Abstract—Web services are useless if they cannot be discovered. So, discovery is the most important task in the Web service model. Recent researchers have focused on performing semantic matching to enhance the accuracy of Web service discovery. In this paper, we present a framework for Web services discovery and selection based on intelligent software agents, OWL-S and domain ontologies. With the help of software agents, information provided by Web services can be made more efficient and more dynamic. With semantics provided by OWL-S and domain concepts, match and discovery can return the most relevant services.

Index Terms—web service discovery, QoS, agents, ontologies.

I. INTRODUCTION

Service-oriented computing (SOC) is an interdisciplinary paradigm that revolutionizes the very fabric of distributed software development. Applications that adopt service-oriented architectures (SOA) can evolve during their lifespan and adapt to changing or unpredictable environments more easily [1]. SOA is built around the concept of Web Services. Web Services are new forms of Internet software which can be invoked using standard Internet protocols. Web Services, as it is defined by the World Wide Web Consortium (W3C), is a software system designed to support interoperable machine-to-machine interaction over a network. Web services interact with each other, fulfilling tasks and requests that, in turn, carry out parts of complex transactions or workflows.

With the rapid development of web services technologies, discovering web services is becoming the most urgent problem to be resolved [2]. Discovery is the process of finding Web services Provider locations which satisfy specific requirements. Web services are useless if they cannot be discovered. So, discovery is the most important task in the Web service model [3]. There are two challenges facing the practicality of Web services discovery [4]: (a) efficient location of the Web service registries that contain the requested Web services and (b) efficient retrieval of the requested services from these registries with high quality of service (QoS). With ever increasing number of available web services it is problematic to find a service with required functionality and appropriate quality characteristics [5]. The main reason for this problem is that current web services technology is not semantic-oriented. Several approaches have been proposed to add semantics to Web Services descriptions to facilitate discovery and selection of relevant Web services (e.g. DAML-S [6], WSDL-S [7], WSML [8], OWL-S [9]). These so-called Semantic Web services can better capture and disambiguate the service functionality, allowing a logic-based matchmaking to infer relationships between requested and provided service parameters [10]. Recently, we have seen an explosion of interest in ontologies as artefacts to represent human knowledge and as a critical component in several applications; among these the web services [11]. Ontologies have been developed to provide a machine-processable semantics of information sources that can be communicated between different applications. They are also essential to the development and use of intelligent systems, particularly for the interoperation of heterogeneous systems. Ontologies are, thus, responsible for informing about the domain vocabulary and explaining the meaning that interacting systems attribute to terms [12]. In our work, we propose the use of ontologies to describe semantically the various parameters and characteristic of the customer request and the Web services.

Even the use of ontologies, machines or programs is still not efficient enough to take the automatic and dynamic decision for semantic Web service discovery. With the help of software agents, information provided by Web services can be made more efficient and more dynamic. Actually agents are intelligent enough to take the decision according to the changing environment and changing level of available information which cannot be expected from the traditional system[13]. Our work takes the multi-agent approach in which a team of agents, each with local information, collaborates to satisfy Web services discovery objective. The overall behavior of the system emerges through the dynamic interactions between agents.

If the discovery engine returned multiple candidate Web services provide the same functionality, then Quality of Service (QoS) is becoming an important criterion for selection of the best available Web service. The consumers have to pay enough attention to find the service provider who can satisfy their QoS requirements. If they cannot find a provider satisfying all their QoS requirements, they usually have to give up all the candidates or make some tradeoff [14]. Our work aims to provide a more “consumer-centric” approach simplifying service discovery using semantics while satisfying QoS requirements. A major problem in using QoS for service discovery is the specification and storage of the QoS information [15]. In this paper, we propose an ontology-based OWL-S extension to adding QoS to Web service descriptions.

Service consumers have different preferences. For

Manuscript received May 1, 2012; revised June 19, 2012.

R. Benaboud is with the Department of Mathematics and Computer Science, University of Tebessa, Tebessa 12000, Algeria (e-mail: r_benaboud@yahoo.fr).

R. Maamri and Z. Sahnoun are with LIRE Laboratory, Mentouri University of Constantine, Ain El Bey 25000, Algeria (e-mail: rmaamri@yahoo.fr; sahnounz@yahoo.fr).
example, a service consumer may want a service that offers the fastest response time while for another execution price could be his most important parameter. For this reason, we propose that service consumer gives different weights for different QoS attributes when he formulates his request.

The rest of this paper is organized as follows: In section 2, we outline related research, Section 3 presents our framework. In Section 4, we present request and service description. Section 5 describes how to calculate the degree of similarity between request and Web servicesand section 6 offers concluding remarks and future directions.

