INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team EXMO

Computer-mediated communication of structured knowledge

Rhône-Alpes

THEME SYM

Activity Report

2006

Table of contents

1. Team

Head of project-team
Jérôme Euzenat [Research scientist (DR INRIA), HdR]

Administrative assistant
Stéphanie Berger [25% (IA INRIA)]

Research scientist
Jean-François Baget [CR INRIA]

Post doctoral fellow
Jason Jung [INRIA, until September 1st, 2006]
Nikolay Pelov [INRIA, from November 1st to November 30th, 2006]
Seungkeun Lee [AVC, since July 1st, 2006]

Ph. D. students
Faisal Al-Khateeb [Jordan government]
Sébastien Laborie [MESR grant]
Jérôme Pierson [CIFRE (France Telecom R&D)]
Antoine Zimmermann [INRIA]

2. Overall Objectives

Keywords: *Alignment API*, *OLA*, *OWL*, *PSPARQL*, *RDF*, *RDF Path*, *XSLT*, *content representation*, *context*, *knowledge representation*, *knowledge transformation*, *multimedia document adaptation*, *ontologies*, *ontology alignment*, *property preservation*, *semantic web*, *semantics of knowledge representation*, *semiotics*.

We assume that expressing formalised knowledge on a computer is useful, not especially for the need of the computer, but for those of communication. In future information systems, formalised knowledge will be massively exchanged. *Exmo*'s goal is the development of theoretical and software tools for enabling interoperability in formalised knowledge exchange. *Exmo* contributes to an emerging field called the semantic web which blends the communication capabilities of the web with knowledge representation.

However, there is no reason why this knowledge should be expressed in the same format or by reference to the same vocabulary (or ontology). In order to interoperate, these representations will have to be matched and transformed. Moreover, in the communication process computers can add value to their memory and medium role by formatting, filtering, classifying, consistency checking or generalising knowledge.

We currently build on our experience of the alignment structure as representing the relationships between two ontologies on the semantic web. Ontology alignments express the correspondences between the entities in two ontologies. They allow maximising sharing on the semantic web: various algorithms can produce alignments and various uses can be made of these alignments. Such alignments can be used for generating knowledge transformation (or any other kind of mediators) that will be used for interoperating. In order to guarantee properties of these transformations, we can consider the properties of alignments and generate transformations preserving them.

Our current roadmap focusses on the design of an alignment infrastructure and on the investigation of alignment properties (and especially semantic propertied) when they are used for reconciling ontologies.

On a longer term, we want to explore "semiotic" properties, i.e., properties which concern the interpretation of the communicated representation by a human user. This goal should require an analysis of the extra-semantic rules that govern the choice of subsets of models.

Anticipated applications are in transformation system engineering (in which the information system is seen as a transformation flow) and semantic web infrastructure.

3. Scientific Foundations

3.1. Knowledge representation semantics

We usually work with semantically defined knowledge representation languages (like description logics [27], conceptual graphs and object-based languages). Their semantics is usually defined within model theory initially developed for logics. The languages dedicated to the semantic web (RDF and OWL) follow that approach.

We consider a language L as a set of syntactically defined expressions (often inductively defined by applying constructors over other expressions). A representation (o) is a set of such expressions. It is also called an ontology. An interpretation function (I) is inductively defined over the structure of the language to a structure called interpretation domain (D). This expresses the construction of the "meaning" of an expression in function of its components. A formula is satisfied by an interpretation if it fulfills a condition (in general being interpreted over a particular subset of the domain). A model of a set of expressions is an interpretation satisfying all these expresions. An expression (δ) is then a consequence of a set of expressions (o) if it is satisfied by all of its models (noted $o \models \delta$).

A computer must determine if a particular expression (taken as a query, for instance) is the consequence of a set of axioms (a knowledge base). For that purpose, it uses programs, called provers, that can be based on the processing of a set of inference rules, on the construction of models or on more classical programming. These programs are able to deduce theorems (noted $o \vdash \delta$). They are said to be sound if they only find theorems which are indeed consequences and to be complete if they find all the consequences as theorems. However, depending on the language and its semantics, the decidability (i.e., the ability to create sound and complete provers) is not warranted. Even for decidable languages, the algorithmic complexity of provers can prohibit their exploitation.

To solve this problem a trade-off between the expressivity of the language and the complexity of its provers has to be found. These considerations have led to the definition of languages with limited complexity - like conceptual graphs and object-based representations - or of modular families of languages with associated modular prover algorithms - like description logics.

Exmo mainly considers languages with well-defined semantics (such as RDF and OWL that we contributed to define), and defines the semantics of some languages such as multimedia specification languages, in order to establish the properties of computer manipulations of the representations.

3.2. Ontology alignment

When different representations are used, it is necessary to identify their correspondences. This task is called ontology matching and its result is an alignment. It can be described as follows: given two ontologies, each describing a set of discrete entities (which can be classes, properties, rules, predicates, etc.), find the relationships (e.g., equivalence or subsumption) that hold between these entities.

An alignment between two ontologies o and o' is a set of correspondences $\langle e, e', r, n\rangle$ in which

- e and e' are the entities between which a relation is asserted by the correspondence (e.g., formulas, terms, classes, individuals);
- r is the relation asserted to hold between e and e'. This relation can be any relation applying to these entities (e.g., equivalence, subsumption).
- n is a degree of confidence in this particular correspondence (which will be omitted here).

Given the semantics of the two ontologies provided by their consequence relation, we define an interpretation of the two ontologies as a triple made of an interpretation for each ontology and an equalising function which maps the domain of each of the models to a common domain on which the relations are interpreted. Such a triple $\langle m, m', \gamma\rangle$ is a model of the aligned ontologies o and o' if and only if, for each correspondence $\langle e, e', r\rangle \in A$ of the alignment, then $m \models o$, $m' \models o'$ and $\langle \gamma(m(e)), \gamma(m'(e'))\rangle \in r^{\gamma}$.

