

Web 服务流程相容性和相似性分析

李喜彤 范玉顺

(清华大学自动化系 北京 100084)

摘 要 服务组合和服务替换是面向服务计算的研究热点,服务流程的相容性和相似性分析是其中的两个密切相关的问题,具有较大实用价值.基于着色 Petri 网建模 Web 服务流程,定义服务流程的正确性和相容性.提出服务良构性的概念,证明良构性能够保证组合服务可达终止状态的正确性.在相容性分析的基础上,提出服务流程相似性的定义,证明若新服务与要被替换的服务流程相似,则所进行的替换是上下文无关的,替换后无须再做组合正确性验证,给出相似性的判定算法.文中结论和算法改进了现有服务组合验证和服务替换方法的不足.

关键词 Web 服务;服务组合;服务替换;相容性;相似性

中图法分类号 TP311 **DOI 号**: 10.3724/SP.J.1016.2009.02429

Analyzing Compatibility and Similarity of Web Service Processes

LI Xi-Tong FAN Yu-Shun

(Department of Automation, Tsinghua University, Beijing 100084)

Abstract Service composition and substitution are key research issues in Service-Oriented Computing (SOC). Among them, analyzing compatibility and similarity of Web service processes are two closely-related issues and of great importance. This paper models Web service processes using Colored Petri nets (CPN) and presents the definitions of correctness and compatibility of Web service processes. Then, the notion of well-structuredness of Web services is developed. The paper proves that all reachable final markings of a composite service composed by well-structured services are correct. Based on the compatibility analysis, the paper develops the definition of similarity of Web service processes in order to investigate service substitution. It is concluded that substituting a service in a composition can be performed independent of the context as long as the new service is similar to the substituted one. There is thus no need to verify the substituted service composition again. The paper develops an algorithm for verifying similarity between two services. The results and algorithm are used to improve the existing methods of service composition verification and service substitution.

Keywords Web service; service composition; service substitution; compatibility; similarity

1 引 言

Web 服务是目前面向服务体系架构最流行的

一种实现形式.单个 Web 服务通常无法满足用户需求,而需要将若干个 Web 服务组合起来协同工作.服务组合是指将若干个已存在的 Web 服务按照一定规则动态发现,并组装成一个增值的、更粗粒度的

服务或系统以满足用户的复杂需求. 服务组合已成为构建面向服务、松散耦合、高适应性的应用系统的主要途径^[1-2].

组合服务的流程能否正确执行直接影响到用户需求是否能够得到满足, 因此需要分析和验证组合服务流程的正确性, 这方面的研究是当前面向服务计算领域的一个热点. 服务组合的正确性可以通过分析参与组合的各成员服务流程的相容性来进行验证^[3-4]. 另一方面, 服务替换是与服务组合密切相关的一个问题^[5]. 如果某个成员服务由于自身的软硬件故障或者外部环境改变(如网络阻塞、病毒威胁)等原因不能继续提供服务, 那么将导致整个组合服务无法正确执行, 用户需求也就无法得到满足. 此时需要发现并选取一个与失效服务的流程相似的服务来替换, 以确保替换后的组合服务仍然能够正确执行, 也即需要保证替换服务与服务组合系统中的其他成员服务的流程相容. 由此可见, 服务流程的相容性和相似性分析是两个密切相关的问题, 对于研究服务组合和服务替换具有较大的应用价值.

由于验证服务流程之间的相容性通常需要对整个组合服务的流程进行分析, 在服务替换频繁发生的情况下, 不断重复的服务组合验证将耗费大量时间和资源, 严重影响系统运行效率^[5]. 因此服务相似性定义要与上下文无关, 即在不需要验证组合服务全局流程的情况下就能判断选取的替换服务如果被替换服务相似, 那么就能够与服务组合系统中的其他成员服务相容, 从而保证替换后的组合服务的流程仍然是正确的.

分析 Web 服务的流程通常需要借助某种形式化方法, 主要有基于有限状态机的方法^[3]、基于进程代数(或 PI-演算)的方法^[4]和基于 Petri 网的方法^[2,6-7]等. 由于 Petri 网适合于对 Web 服务这种松散耦合的分布式系统进行建模, 因此本文采用 Petri 网作为服务流程分析的理论基础. 文献[2]提出了基于着色 Petri 网的 Web 服务模型, 给出了 4 种主要的组合运算. 文献[6]以着色 Petri 网为基础提出了面向消息的 Web 服务模型, 该模型侧重于对 Web 服务元消息的描述及其对服务协同通信的影响. 文献[8]借助工作流理论对 Web 服务流程进行建模, 研究了基于信任关系的 Web 服务工作流服务质量调度方法. 虽然 Petri 网能够对组合服务进行分析和验证, 但是上述研究成果都没有对此进行深入讨论, 没有指出具有哪些性质的两个或多个服务的流程是相容的. 进一步地, 这些研究工作也都没有涉及

到服务流程的相似性和等价性分析.

本文将组合服务看作是全局模型, 而把各个参与组合的成员服务看作是局部模型. 现有研究工作大多是对组合服务的全局模型直接进行建模, 较少考虑单个服务流程的正确性, 因此无法揭示组合服务流程与单个服务流程之间的内在关系. 本文从组合服务的局部模型出发, 基于着色 Petri 网建立单个服务流程模型, 给出服务流程正确性的形式化定义. 该定义不仅适用于单个 Web 服务, 同时也适用于刻画组合服务的流程正确性. 本文提出 Web 服务良构性的概念, 并指出成员服务的良构性能够保证组合服务可达终止状态是正确的, 从而简化了服务流程相容性分析的验证条件. 为了研究具有哪些性质的两个服务可以进行替换, 本文提出服务流程相似性和等价性的定义, 并证明两个流程相似的良好服务可以进行与上下文无关的替换, 替换后的组合服务的流程仍然是正确的. 本文讨论了相似性与相容性之间的关系, 并给出服务流程相似性判定算法.

本文第 2 节给出有关的基本定义; 第 3 节分别讨论服务流程的相容性和相似性; 第 4 节介绍相关工作并进行对比讨论; 第 5 节给出结论和未来的工作.

2 基本定义

Petri 网具有直观的图形符号表示和坚实的理论分析技术, 适合对 Web 服务这种异步并发系统进行建模. 然而基本 Petri 网在实际建模过程中存在一些固有不足^[9], 因此本文采用着色 Petri 网建模 Web 服务的流程. 有关着色 Petri 网的详细内容可参见文献[9].

