ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.20, Supplement, December 2009, pp.138−143

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

面向端用户的服务组合广义决策逻辑模型*

罗元盛[1], 齐 勇[1+], 侯 迪[1], 史 椸[1], 陈 滢[2], 沈林峰[1]

[1](西安交通大学 电子与信息工程学院 计算机科学系,陕西 西安 710049)

[2](IBM中国研究院,北京 100094)

Generalized Decision Logic Based and End-User Oriented Service Composition Formal Model

LUO Yuan-Sheng[1], QI Yong[1+], HOU Di[1], SHI Yi[1], CHEN Ying[2], SHEN Lin-Feng[1]

[1](Department of Computer Science, School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

[2](IBM China Research Laboratory, Beijing 100094, China)

+ Corresponding author: qiy@mail.xjtu.edu.cn, http://www.xjtu.edu.cn

Luo YS, Qi Y, Hou D, Shi Y, Chen Y, Shen LF. Generalized decision logic based and end-user oriented service composition formal model. *Journal of Software*, 2009,20(Suppl.):138−143. http://www.jos.org.cn/1000-9825/09017.htm

Abstract: There's seldom work carried on dealing with the contradiction between the vague and uncertain requirements of end-user and the precise and deterministic process in service composition. A multi-grain formal model for service composition is proposed in this paper. This model considers the requirement of customers in service composition from the end-user view and a formal specification on mapping the Web service description to the generalized decision logic language (GDL) is presented to construct multi-grain service composition views. GDL is a formal logic language proposed in granular computing research community as an expecting specification for definition of granular models. It can be used to define a multi-grain model for service composition and make users and service composition agent work in different information granule level separately. The proposed model is expected to provide a more understandable view for an end-user than traditional service composition model and conforms to the human cognition mode.

Key words: SOA; SOC; service composition; granular computing; generalized decision logic

摘 要: 服务组合中端用户的模糊和不确定的需求与服务组合过程的精确性要求之间存在着矛盾,而几乎没有相关的工作考虑了这个问题.提出一种多粒度的服务组合形式化模型.这一模型从端用户的角度来考虑服务组合中用户的需求.为了构建多粒度的服务组合视图,将 Web 服务描述映射成广义决策逻辑语言(GDL).GDL 是由粒度计算的研究组织提出的一种形式化语言,试图用来作为定义粒度模型的规范.利用 GDL 构建的服务组合多粒度模型,可以使用户和服务组合代理在不同的信息粒层进行组合工作.与传统的模型相比,这一模型可以为

---

* Supported by the National Natural Science Foundation of China under Grant No.60933003 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2006AA01Z101 (国家高技术研究发展计划(863)); the IBM Joint Project of China under Grant No.JLP200906008-1 (IBM共享大学研究项目); the Sci. & Tech. Program of Shaanxi Province of China under Grant No.2008KW-02 (陕西省科学技术研究发展计划); the Sci. & Tech. Key Projects of Shaanxi Province of China under Grant No.2006K04-G23 (陕西省科技攻关项目)

Received 2009-03-05; Accepted 2009-04-03

端用户提供更容易理解的服务组合视图,符合人类感知模式.
关键词: SOA;SOC;服务组合;粒度计算;广义决策逻辑

1 Introduction

Many works carried on semantic service discovery and matching are expected to bring automation to service discovery, selection and invocation[1−6]. The automation can be achieved in web service classification/clustering or service matching. Also, the automation can be achieved in QoS-based selection for web service composition[7]. Although much work has combined the semantic information with syntactic information of service to classify/cluster the services, few studies have distinguished the effects of different information granularities in service composition, for example, the fuzzy and rough requirements from the point of view of end-users on the one hand, while the precise and rigid restriction on business logic and operation on the other hand. An end-user is a user who is unnecessarily to have information technology expertise but just want to provide some value-added services to their potential customers of their websites, their company, or school or barely to themselves.

The purpose of this paper is to put forward an integrated model for semi-automation schema of Web service composition. The information processing methodology of granular computing is used to build the formal service composition model in this paper. More specifically, the generalized decision logic (GDL) language[8] is proposed to define the model and some rules are used for optimization of the model.