This definition is extended to a distributed system which is a set of ontologies and associated alignments, such that a model is a tuple of local models and an equalising function such as each alignment is valid for the models and the equalising function involved in the tuple. In such a system, alignments play the role of model filters which will select the local models which are compatible with all alignments.

So, given a distributed knowledge system, it is possible to interpret it. However, given a set of ontologies, it is necessary to find the alignments between them and the semantics does not tell which one they are. Ontology matching aims at finding these alignments. A variety of methods is used for this task. They perform pair-wise comparisons of entities from each of the ontologies and select the most similar pairs. Most matching algorithms provide correspondences between named entities, more rarely between compound terms. The relationships are generally equivalence between these entities. Some systems are able to provide subsumption relations as well as other relations in the support language (like incompatibility or instanciation). Confidence measures are usually given a value between 0 and 1 and are used for expressing preferences between two correspondences.

3.3. Transformations and properties

A transformation (τ) is an algorithmic manner to generate a representation ($\tau(o)$) from another one (o), not necessarily in the same language. We focus on transformations made by composition of more elementary transformations for which we only know the input, output and assumed properties. These transformations may have been generated from an alignment or by any other means.

A transformation system is characterised by a set of elementary transformations and a set of composition operators. A transformation flow is the composition of elementary transformation instances whose input/output are connected by channels. A transformation flow is itself a transformation.

The design of information systems like transformation flows requires the ability to express such flows and to determine their properties. A property is a boolean predicate about the transformation (e.g., "preserving information" is such a predicate - it is true or false of a transformation - which is satisfied if there exists an algorithmic means to recover o from $\tau(o)$).

We consider more closely preservation properties that can allow the preservation (or anti-preservation) of an order relation between the source representations and the target representations. For instance, one can identify:

- Syntactic properties: based on the organisation of syntactic elements, like the completion ($\tau(o) \leq o$, in which $\leq$ denotes structural subsumption between representations);
- Semantic properties: based on the concepts of model and consequence, like consequence preservation ($\forall o \subseteq L, \forall \delta \in L, o \models_L \delta \Rightarrow \tau(o) \models_{L'} \tau(\delta)$);
- Semiotic properties: based on the interpretation of the manipulated objects as signs, like interpretation preservation (let σ be the interpretation rules and $\models_{L,i}$ be the interpretation of individual i, $\forall o \subseteq L, \forall \delta \in L, \forall i, j, o, \sigma \models_{L,i} \delta \Rightarrow \tau(o), \tau(\sigma) \models_{L',j} \tau(\delta)$).

Our goal is to study transformations based on transformation properties rather than on representations or transformation structures. This does not deal only with semantics but considers various properties (e.g., content or structure preservation, traceability, and confidentiality). However, we more specifically address semantic properties. We also consider properties of transformation systems (given a transformation system, is information preservation decidable and at what cost?). We try to characterise, given a particular type of property, which transformations leave them invariant and what is the action of composition operators.

4. Application Domains

Two application contexts motivate and spur our work: the "semantic web" infrastructure (§4.1) and transformation system engineering (§4.2).

4.1. Semantic web technologies

Internet technologies support organisations in accessing and sharing knowledge, often difficult to access in a documentary form. However, the technologies quickly reach their limits: web site organisation is expensive and full-text search inefficient. Content-based information search is becoming a necessity. Content representation will enable computers to manipulate knowledge on a more formal ground and to carry out similarity or generality search. Knowledge representation formalisms are good candidates for expressing the content.

The vision of a "semantic web" [29] supplies the web, as we know it (informal) with annotations expressed in a machine-processible form and linked together. In the context where web documents are formally annotated, it becomes necessary to import and manipulate the annotations according to their semantics and their use. Taking advantage of this semantic web will require the manipulation of various knowledge representation formats. *Exmo* concerns are thus central to the semantic web implementation. Our work aims at enhancing content understanding, including the intelligibility of communicated knowledge and formal knowledge transformations.

The semantic web idea is essentially based on the notion of ontology (that can be quickly described as conceptual schemes or knowledge bases). Even if a standard knowledge representation language emerges, it will still be necessary to import and exchange ontologies in such a way that the semantics of their representation language is taken care of. We work on finding the correspondences between various knowledge representation languages and ontologies (see §6.1) in order to take advantage of them in ontology merging and bridging or message translation. Bringing solutions to this problem is part of the ambition of *Exmo*.

4.2. Transformation system engineering

Computerisation and networking lead organisations to exchange information in machine-readable form. E-commerce generates a continuous flow of such documents. As transmitted information is neither addressed nor adapted to all the members of the organisation, it is necessary to transform document structure and content. Similarly, web sites are generated from databases or primary funds and e-commerce documents are applied various transformations before goods are shipped. Additionally, the Object Management Group Model-Driven Architecture (MDA) considers the future of software development as a composition of transformations from (platform independent) domain models to other (platform dependent) models in function of platform description models. This is considering any implementation as adaptation.

Interoperability requirements have led to the definition of the structured document representation language XML which helps handling the syntax of documents straightforwardly. Other languages such as XSLT or Omnimark, enable the implementation of standalone transformations.

However, this view of transformations is only partial and local. It seems unavoidable that, in the future, we will have to deal with complex transformation flows automating the combination of transformations, some of which coming from external sources. This will require the global understanding of the behaviour of the flow of transformations. This calls for real "transformation system engineering" which should address the following issues:

- the lack of global consideration of transformations: they are processed in relation with other transformations;
- the need to consider the properties of transformations and especially their semantic properties: this will require the semantic analysis of the transformations;
- the design of transformation flows from external resources (as it is in software engineering): this will require the ability to consider the properties of imported transformations.

