

一种验证业务级服务组合可执行能力的方法^{*})

单保华 韩燕波 李厚福 胡海涛

(中国科学院计算技术研究所软件研究室 北京 100080)

摘 要 在面向服务的环境中,服务组合的可执行能力具有很大的不确定性,在业务级别组合服务更是如此。影响业务级服务组合可执行能力的因素是多方面的,本文针对组合的服务的逻辑结构和业务服务到具体服务的匹配方法对业务级服务组合可执行能力的影响,提出了一种基于着色时间 Petri 网(CTPN)的业务级服务组合可执行能力验证方法,并以制造业网格应用平台 AmGrid 为案例,展示了该方法在平台中的应用效果。

关键词 业务级服务组合,着色时间 Petri 网,可执行能力验证

An Approach to Verify Executable Ability of Business Level Service Composition

SHAN Bao-Hua HAN Yan-Bo LI Hou-Fu HU Hai-Tao

(Institute of Computing Technology, Chinese Academy of Science, Beijing 100080)

Abstract The executable ability of service composition in service-oriented environment has much uncertainty and business level service composition adds more uncertainty. There are many factors affecting the executable ability of business level service composition. Aimed at the influence of composition logic of service and method of matching from business service to concrete service on the executable ability of business level service composition, this paper presents an approach based on colored - timed Petri net to verify the executable ability of business level service composition and then exhibits the usage of this approach in AmGrid.

Keywords Business level service composition, Colored - timed Petri net, Verify of executable ability

1 引言

针对动态、开放的网格环境下业务用户按需创建个性化应用的需求,我们提出了一种面向业务用户的应用即时构造方法—CAFISE^[1],旨在使用户从业务角度透明地使用服务资源并自主地组合服务构造应用,其核心是面向业务用户的服务资源虚拟化和业务级服务动态组合。VINCA 语言^[6]是业务级服务组合的建模语言,它通过对虚拟化后的资源-业务服务进行组合,从而使业务用户能够快速构建满足其业务需求的个性化应用。

使用 VINCA 语言构建个性化应用分为两个阶段:在服务组合阶段,指定业务服务的功能需求、时间约束以及执行权限。在执行阶段,由相关具有权限的人员触发业务服务的执行,通过绑定与业务服务相匹配的具体服务并调用具体服务来完成业务服务的执行。

随着组合的服务数量的增多以及大量的具体服务的存在,业务用户很难直观地判断组合的服务的可执行能力,如果不正确的服务组合已经执行,将会产生很大的影响,如需要重新对服务组合进行修改并需要对已执行的服务所造成的影响进行恢复等。因此需要有一种方法来帮助业务用户判断组合的服务的可执行能力,并根据组合所达到的可执行能力级别,引导用户进行相应的处理以使服务组合可执行。影响业务级服务组合可执行能力的因素是多方面的,如组合的服务的逻辑结构、从业务服务到具体服务的匹配方法以及匹配的具体服务的可用性等。本文针对组合的服务的逻辑结构以及从业务

服务到具体的匹配方法对业务级服务组合可执行能力的影响,提出了一种使用着色时间 Petri 网(CTPN)^[2,3]验证业务级服务组合可执行能力的方法,从不同级别验证业务级服务组合的可执行能力。

本文第 2 节给出了业务级服务组合可执行能力的定义;第 3 节讨论了如何用 CTPN 描述业务级服务组合;第 4 节讨论了使用 CTPN 对业务级服务组合的可执行能力进行验证的方法;第 5 节以制造业网格应用平台 AmGrid 为案例,讨论了方法的应用效果;最后进行了总结。

2 业务级服务组合的可执行能力

在业务级别组合服务,组合的服务的逻辑结构、业务服务到具体服务的匹配方法以及具体服务本身的可用性都会影响业务级服务组合的可执行能力。逻辑结构不正确,如存在死锁等将使得服务组合根本无法执行。业务服务在执行时,根据匹配函数匹配到具体服务,并调用具体服务来完成业务服务的执行。如果业务服务在匹配具体服务时,不存在满足条件的具体服务,那么业务服务也是无法执行的。造成匹配不到具体服务的原因有以下三点:①已注册的服务资源不完备。如企业需要购买笔记本,但不存在企业提供供应笔记本的服务。②匹配函数不正确。③业务服务本身的限定存在不合理的地方。如业务服务的功能为购买汽车,但限定使用 Intel 提供的服务。本文只考虑注册的服务资源的完备性所产生的影响。然而,即使业务服务能匹配到具体服务,还是不能保证业务级服务组合可执行,因为如果匹配到的具体服务不可用,业

