WEB SERVICE DISCOVERY IN A SEMANTICALLY EXTENDED UDDI REGISTRY: THE CASE OF FUSION

Dimitrios Kourtesis[1], Iraklis Paraskakis[1], Andreas Friesen[2], Panagiotis Gouvas[3], Athanasios Bouras[3]

[1] *SEERC - South East European Research Centre,*
A Research Centre of the University of Sheffield and CITY Liberal Studies
Mitropoleos 17, 54624, Thessaloniki, Greece
dkourtesis@seerc.org, iparaskakis@seerc.org

[2] *SAP Research, CEC Karlsruhe*
Vincenz-Prießnitz-Str. 1, D-76131 Karlsruhe, Germany
andreas.friesen@sap.com

[3] *Institute of Communication and Computer Systems,*
National Technical University of Athens
Iroon Polutexneiou 9, 15780 Zografou, Athens, Greece
pgouvas@mail.ntua.gr bouras@mail.ntua.gr

Service-oriented computing is being adopted at an unprecedented rate, making the effectiveness of automated service discovery an increasingly important challenge. UDDI has emerged as a de facto industry standard and fundamental building block within SOA infrastructures. Nevertheless, conventional UDDI registries lack means to provide unambiguous, semantically rich representations of Web service capabilities, and the logic inference power required for facilitating automated service discovery. To overcome this important limitation, a number of approaches have been proposed towards augmenting Web service discovery with semantics. This paper discusses the benefits of semantically extending Web service descriptions and UDDI registries, and presents an overview of the approach put forward in project FUSION, towards semantically-enhanced publication and discovery of services based on SAWSDL.

1. INTRODUCTION

Interoperability among enterprise information systems is key to achieving business agility, especially for enterprises operating within collaborative value networks, as it largely determines their capacity to respond swiftly to changing market conditions and new collaboration opportunities. The paradigm of Service Oriented Architecture (SOA) and its manifestation in the form of Web services promise to alleviate many

enterprise application interoperability barriers, thus significantly reducing the effort to establish collaborative business processes. In a SOA environment, the creation of a new business process necessitates the assembly of reusable services exposed by multiple enterprise applications into new compositions. But for services to be composed, they first need to be discovered. Designers of collaborative business processes must search for services able to perform some specific task within a workflow, and compose them to form a complex business process. When searching in a pool of resources containing hundreds of Web service descriptions, discovery can become a demanding task.

This is why the Universal Description, Discovery and Integration specification (UDDI, 2002) has emerged as a de facto industry standard. UDDI registries are fundamental building blocks within a SOA infrastructure, serving as central cataloguing services for reusable software components. Nevertheless, conventional UDDI registries (v2 or v3) provide very limited means for automated service discovery (Paolucci et al, 2002), (Colgrave, Akkiraju and Goodwin, 2004). This is because indexing and retrieval in UDDI is not grounded on formal specifications of service capabilities with machine-processable semantics, but on natural language descriptions and categorisations retrievable through keyword-based search. Keyword-based annotation and search cannot be relied upon for automated discovery, because it cannot differentiate between services that have totally different names but equivalent functionality, or services that have identical naming but perform totally unrelated operations. To facilitate efficient service discovery during process design, we need to provide an explicit, formal, unambiguous, commonly comprehensible and machine processable representation of service characteristics.

The aim of this paper is to introduce the approach adopted in project FUSION towards semantically-enhanced publication and discovery of services. FUSION is an EU-funded research project aiming to promote semantics-based interoperability among service-oriented business applications and efficient business process integration among collaborating enterprises (Alazeib et al, 2006). The FUSION Semantic Registry is a semantically extended UDDI registry that is a core part of the FUSION system architecture. The rest of this paper is organised as follows. Section 2 defines a scheme of generic service discovery requirements, setting the basis for the discussion to follow. Section 3 discusses the benefits of adding semantics to Web service descriptions and presents the SAWSDL-based approach adopted in FUSION. Section 4 provides a brief overview of the semantic service publication and discovery procedures, as they take place inside the FUSION Semantic Registry.

