

基于抽象服务模型提高 Web 服务匹配的适应性

陈旺虎^{1,2,3} 李 婧¹

(西北师范大学数学与信息科学学院 兰州 730070)¹

(中国科学院计算技术研究所 北京 100190)² (中国科学院研究生院 北京 100039)³

摘要 为提高 Web 服务匹配方法对业务需求变化性和服务差异性的适应性,提出一种抽象服务模型以及该模型驱动的 Web 服务匹配方法。模型通过对多样性和具有差异性的服务的统一描述,刻画了影响服务具体能力的接口和行为属性,可基于模型检验思想,有效支撑适应性的服务匹配。本方法已应用于构建面向科学计算领域的问题求解环境。实验和应用表明基于本模型可以提高服务匹配的适应性。

关键词 服务模型,服务匹配,模型检验

Enabling Adaptive Service Matchmaking with Abstract Service Model

CHEN Wang-hu^{1,2,3} LI Jing¹

(College of Mathematics and Information, Northwest Normal University, Lanzhou 730070, China)¹

(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China)²

(Graduate School of Chinese Academy of Sciences, Beijing 100039, China)³

Abstract To improve the adaptability of the approach to service matchmaking, an abstract service model called ASM was proposed. ASM described service interfaces and behaviors, and thus addressed important factors affecting service capabilities. The adaptive approach to service matchmaking based on ASM were also analyzed. The model and the approach have been partly applied to constructing problem solving environments for bioinformatics researches. Analysis and applications show ASM is effective to enable adaptive service matchmaking.

Keywords Service model, Service matchmaking, Model checking

1 引言

在科学计算领域,存在大量的网络共享 Web 服务。以生物信息领域为例,目前网络上可供使用的 Web 服务就多达 3000 个以上^[1]。同时,仅使用 Blast 工具进行序列相似性查找的 Web 服务通过搜索引擎 Seekda 可检索到 123 个,这个数量还在不断增加。另一方面,许多生物信息实验需要依赖于基因测序仪和基因数据库等稀缺资源。Web 服务为网络环境下的资源共享提供了良好契机,因此使用网络共享 Web 服务构建问题求解环境(Problem Solving Environment, PSE)变得很有必要。

图 1 给出了一个 PSE 系统的概念框架。领域分析人员通过分析本领域基本业务功能,提出本领域的基本业务需求单元。该业务需求单元可作为服务使用者表达业务需求的基本元素,并形成了应用构建的编程元素。业务需求单元最终被关联到具体的 Web 服务,实现真正的业务功能。该过程的关键在于给定一个业务需求,从当前的 Web 服务中查找出满足指定业务需求的服务,这个过程被称为服务匹配。在服务匹配过程中,如果单个的服务无法满足业务需求时,需要通过

多个服务的组合来构造满足需求的组合服务。

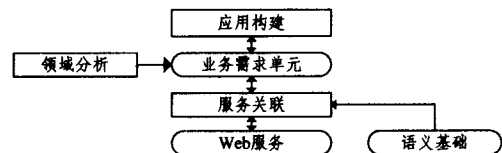


图 1 一个 PSE 系统的概念框架

服务匹配已经成为服务计算领域的一大热点。文献[2, 3]等基于词法匹配技术或者语义推理技术提出了服务相似性的判断方法。文献[4, 5]等基于语义推理,提出了基于目标服务匹配方法。文献[6, 7]等基于 PI 演算和有限状态自动机,从过程互模拟的角度对服务匹配进行了探讨。

然而,根据对生物信息实验和几个权威机构所提供的 Web 服务的分析,在服务匹配的过程中,如下两个因素非常值得关注:

(1)业务需求的变化性:业务功能是表示业务需求的常见手段,而业务需求的变化性^[8]则指刻画业务功能的业务属性的可选性、业务属性在不同情况下取值的可变性、不同业务属性之间的约束以及相应业务活动的时态等约束的变化性。

到稿日期:2008-10-27 返修日期:2009-05-20 本文受国家重点基础研究发展计划(2007CB310805)资助。

陈旺虎(1973—),男,博士生,讲师,主要研究方向为服务计算与软件集成, E-mail: cwhmail@vip.sina.com; 李婧(1975—),女,硕士,讲师,主要研究方向为软件工程和服务工程。

