Improving Collaborations in Neuroscientist Community

Isabelle Mirbel

INRIA Sophia-Antipolis / Laboratoire I3S (CNRS)
2004 route des Lucioles - BP 93
F-06902 Sophia-Antipolis cedex, France
Isabelle.Mirbel@unice.fr

Pierre Crescenzo

Université de Nice Sophia-Antipolis / Laboratoire I3S (CNRS)
930 route des Colles - BP 145
F-06903 Sophia-Antipolis cedex, France
Pierre.Crescenzo@unice.fr

Abstract

*In this paper, we present a new approach, called SATIS (*Semantically AnnotaTed Intentions for Services*), relying on semantic web technologies and models, to assist collaboration among the members of a neurosciences community. The main expected result of this innovative work is to derive and share semantic web service specifications from neuroscientists requirements in order to perform image processing pipelines.*

1. Introduction

Computational neurosciences have for long demonstrated the power of computing techniques to analyse neurological data sets and study the brain functions [6]. The analysis of users' image processing pipelines shows many commonalities in data sets and processing chains. The manipulated data are mostly images completed with clinical information and additional annotations. Basic processing, as for instance intensity corrections or tissue classifications, are common to several image analysis pipelines. But each pipeline also contains pathology specific processing such as brain structure segmentations or image interpretations. In practice, data is spread over many independent data sources and there is no sharing of common basic processing units among the different processing chains. In this context, web services appear to be a privileged mean to assist the federation of distributed basic processing units into dedicated processing pipelines for each targeted application and to share common basic processing units inside a neuroscientist community.

To facilitate the exploitation of web resources (documents, actors or services), the semantic web research community aims at making explicit the knowledge contained into resources. This knowledge is represented by ontologies which structure terms, concepts and relationships of a given domain. Ontologies are often used to extract and represent the meaning of resources. This meaning is expressed through annotations supporting semantic resources indexing in order to explicit and formalise their content. Resource retrieval inside the community relies on the formal manipulation of these annotations and is guided by ontologies. As it is shown
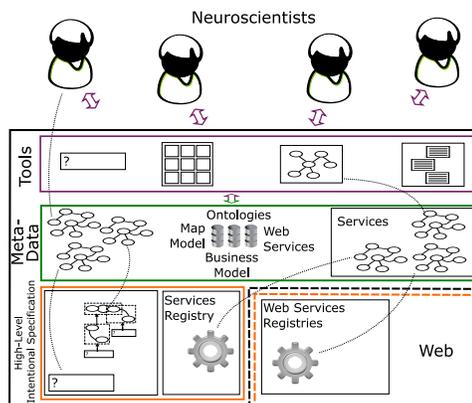

Figure 1. Neuroscientists community

in figure 1, neuroscientists are building image processing pipelines for their targeted application and therefore rely on web services (from their own registry or from a web registry). Web services are annotated by meta-data supporting their manipulation.

In this context, our focus is on how to derive semantic web services (i.e. basic processing units) specification from neuroscientists requirements in order to specify processing pipelines for targeted application. Therefore, we propose an approach to specify high-level domain-oriented activities with the help of an intentional model and to derive web services specification from this high-level description. As one of our aim is to improve collaboration and sharing inside the community, we also propose to consider high-level intentional specification of processing pipelines as resources of the community. Therefore, we provide means to annotate high-level intentional specification in order to assist their retrieval and sharing among the neuroscientists of the community.

Our work takes place in the web services domain and concentrates more specifically on service discovery and selection. It belongs to the family of goal-based service retrieval approaches. These approaches ([8], [9], [10], [1]) aim at specifying the goals which have to be satisfied by the retrieved services. In these proposals, different models are proposed to specify goals but none addresses the problem