2. SERVICE DISCOVERY REQUIREMENTS

A number of solutions for semantically-enhanced service discovery have been proposed in the research literature, each of them satisfying a different set of service discovery requirements. To provide means for comparing with other approaches, and explicating the functionality rationale of the FUSION Semantic Registry, we establish a generic scheme of service discovery requirements. The scheme we present in this section defines the relevance and degree of match between service advertisements and service requests at three distinct levels: (i) functionality-level matching, (ii) message-level matching, and (iii) schema-level matching.
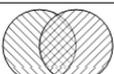
2.1 Functionality-level matching

Functionality-level matching is the most basic interoperability determinant and first aspect of relevance to be considered. Web service descriptions need to categorise a service with respect to some classification scheme, describing the business function being performed. This is one of the factors to be considered in determining if a service advertisement can provide the functionality sought by the service request, or not. Alternative ways to accommodate this type of categorisation and search in conventional UDDI registries (with syntactic means) are quite common in the literature (e.g. for popular taxonomies such as NAICS and UNSPSC, or custom ones). No complex notion of a degree of match exists, since the outcome is boolean.

2.2 Message-level matching

The second aspect of matching we propose is at the level of messages. The goal in this type of matchmaking is to determine the degree to which an advertised service will produce all output data that a business process participant expects to receive, and the degree to which it will accept all input data that a business process participant intends to provide. This involves a comparison between the ontological concepts corresponding to the service messages exchanged by the advertisement, and that of the request. To provide a formalisation of the different degrees of match that could potentially occur at the message level, we provide an intuitive set-theoretic model based on Description Logics (DL), adopted from the work of (Li and Horrocks, 2003), and (Keller et al, 2004).

Table 1 - Set-theoretic model for matchmaking. Advertisement set A is represented with (⊘) and request set R with (⬃)

Degree of Match	Condition	Graphical Representation
Exact Match	The advertisement offers all input and output messages specified by the request, and no irrelevant ones	*(sets A and R are equal)*
Plugin Match	The advertisement offers all input and output messages specified by the request, and some irrelevant ones	*(set A is a proper superset of R)*
Subsumption Match	The advertisement offers only some of the input and output messages specified by the request, and no irrelevant ones	*(set A is a proper subset of R)*
Intersection Match	The advertisement offers only some of the input and output messages specified by the request, and some irrelevant ones	*(sets A and R are intersecting)*
Non Match	The advertisement offers none of the input and output messages specified by the request	*(sets A and R are disjoint)*

2.3 Schema-level matching

The situation where an exact match has been established for all requested inputs and outputs is not necessarily sufficient for the request and the advertisement to be considered as matching. When performing matchmaking under the assumption of a shared base ontology (as in the case of FUSION), which any party can extend by subclassing and applying property restrictions in the form of value or cardinality constraints, we need to also consider matching at the message schema structure level. Collaborating partners in a network may choose to extend concepts in a shared ontology in different ways, best expressing the semantics of the Web services that their enterprise application systems expose.

Figure 1 illustrates this through an example: a case in which the concept of *address,* as defined in the FUSION Ontology, could have been extended in different ways through subclassing in order to be used for modelling two different enterprise applications. Although *System1_Address* and *System2_Address* are subclasses of a common concept, their input and output messages cannot always be interoperable. The message schema of *System2_Address* is more specific than that of *System1_Address*, since the first specifies more attributes than the schema of the latter. In fact, if Web services exposed by *System1* were to exchange *address* related messages with Web services exposed by *System2*, *System2* could consume all of the data provided to it, but require some additional data that would not have been provided, thus leading to potential problems during process execution.
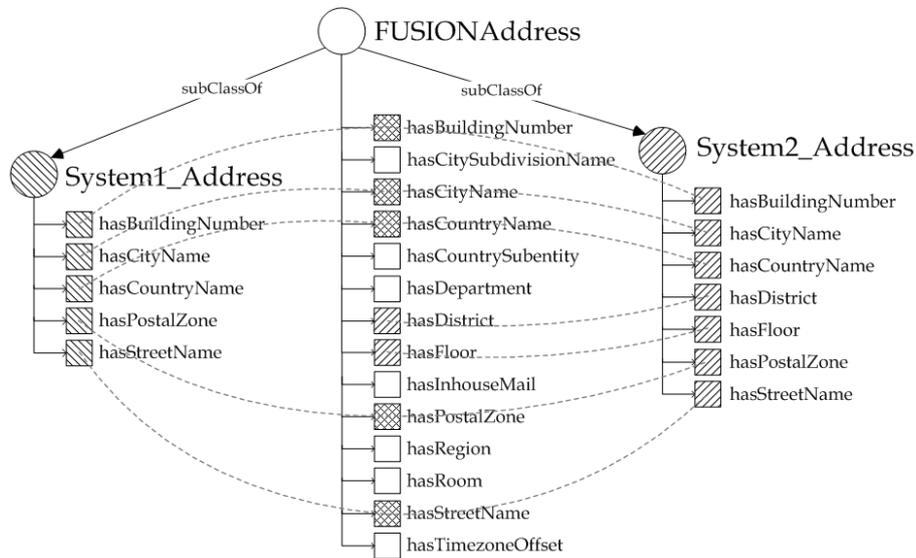

