# Process Oriented Dependency Modelling For Service Identification

Wenge Rong[1,2], Ting Li[3], Yuanxin Ouyang[1,2], Chao Li[1,2], and Zhang Xiong[1,2]

[1] School of Computer Science and Engineering, Beihang University, Beijing, China
[2] Research Institute of Beihang University in Shenzhen, Shenzhen, China
[3] Sino-French Engineering School, Beihang University, Beijing, China
{w.rong@, liting@ecpk., oyyx@, licc@, xiongz@}buaa.edu.cn

**Abstract.** Service-Oriented Architecture (SOA) is important for organisations to achieve dynamic business process and build business agility. One of the first step for service oriented applications implementation is to properly identify a set of fine-grained services. A right service granularity is necessary to satisfy lower coupling and higher cohesion principles for reusable software services. To meet this challenge, a lot of efforts have been attached to support service identification. In this paper, by considering the dependency combined with the idea of graph partition, a service identification method is proposed from the business process's perspective. The illustration example has shown its promising and it is expected that the proposed service identification method can offer researchers further insight into service granularity analysis.

**Keywords:** Service Identification; Semiotics; Semantic Analysis; Granularity

## 1 Introduction

In recent years, the evolving businesses environment has put unpredictable pressure than ever on the way the business is conducted. The success of a business heavily relies on its ability of implementing dynamic business processes in terms of the business model and operations to adjust the changing market, which is one of the major options for effect creation. During the process of building business agility, the concept of Service Oriented Architecture (SOA) has been proposed and widely lauded as an innovative business oriented solution [2].

The implementation of SOA based systems lines in the provision of a set of loosely coupled services. A service can be generally considered as a piece of applications that encapsulate and implement certain business logic for invocation by internal or external partners through well defined interfaces. In reality it is normally scalable in terms of its scope. It can be as simple as a single file compression action or as complicated as a whole software package abstracted for a specific business solution. Though its principle is primary and simple, its analysis, design, realization and implementation are of much importance and difficulty [2, 9]. To solve these problems, a preliminary challenge called service identification must be solved and has attracted much attention from scholars in this domain.

Service identification is an essential process for successful SOA based system implementation since errors made during the procedure of identification will be carried on in next procedures and leads to a chain reaction [8] and thereby influencing the effectiveness of the SOA architecture [3]. In fact, service identification has impact on the goals like the composability of loosely-coupled services and the reusability of individual services in different contexts [11].

The granularity in service identification process generally varies over time and can be classified in term of data, functions, and business value [7]. One of the widely used methods relies on business process decomposition by analysing related tasks and activities [6]. Though process oriented service identification has shown its potential, its potential has not been fully realized [4] and how to analyse in-depth the business process and fulfil the requirements that they represent is still difficult [1].

To deal with demands of examining and articulating possible services with operations of service candidate, in this paper a semantic analysis method inspired by semiotics theory [13] has been proposed to clarifying the meaning and dependency within a business process, thereby supporting the possible service candidate selection.

The remainder of the paper is organized as follows. Section 2 describes the most related works in the domains of service identification. The proposed model of the service identification will be elaborated in section 3. Section 4 will describe the details of proposed approach with an illustration example. Section 5 concludes the paper with some suggestions and recommendations for the future work.

## 2 Related Work

Since SOA is involved in the whole business process, different stakeholders will probably have their own viewpoints on its scope and capacity from different perspectives. As a result, the service identification methods are also diversified. Some of them consider services from lower level implementation, such as data, features and etc. [10], while others consider service granularity from higher level like application domain and business process [12].

In the literature, most approaches of service identification are based on business process due to its intention to realize the reusability to create business value [6]. Wang et al. investigated the service scope analysis by conducting review on inter-/intra-enterprise business processes to identify the qualities which a good service should have [15]. They consider a service as the composition of a set of legacy software components with larger granularity. Based on this, they proposed a service normal form and a normalized method to solve the problem that service are difficult to integrate closely with information systems.

Inaganti and Behara thoroughly studied the handshake between SOA and business process management [8]. Their work conducts value-chain analysis by in-depth analysing process coupled with use case study. Afterwards they identify services with the combination of top-down and bottom-up approaches. Similarly Dwivedi and Kulkarni also proposed a method for service identification by utilizing process map [4].