¹⁾ <http://www.w3.org/Submission/2004/SUBM-OWL-S-related-20041122/#bpel>

(2)服务的多样性和差异性:首先,服务可能被描述为不同的规范(例如 WSDL, WSDL-S^[9], OWL-S^[10], BPEL¹⁾等);其次,完成相同业务功能的服务在接口描述和行为属性方面存在差异性,而这些差异性往往会影响到服务的具体能力。例如,同样是完成序列相似性查找的服务,EBI提供的 WSWU-Blast 服务使用 blast 算法,但该约束是隐含的,并未通过服务接口对外暴露,并且由 4 个操作的协作完成序列相似性查找任务;另外,服务的消息之间可能存在约束;最后,有些服务由单一操作构成,而有些是多操作的协作。

综上所述,如何提高服务匹配,对于业务需求的变化性和服务差异性的适应能力来说是一个有待解决的关键问题。

本文的研究目标正是基于服务的抽象描述模型,提高服务匹配对于业务需求变化性和服务差异性的适应性。主要工作包括:1)提出一种抽象服务模型,该模型具有对差异性服务的统一描述能力,同时,可为服务匹配等推理工作提供良好基础;2)提出了基于该模型的服务匹配方法,方法可提高服务匹配的适应性。

本文第 2 节对本文的研究问题进行了严格定义和分析;第 3 节介绍了本文提出的抽象服务模型,并分析了它的逻辑语义;第 4 节阐述了基于此模型的服务匹配方法;第 5 节结合实验和应用,对本文的方法进行了评价;第 6 节对相关研究工作进行了分析和比较;最后给出本文的结论及其下一步的工作展望。

2 问题定义和分析

定义 1(服务匹配) 给定一个服务 s 以及服务请求 r ,判断 s 是否满足 r 的约束。

对于服务请求,根据匹配的要求不同会有不同的定义,当前的研究大多以服务的输入和输出信息作为主要的描述手段。根据前面的分析,在实际应用中,服务请求通常对应一种业务需求。同时,由于业务需求的变化性,服务请求对于所需服务的要求还会体现在对服务行为属性的约束上。这点主要体现在两个方面,即服务的消息之间的取值约束以及当服务由多操作构成时消息之间所表现出的时态约束。

定义 2(服务匹配的适应性) 服务匹配的适应性是指在服务请求具有变化性以及服务具有差异性的情况下服务匹配方法保持其有效性的能力。

如图 2 所示,为了使得服务的匹配更加适应请求的多样性和服务的差异性,一种抽象的服务统一描述模型是非常重要的。

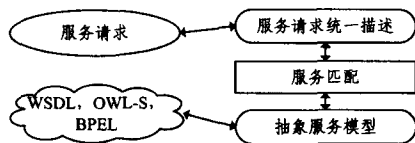


图 2 适应性服务匹配的基本思想

下面将提出一种支撑适应性服务匹配的抽象服务模型,并分析其逻辑语义,探讨基于该模型的服务匹配方法及其效果。

3 抽象服务模型

本文提出了一种抽象服务模型(ASM)。从概念视图的

角度来看,一个服务是由一组操作及其相互交互构成的。

3.1 模型定义

定义 3(ASM 模型) ASM 模型对应一个 3 元组 $ASM = (PF, OP_s, NFP_s)$ 。其中 PF 表示服务的轮廓; OP_s 表示操作的集合; NFP_s 表示服务的非功能属性的集合。服务的轮廓包含诸如服务的名字、功能描述、分类和服务提供者等信息。

服务参数定义为 $Parameter = (N, T, S)$,其中 N, T 和 S 分别对应参数的名称、数据类型及其语义。服务操作的输入和输出均是参数的集合。

定义 4(服务操作) 服务操作定义为 $Operation = (N, I, O, C, CO, NFP_s)$ 。其中, N, I, O, C, CO 和 NFP_s 分别表示操作的名称、输入、输出、约束以及交互代理(Coordinator)和非功能属性。一个服务操作的约束是有关该操作调用及其内部满足的条件,通常基于参数的语义来表达。Coordinator 被关联到每个服务操作,代表该操作与其它的操作或者用户进行交互。