^{*}) 本文得到了国家“八六三”高科技研究发展计划(2003AA414330)、国家自然科学基金(90412005)、中国科学院百人计划(20024040)的支持。

单保华 博士;韩燕波 博导;李厚福 博士;胡海涛 博士。

务级服务组合也是无法执行的。

基于上面的分析,我们定义业务级服务组合可执行能力如下:

定义 1 业务级服务组合可执行能力分为三个级别:

- 结构良好:满足此级别的组合表示结构上不存在逻辑错误。
- 可接受:业务级服务组合结构良好,并且对于组合的每个业务服务都可匹配到一个不为空的具体服务集合。
- 可执行:组合满足可接受约束,并且对于每一个业务服务匹配到的具体服务集合,至少存在一个可执行的具体服务。

3 业务级服务组合的 CTPN 表示

为了使用 CTPN 描述业务级的服务组合,我们定义如下映射规则:

- 对于每一个业务服务 BS, 我们使用两个变迁 t_i 和 t_f 来分别表示业务服务执行开始和业务服务执行结束。
- 添加两个特殊的库所 i 和 e , 其中 $i = \Phi, e = \Phi$ 。 i 的输出变迁是所有的初始变迁(输入库所为空的变迁), e 的输入变迁是所有的终止变迁(输出库所为空的变迁)。同时在初始状态向库所 i 中放入足够所有初始变迁发生所需要的令牌, 令牌的时间戳为 0。
- 业务服务执行过程中需要的资源可分为两部分:动态绑定资源和非动态绑定资源。动态绑定资源在运行时才会确定具体使用哪一个资源,而非动态绑定资源是在组合时已经明确的资源。我们分别用两种不同的符号来表示这两种资源,使用圆圈表示非动态绑定资源,引入新的符号方框表示动态绑定资源。
- 为了反映业务服务的动态绑定特性,CTPN 中从变迁到方框库所以及从方框库所到变迁的边的定义分别为一个匹配函数和一个释放函数。

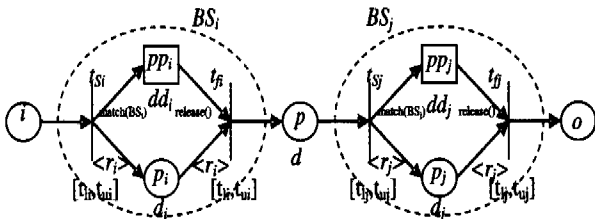


图 1 两个业务服务组合的 CTPN 表示

图 1 描述了两个业务级服务组合的 CTPN 表示,图中存在两个业务服务 BS_i 和 BS_j , 根据上面的映射规则,将业务服务 BS_i 表示成为两个变迁(t_i, t_f)和两个库所(p_i, p_j), BS_j 也相应地表示为两个变迁(t_j, t_f)和两个库所(p_j, p_j)。 t_i 和 t_f 是与同一个业务服务相关联的,因此具有相同的时间间隔约束。 $match(BS_i)$ 是一个服务匹配函数,执行后将根据从注册中心匹配到的具体服务数量决定放入到库所 p_i 中的令牌数量,放入的令牌数量等于匹配到的具体服务的数量。变迁 t_f 发生时,从 p_i 中消耗一个令牌,同时释放 p_j 中的所有其他令牌,如果 p_i 中的令牌数量为 0,则变迁 t_f 无法发生。业务服务具有自身的特性,即只要存在一个与业务服务匹配的具体服务,就可以使变迁 t_f 被允许,因此,我们重新定义变迁的发生规则。

定义 2 给定变迁 t_{ij} 和 t_{fj} , 以及相关联的时间间隔 $IN(t_{ij}) = IN(t_{fj}) = [\tau_l(t), \tau_u(t)]$, $\bullet t_{ij} = \{p_i\}$, $t_{ij} \bullet = \{pp_i, p_j\}$, $\bullet t_{fj} = \{pp_j, p_j\}$, $t_{fj} \bullet = \{p_k\}$, $D(p_i) = d_i$, $D(pp_i) = dd_i$, D