Transformation system engineering will require tools, methodologies and formal methods. As a matter of fact, it will be necessary to check that a particular transformation system does not export sensitive information or that the transformation process terminates. For that purpose, the transformation flow must be expressed in a parsable way and the expected properties of the flow must be expressed (see §3.3). *Exmo* is concerned by tools and formal methods and aims at combining them in solutions for transformation flow design environments.

5. Software

Exmo's work can be implemented in software: we have proposed an API for expressing ontology alignment (§5.1) and we have designed and developed a query evaluator for the PSPARQL query language (§5.2).

5.1. Alignment API: manipulating ontology alignments

Participants: Jérôme Euzenat [Contact], Seungkeun Lee, Jérôme Pierson, Antoine Zimmermann.

We have designed a format for expressing alignments in a uniform way [1]. The goal of this format is to be able to share on the web the available alignments. It should help systems using alignments (e.g., mergers, translators) to take advantage of any alignment algorithm and it will help alignment algorithms to be used in many different tasks. This format is expressed in RDF, so it is freely extensible, and has been defined by a DTD (for RDF/XML), an OWL ontology and an RDF Schema. Aligned entities are identified by their URIs.

The API itself is a *Java* description of tools for accessing the common format. It defines four main interfaces (Alignment, Cell, Relation and Evaluator) and proposes the following services:

- Storing, finding, and sharing alignments;
- Piping alignment algorithms (improving an existing alignment);
- Manipulating (thresholding and hardening);
- Generating processing output (transformations, axioms, rules);
- Comparing alignments.

The API can be used for producing transformations, rules or bridge axioms independently from the algorithm which produced the alignment. The proposed API implementation is based on the OWL-API and offers floating point measures between 0 and 1. It features:

- a base implementation of the interfaces with all useful facilities;
- a library of sample aligners;
- a library of renderers (XSLT, SWRL, OWL, C-OWL, SEKT mapping language);
- a library of evaluators (precision/recall, generalised precision/recall, precision/recall graphs and weighted Hamming distance);
- a parser for the format.

To instanciate the API, it is sufficient to refine the base implementation by implementing the `align()` method. Doing so, the new implementation will benefit from all the services already implemented in the base implementation.

We are developing the *Alignment API* as an Alignment server that can be used by remote clients. It will be added web service and agent communication capabilities in order to be used as an alignment service on the semantic web.

The *Alignment API* is used in the Ontology Alignment Evaluation Initiative data and result processing. It is also used as input or output by a number of alignment tools (among which OLA that we develop jointly with the University of Montréal).

The *Alignment API* is freely available since december 2003 under the LGPL licence at http://alignapi.gforge.inria.fr

5.2. PSPARQL Query evaluator

Participants: Faisal Al-Khateeb [Contact], Jean-François Baget, Jérôme Euzenat.

PSPARQL is a query language for RDF (see §6.2.1) that we designed by extending SPARQL. We have implemented a PSPARQL query evaluator in Java. This evaluator can parse SPARQL and PSPARQL queries, parse RDF documents written in Turtle language, evaluate the query and then return the answer set.

The algorithm follows the backtrack technique developed in our work. The evaluation of regular expression patterns is used for calculating the satisfiability set of a given regular expression, to take into account the multiple appearances of a given blank in different places of the query, i.e., to take into account the current mappings.

This evaluator passed 430 test cases out of the 440 in the W3C Data Access Working Group SPARQL test base. The missed tests are the 10 tests containing the GRAPH keyword which is not yet implemented.

The evaluator is currently being thoroughly tested for performances and practical hard problem detection.

This query evaluator is currently available as a servlet under http://psparql.inrialpes.fr. It will be made available to other parties later.

6. New Results

The results in 2006 are mainly related to the semantic of alignments (§6.1), the development of a PSPARQL interpreter (§6.2), the argumentation about alignments (§6.3.2) and the adaptation of the interactive structure in multimedia documents (§6.4).

6.1. Ontology alignment

Pursuing the work on alignment and ontology matching that is done in the framework of the Knowledge web network of excellence, we develop independent contributions about the semantics of alignment.

6.1.1. Benchmarking

Participant: Jérôme Euzenat [Contact].

In order to evaluate the performance of ontology matching algorithms it is necessary to confront them with test ontologies and to compare the results. Since three years, we run the Ontology Alignment Evaluation Initiative which organises evaluation campaigns for assessing the degree of achievement of actual ontology matching algorithms. This year's event has been held in Athens, Georgia, USA [5][13]. Ten different teams entered the evaluation which consisted in 6 different sets of tests. The participation in this effort is thus increasing both in the number of participants and challenges.

The prominent evaluation criteria are precision and recall originating from information retrieval. We have pursued our investigations on generalising precision and recall by using the measures designed last year in the current campaign (and demonstrating that they have the expected behaviour) and designing new measures based on the new semantics of alignments (see §6.1.2).

We also have attempted to go beyond these raw evaluation results and try to provide guidance to users with a specific ontology matching problem to solve. This amounts to characterising matching problems and matching solutions over specific dimensions and requirements (e.g., the kind of input or the properties expected from the output) and applying decision support procedures in order to find the most adapted solutions [23][26].

6.1.2. Semantics of alignment and distributed systems

Participants: Antoine Zimmermann [Contact], Jérôme Euzenat.

When dealing with alignments, it is important, both for generating them and for using them to know their interpretation. This is even more important when users are dealing with a whole network of ontologies related by alignments. Such a structure composed of a set of ontologies, interconnected with ontology alignments is called a distributed system. A legitimate question is: given the semantics of these ontologies, what are the consequences of a distributed system.

So far, alignments have been given semantics only related to a precise logical framework (e.g., first-order logic). In the continuation of our work on categorical definition of alignment [24], we aimed at an alignment semantics independent from the ontology semantics.