定义 1(Web 服务). Web 服务是一个带标记的着色 Petri 网 $S = (N, A, L)$, 其中 N 为一着色 Petri 网, 且满足

(1) P 是有限库所集, $P = P^{IC} \cup P^{IM} \cup P^{EM}$, 其中: P^{IC} 是服务流程内部控制逻辑库所集, P^{IM} 是内部通信库所集, P^{EM} 是服务与外界通信库所集, P^{IC} 、 P^{IM} 和 P^{EM} 两两互不相交;

(2) P 中有两个特殊的库所, 即起始库所 i 和终止库所 o , 满足 $\bullet i = \emptyset$, $o \bullet = \emptyset$;

(3) 如果去除库所集 P^{EM} 以及所有与 P^{EM} 关联的弧, 由此得到的 Petri 网称为服务 S 对应的工作流网 N , 记为 $N = \Phi(S)$. 若在 N 上添加一个连接库

所 o 到库所 i 的变迁 t^* (即 $\bullet \cdot t^* = o, t^* \cdot = i$), 则 N 是强连通的;

(4) $\forall p \in P, p \neq i$, 满足 $M_i(i) \neq \emptyset$ 且 $M_i(p) = \emptyset$, 其中 M_i 是初始标识;

(5) A 是标记的非空有限集, $\forall a \in A$ 为一字符串型的标识, 其中 $\exists \tau \in A$ 为一空标记;

(6) L 是标记函数, 满足 $L: T \rightarrow A$.

在定义 1 中, P^{EM} 中的每个库所都对应一个与外界通信的消息缓存区, 每个与 P^{EM} 连接的变迁称为通信变迁, 对应一个与外界环境通信的操作 (如发送或接收消息); 服务流程内部库所包括 P^{IC} 和 P^{IM} 两部分, 其中 P^{IC} 表示内部控制逻辑的状态, 而 P^{IM} 表示内部的消息缓存区; 标记集 A 表示该服务所有的操作语义, A 中的每一非空标记字符串对应该服务的一个操作命名; 标记函数 L 将服务模型中的每一个变迁映射到一个服务操作命名. 而映射到空标记 τ 上的变迁所代表的操作行为是服务的内部行为, 即不被外界观察到的行为.

目前工业界已存在一些支持 Web 服务的业务流程建模语言, 其中应用最多的是业务流程执行语言 (Business Process Execution Language, BPEL), 该语言已被 OASIS 组织宣布为描述 Web 服务流程的标准. 需要指出的是, 基于 BPEL 标准描述的 Web 服务流程可以方便地转换为本文定义 1 中的形式化模型. 文献[10-11]提出了基于 BPEL 的服务流程向 Petri 网服务模型转换的具体方法, 在此不再复述. 本文主要研究如何利用定义 1 中的服务模型来形式化地分析 Web 服务流程的相容性和相似性.

一个变迁 t 在标识 M 下使能, 记作 $M[t]$. 若服务在标识 M 下通过实施变迁 t 后到达标识 M' , 记作 $M[t]M'$. 若服务在标识 M 下经过实施一个变迁序列 $\sigma \in T^*$ 到达标识 M' , 记作 $M[\sigma]M'$. 所有从标识 M 可达的标识集记为 $R(M) = \{M' \mid \exists \sigma \in T^*, M[\sigma]M'\}$. 若标识 M_f 满足 $M_f \in R(M_i) \wedge M_f(o) \neq \emptyset$, 则称 M_f 为终止标识, 所有终止标识的集合记为 M_F .

定义 2(正确的终止标识). 服务 S 的终止标识 M_f 若满足 $\forall p \in P^{IC}, p \neq o; M_f(p) = \emptyset$, 则称 M_f 为正确的终止标识, 记为 M_{CF} . 所有可达的正确终止标识的集合记为 M_{CF} .

由定义 2 可知, 正确的终止标识要求终止库所 o 中存在令牌, 且其它的内部控制逻辑库所内不能含有令牌. 反之, 如果终止库所 o 与某个内部控制逻辑

库所同时含有令牌, 则表明该服务流程并没有正确结束. 定义 2 中的正确终止状态并不要求所有非终止库所都不能含有令牌, 而可以在外部或内部通信库所中留有未被消耗的令牌. 这样的要求符合 Web 服务基于消息通信且彼此松散耦合的本质特性.

定义 3(可达变迁序列). 若变迁序列 $\sigma = \langle t_1, t_2, \dots, t_n \rangle \in T^*$ 满足 $\exists M; M_i[\sigma]M$, 则称 σ 为服务 S 的一个可达变迁序列. 所有可达变迁序列的集合记为 $\Sigma(S)$.

定义 4(流程正确性). 服务 S 的流程是正确的, 当且仅当在外界环境总是能够满足服务 S 对外部消息的需求的情况下, 下列两个条件同时满足:

(1) $\forall M \in R(M_i)$, 如果 $M \notin M_F$, 则 $\exists \sigma \in \Sigma(S)$, 使得 $M[\sigma]M_f \in M_F$;

(2) $M_{CF} = M_F$.

定义 4 要求外界环境总是能够满足服务 S 对外部消息的需求, 也即: 当服务需要接收 (或发送) 消息时, 外界环境总是能够适时地发送 (或接收) 该消息. 在这种情况下, 排除了因外界环境导致服务流程出现死锁的情况. 这样的前提是合理且可行的^[12]. 在实际分析过程中, 只需要将服务的外部通信库所 P^{EM} 以及所有与 P^{EM} 关联的弧去除即可, 即只需要考察服务 S 对应的工作流网 $N = \Phi(S)$. 根据定义 4, 服务流程正确需要满足两个条件. 第 1 个条件要求服务流程总是能够到达终止标识, 不会发生死锁; 第 2 个条件则要求服务流程到达的终止标识都必须是正确的终止标识.

3 相容性和相似性分析

3.1 相容性分析

如果参与组合的若干成员服务中某个成员服务的流程不正确, 那么该服务必将导致整个组合服务的流程无法正确执行. 因此, 本文在后续讨论中均默认所有参与组合的成员服务的流程是正确的. 此时组合服务的流程能否正确执行取决于各成员服务的流程是否相容. 本组已有工作提出了 Web 服务组合的形式化定义^[11,13], 本文在此基础上给出 Web 服务流程相容性的定义.

