III. ①网络编程—研究 IV. ①TP393.09

中国版本图书馆 CIP 数据核字(2014)第 295684 号

书 名: 语义 Web 服务匹配及其应用

著作责任者: 彭 晖 著

责任编辑: 刘春棠

出版发行: 北京邮电大学出版社

社 址: 北京市海淀区西土城路 10 号(邮编: 100876)

发行部: 电话: 010-62282185 传真: 010-62283578

E-mail: publish@bupt.edu.cn

经 销: 各地新华书店

印 刷: 北京联兴华印刷厂

开 本: 720 mm×1 000 mm 1/16

印 张: 10

字 数: 192 千字

版 次: 2015 年 1 月第 1 版 2015 年 1 月第 1 次印刷

ISBN 978-7-5635-4252-9

定 价: 25.00 元

• 如有印装质量问题,请与北京邮电大学出版社发行部联系 •

前 言

在开发大规模分布式应用的过程中，如何有效地集成和共享已有的软件和本地计算、提高系统的开发效率、降低系统的开发成本，一直是分布式系统研究和关注的重点问题。为了解决这一问题，面向服务的计算（Service Oriented Computing, SOC）和面向服务的架构（Service Oriented Architecture, SOA）将软件封装成服务，应用系统通过服务组合和调用来完成大规模跨平台系统的开发。这种新的业务集成和软件开发模式，以其标准化、跨平台、透明性等特点，已成为标识分布式系统和软件集成等方向技术进步的一个新的里程碑。

Web 服务作为面向服务计算的一种最主要的实现技术，其特点是它能被标准的 Internet 应用层协议调用，可以很方便地集成到各种 Web 应用系统中。Web 服务的基本架构是服务提供者、服务使用者和服务代理。服务提供者实现服务，描述服务的接口并将它发布到服务代理端；服务使用者查询和调用服务；服务代理提供服务注册，帮助服务提供者发布服务，帮助服务使用者查找所需的服务。如服务注册机制 UDDI（Universal Description, Discovery and Integration）即完成服务代理功能。Web 服务具有完好的封装性，松散耦合，并采用标准的 Internet 协议进行通信，使得 Web 服务成为分布式计算的新模式，并在工业界得到广泛应用。

Web 服务的大量涌现对服务发现提出了挑战，UDDI 上基于关键词和简单分类的服务发现机制已经不能很好地满足需要，怎样在大量已发布的服务中快速、准确和全面地发现满足服务使用者需求的服务成为面向服务的计算中的关键问题和难点问题。语义 Web 服务匹配旨在通过定义标准的服务描述语言，结合领域本体，准确全面地描述已有服务和服务需求的信息，根据服务供需双方的语义描述，采用适当的匹配方法来判断双方描述的匹配程度，返回给服务使用者最合适的服务。

第 1 章介绍了语义 Web、语义 Web 服务、语义 Web 服务匹配等概念及背景

知识,接下来总结了语义 Web 服务的研究现状,分析了语义 Web 服务匹配中需要解决的关键问题。第 2 章介绍了 Web 服务的语义描述。第 3 章介绍了 Web 服务的相似度匹配算法。第 4 章介绍了基于动态描述逻辑的 Web 服务匹配算法。第 5 章是综合的服务组合策略。第 6 章介绍了 Web 服务匹配与组合平台的开发与实现。第 7 章通过网络教学 Web 服务组合实例,补充说明了基于动态描述逻辑的 Web 服务匹配方案的特点和应用方式。本书第 2~6 章都采用了旅游 Web 服务的实例来介绍具体的服务描述、匹配与组合方法。

本书的主要内容及特点如下。

(1) 提出了包含语境信息的 Web 服务需求语义描述模型。现有的语义 Web 服务描述语言不包含服务使用方的语境信息,本书提出了包含服务使用方语境信息的服务需求描述模型,将用户的语境信息用静态语境信息和动态语境信息来描述,从不同的侧面描述用户的需求及其特征,以便对需求的服务以及需求方本身的特征做出尽可能详细的描述,在充分挖掘用户显式的或隐式的需求的基础上,获取与需求最相匹配的服务。本书第 2 章在包含语境信息的服务需求描述与发布的服务语义描述之间建立各部分的匹配关系,并在后续的章节中采用匹配算法判断各对应部分之间的匹配程度。