$(p_i) = d_i$, $D(p_k) = d_k$ 。

(a) 如果 $E_{f(p_i, t_{ij})} \subseteq m(p_i)$ 并且 $\max\{x + D(p_i)\}$ 在 $IN(t_{ij})$ 中, 其中 $x = \max\{x_i | x_i \text{ 是所有 } m(p_i) \text{ 的时间戳}\}$, 则 t_{ij} 是被允许的。被允许的变迁能够发生。

(b) 假定变迁 t_{ij} 在时间 t 发生, 产生一个新的标记状态 M' , 则 $m'(p_i) = m(p_i) - E_{f(p_i, t_{ij})}$, $m'(p_j) = m(p_j) + E_{f(t_{ij}, p_j)}$, $m'(pp_j) = m(pp_j) + E_{match(BS_j)} match(BS_j)$ ($match(BS_j)$ 是一个函数, 返回与业务服务 BS_j 匹配的具体服务集合, E 将每一个具体服务映射为一个令牌), 并且 $m'(p_i)$ 和 $m'(pp_j)$ 中的任何一个令牌时间戳为 t 。

(c) 如果 $E_{f(p_j, t_{fj})} \subseteq m(p_j)$ 并且 $\max\{x + D(p_j)\}$ 在 $IN(t_{fj})$ 中, 其中 $x = \max\{x_i | x_i \text{ 是所有 } m(p_j) \text{ 的时间戳}\}$, 同时 $|m(pp_j)|^2 \geq 1(|m(pp_j)| \text{ 表示集合中元素的数量})$, 则 t_{fj} 是被允许的。被允许的变迁能够发生。

(d) 假定变迁 t_{fj} 在时间 t 发生, 产生一个新的标记状态 M' , 则 $m'(p_k) = m(p_k) + E_{f(t_{fj}, p_k)}$, $m'(p_i) = m(p_i) - E_{f(p_i, t_{fj})}$, $m'(pp_j) = m(pp_j) - E_{release}$ ($release$ 函数在变迁消耗一个令牌后, 释放所有其他令牌)。

4 业务级服务组合可执行能力验证

在使用 CTPN 描述业务级服务组合的基础上, 可以应用 Petri 网的一些特性来分析组合的可执行能力。为此, 我们定义 $\overline{CTPN}$ 。 $\overline{CTPN}$ 是在 CTPN 基础上增加一个变迁, 连接初始库所和终结库所, 如图 2 所示。

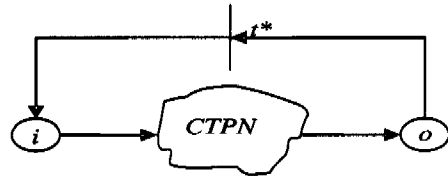


图 2 $\overline{CTPN}$ 图示

定义 3 $\overline{CTPN}$ 是一个颜色时间 Petri 网, $\overline{CTPN} = \langle \overline{PN}, \Sigma, C, E, D, IN \rangle$,

(a) $\overline{PN} = \langle P, T, F, H, M \rangle$

(b) $\overline{T} = T \cup \{t^*\}$

(c) $\overline{F} = F \cup \{f(0, t^*), f(t^*, i)\}$

(d) $E_{f(o, t^*)}$ 是库所 o 中所有令牌颜色的和(和的意义如下, 如果 $C(o) = \{\langle r \rangle, \langle b \rangle\}$, 则 $E_{f(o, t^*)} = \langle r + \langle b \rangle \rangle$), $E_{f(t^*, i)}$ 是初始状态时库所 i 中所有令牌颜色的和。

(e) $IN(t^*) = [*, *]$

(f) 变迁 t^* 发生后, 产生的令牌时间为 0。

上述定义中的条件(c)保证变迁 t^* 会消耗库所 o 中的所有令牌, 同时会向库所 i 中放入足够的令牌, 保证 i 之后变迁的发生。条件(e)保证不会因为时间而影响 CTPN 中变迁 t^* 的发生。

