

一种利用描述逻辑的语义 Web 服务 前提/效果匹配方法

王海¹, 陈妍¹, 范琳², 李增智^{1,3}

(1. 西安交通大学计算机系统结构与网络研究所, 710049, 西安; 2. 西安邮电学院计算机系, 710061, 西安;
3. 北京邮电大学网络与交换技术国家重点实验室, 100876, 北京)

摘要: 为了解决语义 Web 服务匹配方法主要针对服务输入/输出匹配而导致功能匹配结果较为片面的问题, 提出了一种基于描述逻辑的语义 Web 服务前提/效果匹配方法. 利用描述逻辑知识库一致性检查推理, 判断服务请求与广告的前提/效果之间逻辑蕴含关系, 建立类匹配度概念, 并由此度量和比较前提/效果匹配结果的优劣, 从而对匹配结果进行有意义的排序及分类. 实验结果表明, 所提方法在查准率与现有方法相当的情况下, 具有更高的查全率, 特别是将输入/输出与前提/效果匹配一并考虑的情况下, 能够更好地区分匹配结果.

关键词: 语义 Web 服务; 前提/效果; 服务匹配

中图分类号: TP393 **文献标志码:** A **文章编号:** 0253-987X(2010)04-0039-04

A Description-Logic-Based Preconditions/Effects Matchmaking Method for Semantic Web Services

WANG Hai¹, CHEN Yan¹, FAN Lin², LI Zengzhi^{1,3}

(1. Institute of Computer System Structure and Networks, Xi'an Jiaotong University, Xi'an 710049, China; 2. Department of Computer Science and Technology, Xi'an University of Posts and Telecommunications, Xi'an 710061, China;
3. State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: Since the results of the existing matchmaking methods for semantic Web service are more one-sided by only focusing on the matchmaking of inputs/outputs descriptions, a description logic based semantic Web service preconditions/effects matchmaking method is proposed. The knowledge base consistency checking function of description logic reasoning is used to determine the logical implication between preconditions/effects of service requests and advertising. A compatible matching degree is proposed to measure and to compare the pros and cons of the results of preconditions/effects matchmaking. Experimental results show that the proposed method has a higher recall rate with a considerable precision rate. In particular, better matching results are obtained when both the inputs/outputs matchmaking and the preconditions/effects matchmaking are combined.

Keywords: semantic Web service; preconditions/effects; service matchmaking

Web 服务匹配实质上是服务描述的匹配. 服务描述是以形式化的方法对服务进行刻画, 通常包括

对服务功能性属性和非功能属性的刻画. 功能性描述主要指服务的输入、输出、前提和效果的描述, 即

所谓的 IOPE 描述^[1-2].

迄今为止,前提/效果(P/E)匹配的方法还比较鲜见,如:使用逻辑表达式进行 P/E 描述,而表达式中的参数来自本体概念或个体,匹配的依据为表达式的结构等价性与参数的相容性^[3];使用 OWL-S 推荐标准中常用的语义 Web 规则语言(SWRL)进行 P/E 描述,其将出现的术语抽取出来,并仿照 I/O 匹配法来建立广告描述与请求描述中术语之间的映射^[4]. 以上 2 种 P/E 匹配方法本质上不是以逻辑推理为基础的,所以匹配的准确率不够理想.

本文提出以基于描述逻辑断言构成的知识库对服务的 P/E 进行描述,即将 P/E 匹配问题转化为描述逻辑知识库之间的逻辑蕴含判定问题. 为了对匹配结果进行有意义的排序及分类,建立了类匹配度的概念及计算公式.

1 基于描述逻辑知识库一致性检测的 P/E 匹配方法

为了清楚地说明如何利用描述逻辑知识库一致性检测进行服务的 P/E 匹配,需要以形式化的方式定义 Web 服务的 P/E,并在此基础进行分析. 文献[5]将服务形式化为具有前提和效果的动作,从抽象意义上,服务被形式化为一对描述逻辑 $\xi^+(D)$ 知识库. 在此基础上,为了简单起见,本文描述只考虑了原子基础服务,即服务的输入和输出都被实例化为个体.