定义 5(服务流程相容). 设两个服务 S_1 和 S_2 能够组合, 其组合服务 $S = S_1 \oplus S_2$. 服务 S_1 和 S_2 的流程是相容的, 记为 $S_1 \sim S_2$, 当且仅当组合服务 S 的流程是正确的.

不难看出上述服务流程的相容性定义满足对称性,即 $S_1 \sim S_2 \Leftrightarrow S_2 \sim S_1$.

为了简化相容性的验证条件,本文考察一类具有良好结构的 Web 服务,即“良构”的 Web 服务. 工作流理论通常要求所研究的工作流网是良构的,即要求网结构中由 And-Split 发出的两个并行分支必须由 And-Join 汇合,由 Or-Split 发出的两个选择分支必须由 Or-Join 汇合. 在工作流网“良构性”定义的基础上^[14],本文提出 Web 服务良构性的定义.

定义 6(服务良构性). 服务 S 是良构的,当且仅当 S 对应的工作流网 $N = \Phi(S)$ 是良构的.

根据上述定义,易知图 1 所示的服务 S_1 是良构的,而图 2 所示的服务 S_2 则是非良构的.

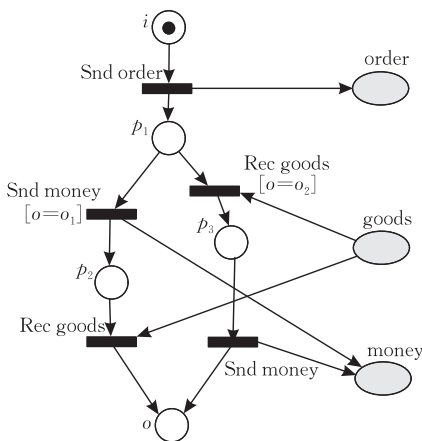


图 1 一个良构的服务 S_1

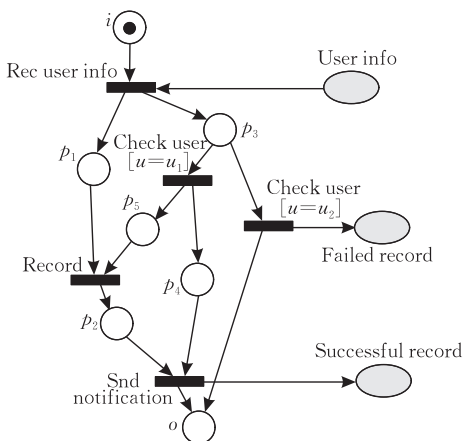


图 2 一个非良构的服务 S_2

下面给出一个关于良构的 Web 服务及其组合的重要结论.

定理 1. 设两个服务 S_1 和 S_2 能够组合,其组合服务 $S = S_1 \oplus S_2$. 如果服务 S_1, S_2 是良构的,那么服务 S 从初始标识可达的任一终止标识都是正确的,即 $M_{CF} = M_F$.

证明(反证法). 由于 $M_{CF} \subseteq M_F$,若 $M_{CF} \neq M_F$,则 $\exists M_f \in M_F$,但 $M_f \notin M_{CF}$. 由 $M_f \in M_F$ 可知, $M_f \in R(S, M_i)$ 且 $M_f(o) \neq \emptyset$. 由于 $M_f \notin M_{CF}$,根据定义 2 可知, $\exists p \in P^{IC}, p \neq o$ 且 $M_f(p) \neq \emptyset$. 根据服务组合的定义可知^[11], $P^{IC} = P_1^{IC} \cup P_2^{IC} \cup \{i, o\}$,若 $p = i$,则 $M'_f \in M_F$,使得 $M'_f(i_1) \neq \emptyset$ 且 $M'_f(i_2) \neq \emptyset$. 因此不妨假设 $p \in P_1^{IC}$. 由 $M_f \in M_F$ 可知,服务 S_1 存在某个可达的终止标识 M_{1f} 使得 $M_{1f}(o_1) \neq \emptyset$ 且 $M_{1f}(p) \neq \emptyset$. 因为服务 S_1 对应的工作流网 $N_1 = \Phi(S_1)$ 是强连通的,那么网 N_1 中存在一条从 i_1 到 o_1 的基本路径 C_1 使得 $p \notin C_1$,并且存在一条从 i_1 到 o_1 的基本路径 C_2 使得 $p \in C_2, C_2 \neq C_1$. 进而可知,存在变迁 $t \in C_1 \cap C_2$,而库所 $o_1 \in C_1 \cap C_2$. 这就与服务 S_1 是良构的前提产生矛盾. 因此原命题得证. 证毕.

根据定义 4 和定义 5,判定两个服务的流程是否相容需要验证两个条件,即组合服务的流程总是能够到达终止标识,并且所有可达的终止标识都是正确的. 而定理 1 的结论实际上指出,由良构的服务组合构成的服务,如果从初始标识能够到达某个终止标识,那么该终止标识一定是正确的. 由此可见,成员服务的良构性有效地保证了组合服务可达终止标识的正确性. 因此,不难得到下面的结论.

定理 2. 设两个良构的服务 S_1 和 S_2 能够组合,其组合服务 $S = S_1 \oplus S_2$. 服务 S_1 和 S_2 的流程是相容的,即 $S_1 \sim S_2$,当且仅当在外界环境总是能够满足服务 S 对外部消息的需求的情况下, $\forall M \in R(M_i)$,如果 $M \notin M_F$,则 $\exists \sigma \in \Sigma(S)$,使得 $M[\sigma] M_f \in M_F$.

证明. 由定义 4 和定理 1 可直接得证.

利用定理 2 的结论,在判定两个良构的服务是否流程相容时,只需要验证其组合服务从初始标识是否总是能够到达终止标识,而不会在某个标识下发生死锁的现象. 需要指出的是,定理 1 和定理 2 的结论可以直接推广到 3 个或 3 个以上服务组合的情况. 据我们观察,现实中独立开发的大多数服务(如基于 BPEL 的服务流程)都能够遵循良构性,因此利用定理 2 的结论可以大大简化服务流程相容性的分析过程.

