Semantic Web Service Discovery and Selection: a Test Bed Scenario

Alessio Carenini[2], Dario Cerizza[2], Marco Comerio[1], Emanuele Della Valle[2], Flavio De Paoli[1], Andrea Maurino[1], Matteo Palmonari[1], Matteo Sassi[2] and Andrea Turati[2]

[1]University of Milano - Bicocca, viale Sarca 336, 20126 Milano, Italy
{comerio,depaoli,maurino,palmonari}@disco.unimib.it
[2]CEFRIEL - Politecnico di Milano, via Fucini 2, 20133 Milano, Italy
{carenini,cerizza,dellavalle,sassi,turati}@cefriel.it

Abstract. The Semantic Web Service Challenge is one of the major initiative dedicated to work on Semantic Web Service (SWS) discovery and selection. It represents an effective manner for evaluating the functionality of SWS technologies. In this paper, we provide a new SWS-Challenge scenario proposal with new interesting problems to be solved on the basis of an analysis of a real world shipment scenario in the logistic operators domain. In the discussion a number of aspects concerning the discovery and selection processes are emphasized. In particular, we focus on the problem of considering the heterogeneity between the provider and the requester perspectives, and on the differences between functional and non functional specifications both on the requester and provider side.

Key words: Semantic Web Services Discovery, Semantic Web Services Selection, Non-functional properties, Semantic Matching, SWS-Challenge.

1 Introduction

Currently, Web services (WSs) technology enables the implementation of the Service Oriented Computing paradigm to develop Web processes that are accessible within and across organizational boundaries [6]. Web services provide for the development and execution of business processes that are distributed over the network and available via standard interfaces and protocols. Nevertheless, there is a growing consensus that pure functional descriptions of WSs are inadequate to develop valuable processes due to the high degree of heterogeneity, autonomy, and distribution of the Web [1]. Moreover, the increasing availability of Web services that offer similar functionalities with different characteristics increases the need for more sophisticated discovery processes to match user requests.

A way to address this problem is to combine Semantic Web and Web service technologies. The resulting Semantic Web service (SWS) technology supports the (semi)-automatic development of Web processes that go beyond simple interactions between organizations. The focus is on the discovery and selection activity that is requested to identify the needed Web services and supply the business information to include them in a business process. The idea is that a number of SWSs can meet some basic requirements specified by a requester. Then, a requester needs to be supported in choosing which one of the above SWSs better fulfills a set of desiderata; the selection can be automatic, e.g. automatically proceeding to invoke the "best" service ("best" according to a set of chosen criteria), or left to the discretion of a human requester.

Today, there is not yet a clear agreement on the definition of the terms discovery and selection. In this paper we call *discovery* the activity of locating a machine-processable description of a SWS that meets certain functional criteria [8]; and *selection* the activity of evaluating the discovered services to identify the ones that better fulfill a set of non-functional properties (NFPs)requested by the actual user [4], whatever the final invocation step is automatized or user-driven. In our vision, the selection process can be considered an enrichment of the discovery process.

A research topic that has been recently addressed concerns the support to NFP-based service selection (see [4, 7] and references therein); with respect to this problem, different models for NFP request and offer specification have been proposed. Typically, models based on attribute-value specifications are used [9]. An attribute identifies the involved NFP and the value quantifies it. This approach is effective and intuitive when simple specifications are represented (i.e. the value is known, fixed, and not dependent on any condition). However, the attribute-value approaches have two main shortcomings. First, an explicit representation of joint offers of NFP is not natively supported; in order to represents such joint offers there is the need to use a set of axioms establishing same conditions for set of values. Second, the expressiveness of the attribute-value approach is quite poor. NFP requests and offers make use of constraint expressions involving constraint operators that are not usually considered. In order to address such kinds of requests, new language constructs for NFP requests should be introduced. In the literature NFP models overcoming some limits of the attribute-value approach have been proposed (e.g. [5, 3]). Some of the authors of this paper proposed a NFP model in order to handle some of the above shortcomings of the attribute-value approach [2, 8]; this model is based on the concept of Policy (i.e., a joint offer or request of NFPs) and on the explicit representation of offers and requests through constraints.

The Semantic Web Service Challenge is one of the major initiative dedicated to SWS discovery and selection. The SWS-Challenge provides test bed scenarios and consists of a set of problems to be solved by the participants. The SWS-Challenge is an effective manner for evaluating the functionality of SWS technologies. One of the scenarios of the SWS-Challenge is about the problem of discovering a shipment service able to satisfy a set of user requirements (e.g., the pick-up location and time). This scenario points out the necessity to consider both functional and non-functional characteristics of SWSs.