For that purpose, we have defined a parameterised family of model-theoretic semantics for alignments and knowledge-based distributed systems. This semantics is parameterised by the interpretation of the set of relation it uses and rely transparently on the semantics of the ontologies (which is only supposed to define the consequence relation). This means that the models (i.e., satisfying interpretations) of an alignment or a distributed system are defined in function of the models of the local ontologies, even when different ontologies are written in different languages.

We have investigated three different variations of this semantics, offering different levels of integration and supporting different paradigms. The three types of semantics use different techniques to obtain commensurate interpretation of formulas: either by constraining interpretations on a common domain, mapping the domains to a common domain or relating the entities of each pairs of domains. We studied the semantic properties of ontology alignment composition according to these three variants [25]. It appears that only the first two types of distributed semantics are sound with respect to alignment composition, while the last one, which corresponds to the paradigm of Distributed First Order Logics (DFOL), Distributed Description Logics (DDL) and C-OWL, is not.

This work is now used in designing ontology modules in cooperation with Frederico Freitas (see §8.3.2) and in the framework of the NeOn project (see §8.2.2). It is also used for defining precisely semantic extensions of precision and recall.

6.2. Algorithms for graph-based knowledge representation languages

We are interested in various knowledge representation languages whose formulas can be represented by graphs and where the truth value of a formula in a model-theoretic interpretation can be expressed by some kind of graph homomorphism. Such languages naturally encompass conceptual graphs or RDF, but also propositional logics, constraint networks or positive Datalog.

RDF (Resource Description Framework) is a knowledge representation language dedicated to describe the meta-data about resources on the web recommended by the World Wide Web. This year we especially made progress in extending the SPARQL query language for RDF.

6.2.1. Path RDF as a query language for RDF

Participants: Faisal Al-Khateeb [Contact], Jean-François Baget, Jérôme Euzenat.

Querying RDF graphs can be reduced to computing entailment. Entailment or semantic consequence in RDF (i.e., determining whether or not a query which is an RDF graph is a logical consequence of the RDF graphs representing the database or semantic web) can be computed by a sort of homomorphism of directed labeled graphs, known as conceptual graph projection. Another way to query RDF graphs, which has been used in graph databases, is to query for path expressed by regular expressions holding between nodes (the former allows for full graph branching and cycling as queries, the latter allows for indetermined lengths of paths).

However, some queries that can be expressed in one approach cannot be expressed in the other: a query whose homomorphic image in the database is not a path cannot be expressed by a regular expression, while RDF semantics is not meant to express paths of unknown length. The two kinds of queries do not identify the same set of queries.

To benefit from both approaches, we have defined PRDF, for Path RDF, an extension of RDF that allows regular expressions over relations as labels to the arcs of RDF graphs. This expresses that there is a path combining relations in a particular way between two resources.

We extended the RDF triples syntax by allowing non-atomic regular expressions (i.e., regular expressions generated inductively over URIrefs) to occur in the place of the predicate. We extended the semantics to make two resources of an interpretation support a regular expression in the interpretation. Then we gave a sound and complete inference mechanism for computing the PRDF graph projection over RDF graphs.

We have defined a language (called PSPARQL) for querying RDF documents. PSPARQL extends SPARQL, the query language for RDF currently designed by the W3C, by replacing the basic graph patterns of SPARQL, which are RDF graphs with variables, by PRDF graph patterns. For instance, one can query if there are transportation means by any combination of bus or train between two cities. We have defined the syntax and semantics of PSPARQL and shown how to use PRDF projection for answering PSPARQL queries. We provide the necessary algorithms for computing the answer set to a given PSPARQL query and have implemented them (see §5.2).

6.3. Dynamic aspects of alignments

We apply the theoretical results obtained on alignments in various contexts where semantic web technologies and alignements are useful.

6.3.1. Context management in pervasive computing

Participants: Jérôme Pierson [Contact], Seungkeun Lee, Jérôme Euzenat.

In a pervasive computing environment, the environment itself is the interface between services and users. Using context information coming from sensors, location technologies and aggregation services, applications adapt their run-time behaviour to the context in which users evolve (e.g., physical location, social or hierarchical position, current tasks as well as related information). These applications have to deal with the dynamic integration in the environment of new elements (users or devices), and the environment has to provide context information to newly designed applications. We study and develop a dynamic context management system for pervasive application. It must be flexible enough to be used by heterogeneous applications and to run dynamically with new incoming devices.

We have designed an architecture in which context information is distributed in the environment. Each device or service implements a context manager component in charge of maintaining its local context. It can communicate with other context manager components: some of them are context information producers, some of them are context information consumers and some of them are both. We have defined a simple protocol to allow a consumer to identify and determine the producer for the information it needs. Context manager components express their context information using an OWL ontology, and exchange RDF triples with each other. The openness of ontology description languages makes possible the extension of context descriptions and ontology matching helps dealing with independently developed ontologies. Thus, this architecture allows the introduction of new components and new applications without interrupting what is working [10][11][12].

In 2006, we have developed a library with several interfaces to build the distributed context management system and we provide an implementation of those interfaces using the multi-agent system called JADE. This implementation provides an agent managing (researching, diffusing and updating) context information. This agent has to be integrated in the services which will be integrated in a dynamic architecture. Developers will have to describe the functionalities of their services, sensors or applications, i.e., to provide an OWL class description of information which they can provide and/or information that they seek.

The next step will be to develop and use an alignment service (in particular, the JADE side of the service). This service will help the context information manager component to find the correspondences between various ontologies with which they are confronted and thus to match application needs in terms of context information with the information provided by the other devices.

Our future works will focus on using this implementation in a real home environment and to test it with real applications and real users.

This work is developed in collaboration with France Telecom R&D and more specifically Fano Ramparany.