(2) 提出基于本体概念语义相似度的 Web 服务匹配方法。服务描述语言提供服务描述的上位本体,约定服务中需要描述的信息,而领域本体则对上位本体中约定的信息做具体的说明与描述。本书根据领域本体概念之间的包含关系,定义了本体概念之间的非对称相似度,利用这种非对称的相似度可以计算两个服务的匹配程度。与经典的 OWL-S/UDDI 算法比较,本书的算法通过实数的取值细分了 OWL-S/UDDI 算法的等级,与通常的本体概念语义相似度的算法比较,本书分析了本体概念语义相似度与服务相似度的区别与联系,定义了符合服务匹配等级的相似度计算方法。

(3) 提出基于动态描述逻辑的 Web 服务匹配方法。动态描述逻辑 (Dynamic Description Logics, DDL) 是描述逻辑的一种动态扩展,支持语义 Web 环境下对动作的描述和推理。Web 服务可以看作是一个动作,它的执行可以改变 Web 信息和状态空间。根据 Web 服务的这一特点,可以将 Web 服务描述为 DDL 的动作,基于动态描述逻辑的语义 Web 服务匹配方法将服务需求描述为 DDL 的目标公式,将每个原子服务描述为一个 DDL 动作,将服务匹配问题归结为 DDL 公式的可满足性判定问题。利用 DDL 对动作的推理功能,采用 DDL 推理机对公式可

满足性的判定,发现与目标服务相匹配的原子服务组合。

(4) 在理论研究的基础上,结合语义 Web 服务匹配流程,设计和开发了语义 Web 服务匹配器,采用综合的匹配算法,将服务的语法匹配和语义匹配相结合,提高服务匹配的质量和效率。修改已有的语义 Web 服务组合原型系统,将服务匹配器嵌入语义 Web 服务开发平台 (Semantic Web Service Broker, SWS-Broker) 中,设计和实现了服务匹配双方语义的自动标注、语义 Web 注册、语义 Web 匹配器这三个主要功能。同时,语义 Web 服务匹配器嵌入了本体知识管理系统 (Knowledge Management Sphere, KMSphere, 用于本体的编辑和管理、服务语义的自动标注) 和 DDL 推理机 (用于公式的可满足性推理),整个系统初步实现了以领域本体描述的服务匹配,系统运行结果表明了本书工作的可行性和有效性。

本书有关研究得到北京市教委科技规划课题 (编号: SQKM201410031001) 的资助,北京邮电大学出版社的领导和编辑对本书的出版给予了大力支持,作者在此深表谢意。

本书的编写得到了史忠植老师的悉心指导,在此,对史老师表示衷心感谢。本书第 5 章由作者和邱莉榕老师共同研究讨论,邱莉榕老师执笔,作者进行修改并添加旅游服务组合模板的内容,在此,作者对邱莉榕老师表示衷心的感谢。感谢常亮、陈立民、姚二林在论文撰写时提供的宝贵意见和帮助。

由于作者水平有限,书中难免有不足之处,恳请读者指正。

彭 晖

北京第二外国语学院

2014 年 6 月

目 录

第 1 章 绪论	1
1.1 语义 Web 服务概述	1
1.2 语义 Web 服务的研究背景及相关概念	2
1.2.1 Web 服务	2
1.2.2 语义 Web	6
1.2.3 语义 Web 服务	11
1.3 语义 Web 服务研究综述	15
1.3.1 语义 Web 服务匹配项目	15
1.3.2 Web 服务组合项目	18
1.3.3 语义 Web 服务匹配方法	21
1.3.4 语义 Web 服务匹配研究中的关键问题	22
1.3.5 主体技术在语义 Web 服务中的应用	23
1.4 本书研究内容	25
本章参考文献	28
第 2 章 语境相关的 Web 服务描述	36
2.1 服务匹配流程和服务描述语言	36
2.1.1 Web 服务匹配流程	36
2.1.2 已有的 Web 服务描述语言分析与比较	38
2.2 语境信息相关概念	40
2.2.1 语境敏感计算	40
2.2.2 语境信息类型	40
2.2.3 语境信息描述	41