II. RELATED WORK

The problems pertaining to Web service discovery have long been taking attention of both academia and industry. Many researches have investigated the discovery of semantic web services, QoS-aware discovery or the use of agents for semantic web services discovery. We provide an overview of some of this work as a context for the research discussed in the remainder of the paper.

A. Semantic Web Services Discovery

Most current approaches for web service discovery cater to semantic web services, i.e., web services that have associated semantic descriptions. The work presented in [16] proposes a DAML-S matchmaking algorithm which is used to match a requested service with a set of advertised ones. This matching algorithm compares the input and output concepts of user request to the service description in registry and defines four levels of matching: Exact, Plug in, Subsumes, Fail. In our work, we don’t use only the subsumption relationships between concepts to calculate their similarity but we also take into account common properties between them.

Reference [17] presents an approach for web service discovery that combines semantic and statistical association metrics. Semantic metrics are based on the semantic aspects of relevant ontology. Statistical association metrics are based on the association aspects of web services instances (their inputs and outputs). The proposed approach exploits semantic relationship ranking for establishing semantic relevance, and a hyperclique pattern discovery method for grouping web service parameters into meaningful associations. These associations combined by the semantic relevance are then leveraged to discover and rank web services.

B. QoS and Web Services Discovery

The QoS requirements for Web Services have become vital for both service providers and consumers as several Web Services offers similar functionality [18].

Kokash [19] proposes a QoS-aware discovery and subscription approach. The core idea of this approach is to build up a “virtual service” grouping function similar services together (called service pool) and dispatching customer requests to the proper service in terms of QoS requirements. After investigating the structure and the underlying semantic similarity of Web services, it employs a similarity matching algorithm to cluster the function-similar services and generate a virtual WSDL so that the service pool can be accessed as a Web service. Assuming consumers’ QoS requirements are compliant with Web Service Quality Model (WSQM) [20], it designs algorithms to automatically finish the QoS negotiation between consumers and providers. However, using traditional standards such as UDDI and WSDL in this approach is insufficient, since the underlying semantics of Web services are not exploited enough. Also, the user requirements are not described explicitly and consistently.

The work described in [21] refers to the need for an extensible QoS model that contains domain-specific QoS criteria. It sustains that QoS must be represented to users according to user preferences and users should express accurately their preferences with this QoS model without resorting to complex coding of user profiles. It also suggests that QoS computation must be fair and open for providers and requesters. Then it proposes an extensible QoS model.

C. Agents and Web Services Discovery

Towards incorporating Web Service into agents has been the subject of many research projects. Reference [22] presents an agent based method for Web service selection, from the information that is given in the WSDL file by the Web service provider. Data mining is done on those data that are collected from WSDL files and feedback taken from the Web service users, by the agent to discover some interesting patterns for further users of the Web service. The author of this work doesn’t provide the details of the matching mechanism and his impact on service selection.

In [23], Zhang et al. propose a multiagent approach for a distributed information retrieval task. In their work, each agent has a view of its environment called agent view. The agent-view structure of an agent contains information about the language models of documents owned by each agent. An agent-view reorganization algorithm is run to dynamically reorganize the underlying agent-view topology. Zhang et al.’s protocol does not use ontologies during information retrieval.

III. AGENT-BASED FRAMEWORK FOR SERVICE DISCOVERY

Our proposed framework has two types of agents are devised namely, Consumer Agent and Provider Agent (Fig. 1). Like in [24], we use a central base of OWL Ontologies as a reference to develop the various local ontologies and semantic descriptions of the different Web services. Each Provider Agent implements a number of Web services described semantically with OWL-S enhanced with QoS attributes. The central ontologies base contains the different concepts used in diverse fields of proposed Web services and will be consulted periodically by Provider Agents to develop or enrich local ontologies.

When a service consumer wants to insert his request, An Ontology-Guided Interface is offered by the Consumer Agent. In order to input the request, service consumer must select the desired terms they want to use in his request from the list of terms provided by the interface in a pop-up. This list of terms is generated by Provider Agents using terms in local ontologies andsent to the Consumer Agent.

When receiving the request from Consumer Agent, each Provider Agent matches the request to the services in the Register Depository using OWL-S description and OWL

468
local ontologies. Then, Provider Agent returns to Consumer Agent a set of candidate services.

When receiving all responses from Provider Agents, the Consumer Agent sorts all candidate services according to the degree of functional similarity and the QoS score. Then, it returns to the service consumer the services that have the highest overall score.