6.3.2. Argumentation over ontology alignments

Participants: Jérôme Euzenat [Contact].

When two independently developed agents want to interact they may not share the same ontologies. In order to reconcile their ontologies, they can take advantage of an alignment service which will provide alignments for the two ontologies. But if it is not possible to obtain an alignment that suits both parties, it is necessary for these parties, if they want to interact, to negotiate the meaning of terms, or, more modestly, to negotiate the correspondences in alignments. For that purpose, we have introduced a novel argumentation framework for arguing for and against correspondences found in alignments [20][21][22]. This framework is based on previous work on argumentation in multi-agent systems, and especially value-based argumentation, but adapts it to the specific case of arguing about alignments and correspondences. It provides a first typology of arguments that can be applied to correspondences between ontology entities (based on the way the correspondences have been obtained). A preference relation among arguments can be defined with regard to this typology. This relation can be different from agent to agent so that they do not all prefer the same arguments. We have used classical multi-agent argumentation theory in order to characterise what is an acceptable argument for an agent as well as the prefered extensions (of a set of arguments) for a set of agents having different preference relations. We also designed an argumentation protocol for reaching these preferred extensions. We provide strategies for evaluating arguments during the unfolding of the negotiation dialogue.

This work is developed in collaboration with the Computer Science Departement of the University of Liverpool and more specifically, Loredana Laera, Valentina Tamma and Trevor Bench-Capon.

6.3.3. Query mediation in peer-to-peer systems

Participants: Jérôme Euzenat [Contact], Arun Sharma, Jason Jung.

We have designed an experimental picture annotation and sharing system based on semantic web technologies. The main characteristics of this system is that pictures are annotated with ontologies that can be modified by the people annotating pictures, e.g., by adding a class that is missing. In consequence, people have annotations based on different ontologies that must be reconciled if they want to share annotations or to query the pictures of each other. Alignments between these local ontologies must be provided, before sharing picture annotations.

This system has been implemented and experimented by 7 people on a common corpus of 50 pictures. Each participant was free to use whatever ontology for annotating these pictures. They have been asked to align their ontology with that of another participant. We have used these results in order to compare basic similarity measures among ontologies. We have also used this experiment as a basis for our work on semantic social network analysis [14][15].

This work has so far remained preliminary, but it has shown us that this approach is practicable.

6.4. Semantic adaptation of multimedia documents

Participants: Sébastien Laborie [Contact], Jérôme Euzenat, Nabil Layaïda [WAM team], Jean-François Baget.

When a multimedia document is played on platforms with limited resources (e.g., a mobile phone that can only display one image at once or an interactive display without keyboard), it is necessary to adapt the document to the target device. In order to assess the meaning of adaptation, we have defined a semantic approach [3], which considers a model of a multimedia document as one of its potential execution (an execution satisfying its specification). In a first approximation, adaptation reduces the set of models of a specification by selecting those satisfying the adaptation constraints.

Adapting amounts to finding this subset of models or, when it is empty, finding a compatible execution as close as possible to the initial execution. For that purpose, we proposed to express the set of possible interpretations by a resolved relation graph. Each relation of this graph could be a temporal, spatial, or spatio-temporal relation. This approach has been applied to the temporal and spatial dimensions based on Allen and RCC algebras respectively.

This year, we have followed on our work on semantic adaptation for the SMIL 2.0 language [16].

On one hand, we have shown that our adaptation framework is generic. We had already shown that it was possible to adapt SMIL documents in the temporal, spatial [18] and spatio-temporal [17] dimensions. This year, we have extended the approach to the hypermedia dimension and mixed it with the temporal one [19]. Moreover, we have completed preliminary work to include the logical dimension, i.e., group together some objects in one single element. Thus, our approach uses all multimedia dimensions.

On the other hand, we have implemented our approach in an interactive adaptation system for SMIL documents. Moreover, we have studied the computation time of adaptating solutions. We want to increase the efficiency of this search in order to adapt real SMIL documents, i.e., documents containing many multimedia objects. For this purpose, we have implemented several basic optimisation schemes such as Branch and Bound, Backmarking, Forward Checking and Backjumping. This search can also be applied on non-complete graphs, i.e., without explicit relation between some multimedia objects.

7. Contracts and Grants with Industry

7.1. France Telecom R&D

Participants: Jérôme Pierson [Contact], Jérôme Euzenat.

The work of Jérôme Pierson is developed in collaboration with the Centre Norbert Segard of France Telecom R&D. Jérôme Pierson's thesis is jointly supervised by Fano Ramparany and Jérôme Euzenat under a CIFRE contract. The topic of this thesis is to investigate the notion of context in ambient computing and the ability to match ontologies when shifting context (see §6.3.1).

8. Other Grants and Activities

8.1. National grants and collaborations

8.1.1. WebContent RNTL platform

Participant: Jérôme Euzenat [Contact].

Exmo is involved in the *WebContent* platform development subsidised by the Agence Nationale de la Recherche (ANR). Its goal is to build a national platform for knowledge retrieval involving natural language and semantic web technologies. *Exmo* is co-responsible with Gémo (Chantal Reynaud) of work-package 3.2 : Ontology matching. It aims at integrating ontology matching solutions from several partners on the platform. We plan to use an Alignment server for that purpose.

More information on *WebContent* can be found at http://www.webcontent.fr.

8.1.2. Réseau régional Web Intelligence

Participants: Jérôme Euzenat [Contact], Sébastien Laborie.

Exmo is involved in the *Web intelligence* project supported by the Rhône-Alpes region. Jérôme Euzenat is in charge of the research area of the project. *Exmo* is more specifically involved in ontology matching in peer-to-peer systems with Marie-Christine Rousset (LSR) and Jean-Marc Petit (LIRIS).

More information on *Web intelligence* can be found at http://eric.univ-lyon2.fr/wi.