2 Service Composition Model

2.1 System model

At first, the definition on information tables with added semantics[8] is introduced to prepare for the formal definition of services and granules. An information table is expressed as:

$$S=(U,At,\{V_a\mid a\in At\},\{f_a\mid a\in At\}) \quad (1)$$

Where:

U is a finite non-empty set of objects,

At is a finite non-empty set of attributes,

V_a is a non-empty set of values for $a\in At$,

f_a: $U\rightarrow V_a$ is an information function.

Each information function f_a is a total function that maps an object of U to exactly one value in V_a. In the context of this paper, the objects in U are the Web services in considering service repository. The services are only perceived, observed or measured by using a finite number of attributes.

An information table with added semantics can be described by a pair:

$$S^{+}=(S,\{L_a\mid a\in At\}) \quad (2)$$

where S is a standard information table, L_a is a set of labels used to name granules of V_a. Each member of L_a can be a subset of V_a.

A description of service repository can be easily mapped to a semantic information table by following definition:

U is a finite non-empty set of services,

At is a finite non-empty set of attributes of services,

V_a is a non-empty set of values for $a\in At$, for the functionality attributes, the values will be some concepts in domain ontologies correlated to the service depository. For the non-functionality attributes, the values will be some real value figures.

f_a: $U \rightarrow V_a$ is an information function.

L_a is a set of labels. Each label in L_a is a set of semantic concepts to describe the granules for comprehension of users.

2.2 Generalized decision logic language

The basic alphabet of GDL for an information table consists of the following three types of symbols[8]:

At: A finite set of attribute symbols,

R_a: A finite set of relation symbols for each attribute $a \in At$, and

L_a: A set of label symbols for each attribute $a \in At$.

Formulas of GDL are defined by the following two rules:

1. An atomic formula of GDL is a descriptor (a,r,l), where $a \in At$, $r \in R_a$, $l \in L_a$,

2. The well-formed formulas (wff) of GDL is the smallest set containing the atomic formulas and closed under $\neg, \cap, \cup, \rightarrow, \equiv$.

Given a GDL and an information table S^+, the satisfiability of a formula ϕ by an object x, written $x \models_{S+} \phi$ or in short $x \models \phi$ if S^+ is understood, is defined as follows[8]:

(1) $x \models (a,r,l)$ iff $f_a(x)\, r\, l$,

(2) $x \models \neg\phi$ iff not $x \models \phi$,

(3) $x \models \phi \wedge \psi$ iff $x \models \phi$ and $x \models \phi$,

(4) $x \models \phi \vee \psi$ iff $x \models \phi$ or $x \models \phi$,

(5) $x \models \phi \rightarrow \psi$ iff $x \models \neg\phi \vee \psi$,

(6) $x \models \phi \equiv \psi$ iff $x \models \phi \rightarrow \psi$ and $x \models \psi \rightarrow \phi$

If Φ is a formula, the set $m_{S+}(\Phi)$ defined by:

$$m_{S^+}(\phi) = \{x \in U \mid x \models \phi\},$$

is called the meaning of the formula Φ in S. If S^+ is understood, we simply write $m(\Phi)$.

2.3 Granulation for end-user oriented service composition

A semi-automated service composition model should have two properties for end-users to participate the design procedure easily:

1. The model should provide multi-grain view on the universe of candidate services. It is unreasonable for end-users to go through the details of each service descriptions. Instead, end-users usually used to consider the problem in a coarse-grain knowledge level and think on abstract, rough and uncertain concepts. On the other hand, an automated planner can worked in a fine-grain level with rigid and precise knowledge description, but it is difficult for the search algorithms to deduce the solution in a coarse-grain level with fuzzy and imprecise knowledge.

2. Appropriate and comprehensive semantic labels should be assigned to each granule for understanding of users.

Constructing a granule, that is granulation, is actually a clustering procedure. A formal description of a granule is defined by:

$$G = (\phi, m(\phi)) \tag{3}$$

Where Φ is a *wff*. More specifically, Φ is a description of $m(\Phi)$ in S^+, the intension of granule G, and $m(\Phi)$ is the set of objects satisfying Φ, the extension of granule G.