Another interesting work is the one by Kim et al., who tried to build a formal approach with the right granularity from the business process model [11]. With the concept of graph partitioning, they distinguish the distance of activities within a role and those belong to different roles so that they can minimise the network round-up costs incurred during the service execution.

Mani et al. proposed a novel method by focusing on the performance of users on the interfaces and use the interface design as an input to identify service [14]. They captured the appropriate references to data and process models, and analysis the requirements from data displayed in the user interface, identify business service requirements from the UI navigation flow and links between the UI and the business process model.

## 3 Methodology

In this research, the semantic analysis method proposed in organisational semiotics theory is employed [13] to get service candidates for an organization by analysing the business process. The proposed methodology is organized into four main phases: 1) map the process onto a semantic chart according to the rules of constructing ontology chart, 2) calculate the similarity between every two affordances in the ontology chart, 3) cluster the affordances based on their similarities, 4) choose an appropriate criterion to partition the ontology chart to cut the service into sub-services with right-granularities. Publishing the services with right-granularities can lower the cost and can be convenient for the consumers.

### 3.1 Semantic Chart Annotation

The first step is to map the business process onto a semantic chart, which provides a graphic representation describing the ontological dependencies between the concepts [13]. An semantic chart is comprised of a set of semantics units. According to the ontological relationships, these units are set in the ontology chart at different places. The ontology chart presents the essential elements and the atomic-level functions during the business process, that is why a ontology chart can explain the participants in a process and what they do exactly. Fig. 2 is an example of semantic chart which consists of agent, affordance, role and determiner [13].

- *Agent*: an agent is represented by rounded rectangle in the ontology chart. It means the the stakeholder(autonomous individuals) who is in charge of some operations in a certain field, and it can be an a group, an organization, an individual and etc.

- *Affordance*: an affordance is represented by rectangle in the chart. It describes action possibilities of which an actor is aware, and it can exhibit the connections between entities and some behaviour patterns in a field. [5].
- *Role*: a role is represented by ellipse in the chart. While the agent describes the class of individuals, the role defines the instance of an agent, and some of the behaviours that an agent has, which means, the role allows explicate the meaning of the concepts in the ontology by designating the agent to whom this authority is released.
- *Determiner*: an determiner is prefixed with # in the chart. It is used to describe the properties of semantic units. The determiners of each affordance are very important in our approach of service identification.

### 3.2 Service Similarity Calculation

To calculate the similarity, two hypotheses are employed, i.e., $h_1$: the interaction between two affordances provided by different places will cause a lot of long-distance communication cost. $h_2$: two affordances who have tighter ontological dependency are more relative to each other, which means, the similarity between them is higher. Based on these two hypotheses, it is then possible to empirically define the similarity between affordance $a_i$ and $a_j$ as:

$$S_{i,j} = w_1 * S^1 + w_2 * S^2$$

where $S^1$ and $S^2$ denote two indexes who measure the similarity associated to the former two hypotheses and $w_1$ and $w_2$ refer to the weight assigned to each similarity factor.

According to $h_1$, $S^1$ depends on whether these two affordances have the same provided physical place. As presented before, the first determiner of every affordance is noted as $\#p$. As such in this paper it is determined that if two affordances have the same $\#p$, their similarity is set to 1, otherwise their similarity is 0, as shown below:

$$S^1 = \begin{cases} 1 & if \;\; a_i.\#p = a_j.\#p \\ 0 & if \;\; a_i.\#p \neq a_j.\#p \end{cases}$$

According to $h_2$, the ontological dependencies in the ontology chart have a very important influence. Accordingly the similarity $S^2$ is associated to hypothesis $h_2$. The ontological dependencies are denoted by lines in the ontology chart. If two affordances are connected by one single line, they have a first-order ontological dependency; if they are connected by two lines, they have a second-order ontological dependency; and if they are connected by three or more than three lines, they have a multi-order ontological dependency. It is obvious that the smaller the order of ontological dependency is, the more similar the two affordances are. Formally, the similarity $S^2$ is assigned as below:

$$S^2 = \begin{cases} 1 & if\ a_i\ and\ a_j\ are\ the\ same \\ 0.75 & if\ a_i\ and\ a_j\ have\ a\ first-order\ ontological\ dependency \\ 0.5 & if\ a_i\ and\ a_j\ have\ a\ second-order\ ontological\ dependency \\ 0 & if\ a_i\ and\ a_j\ have\ a\ multi-order\ ontological\ dependency \end{cases}$$