8.1.3. Collaboration with LIRMM

Participant: Jean-François Baget [Contact].

We collaborate with the "Représentations de connaissances et raisonnement" team of the LIRMM (Montpellier) on graph-based representation languages [7][8][9].

Jean-François Baget has been part-time at LIRMM.

8.2. European initiatives

8.2.1. Knowledge web network of excellence: realising the semantic web

Participants: Jérôme Euzenat [Contact], Faisal Al-Khateeb, Jean-François Baget, Antoine Zimmermann.

Exmo, as part of INRIA, is a founding and active member of the Knowledge web Network of Excellence. The INRIA node comprises the *Acacia*, *Exmo* and *Orpailleur* teams. *Exmo* is in charge of the work package 2.2 on "Heterogeneity" (see §6.1). Jérôme Euzenat is vice-scientific director of *Knowledge web* and represents INRIA in the executive management board.

More information on *Knowledge web* can be found at http://knowledgeweb.semanticweb.org.

8.2.2. NeOn integrated project: Networked ontologies

Participants: Jérôme Euzenat [Contact], Nikolay Pelov, Antoine Zimmermann.

Exmo contributes to the NeOn integrated project considering all the aspects of "Networked ontologies", i.e., the ontologies considered with their links to other ontologies. We work on the aspects of providing semantics to these networked ontologies (through alignments and ontology modules) and using ontology matching to recontextualise ontologies.

More information on *NeOn* can be found at http://www.neon-project.org.

8.2.3. Collaboration with University of Trento

Participant: Jérôme Euzenat [Contact].

Our collaboration with Pavel Shvaiko at University of Trento continues on all aspects of ontology matching and in particular monitoring new techniques, applications and evaluating their adequation.

8.3. International Initiatives

8.3.1. Collaboration with Université du Québec à Montréal

Participant: Jérôme Euzenat [Contact].

Exmo collaborates with Petko Valtchev at the Université du Québec à Montréal on the topic of "Ontology matching". We develop a matching algorithm, OLA together.

8.3.2. Collaboration with Universidade Federal de Pernambuco

Participants: Jérôme Euzenat [Contact], Antoine Zimmermann.

We have started collaborating with Frederico Freitas of the Universidade Federal de Pernambuco (Recife) on ontology modules and composition.

Frederico visited *Exmo* in May and October 2006 while Antoine visited Recife in August 2006.

9. Dissemination

9.1. Leadership within scientific community

- Jérôme Euzenat is founding member of the "Semantic web science association" (steering committee for the ISWC conference series), founding member of the "Scientific Steering Committee" of the "European Semantic Web Conference Series" (SSCESWC), member of the "Scientific advisory board" of the "European Academy for Semantic web Education" (EASE) and foundig member of EASE, member of the steering committee of the LMO conference series and member of the steering committee for the RFIA 2006 conference.
- Jérôme Euzenat is vice-scientific director and coordinator of "Heterogeneity" work package of the *Knowledge web* network of excellence (see §8.2) (2004-2007).

9.2. Editorial boards, conference and workshop committees

- Jérôme Euzenat has been the programme chairman of the "Artificial intelligence : methodology, systems and applications conference 2006 (AIMSA)" (Varna, BG), September 1st, 2006 [6].
- Jérôme Euzenat is the editor-in-chief of the "Journal électronique d'intelligence artificielle (JEDAI)".
- Editorial board of the journal "Journal of Web Semantics" and "Journal on Data Semantics" (Jérôme Euzenat).
- Programme committee for the 2006 issues of the conferences "International Semantic Web Conference (ISWC)", "European Semantic Web Conference (ESWC)", "World-wide web conference (WWW)", "(US) National conference on AI (AAAI)", "International Conference on Formal Ontologies for Information Systems (FOIS)", "International Conference on Semantic and Digital Media Technologies (SAMT)", "International conference on knowledge engineering and knowledge management (EKAW)", and "Reconnaissance des Formes et Intelligence Artificielle (RFIA)".

9.3. Conferences, meetings and tutorial organisation

- Organiser (with Pavel Shvaiko, Alain Léger, Deborah McGuinness and Holger Wache) of the "Context and ontologies" workshop of the 16th ECAI, Riva del Garda (IT), 2006 [4].
- Organiser (with Pavel Shvaiko, Natasha Noy, Heiner Stuckenschmidt, Richard Benjamin and Michael Uschold) of the "Ontology matching" workshop of the 5th ISWC, Athens (GA US), 2006 .
- Organiser (with many other colleagues) of the Ontology Alignment Evaluation Initiative 2006 at the "Ontology matching" workshop of the 5th ISWC, Athens (GA US), 2006 .

9.4. Invited conferences and other talks

- Mettre en correspondence des ontologies, pourquoi, comment?, matinées du LIG, Grenoble (FR), 2006 (Jérôme Euzenat)
- Web sémantique et diversité, Séminaire LISTIC, Annecy (FR), 2006 (Jérôme Euzenat)

9.5. Teaching

- Co-ordination of the Artificial intelligence profile of the second year of mathematics and informatics master, Intelligence, Interaction, Information track -Joseph Fourier university and INPG - Grenoble - resp. Yves Demazeau
- Ontology matching lecture and hands-on session (3h): European Summer School on Ontology Engineering and the Semantic Web, Cercedilla (ES), resp. John Domingue (Jérôme Euzenat)
- Semantic web technologies lecture and practice (12h), EDF, Clamart (FR), resp. Jérôme Euzenat (Jérôme Euzenat and Jean-François Baget).
- Web, Connaissance, Sémantique: Lecture (24h), Master "Mathématiques et informatique" research option, "Intelligence, interaction, information" track, Université Joseph-Fourier-INP Grenoble (Jean-François Baget and Jérôme Euzenat)

9.6. Miscellaneous

- Development and maintainance of the web site (http://jedai.afia-france.org) for the Journal électronique d'intelligence artificielle (JEDAI).
- Jérôme Euzenat has been editor-in-chief (with Philippe Morignot) of the Bulletin de l'AFIA until June 2006.