2.3 语境相关的 Web 服务描述	42
2.3.1 Web 服务使用方对服务的描述	43
2.3.2 发布的 Web 服务与需求的 Web 服务各部分之间的对应关系	47
2.3.3 语境信息在 Web 服务匹配中的作用	48
2.4 相关工作	50
2.5 小结	51
本章参考文献	51
第 3 章 基于本体概念语义相似度的 Web 服务匹配	54
3.1 Web 服务参数的匹配	54
3.1.1 Web 服务功能参数的语义描述	54
3.1.2 Web 服务功能参数之间匹配程度的度量	56
3.2 本体概念语义距离和语义相似度	58
3.2.1 语义距离和语义相似度的定义	58
3.2.2 Web 服务参数相似度	61
3.3 基于服务参数语义相似度的 Web 服务匹配算法	62
3.3.1 基于服务参数相似度的 Web 服务匹配算法介绍	62
3.3.2 基于服务参数相似度的 Web 服务匹配实例	65
3.3.3 算法性能分析与测试	68
3.4 相关工作	70
3.5 小结	72
本章参考文献	72
第 4 章 基于动态描述逻辑的 Web 服务匹配	74
4.1 用动态描述逻辑定义 Web 服务的基本思想	74
4.2 语义 Web 服务的逻辑基础	75
4.2.1 描述逻辑	75
4.2.2 动态描述逻辑	77
4.2.3 动态描述逻辑中的推理	79
4.3 基于 DDL 的语义 Web 服务描述	81
4.3.1 用 DDL 描述语义 Web 服务	81

4.3.2 用 DDL 描述服务匹配	81
4.4 基于 DDL 的语义 Web 服务匹配算法	82
4.4.1 基于 DDL 的服务匹配算法	82
4.4.2 Web 服务匹配实例	83
4.5 相关工作	86
4.6 小结	88
本章参考文献	89
第 5 章 语义 Web 服务的组合策略	92
5.1 半自动组合策略	92
5.1.1 服务组合体系	92
5.1.2 服务虚拟化方法	94
5.1.3 领域模板	95
5.1.4 基于语境信息的模板优化	101
5.2 基于主体的服务虚拟化	101
5.2.1 虚拟服务分类	101
5.2.2 主体虚拟服务描述	103
5.2.3 主体虚拟服务的形式语义	107
5.3 物理服务的注册和注销	111
5.3.1 物理服务注册为主体虚拟服务	111
5.3.2 物理服务注册举例	113
5.4 相关工作	116
5.5 小结	117
本章参考文献	117
第 6 章 语义 Web 服务匹配器的设计与实现	121
6.1 语义 Web 服务开发平台 SWSBroker 介绍	121
6.2 SWSBroker 的功能模块	123
6.2.1 Protégé 本体编辑器	124
6.2.2 知识管理模块	125
6.2.3 DDL 推理机	126

6.2.4 Web 服务匹配器	128
6.2.5 图形化的语义标注模块	133
6.2.6 匹配方案编辑与修改模块	136
6.3 语义 Web 服务开发平台 SWSBroker 的安装及使用	136
6.4 小结	138
本章参考文献	138
第 7 章 面向教学过程的语义 Web 服务组合实例	140
7.1 教学过程的定义及本体	140
7.1.1 教学过程的定义	140
7.1.2 教学过程语境本体的定义	141
7.2 教学过程 Web 服务的功能定义	143
7.3 教学过程 Web 服务组合实例	145
7.4 小结	147

第1章 绪论

1.1 语义 Web 服务概述

在面向服务的计算中,软件被封装成服务,采用约定的协议被调用,以此将服务的提供方和服务使用方有效分离,有效地集成和共享已有的软件和本地计算,提高分布式系统的开发效率,因而面向服务的计算(Service Oriented Computing, SOC)作为一种新兴的技术给开放环境下的分布式应用集成问题的解决带来了曙光^[1~3]。

Web 服务作为当前最主要的一种服务实现技术,旨在为 Internet 上跨越不同地域、不同行业的应用提供更强大的交互操作能力。Web 服务依托一系列开放的协议和标准,可以方便地被使用者调用,它将 Web 应用从信息交互的领域扩展到了服务交互的领域,使得基于服务来解决广域网络范围内跨平台的异构问题成为可能。