IV. REQUEST AND WEB SERVICE DESCRIPTION

Request description includes functional and non-functional requirements. The former describes the functional characteristic of the service demand, such as name, textual description, inputs and outputs. The latter mainly focuses on the customer’s preferences, namely quality of service (QoS). In our work, service consumer doesn’t have to give the value of each desired QoS attribute but he should get the means to specify that a QoS attribute is more important than another one. Indeed, he assigns a weight for each QoS attribute. Weights range from 1 to 5 where higher weights represent greater importance. Fig. 2 shows an example of a user request interface with weights of some QoS attributes.

Typically, Web services are described using functional and non-functional properties. Functional properties contain Service Name, Textual description, a set of Inputs and a set of Outputs. Non-functional properties represent the description of the service characteristics (e.g. QoS). We use OWL-S service profile as a model for semantic matchmaking of service descriptions. However, OWL-S mainly focuses on describing functional aspects of a Web service. Based on works presented in [25] [26], we propose an ontology-based OWL-S extension to adding non-functional description, referred to as QoS, to Web service description. In OWL-S service profile we can use a set of ServiceParameter which has name “serviceParameterName” and a value “sParameter” (Fig. 3). For the connection of OWL-S and QoS ontology, the QoSProperty is a subclass of OWL-S ServiceParameter. And QoSParameterName and qosParameter are subproperties of OWL-S ServiceParmaeterName and sParameter property.

V. DEGREE OF SIMILARITY BETWEEN REQUEST AND WEB SERVICES

In this section we present our matching mechanism, which evaluates the similarity between request and Web service advertisements. This is done using functional match and the calculation of QoS score.

A. Functional Match

In this discovery step, name and textual description of request and services are matched using syntactic similarity function whereas inputs and outputs are matched based on conceptual similarity function.

1) Syntactic Similarity

Syntactic similarity function between a request R and a service S is calculated using the equation below:

\[ \text{SyntacticSim}(R, S) = \frac{\text{card}(L_1) \times \text{card}(L_2)}{\text{card}(L_1) + \text{card}(L_2)} \]

To calculate the syntactic similarity of tow strings, we use the algorithm 1, where the function Subset(String) takes as input a string and devies it into sequences of three characters. The function card(L) returns the number of elements in the set L. In line 4, we normalize the SyntacticSim value to have a value in the range of 0 to 1.

<table>
<thead>
<tr>
<th>Algorithm 1 SyntaxicSim(String1, String2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1, L2, L3: sets of string;</td>
</tr>
<tr>
<td>Begin</td>
</tr>
<tr>
<td>1: L1 = Subset(String1);</td>
</tr>
<tr>
<td>2: L2 = Subset(String2);</td>
</tr>
<tr>
<td>3: L3 = L1 \cap L2;</td>
</tr>
<tr>
<td>4: SyntaxicSim = \frac{2 \times \text{card}(L_1) \times \text{card}(L_2)}{\text{card}(L_1) + \text{card}(L_2)};</td>
</tr>
<tr>
<td>5: return SyntaxicSim;</td>
</tr>
<tr>
<td>End</td>
</tr>
</tbody>
</table>
For example, to calculate
SyntacticSim("FindAlgUniversity","FindAlgerianUniversity"):

L1={fin, ind, nda, dal, alg, lgu, gun, uni, niv, ive, ver, ers, rsi, sit, ity}.
L2={fin, ind, nda, dal, alg, lge, ger, eri, ria, ian, anu, nun, uni, niv, ive, ver, ers, rsi, sit, ity}.
L3={fin, ind, nda, dal, alg, uni, niv, ive, ver, ers, rsi, sit, ity}.
SyntacticSim = \frac{2 \times 13}{15+20} = 0.743.

2) Conceptual Similarity
An output in the request must not be considered as similar to a more generic output in the advertised service, while an input in the request must not be considered as similar to a more generic input in the advertised service. We think also that an input in the request could be considered as similar to a more generic output in the advertised service, while a service input must not be considered as similar to a more generic output in the advertised service. The number of inputs in the request must not be considered as similar to a more generic input in the advertised service, while an output in the advertised service could be considered as similar to a more generic output in the request.