Figure 1 - Different ways to extend a shared ontology concept through subclassing

One could reapply the set-theoretic matchmaking model presented in Table 1 to distinguish among cases where the advertisement message schema is isomorphic to that of the request, more generic, more specific, or has no attributes in common. In the first two cases the match could be considered satisfactory, while in the latter two the match may not be sufficient for seamless message exchange and interoperability.

3. ADDING SEMANTICS TO SERVICE DESCRIPTIONS

Web services are highly reusable, self-contained software components having standardised, platform-independent and implementation-neutral interfaces described using the Web Service Description Language (WSDL) (Christensen et al, 2001). A WSDL description contains the technical information required for invoking a Web service, but does not pronounce anything meaningful about the capabilities or the behaviour of the service itself. The semantics to be added to Web service descriptions should advertise service characteristics in terms of the functionality and intrinsic business value they offer (i.e. functional semantics), the messages they exchange (i.e. data semantics), and the observable behaviour they expose within a complex process execution (i.e. behavioural semantics). This facilitates efficient indexing upon publication to a semantically enhanced service registry, and automated service discovery based on explicit, formal, and unambiguous service characteristics.

Numerous frameworks have been proposed in the recent years for constructing semantic descriptions of services, and several have been promoted for standardisation through W3C member submissions: OWL-S (Martin et al, 2004), WSMO (Bruijn et al, 2005), and WSDL-S (Akkiraju et al, 2005). The implementation of the FUSION Semantic Registry is building upon the most recent development in this area: Semantic Annotations for WSDL (SAWSDL) (Farrell and Lausen, 2007). SAWSDL, building on the same principles as WSDL-S, is a simple and generic mechanism for semantically annotating Web service descriptions taking advantage of the WSDL extension mechanism. SAWSDL defines two types of annotations that can be used for attaching "meaning" to WSDL or XML schema components inside a WSDL file: identifiers of concepts, using *modelReference* extension attributes, and identifiers of mappings from concepts to XML Schema elements, using *loweringSchemaMapping* and *liftingSchemaMapping* extension attributes. SAWSDL supports annotations for both WSDL 1.1 and WSDL 2.0 documents. In this discussion we refer to the use of SAWSDL in conjunction with WSDL 1.1. Table 2 summarises the different type of semantics that must be captured for semantically enriching Web service descriptions, by linking concepts from the FUSION Ontology to WSDL components.

Table 2 - Adding semantics to Web service descriptions

Type of semantics	FUSION Ontology	Annotated WSDL 1.1 component	SAWSDL extension attribute
Functional semantics	Service classification taxonomy	portType	modelReference
Data semantics	Data facet	part	modelReference, loweringSchemaMapping liftingSchemaMapping
Behavioural semantics	State ontology	operation	modelReference

The FUSION Ontology (Bouras et al, 2006) serves as a commonly shared semantic model and interoperability vehicle. It is modular, multi-layered and multi-faceted, and has been created with a view to being easily extended. On one hand, the multiple-layering in the FUSION Ontology reflects hierarchical relationships between entities that are: *domain-independent* (foundational ontology concepts), *vendor-independent* (generic concepts in the domain of enterprise applications), *vendor-dependent* (concepts relating to specific ERP or CRM products) and *case-specific* (concepts relating to a specific installation and customisation of some ERP or CRM product). On the other hand, its multiple-facet structure reflects the different types of such entities, i.e. functionality-related, data-related, and process-related concepts, horizontally pertaining across all abstraction layers.