近年来,Web 服务的大量涌现对服务发现提出了挑战,怎样在大量已发布的服务中快速、准确和全面地发现满足服务使用者需求的服务成为面向服务的计算中的关键问题和难点问题。Web 服务匹配是匹配器(或称匹配程序、算法)根据需求方的描述发现和选择满足要求的服务的过程,也就是服务匹配器根据 Web 服务的描述,确定服务使用方描述的服务与服务提供方描述的服务的匹配程度问题。目前,Web 服务的注册机制 UDDI^[4]上的服务匹配是通过对 UDDI 上的服务注册信息进行关键词精确匹配实现的,主要是对服务 ID 或名称,或是服务的有限的属性值进行匹配^[5]。但如同使用搜索引擎一样,人们在感谢 UDDI Web 服务注册中心带来的寻找 Web 服务的便捷的同时,也常常为查准率^①和查全率^②^[6,7]不高所困扰。主要原因在于,基于关键词匹配的 Web

① 查准率:指查询结果中正确的结果占全部查询结果的比率。

② 查全率:指查询结果中正确的结果占全部正确结果的比率。例如有 100 个服务,其中有 40 个与需求是匹配的,某次查询查到 50 个匹配的服务,其中 30 个正确,20 个不正确,则查准率 = $30/50 = 60\%$,查全率 = $30/40 = 75\%$ 。

服务发现具有以下缺陷^[8]。

(1) 对所需查询的目标的语义不能准确描述。因为服务的语义很丰富,仅仅几个关键词很难准确而全面地描述一个服务。

(2) 不能度量候选者和查询目标间的符合程度。

(3) 不能使用细化、泛化、平级扩展等语义操作进行查询。

其中前两点是影响查准率的重要因素,第三点主要影响查全率。怎样在现有服务描述中加入服务的语义信息,通过服务语义的匹配来准确地查找服务成为关注的焦点,而语义 Web 的出现成为解决这一问题的一种途径。

语义 Web 的核心是本体,本体是领域共享概念的规范的形式化描述^[9],共享意味着为概念提供统一的语义,规范意味着在概念上的语义约束是清晰的,形式化意味着便于机器理解和推理^[10]。将本体的概念和相应技术引入 Web 服务技术中,主要是用本体描述语言丰富的语义表达能力来描述 Web 服务的功能、性能等各方面的信息,并在这些描述的基础之上实现语义推理,实现服务匹配的准确性及服务组合的智能性。

本章首先介绍 Web 服务技术及相关基本概念,讨论当前 Web 服务匹配中存在的问题和困难;接着介绍语义 Web 技术及研究进展,明确将语义技术与 Web 服务技术相结合解决 Web 服务匹配问题的研究方向,在此基础上给出本书解决语义 Web 服务匹配的基本思路,概述相关研究现状;最后以上述内容为基础说明作者的研究思路、研究内容及本书的组织结构。

1.2 语义 Web 服务的研究背景及相关概念

1.2.1 Web 服务

从 20 世纪 90 年代到 21 世纪初,Web 应用主要围绕 Internet 上的数据展开。在几乎所有的应用中,这种面向数据的应用总是由使用者通过使用 Web 浏览器等工具的方式完成。XML 技术以及其他相关技术的快速发展改变了 Web 应用发展的格局:它将 Web 应用从信息交互的领域扩展到了服务交互的领域,这种服务交互就是当前引起工业界和学术界关注的 Web 服务^[11]。

下面我们围绕面向服务计算的基本架构,了解 Web 服务的起源及相关概念。

1. 服务

服务,代表了给用户提供的功能,以帮助用户解决其问题。在面向服务的计算中,服务是指具有规范封装的软件或资源,可以对它进行规范地使用。目前服

务的具体形式有 Web 服务^[11]、网格服务^[12,13]等。为了更好地理解服务的概念,我们给出关于服务的一些权威定义。

(1) 第一届面向服务的计算国际会议上的定义^[1]

服务是一类自治的、平台独立的计算元素,采用基于 XML 的方式可以对其进行描述、发布、发现及组合,从而能够支持大规模分布式应用的构造。

(2) W3C 的定义^[14]

从服务提供者和服务使用者角度来看,服务是由不同服务提供者面向不同服务使用者提供的一组遵循标准定义的操作。

(3) Steve Graham 等人的定义^[15]