定义 1 令 T 为无环 TBox,那么 TBox 原子基础服务 $S=(p,e)$,其中 p 表示服务前提, e 表示服务效果. $S=(p,e)$ 包括:①一个由描述逻辑 $\xi^+(D)$ 公理组成的表示前提的有限集(知识库);②一个由描述逻辑 $\xi^+(D)$ 公理组成的表示效果的有限集(知识库).

从上述定义中的服务规范化描述可以看出,一个服务本质上就是一对描述逻辑 $\xi^+(D)$ 知识库.

对于同一个 T ,请求服务 $S_{\text{Req}}=(p_{\text{Req}},e_{\text{Req}})$ 和发布服务 $S_{\text{Adv}}=(p_{\text{Adv}},e_{\text{Adv}})$ 之间是否匹配基于以下 2 个规则.

(1)发布服务的前提条件 p_{Adv} 要比请求服务提供的前提条件 p_{Req} 更容易满足,也就是必须满足 $S_{\text{Req}} \models p$,但不要求 $S_{\text{Adv}} \models p$.

(2)请求服务的效果 e_{Req} 要比发布服务提供的效果 e_{Adv} 更容易满足,也就是必须满足 $S_{\text{Adv}} \models e$,但不要求 $S_{\text{Req}} \models e$.

从以上分析可以看出,2 个服务之间的 P/E 匹配关系实质上是由 2 对知识库之间的逻辑蕴含关系决定的. 文献[6]给出了知识库之间逻辑蕴含的判定算法,即 2 个知识库 K 和 K' 之间存在 4 种逻辑蕴含关系:① $K \models K'$;② $K' \models K$;③ $K \models K'$ 且 $K' \models K$,记作 $K \equiv K'$;④其他情况.

根据知识库之间语义推导的这 4 种情况,结合以逻辑蕴含关系的一致性检测为基础的实现,本文给出了基于描述逻辑知识库一致性检测的语义 Web 服务 P/E 匹配算法.

1.1 前提匹配

对于发布服务和请求服务,依据上述的匹配算法思想及规则,请求服务前提 O_{r-p} 与发布服务前提 O_{a-p} 之间的匹配可分为以下 4 种匹配情况.

(1)完全匹配: $O_{r-p} \models O_{a-p}$ 且 $O_{a-p} \models O_{r-p}$ 均成立,此时也称作等价匹配,即 $O_{r-p} \equiv O_{a-p}$.

(2)正匹配: $O_{r-p} \models O_{a-p}$ 成立,但 $O_{a-p} \models O_{r-p}$ 不成立.

(3)负匹配: $O_{a-p} \models O_{r-p}$ 成立,但 $O_{r-p} \models O_{a-p}$ 不成立.

(4)匹配失败: $O_{a-p} \models O_{r-p}$ 和 $O_{r-p} \models O_{a-p}$ 均不成立.

效果匹配与前提匹配类似,在此不再赘述.

1.2 P/E 综合匹配

根据上述语义 Web 服务的前提匹配和效果匹配的分析,这 2 个服务之间的前提和效果的综合匹配情况可以分为表 1 所示的 16 种情况.

依据前文所述的语义 Web 服务 P/E 匹配的原则及分类,表 2 中第 1 种情况属于 P/E 的完全匹配,第 2、第 5、第 6 为正匹配,第 3、第 7、第 9~第 11 属于负匹配,其他情况属于匹配不成功. 另外,为了方便对服务之间的 P/E 匹配和最终的 IOPE 匹配进行量化分析,本文借鉴了 I/O 匹配中匹配度的思想,即在 P/E 匹配中给出了每一种情况赋予 $[0, 1]$ 之间的匹配度,同时为了与基于 I/O 的匹配度进行区分,本文将其称为类匹配度. 加入类匹配度之后,服务的 P/E 综合匹配分类及赋值如表 2 所示.

从表 2 可以看出,类匹配度可用下式计算

$$\lambda(S_{\text{Req}}, S_{\text{Adv}}) = \begin{cases} 0, & S_{\text{Req}} \Rightarrow S_{\text{Adv}} \text{ and } S_{\text{Adv}} \Rightarrow S_{\text{Req}} \text{ 均不成立} \\ 1, & S_{\text{Req}} \Leftrightarrow S_{\text{Adv}} \\ \frac{\omega}{10}, & \text{其他} \end{cases} \quad (1)$$

式中: ω 为编号.