4. ADDING SEMANTICS TO UDDI

Augmenting a UDDI registry with semantic matchmaking capabilities facilitates automated high-precision retrieval of services to fulfil specific requirements when composing business processes, and constitutes a significant advancement compared to the existing publication and discovery facilities offered in conventional UDDI. The UDDI server module lying at the heart of the FUSION Semantic Registry is a typical implementation of the UDDI v2 specification by OASIS (UDDI, 2002). In order to support concept-driven discovery of Web service descriptions, the FUSION Semantic Registry enhances the purely syntactic search facilities that a UDDI server can offer. This is achieved without any modifications to the UDDI server implementation or the UDDI API, as has been proposed in (Paolucci et al, 2002), and (Akkiraju et al, 2003), but rather, through the incorporation of a number of semantically-enabled processing modules in the FUSION Semantic Registry architecture, such as a publication and a discovery manager with OWL ontology processing and DL reasoning capabilities. As a result, the UDDI server module remains semantics-agnostic, providing FUSION system implementers with UDDI vendor independence. In this section we provide a high-level overview of both publication-time and discovery-time activities in the FUSION Semantic Registry.

4.1. Publication-time activities

Upon delivery to the FUSION Semantic Registry the SAWSDL document is parsed to extract all modelReference annotations attached on wsdl:portType and wsdl:part entities. The concepts that these annotations reference serve as input for creating a new named OWL class definition in an internal index ontology (an extension to the FUSION Ontology), maintained with the sole purpose of assisting in indexing. The registry associates the newly created class with other FUSION Ontology classes to capture functional and data semantics, by assigning concepts from the Service Classification Taxonomy or the FUSION Ontology Data Facet as range values to the newly created class's object properties. Behavioural annotations pointing to State Ontology concepts are not processed since they are only utilised for semi-automated service composition, and not for discovery. The registry uses a DL reasoner to classify the newly created class, computing new subclass relations and identifying

superclasses for the newly created concept. The classification information serves as input to an indexing procedure involving SAWSDL to UDDI mapping. The URIs of the inferred superclasses are used as indexing keys for the service in UDDI. The mapping is based on the OASIS Technical Note for WSDL to UDDI mapping (Clément, Colgrave and Riegen, 2004), which we extended and adapted to suit the FUSION Ontology, and SAWSDL annotations for WSDL 1.1 service descriptions.

4.2. Discovery-time activities

As soon as the hybrid UDDI/OWL indexing procedure has been completed, service descriptions are readily discoverable by business process designers and system developers within a collaborative network. The suitability of an advertised service is evaluated based on a discovery query that comprises a URI, pointing to a semantic description of the capabilities sought, and an optional UUID, specifying a target business application offering the service. The registry examines the concept that the discovery URI points to, and determines if it constitutes a generic capability profile based on concepts defined in the FUSION Ontology, or some custom-built capability profile defined in a third ontology, specialising FUSION Ontology concepts by adding restrictions (e.g. by adding an extra object property, or setting some cardinality constraint). The first is the simplest discovery case, since the registry's index has already established associations from service advertisements to all concepts defined in the FUSION Ontology, and can directly seek tModels containing at least one keyedReference with a keyValue equal to the discovery URI. In the second case the registry must classify the custom-built profile against all concepts in its internal index ontology. The DL reasoner computes new subclass relations and reveals the superclasses that the custom-built profile can be classified under. The registry then seeks for service advertisements indexed with respect to the URIs of the inferred superclass concepts.

5. CONCLUSIONS

Collaborative networks have a lot to gain from the adoption of SOA and Web services, for which automated service discovery is a key enabler. To promote interoperability among service-oriented business applications and efficient business process integration among collaborating enterprises, FUSION necessitates the introduction of semantics to all aspects of the service discovery process. In this paper we presented a generic scheme of service discovery requirements, an SAWSDL-based approach for adding semantics to Web service descriptions, and an overview of the FUSION Semantic Registry publication and discovery functionality, augmenting the capabilities of conventional UDDI registries.