服务是独立于特定平台及实现的软件组件,它能够被一种服务描述语言描述,在一个服务注册库发布,通过标准的机制发现,基于网络调用以及与其他服务组合。

服务包括原子服务和组合服务,原子服务指单个的、没有调用其他服务的简单服务,而组合服务是指通过组合分立的、相对简单的、已经存在的服务来构造复杂的具有新功能的服务。

(4) Chris Preist 的定义^[16]

服务是提供具有某种特定功能的软件实体,它包括三个方面的内容:首先,服务提供的功能是具有价值的,用户为其价值付费;其次,服务包括服务提供方和用户的在线交互过程;最后,服务的提供方能将服务交付需求方。

2. 面向服务架构(Service Oriented Architecture, SOA)

面向服务的计算技术^[17]对分布式应用集成所带来的最明显的好处可以充分体现在面向服务架构中^[18]。在 2002 年, Gartner 组织^[19]就指出面向服务架构将是“现代应用开发领域最重要的课题”,并预计面向服务架构将成为占有绝对优势的软件工程实践方法。

面向服务架构的核心概念集中体现在如图 1.1 所示的服务提供者(Service Provider)、服务使用者(Service Requestor)、服务代理(Service Broker)三个角色和发布(Publish)、查找(Find)、绑定(Bind)三个基本操作上^[20]。在面向服务的架构中,服务提供者负责实现服务,描述服务的接口,并将服务描述发布给服务代理;服务代理作为服务提供者和服务使用者间的中介,负责维护一个服务描述注册中心,以管理服务提供者发布的服务描述以及根据服务使用者的服务请求找到合适的服务;服务使用者通过服务代理查找需要的服务,并得到相应的服务描述,根据服务描述可以与服务提供者建立绑定关系,完成服务的调用。其中,服务实现是指真正完成服务功能的程序体,它由服务提供者负责编写;服务描述是对服务实现的接

口、访问地址及协议等方面内容的描述;服务请求是由服务使用者提出的、对所期望使用的服务的需求描述。

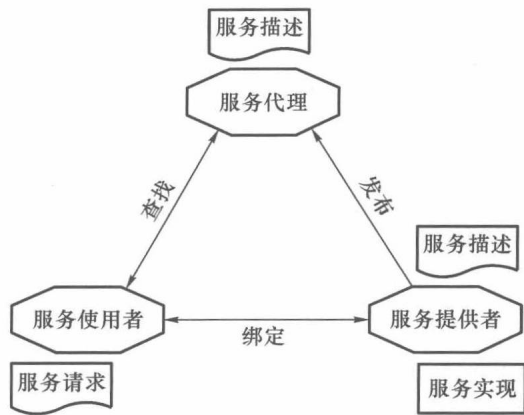


图 1.1 面向服务架构

在面向服务架构下,通过发布、查找、绑定三个基本操作允许在服务使用者和服务提供者间保持一种松散的绑定关系,使得服务提供者对服务使用者具有位置、实现等多方面的透明性,便于在二者间建立动态的绑定关系,从而提供了一种松耦合的编程模型,具备了支持动态应用集成的潜力。

3. 服务匹配

在面向服务的架构下,服务发现、服务匹配、服务组合是几个密切相关的概念,它们互相联系又各有侧重,下面我们给出其定义并说明它们之间的关系与区别。

服务发现是服务代理方根据服务需求方的请求,从已发布的服务中查找满足需求的原子服务或组合服务的过程。服务匹配是在服务发现的过程中,匹配算法或匹配程序根据一定的规则判断需求的服务与发布的某个原子服务或组合服务的匹配程度。服务组合是当单个服务不能满足应用的需求时,通过组合分立的、相对简单的、已经存在的服务来构造复杂的具有新功能的应用的方式。服务发现、匹配与组合三个概念的内容有很多重叠之处,它们的根本目标都是从现有的服务中找到能够满足终端用户需求的服务,可以说服务发现的过程也就是服务匹配和组合的过程,服务匹配主要是强调匹配算法判断服务的相似程度,而服务组合主要强调将多个服务构成一个整体时,服务之间的约束能否满足,服务之间的交互能否顺利完成,服务组合能否实现最终目标等。