3.2 相似性分析

分析服务流程的相似性对于研究服务发现和服务替换技术具有重要意义^[15]. 当参与组合的某个成员服务由于某些原因不能继续提供服务时,将导致组合服务无法正确执行. 此时需要发现并选取一个与失效服务的流程相似的服务来替换. 在服务组合研究中,替换服务与被替换服务的流程相似就必须

要保证替换之后的组合服务仍然是正确的,也即需要保证替换服务的流程与服务组合系统中其他成员服务的流程相容.可见,服务流程的相似性分析与相容性分析密切相关.

由于相容性分析通常需要验证组合服务全局流程的正确性.若替换频繁发生,不断重复的服务组合验证将耗费大量系统时间和资源^[5].因此,服务相似性定义需要与上下文无关,即在不需验证组合服务全局流程的情况下就能判断选取的替换服务如果与被替换服务相似,那么就能够与服务组合系统中的其他成员服务相容,从而保证替换后的组合服务的流程仍然是正确的.

两个服务相似并不一定要求两者的内部流程完全一致,而只需要两者在与外界环境交互通信的过程中表现出一定的相似性.因此,服务相似性的定义和分析主要关注服务表现出来的外部行为特性.下面给出几个关于服务外部行为特性的定义.

定义 7(通信变迁序列). 对于某个可达变迁序列 $\sigma \in \Sigma(S)$,若忽略 σ 中可能存在的内部变迁,则得到相应的通信变迁序列 $\epsilon = \langle t'_1, t'_2, \dots, t'_k \rangle$,其中 $L(t'_l) \neq \tau, l = 1, 2, \dots, k$.记得到相应的通信变迁序列的操作算子为 ϵ .

定义 8(相似变迁序列). 两个可达变迁序列 $\sigma_1, \sigma_2 \in \Sigma(S)$,其对应的通信变迁为 $\epsilon_1 = \epsilon(\sigma_1) = \langle t'_1, t'_2, \dots, t'_k \rangle, \epsilon_2 = \epsilon(\sigma_2) = \langle t''_1, t''_2, \dots, t''_l \rangle$. σ_1, σ_2 相似,记为 $\sigma_1 \approx \sigma_2$,当且仅当 $k = l, \forall j = \{1, \dots, k\}, L(t'_j) = L(t''_j), \varphi(t'_j) = \varphi(t''_j)$ 且 $G(\sigma_1) = G(\sigma_2)$.

定义 9(对偶变迁序列). 两个可达变迁序列 $\sigma_1, \sigma_2 \in \Sigma(S)$,其对应的通信变迁为 $\epsilon_1 = \epsilon(\sigma_1) = \langle t'_1, t'_2, \dots, t'_k \rangle, \epsilon_2 = \epsilon(\sigma_2) = \langle t''_1, t''_2, \dots, t''_l \rangle$. σ_1, σ_2 互为对偶变迁序列,记为 $\sigma_1 = \overline{\sigma_2}$,当且仅当 $k = l, \forall j = \{1, \dots, k\}, L(t'_j) = L(t''_j), \varphi(t'_j) = -\varphi(t''_j)$ 且 $G(\sigma_1) = G(\sigma_2)$.

在定义 8 和定义 9 中,变迁的极性函数 $\varphi: T \rightarrow \{+1, 0, -1\}$ 定义为:若变迁 t 向外界发送消息,则 $\varphi(t) = -1$;若变迁 t 从外界接收消息,则 $\varphi(t) = +1$;若变迁 t 不是通信变迁,则 $\varphi(t) = 0$.另外,本文将着色 Petri 网定义中的警卫(Guard)函数的定义域从单个变迁集扩展到可达变迁序列集.可达变迁序列的警卫函数 $G: \Sigma(S) \rightarrow BoolExpression$,即 $\forall \sigma \in \Sigma(S), \sigma = \langle t_1, \dots, t_n \rangle; G(\sigma) = G(t_1) \wedge \dots \wedge G(t_n)$.在此基础上,本文给出两个服务流程相似和等价的定义.

定义 10(服务流程相似). 给定两个服务 $S_1、$

S_2 ,服务 S_2 的流程与服务 S_1 的流程相似,记为 $S_2 > S_1$,当且仅当 $\forall \sigma_2 \in \Sigma(S_2), M_{2i}[\sigma_2]M_2$,则 $\exists \sigma_1 \in \Sigma(S_1), M_{1i}[\sigma_1]M_1, \sigma_1 \approx \sigma_2$,并且满足:

(1) 如果 $M_1 \in M_{1F}$ 或 $M_1[\tau^*]M_{1f} \in M_{1F}$,那么 $M_2 \in M_{2F}$ 或 $M_2[\tau^*]M_{2f} \in M_{2F}$;

(2) 如果 $\exists t_1 \in T_1, t_1 \neq \tau$,使得 $M_1[\tau^*]M'_1 \notin M_{1F}, M'_1[t_1]$,那么 $\exists t_2 \in T_2, t_2 \neq \tau, L(t_2) = L(t_1), \varphi(t_2) = \varphi(t_1)$,使得 $M_2[\tau^*]M'_2 \notin M_{2F}, M'_2[t_2]$.

定义 11(服务流程等价). 两个服务 S_1 和 S_2 的流程等价,记为 $S_1 \approx S_2$,当且仅当 $S_1 > S_2$ 且 $S_2 > S_1$.

如定义 10 所述,服务 S_2 的流程与服务 S_1 的流程相似并不能得出服务 S_1 的流程与服务 S_2 的流程相似.可见两个服务流程的相似关系并不满足对称性.而定义 11 所述的服务流程等价关系则满足对称性,即 $S_1 \approx S_2 \Leftrightarrow S_2 \approx S_1$.需要指出的是,根据定义 10 和定义 11 不难证明服务流程的相似关系和等价关系均满足传递性,即 $S_1 > S_2, S_2 > S_3 \Rightarrow S_1 > S_3; S_1 \approx S_2, S_2 \approx S_3 \Rightarrow S_1 \approx S_3$.

如前所述,服务流程的相似性定义需要与上下文无关.下面将证明给出的定义 10 对于良构的服务是上下文无关的.