10. Bibliography

Major publications by the team in recent years

[1] JÉRÔME EUZENAT. *An API for ontology alignment*, in "Proc. 3rd conference on international semantic web conference (ISWC), Hiroshima (JP)", FRANK VAN HARMELEN, SHEILA MCILRAITH, DIMITRIS PLEXOUSAKIS (editors). , 2004, in "Lecture notes in computer science", Springer Verlag, Heildelberg (DE), vol. 3298, p. 698-712, http://www.springerlink.com/index/DY8Y9F31A9GT9762.

[2] JÉRÔME EUZENAT. *Représentations de connaissance : de l'approximation à la confrontation*, Habilitation à diriger des recherches, Université Joseph Fourier, Grenoble (FR), janvier 1999, ftp://ftp.inrialpes.fr/pub/sherpa/theses/hdr-euzenat.ps.gz.

[3] JÉRÔME EUZENAT, NABIL LAYAÏDA, VICTOR DIAS. *A semantic framework for multimedia document adaptation*, in "Proc. 18th International Joint Conference on Artificial Intelligence (IJCAI), Acapulco (MX)", p. 31-36, Morgan Kauffman, San-Mateo (CA US), 2003.

Year Publications

Books and Monographs

[4] PAVEL SHVAIKO, JÉRÔME EUZENAT, ALAIN LÉGER, DEBORAH MCGUINNESS, HOLGER WACHE (editors). *Context and ontologies: theory and practiceProc. workshop on Context and ontologies: theory and practice, Riva del Garda (IT)*, 2006, 2006, http://ceur-ws.org/Vol-210/.

[5] PAVEL SHVAIKO, JÉRÔME EUZENAT, NATALYA NOY, HEINER STUCKENSCHMIDT, RICHARD BENJAMINS, MICHAEL USCHOLD (editors). *Ontology matchingProc. 1st international workshop on ontology matching, Athens (GA US)*, 2006, 2006, http://ceur-ws.org/Vol-225/.

[6] JÉRÔME EUZENAT, JOHN DOMINGUE (editors). *Artificial intelligence: methodology, systems and applications*, 2006, in "Lecture notes in computer science", Springer verlag, Berlin (DE), vol. 4183, 2006, http://www.springeronline.com/3-540-40930-0.

Publications in Conferences and Workshops

[7] JEAN-PIERRE AUBERT, JEAN-FRANÇOIS BAGET, MICHEL CHEIN. *Simple concept graphs and simple conceptual graphs*, HENRIK SCHÄRFE, PASCAL HITZLER, PETER OHRSTROM (editors). , 2006, in "Lecture notes in computer science", Springer Verlag, Heildelberg (DE), vol. 4068, p. 87-101, 2006.

[8] JEAN-FRANÇOIS BAGET, OLIVIER CARLONI, MICHEL CHEIN, DAVID GENEST, ALAIN GUTIERREZ, MICHEL LECLÈRE, MARIE-LAURE MUGNIER, ÉRIC SALVAT, RALLOU THOMOPOULOS. *Towards benchmarks for conceptual graphs tools*, in "Proc. 1st workshop on Conceptual Structures Tool Interoperability (CS-TIW), Aalborg (DK)", ALDO DE MOOR, SIMON POLOVINA, HARRY DELUGACH (editors). , 2006, p. 72-86, 2006.

[9] JEAN-FRANÇOIS BAGET, ÉRIC SALVAT. *Rules dependencies in backward chaining of conceptual graphs rules*, HENRIK SCHÄRFE, PASCAL HITZLER, PETER OHRSTROM (editors). , 2006, in "Lecture notes in computer science", Springer Verlag, Heildelberg (DE), vol. 4068, p. 102-116, 2006.

[10] JÉRÔME EUZENAT, FANO RAMPARANY, JÉRÔME PIERSON. *Gestion dynamique de contexte pour l'informatique pervasive*, in "Proc. 15th conférence on reconnaissance des formes et intelligence artificielle (RFIA), Tours (FR)", 113, 2006, ftp://ftp.inrialpes.fr/pub/exmo/publications/euzenat2006a.pdf.

[11] JÉRÔME EUZENAT, FANO RAMPARANY, JÉRÔME PIERSON. *A context information manager for pervasive environments*, in "Proc. 2nd workshop on contexts and ontologies (C&O), Riva del Garda (IT)", p. 25-29, 2006, ftp://ftp.inrialpes.fr/pub/exmo/publications/euzenat2006b.pdf.

[12] JÉRÔME EUZENAT, FANO RAMPARANY, JÉRÔME PIERSON. *A context information manager for dynamic environments*, in "Advances in pervasive computing", TOM PFEIFER, ALBRECHT SCHMIDT, WOONTACK WOO, GAVIN DOHERTY, FRÉDÉRIC VERNIER, KIERAN DELANEY, BILL YERAZUNIS, MATTHEW CHALMERS, JOE KINIRY (editors). , 2006, n° 207, Österreichische computer geselschaft, Wien (OS), p. 79-83, 2006, ftp://ftp.inrialpes.fr/pub/exmo/publications/euzenat2006c.pdf.

[13] JÉRÔME EUZENAT, MALGORZATA MOCHOL, PAVEL SHVAIKO, HEINER STUCKENSCHMIDT, ONDREJ SVAB, VOJTECH SVATEK, WILLEM ROBERT VAN HAGE, MIKALAI YATSKEVICH. *Results of the Ontology Alignment Evaluation Initiative 2006*, in "Proc. 1st international workshop on ontology matching, Athens (GA US)", 2006, http://ceur-ws.org/Vol-225/paper7.pdf.