### 3.3 Affordances Clustering and Service Refinement

A service is essentially a combination of affordances who satisfy some given conditions, as shown below:

$$Service = \{ \sum affordance | constraints \}$$

According to this definition, this service identification approach is substantially about how to determine the constraints who differentiate services. That is why we choose to identify the service with the idea of cluster analysis.

## 4 Case Study

In this research an illustration example is given in form of a basic library service process, which is presented by flowcharts in the Fig. 1, where some basic library functions are presented.

### 4.1 Semantic Chart Generation

With the business process chart, it is able to obtain the service semantic chart. The mapping rules are listed below and part of the semantic chart of this process is shown in Fig. 2.

1. The functions in the process chart are mapped as affordances.
2. The role (provider or consumer) relative to the function is placed in a certain place in the semantic chart based on the link between them, and each role must have an antecedent agent.

The determiners of affordances is listed in the table I and each of affordance is numbered to simplify the expression in the rest of this paper. In the column of 'deternimers' in this table, it is found that the first determiner is always the physical place where the affordance is provided. In this library service process, the physical places are online, book management department division, reading room management division, card management division and logistics. The rest determiners describe the dependencies among all these affordances.

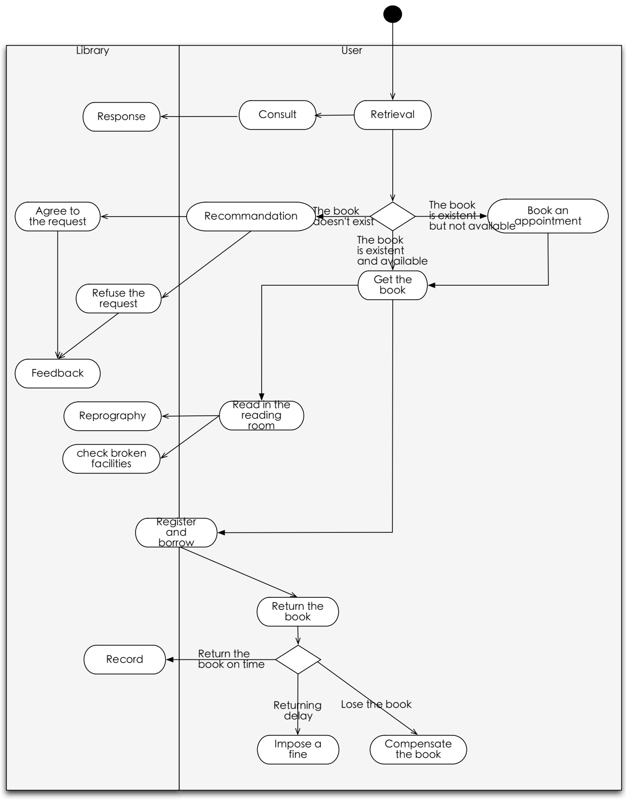


**Fig. 1.** The basic library process.

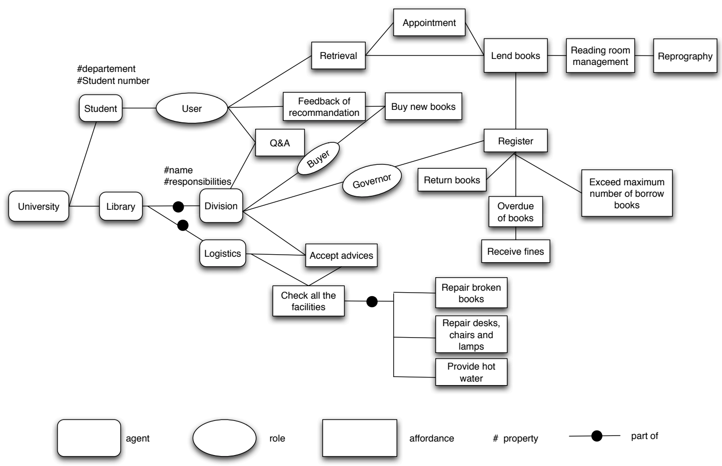


**Fig. 2.** The library service ontology chart.

**Table 1.** Determiners of all the affordances.

| Number | Affordance | Determiners |
|---|---|---|
| 1 | Q&A | #Online, #Division, #User |
| 2 | Retrieval | #Online, #user, #3, #4 |
| 3 | Appointment | #Online, #2, #4 |
| 4 | Lend books | #Book management division, #2, #5 |
| 5 | Reading room management | #Reading room management division, #4, #6 |
| 6 | Reprography | #Reading room management division, #5 |
| 7 | Register | # Book management division, #Division, #4, #8, #9, #10, #11 |
| 8 | Overdue of books | #Card management division, #7, #9 |
| 9 | Receive fines | #Card management division, #8 |
| 10 | Return books | #Book management division, #7 |
| 11 | Exceed maximum number of borrow books | #Card management division, #7 |
| 12 | Buy new books | #Book management division, #Division, #13 |
| 13 | Feedback of recommendation | #Online, #User, #12 |
| 14 | Accept advices | #Logistics, #Division, #15 |
| 15 | Check all the facilities | #Logistics, #14, #16, #17, #18 |
| 16 | Repair broken books | #Logistics, #15 |
| 17 | Repair desks, chairs and lamps | #Logistics, #15 |
| 18 | Provider hot water | #Logistics, #15 |