of how to capture goals. They all consider that goals have already been identified and specified. On the contrary, our aim is to provide means to assist neuroscientists in querying the web services registry to find web services to operationalise a processing pipeline; and we are particularly interested in providing means to elicit and specify neuroscientists requirements in terms of services, upstream of the previously cited approaches. The GODO approach [4] also addresses this issue by proposing models and tools to capture user's goals with the help of an ontology or in natural language. What sets us apart from this approach is that we propose an incremental process to refine neuroscientists' requirements in order to specify the features required for the web services under retrieval, as it is also the case in [5]. Our approach distinguishes itself from [5] by the fact that we rely on semantic annotations and semantic web models and techniques to enrich the goal (or intention) specification, in order to provide reasoning and explanation capabilities. Beyond an alternative way to discover and retrieve web services, we also provide means to capitalise know-how about web service discovery and search processes themselves. Another novelty of our approach is to operationalise goals by rules in order to promote both mutualisation of high-level intentional specification and cross fertilisation of know-how about web services discovery and search processes among the community members.

The paper is organised as follows. In section 2, we discuss the different collaboration means provided in our approach. In section 3, we present the different steps of our refinement process allowing to derive web services specification (i.e. basic process units specification) from high-level intentional specification. Then, in section 4, we conclude and give some perspectives of our work.

2. Collaborations among neuroscience community members

When setting up a new procedure for medical image processing or when there is a need to change an existing procedure, neuroscientists may want to rely on existing web services to implement some of the image processings steps. By providing support to web service discovery and retrieval for non-computer scientist users, our main objective is to promote know-how sharing among community members. More specifically, we assist the know-how *transfer* from expert members to novice ones by providing means to incrementally specify high-level domain-oriented activities with an intentional modeling technique. Indeed, we provide means to populate a library of high-level intentions defined at different abstraction levels and allowing a novice user to start his/her web services discovery and search process at the level of specification s/he is comfortable with. S/he is then guided by the know-how previously entered by an expert member into the community memory to derive from the

initial requirements a set of web services (or basic process units) specification.

Beyond a way to discover and retrieve web services, our approach aims at providing means to promote mutualisation of high-level intentional specification and cross fertilisation of know-how about search processes among the community members. Indeed, our second objective is to assist the know-how *sharing* among expert members. Therefore, high-level incremental specifications of intentions are decomposed into fragments, highlighting the reusable dimension of high-level specifications out of the scope of the targeted application. We also propose an operationalisation of fragments by rules to take advantage of inference capabilities to discover alternative know-how to operationalise a web service discovery and search process.

Finally, in such communities, web services are provided by computer scientists and high-level intentional specification are specified by neuroscientists. In this context, our aim is to assist the derivation of web services specification from high-level intentional specification. In addition to assist know-how transfer between novices and experts and to share know-how between experts, our aim is to support *collaboration* between service designers (computer scientists) and service final users (neuroscientists). Indeed, our last objective is to provide both : (i) means for service final users to identify and specify their requirements and transmit them to service designers and (ii) means for service designers to disseminate information about available services.

By relying on a rule based specification to derive web services specification and by providing distinct and dedicated modeling techniques to both service designers and service final users as well as mapping mechanisms between them, we assist the bidirectional collaboration between neuroscientists and computer scientists.

To support the different collaboration means discussed above, we propose an approach based on:

- the map formalism [7] to identify and specify high-level intentional specification,
- W3C standards RDF, RDFS and SPARQL to provide means to define a common vocabulary, to annotate both web services (i.e. basic process units) and intentions, and to query the intention library as well as the service registry,
- CORESE [2] rules to store intention operationalisation in a reusable way in order both to promote cross fertilisation of know-how about web service discovery and retrieval process and to support dissemination of available services inside the community.

The approach is further detailed in the following section.

3. SATIS Approach

SATIS (*Semantically AnnotaTed Intentions for Services*) is the name of our new integrated approach. Its main aim

is to provide to neuroscientists which are not familiar with computer science, a complete solution to easily use a set of web services. Figure 2 describes the five phases of the SATIS approach, which are further discussed in the following.
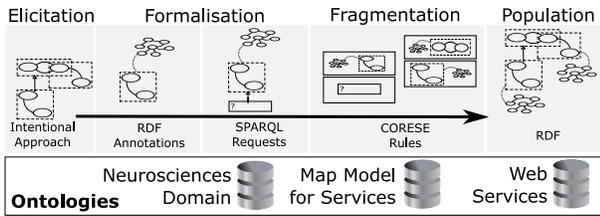

Figure 2. The SATIS approach

During all its five phases, our approach relies on common vocabulary and semantics. So, we reuse or define : (i) an ontology dedicated to the Neurosciences Domain, (ii) an ontology of the Map Model[3] (extended for web services) and (iii) an ontology for web services description themselves, as OWL-S for instance.