[14] JASON JUNG, JÉRÔME EUZENAT. *From Personal Ontologies to Socialized Semantic Space*, in "Proc. 3rd ESWS poster session , Budva (ME)", 2006, ftp://ftp.inrialpes.fr/pub/exmo/publications/jung2006a.pdf.

[15] JASON JUNG, JÉRÔME EUZENAT. *Measuring semantic centrality based on building consensual ontology on social network*, in "Proc. 2nd workshop on semantic network analysis (SNA), Budva (ME)", p. 27-39, 2006, ftp://ftp.inrialpes.fr/pub/exmo/publications/jung2006b.pdf.

[16] SÉBASTIEN LABORIE, JÉRÔME EUZENAT, NABIL LAYAÏDA. *Adaptation spatiale efficace de documents SMIL*, in "Proc. 15th conférence on reconnaissance des formes et intelligence artificielle (RFIA), Tours (FR)", 127, 2006, ftp://ftp.inrialpes.fr/pub/exmo/publications/laborie2006a.pdf.

[17] SÉBASTIEN LABORIE. *Spatio-temporal proximities for multimedia document adaptation*, in "Artificial intelligence: methodology, systems and applications", JÉRÔME EUZENAT, JOHN DOMINGUE (editors). , 2006, in "Lecture notes in computer science", Springer verlag, Berlin (DE), vol. 4183, p. 128-137, 2006.

[18] SÉBASTIEN LABORIE, JÉRÔME EUZENAT, NABIL LAYAÏDA. *A spatial algebra for multimedia document adaptation*, in "Proc. 1st International Conference on Semantic and Digital Media Technologies (SAMT), Athens (GR)", p. 7-8, 2006.

[19] SÉBASTIEN LABORIE, JÉRÔME EUZENAT. *Adapting the hypermedia structure in a generic multimedia adaptation framework*, in "Proc. 1st International Workshop on Semantic Media Adaptation and Personalization (SMAP), Athens (GR)", p. 62-67, 2006.

[20] LOREDANA LAERA, VALENTINA TAMMA, TREVOR BENCH-CAPON, JÉRÔME EUZENAT. *Agent-based argumentation for ontology alignments*, in "Proc. 6th workshop on Computational models of natural argument (CMNA), Riva del Garda (IT)", p. 40-46, 2006, ftp://ftp.inrialpes.fr/pub/exmo/publications/laera2006a.pdf.

[21] LOREDANA LAERA, VALENTINA TAMMA, JÉRÔME EUZENAT, TREVOR BENCH-CAPON, TERRY PAYNE. *Reaching agreement over ontology alignments*, in "The semantic web - ISWC 2006", ISABEL CRUZ, STEFAN DECKER, DEAN ALLEMANG, CHRIS PREIST, DANIEL SCHWABE, PETER MIKA, MIKE USCHOLD, LAURA

AROYO (editors). , 2006, in "Lecture notes in computer science", Springer Verlag, Heildelberg (DE), vol. 4273, p. 371-384, 2006, http://iswc2006.semanticweb.org/items/Laera2006oz.pdf.

[22] LOREDANA LAERA, VALENTINA TAMMA, JÉRÔME EUZENAT, TREVOR BENCH-CAPON. *Arguing over ontology alignments*, in "Proc. 1st workshop on ontology matching (OM), Athens (GA US)", p. 49-60, 2006, http://ceur-ws.org/Vol-225/paper5.pdf.

[23] MALGORZATA MOCHOL, ANJA JENTZSCH, JÉRÔME EUZENAT. *Applying an analytic method for matching approach selection*, in "Proc. 1st workshop on ontology matching (OM), Athens (GA US)", p. 37-48, 2006, http://ceur-ws.org/Vol-225/paper4.pdf.

[24] ANTOINE ZIMMERMANN, MARKUS KRÖTZSCH, JÉRÔME EUZENAT, PASCAL HITZLER. *Formalizing ontology alignment and its operations with category theory*, in "Proc. 4th International conference on Formal ontology in information systems (FOIS), Baltimore (ML US)", BRANDON BENNETT, CHRISTIANE FELLBAUM (editors). , 2006, IOS Press, Amsterdam (NL), p. 277-288, 2006.

[25] ANTOINE ZIMMERMANN, JÉRÔME EUZENAT. *Three semantics for distributed systems and their relations with alignment composition*, in "The semantic web - ISWC 2006", ISABEL CRUZ, STEFAN DECKER, DEAN ALLEMANG, CHRIS PREIST, DANIEL SCHWABE, PETER MIKA, MIKE USCHOLD, LAURA AROYO (editors). , 2006, in "Lecture notes in computer science", Springer Verlag, Heildelberg (DE), vol. 4273, p. 16-29, 2006, http://iswc2006.semanticweb.org/items/Zimmermann2006jw.pdf.

Internal Reports

[26] JÉRÔME EUZENAT, MARC EHRIG, ANJA JENTZSCH, MALGORZATA MOCHOL, PAVEL SHVAIKO. *Case-based recommendation of matching tools and techniques*, deliverable, 2006, n° 1.2.6, Knowledge web NoE, , IST Knowledge web NoE, 2006.

References in notes

[27] FRANZ BAADER, DIEGO CALVANESE, DEBORAH MCGUINNESS, DANIELE NARDI, PETER PATEL-SCHNEIDER (editors). *The description logics handbook*, Cambridge university pressCambridge (UK), 2003.

[28] JAMES CLARK (editor). *XSL Transformations (XSLT)*, Recommendation, 1999, W3C, http://www.w3.org/TR/xslt.

[29] TIM BERNERS-LEE, JAMES HENDLER, ORA LASSILA. *The semantic web*, in "Scientific american", vol. 279, n° 5, p. 35-43, 2001.