表 1 前提和效果匹配情况列表

编号	p	e
1	$p_{Req} = p_{Adv}$	$e_{Adv} = e_{Req}$
2	$p_{Req} = p_{Adv}$	$e_{Adv} = e_{Req}$
3	$p_{Req} = p_{Adv}$	$e_{Req} = e_{Adv}$
4	$p_{Req} = p_{Adv}$	失效
5	$p_{Req} = p_{Adv}$	$e_{Adv} = e_{Req}$
6	$p_{Req} = p_{Adv}$	$e_{Adv} = e_{Req}$
7	$p_{Req} = p_{Adv}$	$e_{Req} = e_{Adv}$
8	$p_{Req} = p_{Adv}$	失效
9	$p_{Adv} = p_{Req}$	$e_{Adv} = e_{Req}$
10	$p_{Adv} = p_{Req}$	$e_{Adv} = e_{Req}$
11	$p_{Adv} = p_{Req}$	$e_{Req} = e_{Adv}$
12	$p_{Adv} = p_{Req}$	失效
13	失效	$e_{Adv} = e_{Req}$
14	失效	$e_{Adv} = e_{Req}$
15	失效	$e_{Req} = e_{Adv}$
16	失效	失效

表 2 前提与效果综合匹配分类

编号	对应表 1 中的编号	分类	类匹配度
0	16	匹配失效	0. 0
1	12,15	匹配失效	1/10=0. 1
2	8,14	匹配失效	2/10=0. 2
3	4,13	匹配失效	3/10=0. 3
4	11	负匹配	4/10=0. 4
5	7,10	负匹配	5/10=0. 5
6	3,9	负匹配	6/10=0. 6
7	6	正匹配	7/10=0. 7
8	2,5	正匹配	8/10=0. 8
9	1	完全匹配	1. 0

在此需要说明的是,本文在基于描述逻辑知识库一致性检测的 P/E 匹配算法中,对 P 和 E 的综合类匹配度的计算不是 I/O 匹配中取 I 和 O 的匹配度的算术平均值. 这是因为:在某种意义上 I 和 O 之间没有必然的联系,其关系比较松散;P 和 E 则不同,二者反映的是现实世界状态的变化,前提不满足就肯定得不到相应的效果,反之如果前一个动作的效果没有达到,那么后一个动作的前提也不可能达

到,所以 P 和 E 之间依赖关系非常强. 因此,如果简单地将 P 和 E 的匹配度按照算术平均值来计算,就不能反映二者之间的逻辑关系,所以必须对 P/E 的综合类匹配度重新定义.

2 实验及分析

为了验证方法的可行性及效果,本文实现了一个语义 Web 服务 IOPE 综合匹配的原型系统 Matchmaker1. 0,它可以为服务的 I/O 匹配、P/E 匹配和 IOPE 综合匹配提供一个集成匹配环境,其中的 I/O 匹配算法是以本实验室在文献[7]算法基础上改进的弹性匹配方法,其匹配度可连续地映射到 [0-1]区间.

2.1 用户请求服务描述

S_{Req1} : 用户请求航班预定服务进行航班预定. 用户输入信用卡号、起始地、目的地和出发时间,希望输出航班号、票价. 前提是信用卡对应的信誉度为“好”,效果是“有票”.

S_{Req2} : 用户请求航班预定服务进行航班价格查询. 用户输入信用卡号、目的地、航班号和出发时间,希望输出票价. 前提是信用卡对应的信誉度为“一般”,效果为“空”.

2.2 航班预定服务描述

S_{Adv1} : 要求用户提供的输入包括信用卡号、起始地、目的地、出发时间;输出包括航班号、票价;前提条件要求用户提供的信用卡对应的信誉度为“一般”级别以上;效果是“有票”.

S_{Adv2} : 要求用户提供的输入包括信用卡号、起始地和目的地;输出包括航班号、票价和起飞时间;前提条件要求用户提供的信用卡对应的信誉度为“好”级别以上;效果是“有票”.

S_{Adv3} : 要求用户提供的输入包括信用卡号、起始地和目的地;输出包括航班号、票价;前提条件要求用户提供的信用卡对应的信誉度为“优”;效果是“有票”.