By starting from that fictitious scenario, in this paper we propose a new similar scenario that is more realistic because it is based on a real domain. For this reason, we provide a detailed description of the activities and characteristics of logistic operators in order to acquire the needed knowledge about the domain. We will focus on a number of aspects concerning the discovery process, such as the problem of considering the heterogeneity between the provider and the requester perspectives, and the differences between functional and non-functional specifications both on the requester and provider side. This new test bed scenario will introduce new interesting problems to be solved.

The paper is organized as follows. Section 2 describes the logistic operator domain, while Section 3 highlights important aspects that should be considered in modeling such domain. Then, Section 4 proposes a novel shipment discovery scenario for the SWS-Challenge. Finally, Section 5 draws conclusions.

2 The Logistic Operator Domain

In order to provide a fully featured test for semantic Web services, a structured knowledge-intensive description of a real domain is necessary. The logistic operator domain has all the features to attain a complete testing for semantic Web services: (i) well defined terminology; (ii) well defined processes; (iii) complex normative; (iv) different points of views between the service provider and the requester.

This realistic description has been realized within the Networked Peers for Business (NeP4B) project [1] through (i) the analysis of several real-world logistic operator prices and service offer lists, (ii) several phone and face to face interviews with logistic operators and client companies and (iii) the analysis of the normative that regulates the logistic operator activities. In the analyzed domain there are different kinds of actors (logistic operator, client company and broker) each one with his features and objectives. The primary objective of a logistic operator is to provide logistic support to third parties. Client companies, instead, need a logistic operator for their transportation, in order to fulfill a single or a continuous need, i.e. an organizational emergency or the outsourcing of the logistic coordination. A broker is a company with the intent of collect the service offers of several small logistic operators, in order to offer a more complete service. For a client company there's no real difference between a broker or a single logistic operator.

The logistic support offered by logistic operators can concern freight transport and warehousing which are complex and structured services. In freight transport service, the logistic operator makes a delivery of a freight, in an agreed location, within some time constraints. In warehousing service, the logistic operator stores a certain amount of goods for some time in one of its warehouse. A logistic operator can provide only freight transportation service or only warehousing service or both the services, even in combination.

Freight transport and warehousing regard different kinds of goods. Each kind must be treated in different ways and needs a specific means of transportation i.e. classes of trucks. Mainly, the categories of goods are: (i) ordinary; (ii) perishable; (iii) dangerous and (iv) cumbersome. Ordinary goods do not need a

[1] http://www.dbgroup.unimo.it/nep4b/en/index.htm

specific way or particular attention to be treated. They can be transported by ordinary trucks, with no particular features. An ordinary load can be transported aggregated with other kinds of ordinary goods (groupage). Perishable goods, instead, need a particular temperature that must be maintained for all the transport duration. Perishable goods can be fresh, frozen or deep-frozen and need to preserve different temperature during the transportation. This means that not every perishable goods can be transported together with other ones. Dangerous goods, need a particular planning of the freight and particular treatment. Cumbersome goods are particular loads characterized by big dimensions and weight. The transport of this kind of good is regularized by the local rules of the road (exceptional transport).

In order to provide a logistic support service, a logistic operator must have a fleet and/or a logistic platform, i.e. a set of means of transport and/or a set of warehouses in different locations. The features of every kind of truck are stated by a specific normative. For example, the Accord Transport Perissable (A.T.P.) normative describes several classes of trucks with their temperature range. Warehouses are characterized by several properties as the kind of goods that can be stored, the minimum and/or maximum dimension and weight of a load and a weekly based calendar with its opening time.

3 Discussion: Aspects to be Considered

A number of aspects need to be considered in the above described scenario. In particular in what follows we focus on the need to consider different perspectives for the providers and the requesters of the service, and on the necessity to address both functional and non-functional properties of the service.