规则 1 业务级服务组合是结构良好的当且仅当组合后的 CTPN 表示是强连通的和易处理的(Well-handled)。

Aalst 在文[4]中给出了易处理(Well-handled)的定义, 该性质可以保证组合的服务不存在结构上的错误, 如以并行节点(and-split)开始, 而以选择并(or-join)结束。同时, 强连通性也保证服务组合的完整性。结构良好的服务组合是可执行的基础, 而能否执行还取决于业务服务能否匹配到具体服务以及具体服务的可执行性。

规则2 业务服务组合是可接受的当且仅当满足结构良好定义,并且组合CTPN的表示满足活性。

根据活性的定义,一个 Petri 网是活的当且仅当对于每个任意可达的状态 M' 和变迁 t ,都存在一个状态 M'' ,变迁 t 发生导致从状态 M' 到状态 M'' 。也就说从初始状态 i 一定存在一个变迁序列能到达终止状态 o 。即对于每一个业务服务来说,根据映射规则,都可以匹配到一个至少包含一个具体服务的集合。

对于业务级服务组合的可执行级别,很难使用 Petri 网进行验证,因为业务服务在执行时是一个动态绑定的过程,即使存在一个可使用的具体服务,在绑定时不一定就选择这个可用的服务。目前在服务可用性方面已经有了一定的研究^[5],将预处理与服务可用性相结合,执行之前先判断所调用服务的可用性,可以在一定程度上解决服务组合的可执行验证问题。然而,一个可执行级别的服务组合一定是满足可接受约束的。

5 在 AmGrid 中的应用示例

AmGrid^[7]是针对网络化制造的资源的灵活协同的需求,提出的一个旨在提供动态、跨组织边界的资源整合的支撑环境,解决虚拟组织这类由个体成员、组织结构和资源组成的动态联合体所面临的灵活的资源共享和基于各类资源的协同工作问题的网格应用平台。图3是 AmGrid 中的业务级服务组合工具及虚拟组织部分。

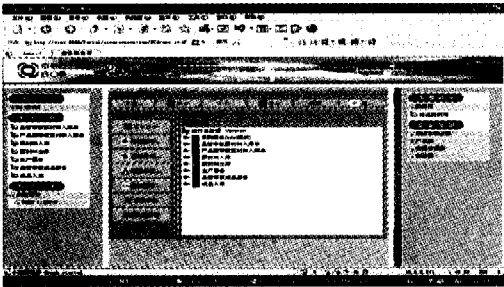


图3 AmGrid 中的业务级服务组合工具

为了展示本文中的方法的应用,我们假定市场上某种型号笔记本电脑热销,企业 A 为了能够抓住市场机会,决定生产笔记本,但由于自身条件的限制,决定先从联盟成员购买硬盘和 CPU,组装后投向市场。企业的业务人员根据自身的条件及对联盟成员的要求,使用业务级服务组合语言 VINCA^[6]组合的服务如图4所示。

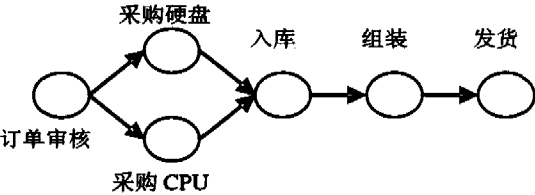


图4 业务级服务组合示例

在组合的服务中,采购硬盘和采购 CPU 需要调用硬盘提供商和 CPU 提供商的服务,发货需要调用供应商提供的服务。具体的供货商和物流供应商在流程运行中根据候选企业的情况再进行选择。各个任务的输入、输出以及执行人员如表1所示。

表1 服务输入、输出执行时间以及参与人员

编号	任务	输入	输出	执行时间	参与者
1	订单审核	订单	采购单	[0,2]	a
2	采购硬盘	采购单	回执	[3,6]	b
3	采购 CPU	采购单	回执	[3,6]	c
4	入库	回执	出库单	[8,11]	d
5	组装	出库单	成品单	[12,14]	e
6	发货	成品单	发货单	[15,16]	f