S_{Adv4} : 要求用户提供的输入包括信用卡号、起始地和目的地;输出为票价;前提条件为“空”;效果是“有票”.

S_{Adv5} : 要求用户提供的输入包括目的地和航班号;输出包括票价和起飞时间;前提条件为“空”;效果是“空”.

2.3 匹配结果及分析

依据 2.1 节、2. 2 节内容的请求服务与发布服务匹配结果如表 3、表 4 所示.

表 3 S_{Req1} 请求服务与发布服务匹配结果

服务 组对	I/O 匹配度			P/E 匹配情况						IOPE 综合匹配度	
	I	O	I/O	SWRL			一致性			SWRL	一致性
				P	E	P/E	P	E	P/E		
S_{Req1}, S_{Adv1}	1.000	1.000	1.000 0	完全	完全	完全	正匹配	完全	正匹配	1.000 00	0.900 00
S_{Req1}, S_{Adv2}	0.850	1.000	0.925 0	完全	完全	完全	正匹配	完全	正匹配	0.962 50	0.862 50
S_{Req1}, S_{Adv3}	0.925	1.000	0.962 5	失效	完全	失效	失效	完全	失效	0.481 25	0.631 25
S_{Req1}, S_{Adv4}	0.850	0.750	0.800 0	完全	完全	完全	正匹配	完全	正匹配	0.900 00	0.800 00
S_{Req1}, S_{Adv5}	0.250	0.750	0.500 0	完全	失效	失效	正匹配	失效	失效	0.250 00	0.350 00

表 4 S_{Req2} 请求服务与发布服务匹配结果

服务 组对	I/O 匹配度			P/E 匹配情况						IOPE 综合匹配度	
	I	O	I/O	SWRL			一致性			SWRL	一致性
				P	E	P/E	P	E	P/E		
S_{Req2}, S_{Adv1}	0.800	1	0.900 0	失效	完全	失效	负匹配	正匹配	失效	0.450 00	0.700 00
S_{Req2}, S_{Adv2}	0.775	1	0.887 5	失效	完全	失效	负匹配	正匹配	失效	0.443 75	0.693 75
S_{Req2}, S_{Adv3}	0.850	1	0.925 0	失效	完全	失效	负匹配	正匹配	失效	0.462 50	0.712 50
S_{Req2}, S_{Adv4}	1.000	1	1.000 0	完全	完全	完全	完全	正匹配	正匹配	1.000 00	0.900 00
S_{Req2}, S_{Adv5}	1.000	1	1.000 0	完全	完全	完全	完全	完全	完全	1.000 00	1.000 00

从表 3 可以看出:对于 S_{Req1} 请求服务来讲,采用基于 SWRL 规则的 P/E 匹配算法和采用基于描述逻辑知识库一致性检测的 P/E 匹配算法,两者均匹配成功的是 S_{Adv1} 、 S_{Adv2} 和 S_{Adv4} ,并且这 3 个服务的匹配度相同;进行 I/O 匹配时,匹配度最高的服务是 S_{Adv1} ,其次是 S_{Adv2} 、 S_{Adv3} 和 S_{Adv4} ;最终的 IOPE 匹配度最高的是 S_{Adv1} ,其次是 S_{Adv2} 和 S_{Adv4} .

从表 4 可以看出:对于 S_{Req2} 请求服务来讲,I/O 匹配度最高的服务是 S_{Adv4} 和 S_{Adv5} ;采用基于 SWRL 规则的 P/E 匹配算法,匹配成功的服务是 S_{Adv4} 和 S_{Adv5} ;采用综合 IOPE 匹配,匹配度最高的也是 S_{Adv4} 和 S_{Adv5} . 到此为止,用户还不能判别出 S_{Adv4} 和 S_{Adv5} 到底哪一个更适合自己的要求,因为它们有相同的匹配度. 如果采用基于描述逻辑知识库一致性的 P/E 匹配算法,可以得到匹配成功的有 S_{Adv4} 和 S_{Adv5} ,但两者的匹配度却不同, S_{Adv5} 的匹配度比 S_{Adv4} 的高,由此很容易选出 S_{Adv5} 作为最终的匹配服务.