定义 5(交互代理) 服务操作的一个交互代理(Coordinator)定义为 $Coordinator = (PEER_s, PRE, POST)$ 。其中, $PEER_s$ 表示当与其关联的操作完成后与其进行交互的其它 Coordinator 的集合; PRE 和 $POST$ 分别表示进入和离开该 Coordinator 是应该满足的前置和后置条件。这样,在服务的 Coordinator 集合上存在一个关系 R ,其每个元素 (c_1, c_2) 表示 c_2 是 c_1 的 $PEER_s$ 的元素。也就是说,当 c_1 所关联的操作执行完成后,并且其后置条件满足时, c_2 所关联的操作将会被调用。

例 1(服务 WSEmboss 的抽象模型描述) 图 3 是 EBI (European Bioinformatics Institute) 所提供的 Web 服务 WSEmboss²⁾ 的概念视图的示意。如图 3 所示, WSEmboss 提供 4 个操作 Run, CheckStatus, getResults 和 Poll。图 3 中的实线箭头表示操作之间的交互关系,虚线箭头则表示操作之间的数据流。图 3 中并未对操作的参数等细节进行刻画。每个操作关联的 Coordinator 描述如下:

$$\begin{aligned}
 C_0 &= (\{C_1\}, TRUE, TRUE) \\
 C_1 &= (\{C_1, C_2\}, status = false, status = true) \\
 C_2 &= (\{C_3\}, status = true, TRUE) \\
 C_3 &= (Null, TRUE, TRUE)
 \end{aligned}$$

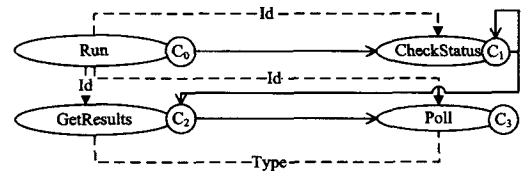


图 3 ASM 模型举例

其中, $status$ 表示该操作是否成功执行; $TRUE$ 表示一种特殊的状态,即在任何条件下该状态均满足。

从 ASM 的定义可以看出,ASM 对服务的接口和行为两方面的属性均进行了刻画。特别在服务的行为属性方面,具有描述操作的交互约束,以及依赖于操作的约束条件以及其前置和后置条件表示服务的内部约束和调用约束的能力。这些特性对于准确刻画服务的具体能力非常重要。更为重要的是 ASM 可对多样性和差异性的服务进行统一描述。

²⁾ <http://www.ebi.ac.uk/Tools/webservices/>

3.2 模型的逻辑语义

ASM 的一大特点是其逻辑结构,可被看作是一个基于 Kripke 结构^[11]的虚拟过程模型。

定义 6(ASM 逻辑结构) ASM 的逻辑结构表示为 $K=(S, S_0, R, AP, L, T, M)$ 。其中 S 是一个描述 Web 服务行为的状态集合, S 的每个元素 s 对应服务的一个操作的(虚拟)运行状态; $S_0 \subseteq S$ 表示初始状态的集合; $S_t \subseteq S$ 表示终止状态的集合; $R \subseteq S \times S \times M$ 是集合 S 上的二元关系,表示状态的转移关系,其每个元素 $(s_i, s_j, M_k) \in R$ 表示在从 s_i 状态转移到 s_j 的过程中,传递消息集合 M_k ; AP 是一个命题的集合,用以表示刻画 Web 服务行为的约束; $L: S \rightarrow 2^{AP}$ 是一个二元函数,给定 $s \in S, L(s)$ 表示在集合 AP 中状态 s 满足的命题的集合; M 是服务的消息集合。

从上述定义 6 可以看出,ASM 在逻辑上表示了服务在一次虚拟执行过程中的状态迁移、消息传递和时态约束等特性,因此可作为服务的匹配、组合和验证等推理工作的基础。

所要说明的是,从 ASM 模型到其逻辑结构的转换是直接的,具体的算法在本文中不做详述。

4 基于 ASM 的服务匹配

服务描述经过了语义标注,其语义来自于一个统一的本体。如果 Web 服务采用 OWL-S 描述,需要通过上述的规则对其过程进行解析。这样,所有的服务可被按照一个时态过程模型来对待。由于基于分支树时态逻辑(CTL)^[11]的模型检验被广泛应用于验证并发系统的性质^[12],因此,将服务请求描述为时态逻辑条件,服务匹配便是基于模型检验过程判断一个服务的逻辑视图是否满足服务请求约束。