引理. 设两个良构的服务 S_1 和 S_2 能够组合,服务 S_1 和 S_2 的流程是相容的,即 $S_1 \sim S_2$,当且仅当:对于 $\forall \sigma_1 \in \Sigma(S_1), \forall \sigma_2 \in \Sigma(S_2), M_{1i}[\sigma_1]M_1, M_{2i}[\sigma_2]M_2$,如果从 $\sigma_1(\sigma_2)$ 中去除那些与 $\sigma_2(\sigma_1)$ 无关的通信变迁之后,分别得到变迁序列 σ'_1 和 σ'_2, σ'_1 和 σ'_2 互为对偶变迁序列,那么下列两个条件必有一个满足

(1) $M_1 \in M_{1F}$ 或 $M_1[\tau^*]M_{1f} \in M_{1F}$,且 $M_2 \in M_{2F}$ 或 $M_2[\tau^*]M_{2f} \in M_{2F}$;

(2) $\exists t_1 \in T_1, t_1 \neq \tau, \exists t_2 \in T_2, t_2 \neq \tau, L(t_1) = L(t_2), \varphi(t_1) = -\varphi(t_2)$,使得 $M_1[\tau^*]M'_1 \notin M_{1F}, M'_1[t_1], M_2[\tau^*]M'_2 \notin M_{2F}, M'_2[t_2]$.

证明. 记服务 S_1 和 S_2 的组合服务为 S .一方面,如果已知服务 S_1 和 S_2 的流程相容, σ'_1 和 σ'_2 互为对偶变迁序列,那么此时组合服务 S 到达某个可达标识 M .如果 M 是终止标识,则 $M_1 \in M_{1F}$ 且 $M_2 \in M_{2F}$.反之,根据定理 2 可知,服务 S_1 和 S_2 均没有到达终止标识,并且在标识 M 下存在某个变迁可以使能触发.因此命题的必要性得证.

另一方面,如果 σ'_1 和 σ'_2 互为对偶变迁序列,不妨设组合服务 S 到达某个可达标识 M .如果 M 不是一个终止标识,那么或者 $M_1[\tau^*]M_{1f} \in M_{1F}$ 且

$M_2[\tau^*]M_{2f} \in M_{2F}$, 或者 $\exists t_1 \in T_1, t_1 \neq \tau, \exists t_2 \in T_2, t_2 \neq \tau, L(t_1) = L(t_2), \varphi(t_1) = -\varphi(t_2)$, 使得 $M_1[\tau^*]M'_1 \notin M_{1F}, M'_1[t_1] \in M_{1F}, M_2[\tau^*]M'_2 \notin M_{2F}, M'_2[t_2] \in M_{2F}$. 无论哪种情况, 都可知 $\exists \sigma \in \Sigma(S)$, 使得 $M[\sigma]M_f \in M_F$. 根据定理 2 可知, 命题的充分性得证. 证毕.

定理 3. 设服务 S_1, S_2 和 S 是良构的, 如果 $S_2 > S_1, S \sim S_1$, 那么 $S \sim S_2$.

证明. 对于 $\forall \sigma \in \Sigma(S), \forall \sigma_2 \in \Sigma(S_2), M_i[\sigma]M, M_{2i}[\sigma_2]M_2$, 从 $\sigma(\sigma_2)$ 中去除那些与 $\sigma_2(\sigma)$ 无关的通信变迁之后, 分别得到变迁序列 σ' 和 σ'_2 , σ' 和 σ'_2 互为对偶变迁序列. 因为 $S_2 > S_1$, 根据定义 10 可知, $\exists \sigma_1 \in \Sigma(S_1), M_{1i}[\sigma_1]M_1, \sigma_1 \approx \sigma_2$. 因此, 当 σ_1 中去除那些与 σ 无关的通信变迁之后得到变迁序列 σ'_1 , 可知 σ'_1 和 σ' 互为对偶变迁序列. 又因为 $S \sim S_1$, 根据引理结论可知, $M \in M_F$ 或 $M[\tau^*]M_f \in M_F$, 且 $M_1 \in M_{1F}$ 或 $M_1[\tau^*]M_{1f} \in M_{1F}$; 或者 $\exists t \in T, t \neq \tau, \exists t_1 \in T_1, t_1 \neq \tau, L(t) = L(t_1), \varphi(t) = -\varphi(t_1)$, 使得 $M[\tau^*]M' \notin M_F, M'[t], M_1[\tau^*]M'_1 \notin M_{1F}, M'_1[t_1] \in M_{1F}$. 因为 $S_2 > S_1$, 根据定义 10 可知, 若 $M_1 \in M_{1F}$ 或 $M_1[\tau^*]M_{1f} \in M_{1F}$, 则 $M_2 \in M_{2F}$ 或 $M_2[\tau^*]M_{2f} \in M_{2F}$. 反之, 若 $\exists t_1 \in T_1, t_1 \neq \tau$ 使得 $M_1[\tau^*]M'_1 \notin M_{1F}, M'_1[t_1] \in M_{1F}$, 则 $\exists t_2 \in T_2, t_2 \neq \tau$ 使得 $M_2[\tau^*]M'_2 \notin M_{2F}, M'_2[t_2] \in M_{2F}$, 且 $L(t_2) = L(t_1), \varphi(t_2) = \varphi(t_1)$. 根据引理结论可知, S 和 S_2 的流程是相容的, 即 $S \sim S_2$. 证毕.

定理 3 的结论表明, 在服务都是良构的前提下, 如果某个服务 S_2 与另一个服务 S_1 相似, 那么对于所有与服务 S_1 流程相容的服务 S , 服务 S_2 与服务 S 的流程必然也是相容的. 这一点在进行服务替换的过程中非常有用. 如果服务 S_1 是某个组合服务系统中的一个成员服务, 当 S_1 由于某些原因无法继续提供服务或者无法保证服务质量要求而需要被替换时, 只需要找到一个与 S_1 相似的服务 S_2 进行替换即可, 而不需要再对替换后的组合服务进行正确性验证. 这是因为定理 3 的结论保证了用 S_2 替换 S_1 之后的服务组合仍然是正确的. 由此可见, 本文给出的服务相似性的定义是与上下文无关的.

一些分析服务相容性的方法用到了“反(opposite)”服务的概念^[16]. 对于一个服务 S , 改变其中所有与外部通信库所关联的弧的方向即可得到服务 S 的反服务, 记为 \bar{S} . 可见得到反服务的操作非常简单, 并且易知一个服务与其反服务总是流程相容的, 即 $S \sim \bar{S}$. 下面给出与反服务有关的两个结论. 这两个结论对已有的基于反服务概念来分析服务相容性的方法具有一定应用价值.