根据第4节定义的映射,我们得到组合后服务的 CTPN 表示如图5所示。

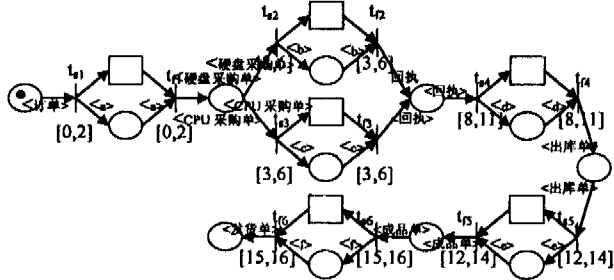


图5 组合服务的 CTPN 表示

我们再假定注册了 B、C、D、E 四个企业想参与到动态联盟中,B 和 C 企业注册服务提供硬盘,而 D 和 E 两个企业注册服务提供 CPU,如表2所示。

表2 注册的企业、服务以及提供的产品

注册企业	服务	产品
B	S_b	硬盘
C	S_c	硬盘
D	S_d	CPU
E	S_e	CPU

根据上一节的讨论,该组合后的服务是满足结构良好的,但不可接受,因为组合服务的 CTPN 表示不满足规则2。任务发货需要调用物流供应商的服务,然而,在已经注册的企业中,不存在企业能够提供物流服务,因此,在变迁 t_{f6} 发生后,方框形库所中的令牌数量为0,导致变迁 t_{f6} 是不被允许的,无法继续执行。如果这时新注册了一个企业 F,提供物流服务 S_f ,则可以满足可接受约束。

总结 关于服务组合的研究,已经取得了一定的成果。然而大部分工作都关注服务之间是否可组合^[8],在一个横向的层面讨论服务组合问题。本文从纵向方面讨论了业务级服务组合的可执行能力问题,对业务级服务组合的可执行能力给出了定义,并讨论了如何使用着色时间 Petri 网对业务级服务组合的不同级别的可执行能力进行验证。本文主要做了以下三方面的工作:①给出了业务级服务组合的可执行能力定义;②以两个变迁反映业务服务的执行过程,将动态绑定资源和非动态绑定资源分离,在着色时间 Petri 网中引入一种新的库所来描述动态绑定资源,并对 CTPN 中的边及变迁的发生规则重新进行了定义;③提出了一种使用 CTPN 对业务级服务组合的可执行能力进行验证的方法。文中提出的方法可以在两个方面协助业务用户进行业务级服务组合:

(1)判断流程结构的合理性,不满足规则1的业务级服务组合说明用户在组合服务时使用了不正确的组合逻辑,此时必须修改组合逻辑才能使组合的服务可执行。

(下转第128页)

AND PayRate.Rank=WorksOn.ProjRank.

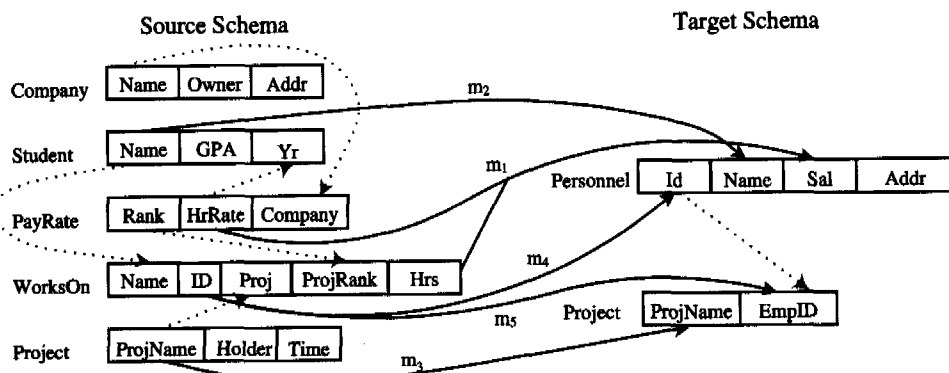


图3 数据源模式与目标模式示例

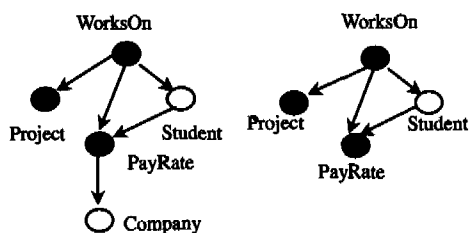


图4 例3.3中的有向图G

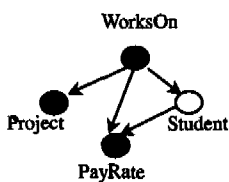


图5 例3.3中的导出子图H

结束语 模式集成和模式映射是当前异构数据源集成采用的两种主要方法。本文提出的基于模式映射的查询计划生成算法能够根据不同的查询条件和不同的数据源模式,自动构造查询计划,并保证结果数据满足目标模式结构与引用完整性要求,从而完成异构模式下的数据转换与集成。