3.1 The Requester and Provider Perspectives in the Logistic Operator Domain

In the process of service discovery one must keep in mind that there are different points of view when describing a logistic operator service offer or a client company request. In fact the logistic operator has its terminology for the description of the services that a client company may not know or understand. A logistic operator: (i) has a fleet; (ii) has a logistic platform; (iii) covers a particular geographic area; (iv) has time constraints; (v) has a base price; (vi) accepts particular payment methods; (vii) fulfills particular kinds of transportation; (viii) provides some additional services. The logistic operator fleet identifies the kinds of goods that a logistic operator can treat since specific normative states particular procedures and freight truck for every kind of goods. Other features such as carrying capacity of the trucks in the fleet may attest the ability of a logistic operator to fulfill a client company request. The logistic platform identifies the kinds of goods that a logistic operator can store, the minimum and maximum load dimensions and the location of the warehouse. The covered geographic area is where a logistic operator can pick-up or deliver a load. Freight transport and warehousing have time constraints (e.g., the minimum quantity of hours to plan the freight) and a base price (i.e., the cost for the service invocation). A freight can be pre-paid or collect, i.e. the freight bill is at sender's or addressee's expense. The kinds of transportation identifies the kinds of locations that a logistic operator is willing to reach as pick-up or delivery location. The first distinction is between a specific address (domicile) or a warehouse of the logistic operator as start/end location of the freight (even in combination). Other particular kinds of transportation are the delivery to fair, to supermarket, to installation site or to remote location. Each of the particular kinds of locations requires particular procedures or means of transport and not all the logistic operators provide these kinds of services. Every logistic operator can provide additional services such as freight insurance, COD (Cash On Delivery), express deliveries or deliveries with precise and mandatory delivery instant. In addition, based upon load weight and the kind of transportation, a logistic operator may choose to do a Full-Truckload (FTL) or Less-than-Truckload (LTL) transport, i.e. aggregate several loads in one freight (cheaper) or use a means of transport for only one load (faster). Usually, LTL transport is performed only in certain locations (warehouses) in relation with certain time intervals (e.g. every Monday morning, delivery in Milan, etc).

A client company may express its need in a whole different way. A request is characterized by: (i) a quantity of a particular kind of goods; (ii) a pick-up location; (iii) a delivery location; (iv) time constraints; (v) a payer; (vi) additional features. The first point identifies the load to be transported. The pick-up and delivery locations identify the address and the kind of location of the freight (warehouse, domicile, etc.). A transport has a start and an end time identifying the interval between the pick-up and the delivery of the load. In a request is even specified at who's expense is the freight bill, i.e. the addressee or the sender. In the additional features, a requester may specify preferences related to the possibility to aggregate the load with other loads, the freight insurance, the delivery time and particular truck features.

3.2 Functional and Non Functional Properties

Typically, *functional properties* (FPs) of services represent the description of the service functionalities and *non-functional properties* (NFPs) represent the description of the service characteristics (e.g., business aspect, QoS) which are not directly related to the functionality provided [2, 8].

Nevertheless, this distinction is ambiguous and does not provide a set rule to qualify a property as FP or NFP. From our point of view this is a consequence of the fact that functional or non-functional is not an intrinsic qualification of a property, but it depends on the application domain and context. For example, the service location could be classified as a FP for a logistic service and as a NFP for a payment service. Moreover, from the requester perspective, the classification of requested properties with respect to FP/NFP is of little interest and counterintuitive. The requested properties represent the user preferences and could be mandatory and optional.

In this paper, we distinguish between the provider and the requester perspectives and we propose the following classification of properties. From the requester perspective, we considered hard and soft constraints to distinguish between the properties that are specified as mandatory or optional in a service request. From the provider perspective, we consider FPs those properties of a service that strictly characterized the offered functionality (e.g., service location for a shipment service) and NFPs those properties of a service that do not affect or affect only marginally the offered functionality (e.g., service location for a payment service).

The above consideration makes the problem of matching between requested and offered properties more difficult to handle. A possible approach consists in mapping FPs and NFPs with hard and soft constraints respectively. Note that, even if NFPs are considered as soft constraints, they could be quite relevant in order to select a service that better matches the user request. The selection should occur among a set of services that satisfy the specified hard constraints. In practice, the role of hard constraints is to define a boolean match that filters eligible services. On the contrary, matching between soft constraints should be considered not crispy and degrees of satisfaction should be evaluated (i.e., soft constraints can be fully, partially or not satisfied). In order to support non-boolean matching a number of characteristics should be considered:

- **Support for sophisticated descriptions**: Requests, and also offers, should refer to a full set of constraint operators, including "$\leq$", "$\geq$" and range. Moreover, *relevance* attributes should specify the importance that requesters give to each requested property. The use of relevance attributes and operators in constraint expressions (e.g., the cost of the service must be ≤ 3 Euro) enhance the expressivity of the descriptions to support non-boolean matching.
- **Support for offer clustering**: A Web service can be offered with different levels of NFPs. For example, a SMS service offered by a telephone company can be characterized by different policies: one offering a price of 0.10 Euro for SMS and a maximum of 100 SMS per day; another offering a price of 0.15 Euro and no restriction on the number of SMS per day. Moreover policies can be associated with *conditions* to state that their applicability depends on the requester's profile or context; e.g. a condition may grant a discount on shipment services to registered customers or for multiple shipment requests.
- **Support for static and dynamic properties description**: *Static* NFPs specify properties of services by fixed values that can be included in service descriptions at publishing time. *Dynamic* NFPs describe properties of services that are strictly related to the context of execution and cannot be stated at publishing time. Moreover, dynamic NFPs require service calls to fix the actual values. For example, the actual price for shipment depends on the shipped object's weight, therefore a call is necessary to fix it.

4 New SWS-Challenge Proposal

The current discovery scenario of the SWS-Challenge is based on shipment services offered by generic logistic operators. These services are defined mainly in terms of the time required to delivery the goods, the covered geographical area and the price. In sections 2 and 3 we have described several features that must be taken into account when modeling the logistic operator domain. In this section we propose a new shipment discovery scenario by defining four logistic operator web services and two user goals. As our logistic operator domain description offers much more details with respect to the one given by the current scenario, we decided to use different names for the services. The reason is that we do not merely add more attributes, but we redefine some concepts that were already present in the previous shipment service descriptions.

4.1 Logistic operators

"Liteworld" is an operator specialized in transcontinental deliveries of small ordinary goods. Its features are listed below.

- Price (flat fee/each kg): Europe(41/6.75), Asia(47.5/7.15), North America(26.25,4.15), Rates for South America like North America, Rates for Oceania like Asia
- Performs LTL delivery
- Maximum weight: 5 kg
- Transports to/from domicile
- Covered geographic area: Africa, Europe, North America, Asia
- Fleet: 20 trucks for ordinary goods
- Freight insurance: refund for loss
- If number of requested shipments ≤ 5:
  - Base price: 50 Euro
  - Minimum hours to delivery: 96 hours
- If number of requested shipments > 5:
  - Base price: 30 Euro
  - Minimum hours to delivery: 72 hours

"Hello" delivers ordinary goods and performs exceptional freight deliveries in three European countries and in Asia:

- Price on request
- Performs LTL and FTL delivery
- Transports to/from domicile
- Minimum time to deliver: 48 hours
- Covered geographic area: Italy, France, Germany, Asia
- Fleet: 10 trucks for ordinary goods, 12 trucks for exceptional transports
- Freight insurance: refund for damage and refund for loss
- Price:
  - Base price: 35 Euro

- Base price discount if number of requested shipments $\geq$ 10: 20%
- Base price discount if number of requested shipments $\geq$ 20: 50%

"Fresh and fast" has a fleet that allows it to deliver fresh goods (ATP IN) as well as frozen goods (ATP IR). Its two warehouses allow this operator to perform transports to and from warehouses:

- Price on request
- Performs LTL and FTL delivery
- Transports to/from domicile and to/from warehouse
- Minimum time to deliver: 24 hours
- Covered geographic area: Italy, France, Germany
- Fleet: 10 trucks for ordinary goods, 5 trucks for ATP IN goods, 5 trucks for ATP IR goods
- Warehouses: see table 1

Table 1. "Fresh and fast" operator: warehouses

	W1	W2
Location	Cannes	Paris
Accepted goods	ATP (IN and IR)	ATP (IN and IR)
Minimum weight	5 q	5 q
Minimum storage time	3 days	5 days

"Safe transports" is a specialized operator that performs transports of frozen and dangerous goods:

- Price on request
- Performs FTL delivery
- Transports to/from warehouse
- Minimum time to deliver: 48 hours
- Covered geographic area: Belgium, France, Spain
- Fleet: 3 trucks for dangerous (ADR) goods, 5 trucks for ATP IR goods
- Warehouses: see table 2

Table 2. "Safe transports" operator: warehouses

	W1	W2	W3	W4	W5
Location	Cannes	Lion	Madrid	Antwerp	Paris
Accepted goods	ATP	ATP	ADR	ADR	ATP
Minimum weight	20 q	10 q	10 q	10 q	20 q
Minimum storage time	3 days	3 days	3 days	3 days	3 days

4.2 Goals

A first goal requires the delivery of fruit ice cream which requires an ATP IN truck. The goods will be brought by the customer to a warehouse that has to be located in Paris and have to be delivered to a warehouse in Cannes. This goal tests both the fleet of the operator and the features of its warehouses.