推论 1. 设服务 S 和 S_1 是良构的, 服务 S 的反服务为 \bar{S} . 如果 $S_1 > S$, 那么 $S_1 \sim \bar{S}$.

证明. 因为 $S \sim \bar{S}, S_1 > S$, 由定理 3 可知 $S_1 \sim \bar{S}$. 证毕.

推论 2. 设服务 S_1, S_2 和 S 是良构的, 服务 S 的反服务为 \bar{S} . 如果 $S_1 > S, S_2 > \bar{S}$, 那么 $S_1 \sim S_2$.

证明. 因为 $S_1 > S$, 由推论 1 可知: $S_1 \sim \bar{S}$. 又因为 $S_2 > \bar{S}$, 由定理 3 可知: $S_1 \sim S_2$. 证毕.

下面给出一个算法用于判断两个 Web 服务流程之间是否存在相似关系, 即判断服务 S_2 的流程是否与服务 S_1 的流程相似. 该算法本质上需要分别遍历服务 S_1 和 S_2 的通信可达树^[11], 并根据定义 10 来判断是否存在相似性. 由于通信可达树的节点个数有限, 且通信变迁个数也有限, 因此该算法是可以终止的.

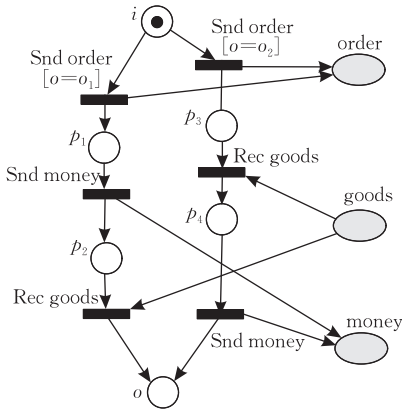
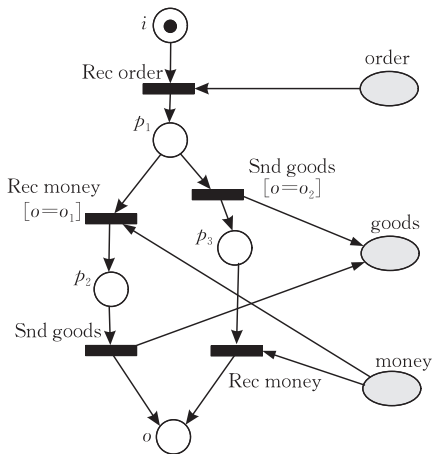
服务流程相似性判定算法如下:

```

Set  $T(S, M)$ ; //for all transitions of service  $S$  that are
//enabled in the marking  $M$ 
Set  $U_2 = \emptyset$ ; //the set of markings of service  $S_2$  that
//has been checked
Set  $M_1 = M_{1i}, M_2 = M_{2i}$ ;
IsSimilar ( $S_2, S_1$ )
1. if  $M_2 \in M_{2F}$ 
   if  $M_1 \in M_{1F}$  or  $M_1[\tau^*]M_{1f} \in M_{1F}$ 
     then return TRUE;
   else return FALSE;
2. for each  $t_2 \in T(S_2, M_2)$  and  $M_2 \notin U_2$ 
   if  $t_2 = \tau, M_2[\tau]M'_2$ 
     then add  $M_2$  to  $U_2, M_2 = M'_2$ ,
     return IsSimilar ( $S_2, S_1$ );
   else  $M_2[t_2]M'_2$ 
     if  $\exists t_1 \in T_1, t_1 \neq \tau, M_1[\tau^*]M'_1, M'_1[t_1]M''_1$ ,
     and  $L(t_2) = L(t_1), \varphi(t_2) = \varphi(t_1)$ 
       then if  $M''_1 \in M_{1F}$  or  $M''_1[\tau^*]M_{1f} \in M_{1F}$ 
         if  $M'_2 \in M_{2F}$  or  $M'_2[\tau^*]M_{2f} \in M_{2F}$ 
           then return TRUE;
         else return FALSE;
       else  $\exists t'_1 \in T_1, t'_1 \neq \tau, M''_1[\tau^*]M_1^{(3)}, M_1^{(3)}[t'_1]M_1^{(4)}$ 
         if  $\exists t'_2 \in T_2, t'_2 \neq \tau, M'_2[\tau^*]M_2''', M_2'''[t'_2]M_2^{(3)}$ ,
         and  $L(t'_2) = L(t'_1), \varphi(t'_2) = \varphi(t'_1)$ 
           then add  $M_2$  to  $U_2, M_1 = M_1^{(4)}, M_2 = M_2^{(3)}$ ,
           return IsSimilar ( $S_2, S_1$ );
         else return FALSE;
     else return FALSE.

```

根据上述判定算法可以判断得到: 图 3 所示的服务 S_3 与图 1 所示的服务 S_1 相似. 实际上可以进一步得到, 服务 S_1 和 S_3 是等价的. 图 4 给出了服务 S_1 的反服务 S_4 , 不难发现服务 S_1 与服务 S_4 是相容的, 服务 S_3 与服务 S_4 也是相容的, 从而验证了前面得到的结论.

图 3 与服务 S_1 等价的服务 S_3 图 4 服务 S_1 的反服务 S_4

4 相关工作及讨论

组合服务流程的正确性验证是当前面向服务计算领域的一个研究热点. 文献[14]以工作流网为基础, 给出了工作流网正确性的定义, 即要求同时满足可终止条件、正确终止条件以及无死变迁条件. 文献[12]指出 Web 服务流程正确可以允许流程内部存在死变迁, 由此提出“弱合理性(weak soundness)”概念. 文献[17]讨论了服务流程相容的 3 种不同的定义, 并指出“完成(completing)相容”能够正确验证出所有异步服务流程死锁的问题. 然而, 这些定义并不能完全反映 Web 服务基于消息通信以及松散耦合的本质特性.