3 结束语

针对语义 Web 服务的前提/效果匹配问题,提出了一种基于描述逻辑知识库一致性检测的 P/E

匹配算法,并深入探讨了算法的理论基础及各种匹配情况的计算方法,对该算法的合理性进行了实例分析. 同时,提出了类匹配度,其可将 P/E 匹配的结果进行量化比较. 最后,根据本文所述方法,实现了一个语义 Web 服务匹配的原型系统. 经实验证明,基于描述逻辑知识库一致性检测的服务匹配方法是正确和有效的.

参考文献:

[1] DE BRUIJN J, LAUSEN H, POLLERES A, et al. The web service modeling language: an overview [C] //Proceedings of the 3rd European Semantic Web Conference. Berlin, Germany: Springer-Verlag, 2006: 590-604.

[2] OWL-S Coalition. OWL-S 1.1 release [EB/OL]. [2009-3-12]. <http://www.daml.org/services/owl-s/1.1/>.

[3] BELLUR U, VADODARIA H. On extending semantic matchmaking to include preconditions and effects [C]//Proceedings of the 2008 IEEE International Conference on Web Services. Los Alamitos, CA, USA: IEEE Computer Society, 2008: 120-128.

[4] BENER A B, OZADALI V, ILHAN E S. Semantic

- matchmaker with precondition and effect matching using SWRL [J]. Expert Syst Appl, 2009, 36(5): 9371-9377.
- [5] BAADER F, LUTZ C, MILICIC M, et al. A description logic based approach to reasoning about web services [C]//Proceedings of the WWW 2005 Workshop on Web Service Semantics. New York, USA: ACM 2005: 636-647.
- [6] WANG Hai, LI Zengzhi, FAN Lin. An unabridged method concerning capability matchmaking of web services web intelligence [C]//2006 IEEE/WIC/ACM International Conference. Piscataway, NJ, USA: IEEE, 2006: 662-665.
- [7] PAOLUCCI M. Semantic matching of web services capabilities, in the semantic web [C]//1st International Semantic Web Conference Proceedings. Berlin, Germany: Springer-Verlag, 2002: 333-347.
- [本刊相关文献链接]**
- 以语义规则刻画服务请求的 Top- k Web 服务发现方法. 西安交通大学学报, 2010, 44(2): 35-38.
- 一种等价类区分的多属性评价指标约简算法. 西安交通大学学报, 2009, 43(4): 26-29.
- 一种支持大规模应用层组播的混合拓扑覆盖网络. 西安交通大学学报, 2010, 44(2): 11-14.
- 一种分段测量保证 QoS 约束的任播通信模型. 西安交通大学学报, 2010, 44(2): 44-49.
- 重配置与周期再生相结合的软件再生模型. 西安交通大学学报, 2010, 44(1): 91-95.
- 无线传感器网络地理位置路由度量方法. 西安交通大学学报, 2009, 43(12): 55-59.
- 可废除并发签名机制. 西安交通大学学报, 2009, 43(12): 45-49.
- 基于隐马尔可夫模型的无线局域网媒体接入控制层入侵检测方法. 西安交通大学学报, 2009, 43(12): 26-39.
- 面向变异分析的协议安全测试方法. 西安交通大学学报, 2009, 43(12): 11-15.
- 面向 Web 服务总线集成的分阶段优先级事件驱动架构. 西安交通大学学报, 2009, 43(12): 16-20.
- 自适应滤波实时网络流量异常检测方法. 西安交通大学学报, 2009, 43(12): 1-5.
- 动态短信通信复杂网络演化模型研究. 西安交通大学学报, 2009, 43(6): 5-9.
- 一种相邻节点协作的无线传感器网络可靠传输方案. 西安交通大学学报, 2009, 43(2): 33-37.
- 基于自愈技术的无线传感器网络基站生存模型. 西安交通大学学报, 2008, 42(8): 950-953.
- 一种基于子网保护的蠕虫传播模型及其应用. 西安交通大学学报, 2008, 42(6): 788-790.
- 一种基于证明树反演的安全漏洞定位方法. 西安交通大学学报, 2007, 41(4): 439-443.
- 应用控制的 Web 服务器磁盘缓冲方法. 西安交通大学学报, 2007, 41(2): 153-157.

(编辑 苗凌)