### 4.2 Similarity Calculation

In this case study, we use the similarity equation presented in previous section to determine the similarity between each service candidate. To simplify the case study, the two weights are set to 0.5 respectively. With these rules, it is able to derive a eighteenth-order matrix where the element in $i$-th row and $j$-th column $M_{i,j}$ represents the similarity of $a_i$ and $a_j$. Obviously, this matrix is a symmetric matrix because $S_{i,j} = S_{j,i}$. With the results of calculating similarities, the affordance can be further clustered to identify possible services.

### 4.3 Cluster the Affordance

After constructing the ontology chart, affordance need to be clustered together to generate proper service candidates and the clustering process contains 4 steps:

1) Consider the whole process as an one-to-one service, which means, every affordance is treated as a single service, and calculate the distance between every two services. The distance is inversely proportional to the similarity because the closer two services are, the more similar they are, and the higher the similarity is. As such the distance between two services can be defined as:

$$D_{i,j} = \frac{1}{S_{i,j}}$$

There is problem that the distance between a service with itself is 1, which is not correspond with the reality. In this paper we set the distance of this case to 0. Also, we admit that if $S_{i,j} = 0$, $D_{i,j}$ is defined as 10 to allow the calculate.

2) Find out the closed two services, put these two services into one service. In this way, the number of services is one less than before.

3) Re-calculate the similarities between the new services derived from step 2.

4) Repeat step 2 and step 3 until an all-to-one service is finally realized, which means all the affordances are regarded as one single service.

These results can be presented in a tree in Fig. 3, with which it is able to obtain the services with the right granularities.

### 4.4 Service Refinement

The graph of the tree which presents the results of clustering is shown in Fig. 3. In this figure, we can see clearly that this bottom-up approach transfer these one-to-one services into a all-to-one service step by step. In this process, we can choose some many-to-many results.

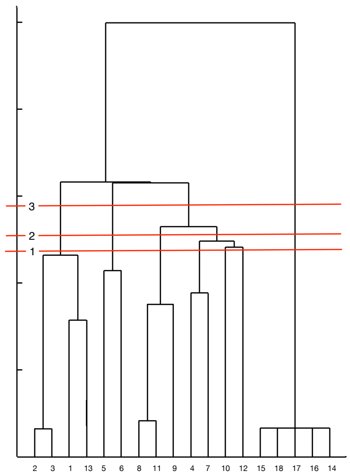


**Fig. 3.** Results of clustering.

The three red lines gives three possible service recommendations. For example, the line 2 gives a result of five services. All the affordances connected by lines below line 2 are placed together like {2, 3, 1, 13}, {5, 6}, {8, 11, 9}, etc. These three kinds of results have their own advantages and disadvantages. The first result is more flexible for consumers to use, while it needs more cost of library to manage and maintain. The third result costs less, but it is less flexible and it loses some reusability at the same time. With these provided results, people should choose their preferable result according to their real needs.