4.1 基于 ASM 的适应性服务匹配方法

基于 ASM,服务匹配的过程转化为判断过程模型是否满足服务请求约束,其实质是基于 Kripke 结构的模型检验。

根据模型检验理论,给定一个规约为 CTL^[11]表达式的服务请求,匹配的关键是确定满足该请求的状态集。令 \emptyset 是一个对应服务请求的 CTL 表达式, $K=(S, S_0, R, AP, L, T, M)$ 对应服务的逻辑结构,则满足条件 \emptyset 的集合 $\psi = \{\omega | M, \omega \models \emptyset\}$ 可以通过扩展标记算法 SAT^[11,12] 获得。SAT 的基本思想是计算满足 \emptyset 的每个原子 CTL 公式的状态集合,然后通过迭代获取满足 \emptyset 的集合。如果 $|K. S_0 \cap \psi| > 0$,则该服务满足 \emptyset 。因为,此时从 K 中可发现一条路径满足 \emptyset 的约束。

在扩展 SAT 算法时,主要解决的是形如 $EF(?msg)$ 和 $EF(!msg)$ 的子公式,即判断存在服务模型的某条路径, E 和 F 是 CTL 的路径量词和时态算子。计算消息 msg_0 和 msg_1 的语义相似度的方法如下:

$$\text{similarity}(msg_0, msg_1) = \begin{cases} 1, & msg_0 \Leftrightarrow msg_1 \\ 1/\alpha, & msg_0 \supseteq msg_1 \\ 0, & \text{else} \end{cases}$$

其中, α 是 msg_0 和 msg_1 在本体中的语义距离。服务消费的消息向量和请求之间的相似性可表示为 $\prod_{i=1}^n m_i$, 其中, m_i 表示消息的一个参数与相应的请求信息的相似性。

事实上,上述的匹配过程既包含了服务的接口匹配,也包含了对于服务行为的匹配,在保证适应请求变化性和服务差异性的同时,使得服务匹配的过程更为准确。

ASM 之所以能够使适应性服务匹配的另一个原因是:当

单一服务不满足服务请求时,可以有效地获取可能满足请求的组合服务。

4.2 构造与业务需求匹配的组合服务

在服务匹配的过程中,当单一的服务不能满足业务需求时,可通过组合手段来满足业务需求。ASM 同样可以有效支持这种需要。其基本思想是首先进行服务模型的正确合并,然后通过模型检验的方法,在合并结果中生成满足业务需求的路径。

规则 1 K_2 能够被合并到 K_1 当且仅当 $\exists (s_1, s_2)(M(s_1, s_2) \wedge R(s_1, s_2))$, 其中 $M(s_1, s_2) = s_1 \in K_1. S \wedge s_2 \in K_2. S \wedge \exists (msg) (|s_2. q_1| > 0 \rightarrow msg \in s_1. q_0 \wedge msg \in s_2. q_1)$ 和 $R(s_1, s_2) = c_1. PRE \supseteq c_2. POST$, c_1 和 c_2 表示 s_1 和 s_2 相应的 Coordinator。

定义 7(2-merged 过程模型) 给定两个服务 $K_i = (S, S_0, R, AP, L, T, M), i=1, 2$, 则 2-merged 过程模型 $K = K_1 \times K_2$, 其中

- 1) $K. S = K_1. S \cup K_2. S$
- 2) $K. S_0 = K_1. S_0 \cup K_2. S_0$
- 3) $K. R = K_1. R \cup K_2. R \cup \{(s_1, s_2) | M(s_1, s_2) \wedge R(s_1, s_2)\}$
- 4) $K. AP = K_1. AP \cup K_2. AP$
- 5) $K. L = K_1. L \cup K_2. L$
- 6) $K. T = K_1. T \cup K_2. T \cup \{(s_1, s_2, msg) | M(s_1, s_2) \wedge R(s_1, s_2) \wedge (msg \in s_1. q_0 \wedge msg \in s_2. q_1)\}$
- 7) $K. M = K_1. M \cup K_2. M$