ConceptSim(A, B) function matches a concept A(A ∈ request inputs or outputs) against a concept B (B ∈ service inputs or outputs) and returns the conceptual similarity of the two concepts. For illustration, let us take the example shown in Fig.5. All inputs and outputs refer to concepts of domain ontology, an example portion of which is shown in Fig.4 in a logic description notions. The function nbprop(A) denotes the number of properties of the concept A. To calculate the ConceptSim(A, B) function, we distinguish several scenarios:

Case 01: if(A and B are inputs and A is subclass of the concept B directly or indirectly) then
ConceptSim(A, B) = 1;
Example: Fig.5(a): ConceptSim(PhdStudent, Person) = 1.
Case 02: if(A and B are inputs and A is subclass of the concept B directly or indirectly) then
ConceptSim(A, B) = nbprop(B) - nbprop(A);
Example: Fig.5(a): ConceptSim(AlgUniversity, University) = 0.80.
Case 03: if(A and B are outputs and B is subclass of the concept A directly or indirectly) then
ConceptSim(A, B) = 1;
Example: Fig.5(b): ConceptSim(Person, PhdStudent) = 0.743.
Case 04: if(A and B are outputs and B is subclass of the concept A directly or indirectly) then
ConceptSim(A, B) = nbprop(B);
Example: Fig.5(b): ConceptSim(Person, AlgUniversity) = 0.60.
Case 05: if(A and B are inputs and B is subclass of the concept A directly or indirectly) then
ConceptSim(A, B) = nbprop(A);
Example: Fig.5(b): ConceptSim(Person, PhdStudent) = 0.5.
Case 06: if(A does not have a parent/child relationship with B, but both concepts have a parent concept C in common) then
ConceptSim(A, B) = \frac{nbprop(A∩B)}{nbprop(A∩B) + nbprop(A∩C)};
Example: ConceptSim(PhdStudent, Employer) = 0.5.
Case 07: if(otherwise) then ConceptSim(A, B) = 0;
Example: ConceptSim(Person, University) = 0.

Algorithm 2 performs an inputs matching. Where R.Inputs and S.Inputs denote the set of inputs in the request R and the set of inputs in the service S respectively, Card(E) denotes the cardinality of the set E, Sort(A) allows to sort the elements of the array A in descending order. In lines 1, 2, 3 and 4, the algorithm matches each request input against all Web service inputs, and keeps the best mapping for each request input. If the number of request inputs is less than the number of service inputs, then we have a miss of information; therefore InputsSim value is decreased (line 10).

The outputs similarity, given by OutputsSim function, is also calculated in the same way as inputs similarity. But when the number of service outputs is less than the number of request outputs, the value of OutputsSim is decreased. Therefore we inverse line 10 with 12 and perform changes in variable names in the algorithm 2.
Algorithm 1: InputsSim(R Inputs, S Inputs)

InSim: array of float; InputsSim = ConceptSim(PhdStudent, Person) = 1.

B. QoS Score

Let us calculate the Inputs and Outputs similarity between request and service shown in Fig5(a).

InputsSim = ConceptSim(PhdStudent, Person) = 1.

OutputsSim =

\[
\text{ConceptSim(Location, Location)} + \text{ConceptSim(AlgUniversity, University)}
\]

\[
= \frac{1+0.9}{2} = 0.9.
\]

InputsOutputs similarity is calculated based on InputsSim and OutputsSim functions as follows:

\[
\text{InputsOutputsSim}(R, S) = \frac{\text{InputsSim(R Inputs, S Inputs)} + \text{OutputsSim(R Outputs, S Outputs)}}{2}
\]

Functional similarity can be calculated using the equation below. Where weights w1 and w2 are real values between 0 and 1 and must sum to 1; they indicate the degree of confidence that the service consumer has in the syntactic similarity and inputs and outputs similarity.

\[
\text{FunctionalSim}(R, S) = w1 \times \text{SyntacticSim}(R, S) + w2 \times \text{InputsOutputsSim}(R, S)
\]

VI. SUMMARIES

Web service discovery is a hot topic in past a few years. Seeking the right service based on user’s search criteria in which the user may be interested is still a problem[28]. This paper has introduced a framework for Web services discovery based on intelligent software agents and ontologies. An advanced feature of our framework is that we perform the service discovery, selection and ranking based on the matching level of service advertisements to user requests both in terms of functionality and QoS. With the use of agents, information provided by Web services can be made more efficient and more dynamic. With the use of OWL-S and domain ontologies, match and discovery engine can return the most relevant services.

To implement our framework, we have used JADE platform [29]. JADE is a very powerful middleware framework built with Java to design a MultiAgent Systems based architecture. Consumer Agent and different Provider Agents are created with JADE and inherent “Agent” JADE class. Each agent has a “match module” which is realized using Jena APIs which provides plenty of methods to access ontology files. Registers Depository of our system are stored in Oracle 10g.

In our future research, we will have to incorporate the Web services composition into our proposed framework, by doing that, we hope that our based agent framework can be further improved and become more practical in real-world applications.

REFERENCES