参考文献

- 1 Popa L, Velegrakis Y, Miller R J, et al. Translating Web Data. In: Proc. VLDB, 2002. 598~609
- 2 Hernández M A, Miller R J, Haas L M. Clio: A Semi-Automatic Tool For Schema Mapping. SIGMOD, 2001

- 3 Fagin R, Kolaitis P G, Miller R J, et al. Data Exchange: Semantics and Query Answering. In: ICDT, 2003. 207~224
- 4 Miller R J, Haas L M, Hernández M. Schema Mapping as Query Discovery. In: Proc. of the Int'l Conf on Very Large Data Bases (VLDB), Cairo, Egypt, 2000. 77~88
- 5 Miller R J, Hernández M A, Haas L M, et al. The Clio Project: Managing Heterogeneity. SIGMOD Record, 2001, 30(1)
- 6 Domenigand R, Ditttrich K R. An Overview and Classification of Mediated Query Systems. SIGMOD Record, 1999, 28(3)
- 7 Li You, Liu Dongbo, Zhang Weiming. A Data Transformation Method Based On Schema Mapping. In: 3rd International Conference on Information Systems Technology and its Application, IS-TA. Salt Lake City, USA 2004
- 8 李瑞轩, 卢正鼎, 等. 多数据库系统中基于 XIDM 的模式映射方法研究. 计算机研究与发展, 2004, 41(3)
- 9 陈彤兵, 胡金化, 等. 分布式自治数据源的联合查询. 计算机研究与发展, 2004, 41(4)
- 10 Silberschats A, Korth H F. 数据库系统概念. 杨冬青, 等译. 北京机械工业出版社

(上接第 116 页)

(2) 判断注册的资源的完备性: 满足规则 2 的组合说明注册的具体服务资源是完备的, 即组合的业务服务都可以匹配到至少一个具体服务。如果不满足规则 2, 那么必须注册新的具体服务资源或修改服务组合以使得组合后的服务可执行。

本文从业务级服务组合的逻辑结构和注册的具体服务资源的完备性角度探讨了业务级服务组合的可执行能力, 但影响业务级服务组合可执行能力的因素是多方面的。组合的服务之间不但存在控制流, 也存在数据流, 只有二者都正确流转才能真正保证组合的正常执行。本文对在组合阶段如何验证业务级服务组合的可执行能力进行了初步探讨, 下一步的工作就在于研究组合中的数据流对业务级服务组合可执行能力的影响。

参考文献

- 1 Han Y, et al. CAFISE: An Approach Enabling On-Demand Configuration of Service Grid Applications. Journal of Computer Science and Technology, 2003, 18(4)

- 2 van der Aalst, W M P. Interval timed coloured petri nets and their analysis. In Application and Theory of Petri Nets 1993, Proc. 14th International of Conference, Chicago, (USA), 1993, 691, 453~472
- 3 Morasca S, Pezzè M, Trubian M. Timed high-level nets. Journal of Real-Time Systems, 1991, 3, 165~189
- 4 van der Aalst W M P. The Application of Petri Nets to Workflow Management. The Journal of Circuits, Systems and Computers, 1998, 8(1), 21~66
- 5 李东来, 韩燕波, 王建武, 喻坚. 面向服务应用中服务可用性及其引发的异常处理研究. 计算机研究与发展, 2004, 41(12)
- 6 Han Yanbo, Geng Hui, Li Houfu, et al. VINCA - A Visual and Personalized Business level Composition Language for Chaining Web-based Services. International Conference on Service Oriented Computing, Italy, 2003. 165~177
- 7 梁英, 虎嵩林, 李厚福, 等. 面向网络化制造的网格应用平台及其核心技术研究. 计算机研究与发展, 2004, 41(12)
- 8 Fang Jun, Hu Songlin, Han Yanbo. A Service Interoperability Assessment Model for Service Composition. The Conference of SCC, 2004