定义 8(n-merged 过程模型) 给定一个服务集 $S = \{K_1, K_2, \dots, K_n\}$, 那么 n-merged 过程模型 $K = K_1 \times K_2 \times \dots \times K_n = (K_1 \times K_2 \times \dots \times K_{n-1}) \times K_n, n \geq 1$ 。

在此基础上,组合过程的算法如下:

算法 1 Goal-based-Composition(r, S)

$K = K_1 \times K_2 \times \dots \times K_n$

$t = SAT_E(r, K)$

If $(t \supseteq K. S_0) Get_Match_Path(r, K)$

算法结束。

在算法 1 中, r 表示以时态逻辑条件描述的服务请求, $S = \{K_1, K_2, \dots, K_n\}$ 是以 ASM 的逻辑结构描述的服务的集合。Get_Match_Path 表示构造组合结果。函数 SAT_E 返回满足 r 的状态的集合。

需要指出的是,该组合方法仍然需要改进以避免多个模型合并时可能产生的副作用。

5 实验结果和应用评价

ASM 通过初步应用到生物信息领域的 PSE 系统中,验证了其在使适应性的服务匹配方面的有效性。在该应用中,使用了由 DDBJ, EBI 等生物信息研究权威机构的 Web 服务。需要指出的是由第三方机构提供的 Web 服务当前均以 WS-DL 的方式描述,而该机制同时具有支持其他类型服务的能力。所有的服务被注册到一个服务社区中。服务请求以业务服务^[13]的形式提供,在服务匹配时被转换为时态逻辑条件。

表 1 给出了服务匹配中的一些结果。BS 和 WS 分别表示业务服务和 Web 服务。Path length 表示一个满足请求的服务包含的参与协作的服务操作的个数。我们发现,用以序列相似性查找和序列比对的服务在具体的能力上具有差异性。同时,以业务服务表示的服务请求也呈现出变化性。

表 1 服务匹配结果片断

BS	WS	Provider	Path length	Constraints
Sequence alignment	ClustalW	DDBJ	1	
	WScLustal	EBI	4	
	WScLustal	EBI	4	
	WS Emboss	EBI	4	
pair sequence alignment	WS Emboss	EBI	4	sequence number= 2
Sequence similarity search	Blast	DDBJ	1	
	WU Blast	EBI	4	
	WS Fasta	DDBJ	4	
Sequence search- blast	Fasta	EBI	1	
	Blast	DDBJ	1	tool=blast
	WU Blast	EBI	4	tool=blast
Sequence search-fasta	WS Fasta	EBI	1	tool=fasta
	Fasta	DDBJ	4	tool=fasta

从表 1 中可以看出,一个序列比对请求可由 4 个不同的服务所满足。然而,当期望进行成对序列比对,只有 WSEM-Boss 能够满足要求。如果期望使用 Blast 进行比对时,只有两个服务是满足要求的。结果表明,ASM 能够提高服务匹配的适应性。

6 相关工作

WSDL 和 WSDL-S^[9] 均是基于消息的服务描述规范,不关心服务的行为属性。OWL-S^[10] 使用 service profile, service model 和 grounding 来描述 Web 服务。近年来提出了一些抽象服务模型。文献[14]提出一种基于描述逻辑(DL)的服务描述的框架,主要包括服务的接口信息和一些特殊约束。文献[15]提出一个基于有限状态自动机的抽象服务模型,该模型可被看作是一个不包含接口描述、消息传递和其他约束的抽象过程模型。文献[16]提出的基于有限状态自动机的服务模型对文献[15]中的模型在一定程度上进行了加强。文献[27]提出一个基于一阶逻辑的抽象服务模型。该模型虽然具有不可判定型,实际应用能力受到限制,但提出了一种比较其他服务模型的基础。文献[18]提出了一种机制 CTR-S,用以描述服务的动态规约。