3.1. Elicitation phase

During the first step, *Elicitation phase*, final users (neuroscientists) define their image analysis pipeline by describing intermediate *intentions* (i.e. goals and subgoals to be satisfied through the processing chain) and *strategies* (i.e. means to reach goals). In order to give a convenient way to express these intentions, we have chosen the graphical Map Model formalism [7].

Figure 3 gives an example of an intentional map. We can see two main intentions specified by the neuroscientists doctor: *Scan Patient* and *Reduce Noise in Images*. The second one is optional in the approach since a cancellation or a direct printing could be selected. Between the intentions, we discover strategies such as *by hospital scanner* or *by image transformation*. Strategies define the way to pass from an intention to a next one. There can be many strategies which link up the same intentions like *by cancellation...* and *by printing* from *Scan Patient* to *Stop*. Strategies may also carry information related to constraints about the kind of web services under consideration, like access rights for instance. Indeed, in a map, each set which is made up by a source intention, a strategy and a target intention is a *section* of the map. Let's precise that an intentional map is neither a state diagram, because there is no data structure, no object, and no assigned value, nor an activity diagram, because there is always a strong context for each section of the map: its source intention and its strategy. We can attach more information to this kind of schema (in order to help the user of the map to choose the adequate strategy, for example), but this is not the goal of this paper to fully describe the Map Model.
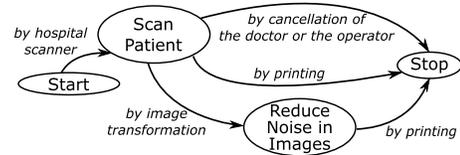

Figure 3. An Intentional Map

An intentional map may be refined. We can refine a section by giving a new map which describes how to reach a target intention in a more detailed way (by using more specific and low-level intentions and strategies). Indeed, by refining an intentional map, experts describe internal procedures to reach the target intention. When this elicitation phase is successfully completed, we have at disposal a general map and one or several level of refinements of its sections.

So we build, during the Elicitation phase, a high-level definition of the image analysis pipeline under consideration, in a very simple formalism for novice users.

3.2. Formalisation phase

The second phase, *Formalisation phase*, consists in generating and/or writing (this is designed to be a semi-automatic transformation) a SPARQL query in order to operationalise each section by an adequate web service or set of web service specifications. Indeed, the SPARQL query aims at retrieving the description of a web service or a set of web services supporting the fulfilment of the target intention of the section the query is associated to.

RDF annotations are associated to each sections of the maps in order to formalize their content (source intention, strategy, target intention, ...). Additional information is also integrated to the annotations, as for instance information about the authors of the map. This knowledge, for instance, aims at distinguishing the responsibilities of the system (chaining rules and proposing web services descriptions to operationalise image processing pipelines) from those of the neuroscientists authoring image processing pipelines.

3.3. Fragmentation phase

In the third step, *Fragmentation phase*, we transform all specifications captured during the Formalisation phase in a set of CORESE [2] rules. This technology choice has three great advantages. First, CORESE is an RDF engine based on Conceptual Graphs. It enables the processing of RDFS, OWL-Lite and RDF statements relying on a Conceptual Graphs formalism. It performs SPARQL Queries and run rules over RDF graphs. The second benefit is that, during this transformation into rules, the original intentions and strategies are naturally modularised and this fact far

improves the reusability of the concerned discovery process fragment. Last but not least, CORESE provides a forward chaining engine and we take advantage of it in order to offer knowledge inference.