As aforementioned in Section 2.2, a *wff* is made up of At, R_a, and L_a. To form a granule, many attributes of services can be used in the granulation, such as the service name, service input/output parameters and service business community category, etc. The metric relation used to granulation can be similarity calculation on some attributes. The semantics labeling on the service granules is a trivial work often omitted in many cases. But it is an important and necessary aspect for end-user oriented service composition. We have following formal description on semantic labeling:

1. Let C_a be the set of all concepts in a concept hierarchy with respect to an attribute a, a label L_a for granules categorized on a could be $L_a=C_a$.

2. If the granules is categorized on a set of attributes $a_1, a_2,\ldots,a_m$, then the label could be:

$$L=\bigcup_{i=1}^{m} C_{a_i} \tag{4}$$

The above description is intuitive and simple. In many cases, a reduction on this semantic label is necessary to avoid the information redundance. We firstly introduce the definition of neighborhood[9]:

$$\begin{aligned}&\forall x\in U, A=(a_1,a_2,\ldots,a_m),\ a_i\in At,\ i=1\ldots m,\ \phi=(A,r_A,l_A),\\&N_A(x)=\{y\in U\mid (x\models\phi)\wedge(y\models\phi)\}\end{aligned} \tag{5}$$

The reduction rules of the semantic label are as follows:

1. Let A is a set of attributes, G_A is the granules categorized on A. If for an $a_i\in A$, for any $x\in U$,

$N_A(x)\equiv N_{A-ai}(x)$ holds,

then the C_{ai} can be removed from L_A. Where A-a_i means the set of attributes in A except the a_i. L_A is the label of G_A.

2. If a set of concepts can be represented by a super concept, then this set of concepts can be substituted by the super concept in the label.

The rule 1) is to remove the redundant concepts from a label of a granule, which don't contribute to the semantic description of a granule anymore. The rule 2) is often called as concept aggregation in information retrieval community[1].

3 Related Works and Comparison

The clustering and search techniques in data mining and information retrieval (IR) field have been widely used in SOC field. Majority of related works are devoted to service discovery and matching. Xin, etc.[1] presented a similarity search technique for WSDL-oriented Web services. Corella, etc.[3−5] presented a heuristic approach to Web services classification and a prototype for service discovery. Klusc, etc.[2] presented their works on hybrid semantic Web service matching with their matchmaker OWL-MX. Kokash, etc.[6] also introduced their hybrid matching approach for Web service discovering and evaluate their approach using WSDL corpus. Although these approaches can be used in similarity clustering for service composition, none of them is totally suitable for service composition for the difference between service composition and service discovery. A service composition procedure is more complicated than service discovery. Not only is an approximated collection of services needed be selected for each task in service composition workflow, but also are the precise business logic and composability between services needed be considered. We adopt a semi-automated service composition selection model which is end-user oriented. An end-user can choose the service composition plan at a coarse-grain level while service composers can select the business logic and composable services for this composition plan at a fine-grain level. To achieve this target, an appropriate clustering in different granular views for service composition is necessary. Although some researchers have taken the grain into consideration, their methods are intuitive and simple. E.g., in Ref.[3], the coarse-grain service discovery means to provide the services groups categorized on commercial categories to end-users. In some cases, the granule has different universe of objects with the one in our paper. E.g., in Ref.[6], the fine-grain approaches are the ones that can't be further decomposed in contrast to hybrid approaches. In our paper, the coarse or fine grains are correlated to clusters with the different number of services included in clusters with different abstract degree of concepts. We use the generalized decision logic (GDL) language for granular computing[8] with some optimization rules to construct an end-user oriented service composition rule. Different with the hybrid approaches presented in service discovery field, which usually are the linear combination or cascade of matching approaches, our strategy involves the functional and non-functional features into clustering in distinct granules view. Although Medjahed, etc.[10,11] put forward their composable rules with multilevel service composition model, they didn't consider the service composition from different granules. Actually, they used the linear combination of composition rules to evaluate the composability between services, which is a typical single dimensional granule way.