在服务匹配方面,文献[2,3]致力于如何基于 WSDL 描述判定服务的相似性。文献[1,4]基于描述逻辑提出了一种判定服务请求和 Web 服务相似性的方法。文献[5,6]将服务的过程属性加入到服务匹配的过程中。文献[7]介绍了一种基于目标的服务匹配算法,其总目标和服务均被描述为有限状态自动机。

与上述的工作相比,本文的目标是提高服务匹配对于服务请求多样性和服务差异性的适应能力。因此,ASM 关注了影响服务匹配的一些重要元素,提供了一个通过刻画状态转移、消息传递和操作交互的时态约束来描述服务行为的内部过程模型以及基于该模型的适应性服务匹配方法。

结束语 ASM 提供了一种服务能力的统一描述途径,可有效支持适应性的服务匹配。将来,我们将对基于该模型的服务组合和聚类进行更为深入的探讨。

参 考 文 献

[1] Hull D, Zolin E, Bovykin A, et al. Deciding Semantic Matching

of Stateless Services[C]// The Twenty-First National Conference on Artificial Intelligence. California, 2006; 16-20

[2] Wang Y, Stroulia E. Semantic Structure Matching for Assessing Web-Service Similarity[C]// ICSOC. 2003; 194-207

[3] Wu J, Wu Z K. Similarity-based Web Service Matchmaking[C] // IEEE International Conference on Services Computing (SCC'05). July 2005, 1; 287-294

[4] Paolucci M, Kawamura; Semantic matching of Web service capabilities[C]// The First International Semantic Web Conference (ISWC'02). Berlin, 2002

[5] Bansal S, Vidal J M. Matchmaking of Web Services Based on the DAML-S Service Model[C]// AAMAS'03. Melbourne, Australia July, 2003

[6] Agarwal S, Ankolekar A. Automatic matchmaking of Web services[C]// The 15th International Conference on World Wide Web (WWW '06). New York, USA, 2006; 1057-1058

[7] Gao X, Yang J, Papazoglou MP. The Capability Matching of Web Services[C] // IEEE Fourth International Symposium on Multimedia Software Engineering (MSE'02). 2002; 56-63

[8] Czarnecki K, Eisenecker U. Generative Programming, Methods, Tools and Applications[M]. New York; Addison-Wesley, 2000

[9] Patil A, Oundhakar S, Sheth A, et al. METEOR-S Web Service Annotation Framework[C]// The World Wide Web Conference. July 2004

[10] Balzer S, Liebig T, Wagner M. Pitfalls of OWL-S; a practical semantic Web use case[C]// The 2nd International Conference on Service Oriented Computing. New York, USA, 2004

[11] Huth M, Ryan M. Logic in Computer Science: Modelling and Reasoning About System Second Edition[M]. England; Cambridge University Press, 2004; 207-243

[12] Manna Z, Pnueli A. The Temporal Logic of Reactive and Concurrent System; Specification[M]. NY, Springer-Verlag, 1992; 179-273

[13] Wang J, Yu J, Falcarin P, et al; An Approach to Domain-Specific Reuse in Service-Oriented Environments[C]// ICSR. 2008; 221-232

[14] Li L, Horrocks L A software framework for matchmaking based on semantic Web technology[C]// The 12th International Conference on the World Wide Web (WWW 2003). Budapest, Hungary, May 2003; 331-339

[15] Berardi D, Calvanese D, Giacomo D, et al. Automata-based representation of humanmachine Web service[C] // ICSOC '03. 2003

[16] Berardi D, Calvanese D, Giacomo G D, et al. Automatic composition of Web services in Colombo[C]// The 13th Italian Symposium on Advanced Database Systems. June 2005

[17] Berardi D, Gruninger M, Hull R, et al. Towards a first-order ontology for Web services[C]// W3C Workshop on Constraints and Capabilities for Web Services. October 2004

[18] Davulcu H, Kifer M, Ramakrishnan I V. CTR-S; A Logic for Specifying Contracts in Semantic Web Services[C] // WWW. 2004

[19] Gerede C E, Hull R, Ibarra O H, et al. Automated Composition of e-services; lookaheads[C]// The 2nd International Conference on Service-oriented Computing. New York, NY, USA, 2004; 252-262