由于 Petri 网适合于对分布式异步系统进行建模, 因此许多研究工作均采用 Petri 网作为服务流程建模和分析的理论基础. 文献[18]给出了一个基于 Petri 网的 Web 服务模型, 提出了 10 种不同的服务组合方式, 并指出可以借助 Petri 网的分析技术对组合服务进行验证. 文献[2]提出了基于着色

Petri 网的 Web 服务模型, 给出了 4 种主要的组合运算, 并讨论了这 4 种组合运算以及组合模型的某些性质. 文献[6]以着色 Petri 网为基础提出了面向消息的 Web 服务模型, 该模型侧重于对 Web 服务元消息的描述及其对服务协同通信的影响. 文献[7]提出了一种基于 Petri 网的语义 Web 服务自动组合方法, 该方法能够同时考虑服务的输入/输出以及服务的行为约束. 虽然 Petri 网能够对组合服务进行分析和验证, 但是上述研究成果都没有对此做深入讨论, 没有指出具有哪些性质的两个或多个服务的流程是相容的. 文献[19]借助“虹(siphons)”理论研究了基于 Petri 网模型的 Web 服务流程相容性. 然而, 上述研究工作大都没有涉及到服务流程的相似性和等价性分析.

服务流程相似性分析是与相容性分析密切相关的一个问题. 文献[15]提到了研究服务流程相似和等价的必要性, 区分了服务发现和服务替换两种不同应用背景下的相似性定义, 但是作者没有讨论相似性与相容性之间的关系, 也没有分析所给出的相似性定义在进行服务替换时是否与上下文无关. 执行路径(trace)相似和互模拟(bisimulation)相似是异步并发系统行为相似性研究中两个最常用的概念^[20]. 但是这两个概念并不能完全反映 Web 服务流程相容的特性和需求. 执行路径相似的要求过于宽松, 而互模拟相似的要求则过于苛刻^[21]. 文献[16]主要研究两个服务流程的相容性和可替换性(substitutability). 作者借助标记变迁系统(LTS)对 Web 服务的行为进行建模, 提出了 3 个不同概念意义下的相容性定义, 并相应地给出了 3 种不同的可替换性定义, 同时指出可替换性定义分为与上下文相关和与上下文无关两种. 但是作者并没有深入讨论与上下文无关的可替换性需要满足的充分条件. 文献[5]采用进程代数建模 Web 服务流程, 给出了与上下文无关的服务流程一致性的定义及其验证算法. 作者指出 Web 服务需要为每种接收到的消息类型设置缓存区, 然而基于进程代数的方法无法显式地建模这些消息缓存区. 文献[22]使用 Martin 类型论(Martin Type Theory, MTT)建模 Web 服务行为, 研究了 Web 服务行为一致性与相容性的判定方法. 但是基于进程代数或 MTT 的方法并不能对服务的内部选择操作进行精细的刻画, 也无法建模和分析 Web 服务的结构性性质(如良构性等).

上述相关工作为本文展开服务流程相容性和相似性研究提供了很好的启示和基础. 与之相比, 本文

研究工作的特色在于:(1)采用着色 Petri 网建模 Web 服务流程,该模型能够对服务流程的消息缓存区和内部选择操作进行精细的刻画,较其他建模方法更为直观.在此基础上给出了服务流程正确性的定义,该定义允许服务流程正确结束时其内部或外部消息缓存区(也即通信库所)中留有未被消耗的消息.与现有正确性定义相比,该定义的要求更为宽松,更符合 Web 服务基于消息通信、松散耦合的本质特性;(2)分析了 Web 服务的结构性性质,提出了 Web 服务良构性的概念,并指出成员服务的良构性有效地保证了组合服务可达终止标识的正确性,从而简化了服务流程相容性分析的条件;(3)基于着色 Petri 网模型提出了服务流程相似性的定义,该定义考虑了服务流程的内部选择操作.证明了该定义在进行服务替换时是与上下文无关的,并且讨论了服务流程相似性与相容性之间的关系.

5 结束语

在服务组合验证研究中,服务流程的相容性和相似性分析是两个密切相关的问题,具有较大的实用价值.本文采用着色 Petri 网建模 Web 服务的流程,给出了服务流程正确性的定义,并在此基础上定义了服务流程的相容性.提出了服务良构性的概念,并指出成员服务的良构性能够保证组合服务可达终止状态是正确的,从而简化了相容性分析的条件.在服务流程相似性分析方面,本文给出了相似性的定义,证明了满足相似关系的两个服务在进行替换时是上下文无关的,并给出了相似性判定算法.本文工作对于研究服务发现、服务组合和服务替换具有重要参考价值.

未来工作主要在 3 个方面展开:(1)本文中相似性判定算法的复杂度随通信可达树节点个数呈指数增长,未来将考虑采用稳固集(stubborn sets)的方法提高算法在产生状态空间方面的性能;(2)将研究当两个 Web 服务的流程不完全相容时,是否可以通过产生中介器(mediator)来帮助两者实现流程相容.目前学术界针对该问题已经有了一些研究成果,但是还没有一个系统化的解决办法;(3)本研究组正在开发和完善一套 Web 服务组合工具,以便自动化或半自动化地实现服务流程的建模、组合以及相容性和相似性分析.