3.4. Population phase

The last step of SATIS, *Population phase*, consists in deriving web services semantic specifications to operationalise a set of intentions and strategies associated to an image processing pipeline. We rely on the CORESE semantic engine for both i) forward chaining on the knowledge base of SPARQL queries (expressed by rules) and ii) matching with the knowledge base of RDF annotations describing available web services. The knowledge base only stores the queries, and not the maps. These are dynamically created when needed all along the forward chaining process, as temporarily subgoals, until web services annotations are found to match all the subgoals and therefore the general goal section. As a result, a neuroscientist looking for a discovery and search process to implement an image processing pipeline will take advantage of all the rules and all the web service annotations stored in the community memory at the time of his/her search. This memory may evolve over the time and therefore the web services retrieved by applying a rule may vary as well. In other words, the instantiation of a full web services discovery and search process is done at runtime and depends on the web services available in the memory.

Let's clarify that the result is composed by specifications of candidate web services, and not by web services themselves. The invocation of the selected (among the candidates) web services is done dynamically and is out of the scope of this work.

4. Conclusion and Perspectives

In this paper, we presented an approach, relying on semantic web technologies and models, to assist collaboration among the members of a neurosciences community. Our main objectives are more precisely to assist the know-how transfer from expert members to novice ones, to promote cross fertilisation of know-how among community members and to support collaboration between computer scientists and neuroscientists. Therefore, starting from an intention based requirement elicitation, we proposed to derive RDF schemas to formalise neuroscientists requirements and CORESE rules to operationalise them through SPARQL queries. As a result, the instantiation of web services discovery and search processes is supported by forward chaining among the rules base and matching with the RDF data set annotating the web services. Beyond a mix between existing intentional requirement modeling techniques and web semantic models and techniques, our main contribution consists in: (i) providing reasoning and query capabilities for intentional requirement modeling, (ii) leveraging domain knowledge from computer sciences related aspects to neurosciences related aspects in an integrated way and (iii) providing means to support contextualised web services retrieval.

Future works will focus on composition and evolution concerns by providing dedicated operators in each step of our approach, in addition to mappings from one step to the other. We also plan to exploit traceability and reasoning capabilities of CORESE engine to improve responsibility considerations handling during the web services discovery and retrieval process. Last, integration with Hospital Information Systems will also be studied and practical work will consist in testing our approach through several case studies.

References

[1] L.O. Bonino da Silva Santos, L. Ferreira Pires, M.J. van Sinderen, *A Goal-Based Framework for Dynamic Service Discovery and Composition*. International Workshop on Architectures, Concepts and Technologies for Service Oriented Computing, July 2008, Porto, Portugal. pp. 67-78.

[2] O. Corby, R. Dieng-Kuntz, C. Faron-Zucker, *Querying the semantic web with the CORESE search engine*. 16th European Conference on Artificial Intelligence (ECAI/PAIS). pp. 705-709. Valencia, Spain. 2008.

[3] O. Corby, C. Faron-Zucker, I. Mirbel, *Implementation of Intention-Driven Search Processes by SPARQL Queries*, International Conference on Enterprise Information Systems, Milan, Italy. 2009.

[4] J.M. Gomez, M. Rico, F. Garcia-Sanchez, *GODO: Goal Oriented Discovery for Semantic Web Services*. 5th International Semantic Web Conference. 2006.

[5] R.S. Kaabi, *Une approche méthodologique pour la modélisation intentionnelle de services et leur opérationnalisation*. PhD Thesis, Université Paris I Sorbonne. February 2007.

[6] J. Montagnat, A. Gaignard, D. Lingrand, J. Rojas Balderrama, P. Collet, P. Lahire, *NeuroLOG: a community-driven middleware design*. HealthGrid'08, IOS Press, Chicago, 2-4 June 2008.

[7] C. Rolland, *Conceptual Modelling in Information Systems Engineering*. Springer-Verlag, 2007.

[8] M. Stollberg, B. Norton *A Refined Goal Model for Semantic Web Services*. Second International Conference on Internet and Web Applications and Services (ICIW'07), 2007.

[9] M. Vukovic, P. Robinson *GoalMorph: Partial Goal Satisfaction for Flexible Service Composition*. International Conference on Next Generation Web Services Practices. 2005.

[10] K. Zhang, Q. Li, Q. Sui *A Goal-driven Approach of Service Composition for Pervasive Computing*. 1st International Symposium on Pervasive Computing and Applications, pp. 593-598, August 2006.