4 Conclusion and Future Work

This paper proposes a formal model for end-user oriented service composition. We consider the vague and ambiguous requirements in the point of view of end-users while the demands on facility and accuracy in the view of service composers. We construct the proposed model using formal method (GDL) deriving from granular computing. As far as we know, it's the first time to put forward the GDL in service composition field.

Our future work is to combine our model with a domain-specific case in a real project.

References:

[1] Dong X, Halevy A, Madhavan J, Nemes E, Zhang J. Similarity search for Web services. In: Proc. of the 30th Very Large Data Bases (VLDB) Conf. Toronto: Morgan Kaufmann Publishers, 2004. 372−383.

[2] Klusch M, Fries B, Sycara K. Automated semantic Web service discovery with OWLS-MX. In: Proc. of the Int'l Conf. on Autonomous Agents and Multiagent Systems (AAMAS). Hokkaido: ACM, 2006.

[3] Lara R, Corella MA, Castells P. A flexible model for Web service discovery. In: Proc. of the 1st Int'l Workshop on Semantic Matchmaking and Resource Retrieval: Issues and Perspectives (SMR 2006) at the 32nd Int'l Conf. on Very Large Data Bases (VLDB 2006). Seoul: Morgan Kaufmann Publishers, 2006.

[4] Corella MA, Castells P. Taxonomy-Based Web service categorization using conceptual parameter descriptions. In: Proc. of the 1st Int'l Workshop on Semantic Matchmaking and Resource Retrieval: Issues and Perspectives (SMR 2006) at the 32nd Int'l Conf. on Very Large Data Bases (VLDB 2006). Seoul: Morgan Kaufmann Publishers, 2006.

[5] Corella MA, Castells P. A heuristic approach to semantic Web services classification. In: Proc. of the 10th Int'l Conf. on Knowledge-Based & Intelligent Information & Engineering Systems (KES 2006). LNCS, Bournemouth: Springer-Verlag, 2006. 598−605.

[6] Kokash N, Heuvel W, Andrea VD. Leveraging Web services discovery with customizable hybrid matching. In: Dan A, Lamersdorf W, eds. Proc. of the Int'l Conf. on Service Oriented Computing (ICSOC 2006). LNCS, Berlin, Heidelberg: Springer-Verlag, 2006. 522−528.

[7] Zeng L, Benatallah B, Ngu A. QoS-Aware middleware for Web service composition. IEEE Trans. on Software Engineering, 2004, 5(30):311−327.

[8] Yao YY, Liau CJ. A generalized decision logic language for granular computing. In: Proc. of the 11th IEEE Int'l Conf. on Fuzzy Systems. Honolulu: IEEE, 2002. 773−778.

[9] Lin TY, Chiang IJ. Formal concept analysis and document clustering via granular computing. In: Proc. of the IEEE Int'l Conf. on Systems, Man, and Cybernetics. Taipei: IEEE, 2006. 4763−4767.

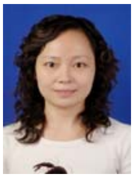
[10] Medjahed B, Bouguettaya A, Elmagarmid AK. Composing Web services on the semantic Web. The VLDB Journal, 2003, 12(4):333−351.

[11] Medjahed B, Bouguettaya A. A multilevel composability model for semantic Web services. IEEE Trans. on Knowledge and Data Engineering, 2005,17(7):954−968.

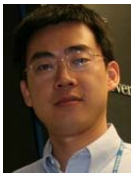
LUO Yuan-Sheng was born in 1980. He is a Ph.D. candidater of Xi'an Jiaotong University. His current research areas are service oriented computing, service composition, etc.

QI Yong was born in 1957. He is a professor of Xi'an Jiaotong University. His current research areas are computer network and distributed system, middleware technology, etc.

HOU Di is an associate professor of Xi'an Jiaotong University. His current research areas are database, formal method, etc.



史椸(1975－),女,博士,讲师,主要研究领域为操作系统,分布式系统,系统仿真.



陈滢(1973－),男,博士,主要研究领域为SOA,分布式计算,云计算.

沈林峰(1979－),男,博士生,主要研究领域为面向服务的计算,服务组合.