4. Web 服务(Web Service, WS)

Web 服务是当前最主要的一种服务实现技术。最初 Web 服务是由 Ariba、

IBM 和 Microsoft 等共同提出的,旨在为 Internet 上跨越不同地域、不同行业的应用提供更强大的互操作能力。W3C 对 Web 服务给出了如下的定义:Web 服务是由一个 URI 标识的软件应用,它的接口和绑定方式可以基于 XML 来定义、描述及发现;基于当前的 Internet 协议,Web 服务支持以 XML 格式的消息交换方式与其他软件进行直接交互^[11]。

Web 服务之所以能被有效地封装、集成和调用,离不开一系列开放和标准的协议。下面我们介绍与 Web 服务相关的主要协议。

(1) 简单对象访问协议 (Simple Object Access Protocol, SOAP)

简单对象访问协议是用于交换 XML 编码信息的协议^[21]。SOAP 包括三个主要部分:SAOP 封包,定义了 SOAP 消息表示框架,如消息的发送方、接收方、消息内容等;SOAP 编码规则,定义了消息的编码格式;SOAP 远程过程调用,定义了用于表示远程调用和响应的约定,例如,如何使用 HTTP 调用按照 SOAP 定义的消息等。

(2) Web 服务描述语言 (Web Services Description Language, WSDL)

Web 服务描述语言是描述服务的接口和调用方式的 XML 文件,用于服务的发布、查找和绑定。在 WSDL 2.0 版本中^[22],用六个元素描述 Web 服务,如图 1.2 所示。这六个元素是 Messages、Operation、Interface、Endpoint、Binding 和 Service。其中 Messages、Operation 和 Interface 描述服务的抽象信息,Messages 描述了服务可以发送和接收的信息,Operation 定义了信息之间交互的形式,不同类型的 Operation 分组构成 Interface。服务的调用信息由 Endpoint、Binding 和 Service 描述,Binding 定义了 Interface 中信息的传输方式,Endpoint 是网络中的传输端点,实现同一个 Interface 的 Endpoint 构成了 Service^[23]。

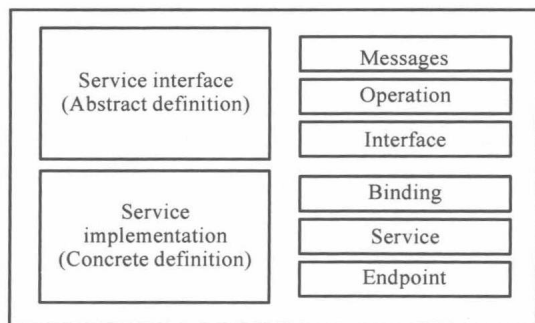


图 1.2 WSDL 2.0 服务描述

在 WSDL 中包含了使用 SOAP 的服务描述的绑定,也包含了使用简单 HTTP

GET 和 POST 请求的服务描述的绑定。

(3) Web 服务注册标准和规范(UDDI^[8])

它使用 XML 文档来描述企业和企业注册的服务。UDDI 注册信息分为白页、黄页和绿页。白页描述了企业的地址、联系方式等企业信息,黄页是基于标准分类的信息,绿页中存放 Web 服务的技术信息。

正是上述被广泛接受的开放标准使得 Web 服务成为当前最主要的服务实现技术,同时也使得基于服务来解决广域网络上应用的异构问题成为可能。在本书中 Web 服务将被用作服务概念的具体实现技术。如不作特殊说明,以下本书中的“服务”均特指“Web 服务”。

随着 Web 2.0 的广泛应用,以及相关技术、模式上的进步与不断成熟,Web 服务不像初始定义的那样仅用于企业应用的集成,Web 服务合成也开始从面向专业用户转变到面向普通的端用户^[24],产生了像 MASHUPS 这样的大众化服务合成工具。任何用户都可以开发和发布自己的服务,或者将 Web 上的开放服务组合以提供新的增值服务。Web 开始成为一个巨大的、开放的可编程平台^[25],在用户需求多样化的驱动下,Web 服务的形式也呈现多样化的趋势。除了标准的基于 WSDL 的 Web 服务,包括 RESTful API 在内的各种 Web API 也开始成为 SOA 中新的服务载体^[26]。