参 考 文 献

- [1] Han Yan-Bo, Wang Hong-Cui, Wang Jian-Wu et al. An end-user-oriented approach to exploratory service composition. *Journal of Computer Research and Development*, 2006, 43(11): 1895-1903(in Chinese)
(韩燕波,王洪翠,王建武等.一种支持最终用户探索式组合服务的方法. *计算机研究与发展*, 2006, 43(11): 1895-1903)
- [2] Guo Yu-Bin, Du Yu-Yue, Xi Jian-Qing. A CP-net model and operation properties for Web service composition. *Chinese Journal of Computers*, 2006, 29(7): 1067-1075(in Chinese)
(郭玉彬,杜玉越,奚建清. Web 服务组合的有色网模型及运算性质. *计算机学报*, 2006, 29(7): 1067-1075)
- [3] Foster H, Uchitel S, Kramer J et al. Compatibility verification for Web service choreography//*Proceedings of the IEEE International Conference on Web Services*. San Diego, CA, USA, 2004: 738-741
- [4] Deng Shui-Guang, Li Ying, Wu Jian et al. Determination and computation of behavioral compatibility for Web services. *Journal of Software*, 2007, 18(12): 3001-3014 (in Chinese)
(邓水光,李莹,吴健等. Web 服务行为兼容性的判定与计算. *软件学报*, 2007, 18(12): 3001-3014)
- [5] Liu Fang-Fang, Shi Yu-Liang, Zhang Liang et al. Substitution analysis of Web service composition via process algebra. *Chinese Journal of Computers*, 2007, 30(11): 2033-2039(in Chinese)
(刘方方,史玉良,张亮等.基于进程代数的 Web 服务合成的替换分析. *计算机学报*, 2007, 30(11): 2033-2039)
- [6] Qian Zhu-Zhong, Lu Sang-Lu, Xie Li. Automatic composition of Petri net based Web services. *Chinese Journal of Computers*, 2006, 29(7): 1057-1066(in Chinese)
(钱柱中,陆桑璐,谢立.基于 Petri 网的 Web 服务自动组合研究. *计算机学报*, 2006, 29(7): 1057-1066)
- [7] Tang Xian-Fei, Jiang Chang-Jun, Ding Zhi-Jun et al. A Petri net-based semantic Web service automatic composition method. *Journal of Software*, 2007, 18(12): 2991-3000 (in Chinese)
(汤宪飞,蒋昌俊,丁志军等.基于 Petri 网的语义 Web 服务自动组合方法. *软件学报*, 2007, 18(12): 2991-3000)
- [8] Hu Chun-Hua, Wu Min, Liu Guo-Peng. QoS scheduling based on trust relationship in Web service workflow. *Chinese Journal of Computers*, 2009, 32(1): 42-53(in Chinese)
(胡春华,吴敏,刘国平. Web 服务工作流中基于信任关系的 QoS 调度. *计算机学报*, 2009, 32(1): 42-53)
- [9] Jensen K. *Coloured Petri Nets: Basic Concepts, Analysis Methods and Practical Use*. Vol. 1. Berlin, Heidelberg, New York: Springer-Verlag, 1997
- [10] Ouyang C, Verbeek E, van der Aalst W M P et al. Formal semantics and analysis of control flow in WS-BPEL. *Science of Computer Programming*, 2007, 67(2-3): 162-198
- [11] Tan W, Fan Y S, Zhou M C. A Petri net-based method for compatibility analysis and composition of Web services in business process execution language. *IEEE Transactions on Automation Science and Engineering*, 2009, 6(1): 94-106

- [12] Martens A. On compatibility of Web services//Robert Lorenz, Ekkart Kindler eds. Petri Net Newsletter. Bonn, Germany, 2003; 12-20
- [13] Li Xi-Tong, Fan Yu-Shun. Modeling and logical correctness verification of Web service processes. Computer Integrated Manufacturing Systems, 2008, 14(4): 675-682(in Chinese) (李喜彤, 范玉顺. Web 服务过程建模及其逻辑正确性验证. 计算机集成制造系统, 2008, 14(4): 675-682)
- [14] van der Aalst W M P. Workflow verification: Finding control-flow errors using Petri net-based techniques//Business Process Management. Berlin: Springer-Verlag, 2000; 161-183
- [15] Hu Hao, Yin Qin, Lu Jian. Service behavior and quality consistency in virtualized computing environment. Journal of Software, 2007, 18(8): 1943-1957(in Chinese) (胡昊, 殷琴, 吕建. 虚拟计算环境中服务行为与质量的一致性. 软件学报, 2007, 18(8): 1943-1957)
- [16] Bordeaux L, Salun G, Berardi D et al. When are two Web services compatible//Proceedings of the TES 2004. LNCS 3324. Toronto, Canada, 2005; 15-28
- [17] Dijkman R. Notions of behavioral compatibility and their im-
- plications for BPEL processes. University of Twente, Netherlands; Technical Report TR-CTIT-06-41, 2006
- [18] Rachid H, Boualem B. A Petri net-based model for Web service composition//Proceedings of the 14th Australian Database Conference on Database Technologies. Adelaide, South Australia, 2003; 191-200
- [19] Xiong P X, Fan Y S, Zhou M C. A Petri net approach to analysis and composition of Web services. IEEE Transactions on System, Man and Cybernetics, Part A, 2009(in press)
- [20] Milner R. Communication and Concurrency. Upper Saddle River, NJ, USA: Prentice Hall, 1989
- [21] Martens A. Simulation and equivalence between Bpel process models//Proceedings of the International Conference on the Design, Analysis and Simulation of Distributed Systems. San Diego, California, 2005
- [22] Yin Yu-Yu, Li Ying, Deng Shui-Guang et al. Determining on consistency and compatibility of Web services behavior. Acta Electronica Sinica, 2009, 37(3): 433-438(in Chinese) (殷昱煜, 李莹, 邓水光等. Web 服务行为一致性相容性判定. 电子学报, 2009, 37(3): 433-438)



LI Xi-Tong, born in 1982, Ph. D. candidate. His research interests include Web services, service oriented computing, and Petri nets theory and application.

FAN Yu-Shun, born in 1962, professor, Ph. D. supervisor. His research interests include enterprise modeling, workflow management, and service oriented architecture.

Background

Service Oriented Computing (SOC) is a new computing paradigm in the area of software engineering and architecture. Service composition and substitution are key research issues in SOC, among which analyzing compatibility and similarity of Web service processes are two closely-related issues and of great importance to both the academia and industry. A few approaches have been developed to verify the correctness of service composition, but their notions of compatibility and similarity of Web service processes are not well-defined. Particularly, the relationship between compatibility and similarity lacks of sufficient research, rendering existing techniques of service substitution very costly and time-consuming. The authors' research work focuses on modeling and analyzing Web service processes based on Colored Petri nets (CPN). The authors present the definitions of correctness and compatibility of Web service processes. In order to alleviate the compatibility verification, the authors develop the notion of well-structuredness of Web services and prove that all reachable final markings of a composite service composed by well-

structured services are correct. Based on the above results of compatibility analysis, the authors explore the definition of similarity of Web service processes with the purpose to improve the existing methods of service substitution. The results conclude that substituting a service in a composition can be performed independent of other partners in the composition, as long as the new service is similar to the substituted one. Therefore, it is no need to verify the substituted service composition again. The authors develop an algorithm for automatically verifying similarity between two services. The results and algorithm concluded in the research can be used to improve the existing methods and develop more efficient techniques for service composition verification and service substitution.

This work is supported by the National Natural Science Foundation of China under grant No.60674080 and No.60704027, and the National High Technology Research and Development Program (863 Program) of China under grant No.2007AA04Z150.