As illustrated before, different disciplines will result in different service candidates generation. Here the second result as mentioned above is chosen to make some further illustrations. According to the second result, this library process service can be separated into five services which can be named as: $Service_1$: online service, $Service_2$: facility service, $Service_3$: library card service, $Service_4$: library service, $Service_5$: logistics service. These five right-granularity services can achieve the objective of high-cohesion and low coupling.

## 5 Conclusion

The concept of SOA has been debated in recent years, the service granularity is a crucial issue in designing the SOA in order to satisfy low coupling, high cohesion and low reuse cost principles. Although the coarse-grained services have their own significance, its important to deal with other possible granularity levels. In this paper, we have attempted to come up with a new service identification approach which takes advantage of features of semantic analysis. This approach takes the business process as handling object to construct an ontology chart and partition the chart to get the identified services to reduce the coupling of remote functions and to increase the local function cohesion. This proposed framework is being instantiated in the library process to be evaluated but it still needs some improvements. The application of this method challenges us to concentrate our work on designing a special service identification tool, and to validate it in other domains in the future.

## Acknowledgements

This work was partially supported by the State Key Laboratory of Software Development Environment of China (No. SKLSDE-2013ZX-25), the National Natural Science Foundation of China (No. 61103095), and the National High Technology Research and Development Program of China (No. 2013AA01A601).

## References

1. Amsden, J.: Modeling soa: Part 1. service identification (2007), `http://www.ibm.com/developerworks/rational/library/07/`
2. Bianchini, D., Cappiello, C., De Antonellis, V., Pernici, B.: P2s: A methodology to enable inter-organizational process design through web services. In: Advanced Information Systems Engineering. pp. 334–348. Springer (2009)
3. Bianchini, D., Pernici, B., Cappiello, C., De Antonellis, V.: Service identification in inter-organizational process design. IEEE Transactions on Services Computing In Press (2013)
4. Dwivedi, V., Kulkarni, N.: A model driven service identification approach for process centric systems. In: Proceedings of 2008 IEEE Congress on Services, Part II. pp. 65–72 (2008)

5. Gibson, J.J.: The ecological approach to visual perception. Routledge (1986)
6. Gu, Q., Lago, P.: Service identification methods: A systematic literature review. In: Proceedings of Towards a Service-Based Internet - 3rd European Conference. pp. 37–50 (2010)
7. Haesen, R., Snoeck, M., Lemahieu, W., Poelmans, S.: On the definition of service granularity and its architectural impact. In: Proceedings of 20th International Conference on Advanced Information Systems Engineering. pp. 375–389 (2008)
8. Inaganti, S., Behara, G.K.: Service identification: BPM and SOA handshake. BPTrends 3, 1–12 (2007)
9. Jamshidi, P., Sharifi, M., Mansour, S.: To establish enterprise service model from enterprise business model. In: Proceedings of 2008 IEEE International Conference on Services Computing. pp. 93–100 (2008)
10. Kang, D., Song, C.y., Baik, D.K.: A method of service identification for product line. In: Proceedings of 3rd International Conference on Convergence and Hybrid Information Technology. vol. 2, pp. 1040–1045 (2008)
11. Kim, Y., Doh, K.G.: Pragmatic granularity decision for right-grained services in service-oriented modelling. International Journal of Web and Grid Services 8(2), 111–133 (2012)
12. Kohlborn, T., Korthaus, A., Chan, T., Rosemann, M.: Identification and analysis of business and software services - a consolidated approach. IEEE Transaction on Services Computing 2(1), 50–64 (2009)
13. Liu, K.: Semiotics in Information Systems Engineering. Cambridge University Press (2000)
14. Mani, S., Sinha, V., Sukaviriya, N., Ramachandra, T.: Using user interface design to enhance service identification. In: Proceedings of 2008 IEEE International Conference on Web Services. pp. 78–87 (2008)
15. Wang, Z., Xu, X., Zhan, D.: Normal forms and normalized design method for business service. In: Proceedings of 2005 IEEE International Conference on e-Business Engineering. pp. 79–86 (2005)