6. ACKNOWLEDGMENTS

7. REFERENCES

1. Akkiraju, R., Goodwin, R., Doshi, P. and Roeder, S. A method for semantically enhancing the service discovery capabilities of UDDI. In Proceedings of the IJCAI-03 Workshop on Information Integration on the Web, 2003.

2. Akkiraju R., Farrell J., Miller J., Nagarajan M., Schmidt M.T., Sheth A., Verma K. Web Service Semantics (WSDL-S). W3C Member Submission, November 2005. Available at: http://www.w3.org/Submission/WSDL-S/

3. Alazeib A., Bauer M., Bouras A., Dyczkowska M., Friesen A., Gouvas P., Jurkowski G., Jurkowski P., Kalaboukas K., Kourtesis D., Lukacsy G., Martinek P., Mentzas G., Pantelopoulos S., Paraskakis I., Szalai B. FUSION project Deliverable D1.2 - FUSION Approach. July 2006. Available at: http://www.fusion-strep.eu/

4. Bouras A., Gouvas P., Friesen A., Pantelopoulos S., Alexakis S., Mentzas G. Business Process Fusion based on Semantically-enabled Service-Oriented Business Applications. In Proceedings of the Second International Conference on Interoperability for Enterprise Software and Applications (I-ESA 2006), Bordeaux France, March 2006.

5. Cardoso, J. Discovering Semantic Web services with and without a Common Ontology Commitment.In Proceedings of the 3rd International Workshop on Semantic and Dynamic Web Processes (SDWP 2006), September 2006.

6. Christensen E., Curbera F., Meredith G., and Weerawarana S. Web Services Description Language (WSDL) Version 1.1, World Wide Web Consortium, 15 March 2001. Available at: http://www.w3.org/TR/2001/NOTE-wsdl-20010315

7. Clément L., Colgrave J., Riegen C.V. Using WSDL in a UDDI Registry, Version 2.0.2 – OASIS Technical Note. Available at: http://www.oasis-open.org/committees/uddi-spec/doc/tn/uddi-spec-tc-tn-wsdl-v2.htm

8. Colgrave J., Akkiraju R., Goodwin R., External Matching in UDDI. In Proceedings of the International Conference on Web Services (ICWS 2004), San Diego, USA, 2004.

9. Farrell J. and Lausen H. (eds) Semantic Annotations for WSDL and XML Schema. W3C Candidate Recommendation, January 2007. Available at: http://www.w3.org/TR/sawsdl/

10. Keller U., Lara R., and Polleres A., Toma I., Kifer M., Fensel D. WSMO Web Service Discovery. Technical Report, DERI, University of Innsbruck, 2004.

11. Lara R., Binder W., Constantinescu I., Fensel D., Keller U., Pan J., Pistore M., Polleres A., Toma I., Traverso P., and Zaremba M. Semantics for Web Service Discovery and Composition. Technical Report, Knowledge Web project, December 2004.

12. Bruijn J.d., Bussler C., Domingue J., Fensel D., Hepp M., Keller U., Kifer M., König-Ries B., Kopecky J., Lara R., Lausen H., Oren E., Polleres A., Roman D., Scicluna J., Stollberg M. Web Service Modeling Ontology (WSMO). W3C Member Submission 3 June 2005. Available at: http://www.w3.org/Submission/WSMO/

13. Li L. and Horrocks I.. A Software Framework for Matchmaking Based on Semantic Web Technology. In Proceedings of the World Wide Web Conference (WWW'03), Hungary, 2003.

14. Martin D., Burstein M., Hobbs J., Lassila O., McDermott D., McIlraith S., Narayanan S., Paolucci M., Parsia B., Payne T., Sirin E., Srinivasan N., Sycara K. OWL Web Ontology Language for Services (OWL-S). W3C Member Submission 22 November 2004. Available at: http://www.w3.org/Submission/OWL-S

15. Paolucci M., Kawamura T., Payne T., and Sycara K. Semantic Matching of Web Service Capabilities. In Proceedings of the First International Semantic Web Conference (ISWC2002), Sardinia, Italy, June 2002.

16. UDDI Version 2.04 API, Published Specification, July 2002. Available at: http://uddi.org/pubs/ProgrammersAPI-V2.04-Published-20020719.htm