- Freight type: Fruit ice cream
- Weight: 10 q
- Source: Paris, France
- Destination: Cannes, France
- Other requirements: Transport from warehouse to warehouse

This goal will be satisfied by the "Fresh and fast" operator. One of the warehouses of "Safe transports" has a minimum weight for incoming freight that is higher than the weight of the ice cream, while neither the fleet of the "Hello" operator nor the one of the "Liteworld" operator can deliver perishable goods.

We can now introduce non functional properties while expressing a second goal, which requires the transport of piles of documents from Singapore to Turin. This goal states preferences about the price, the maximum time to deliver and the freight insurance. This means that if two operators can fulfill the functional requirements, the one that better satisfies the non functional requirements have to be chosen.

- Freight type: Documents
- Weight: 3 kg
- Source: Singapore
- Destination: Turin, Italy
- Base price $<$ 30 Euro (preferred)
- Hours to delivery $\leq$ 72 hours (preferred)
- Freight insurance: refund for damage and refund for loss (preferred)
- Collaboration period: from 2008-03-03 to 2008-05-26
- Transport frequency: weekly

The request expressed by this goal is fulfilled by both "Liteworld" and "Hello". "Safe transports" and "Fresh and fast" do not match the requests, as the former performs only FTL goods delivery and the latter does not cover the source area. As "Hello" better satisfies the optional requests related to Freight insurance, Base price and Minimum hours to delivery, it should be preferred over "Liteworld".

5 Conclusion

The Semantic Web Service Challenge is one of the major initiative dedicated to SWS discovery and selection. The SWS-Challenge provides test bed scenarios and consists of a set of problems to be solved by the participants. In this paper we have proposed a new shipment discovery scenario based on an analysis of the logistic operator domain realized within the Networked Peers for Business (NeP4B) project. The new scenario considers the necessity of: (i) addressing the heterogeneity between the provider and the requester perspectives and (ii) specifying functional and non-functional properties both on the requester and provider side. This new scenario can be considered an effective manner for evaluating the possibility to use SWS technologies for generating requested and offered property descriptions useful for the discovery and selection activities.

Acknowledgements

The work presented in this paper has been partially supported by the European IST project n. 27347 SEEMP - Single European Employment Market-Place and the Italian FIRB project RBNE05XYPW NeP4B - Networked Peers for Business.

References

1. J. Cardoso and A. P. Sheth. Introduction to semantic web services and web process composition. In *Proc. of First Intl Workshop on Semantic Web Services and Web Process Composition (SWSWPC'04)*, San Diego, CA, USA, 2004.
2. M. Comerio, F. DePaoli, A. Maurino, and M. Palmonari. Nfp-aware semantic web services selection. In *Proc. of the Enterprise Computing Conference (EDOC). Accepted to be published.*, Annapolis, Maryland, USA, 2007.
3. E. Giallonardo and E. Zimeo. More semantics in qos matching. In *Proc. of International Conference on Service-Oriented Computing and Application (SOCA07)*, pages 163–171, Newport Beach, CA, USA, 2007.
4. M. Kerrigan. Web service selection mechanisms in the web service execution environment (wsmx). In *Proc. of the 21st Annual ACM Symposium on Applied Computing (SAC2006)*, Dijon, France, 2006.
5. K. Kritikos and D. Plexousakis. Semantic qos metric matching. In *ECOWS '06: Proceedings of the European Conference on Web Services*, pages 265–274, Washington, DC, USA, 2006. IEEE Computer Society.
6. M. P. Papazoglou. Web services and business transactions. *World Wide Web*, 6(1):49–91, 2003.
7. J. Pathak, N. Koul, D. Caragea, and V. Honavar. A framework for semantic web services discovery. In *In WIDM '05: Proceedings of the 7th annual ACM international workshop on Web information and data management*, pages 45–50, New York, NY, USA, 2005. ACM Press.
8. I. Toma, D. Foxvog, F. D. Paoli, M. Comerio, M. Palmonari, and A. Maurino. Non-functional properties in web services. wsmo d28.4 v0.2. Technical report, 2008.
9. WSMO. *The Web Service Modeling Ontology (WSMO). Final Draft.* Available at: http://www.wsmo.org/TR/d2/v1.2/20050413/, 2005.