1.2.2 语义 Web

所谓“语义”就是文本的含义,与之对应的概念有语法、语用。语义 Web 是 Web 的延伸,语义 Web 中的资源、数据和链接方式具有良好的规范的定义,以利于信息与知识的发现、自动处理、集成和重用。语义 Web 的目标是使 Web 资源可以被机器理解 and 处理,能够更好地支持人机协同工作。

1983 年 3 月 Web 的创始人 Tim Berners-Lee 提出分布式超文本系统图,其中所包含的链接都是蕴含语义的,该图最早提出了 Web 信息包含语义的思想。1994 年 5 月 第一届国际 WWW 会议上, Tim Berners-Lee 在会议报告中特别强调了语义的重要性。1997 年 12 月,他在 W3C 报告中指出, Web 的第一个目标是成为人们通过知识共享进行交流的媒介,第二个目标是成为人们协同工作的媒介,这为语义 Web 做了铺垫。1998 年 9 月 Tim Berners-Lee 在文献中提出了语义 Web 的总体框架,并概括介绍了它的各个组成部分。2000 年 12 月, Tim Berners-Lee 在 XML 2000 会议报告中首次提出了语义 Web 功能逐层增强的层次结构图,指明了语义 Web 的研究框架。2001 年 5 月 Tim Berners-Lee 在 *Scientific American* 中系统地给出了语义 Web 的概念与相关技术,至此可以说

语义 Web 正式出现^[27]。W3C 组织是目前语义 Web 的标准化组织,它在语义 Web 的标准、规范的设计制定和开放式、合作式技术研发中起着指导性的作用,对非集中研究和开发的关键技术进行指导和标准化,以确保它们的协调发展。

W3C 提出的最新的语义 Web 的层次模型(亦称协议栈)如图 1.3 所示^[28]。

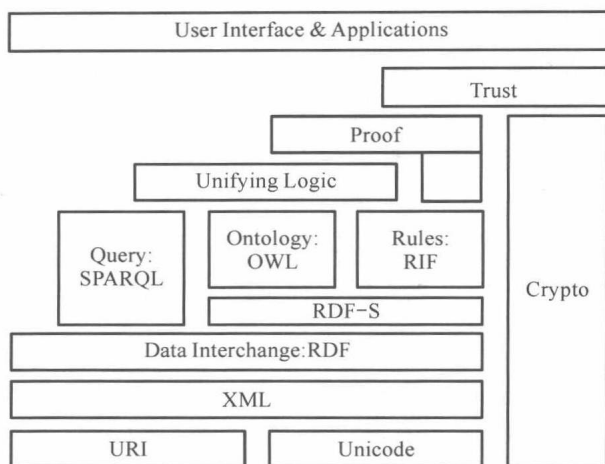


图 1.3 语义 Web 的层次模型

从图 1.3 中可以看出语义 Web 发展的层次结构。第一层是统一编码(Unicode)和统一资源标识符(Uniform Resource Identifier, URI^[29]),它是整个语义 Web 的基础,Unicode 处理资源的编码,URI 负责资源的标识。第二层是可扩展标记语言(eXtensible Marked Language, XML^[23]),用于表示数据的内容和结构。第三层是资源描述框架(Resource Description Framework, RDF^[30]),用于描述资源及其相互关系。第四层是 RDF 模式(RDF-Schema, RDF-S^[31]),为 RDF 提供了类型定义机制,确定了 RDF 所描述的资源所使用的领域词汇。第五层是本体(Ontology)和规则(Rules),本体用于描述领域概念以及概念之间的联系;规则用于描述领域知识中的前提和结论;本体和规则构成领域知识层。本体描述语言(Web Ontology Language, OWL^[32])和规则交换模式(Rule Interchange Format, RIF^[33])是常用的本体和规则描述语言。覆盖第四层、第五层的查询语句 SPARQL 提供了对 OWL 和 RIF 文件的查询功能。第六层是统一的逻辑(Unifying Logic),在下面五层的基础上进行逻辑推理操作。第七层是验证(Proof),根据逻辑陈述进行验证以得出结论。第八层是信任(Trust),在用户间建立信任关系。第九层是应用层(User Interface & Applications),是构建在语义 Web 之上的各种应用。概括地说,最下面两层是语义 Web 的基础设施,中间从元数据发展到本体描述语言及其统一的逻辑是语义 Web 的关键研究所在,上面的证明和信任及各层次