INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team EXMO

Computer-mediated communication of structured knowledge

Rhône-Alpes

THEME SYM C

Activity Report

2007

Table of contents

1. Team

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

Administrative assistant
Helen Pouchot [25% (TR INRIA)]

Research scientist
Jean-François Baget [CR INRIA, until March 31st, 2007]

Post doctoral fellow
Jérôme David [INRIA, from December 1st, 2007]
Chan Le Duc [INRIA, from September 1st, 2007]
Seungkeun Lee [AVC, until June 30th, 2007]

Ph. D. students
Faisal Al-Khateeb [Jordan government until September 30th, 2007 and INRIA from November 1st, 2007]
Sébastien Laborie [MESR grant until September 30th and UJF from October 1st]
Jérôme Pierson [CIFRE (France Telecom R&D)]
Antoine Zimmermann [INRIA until September 1st and UPMF from September 1st, 2007]

2. Overall Objectives

Keywords: *Alignment API*, *CPRDF*, *CPSPARQL*, *OLA*, *OWL*, *PRDF*, *PSPARQL*, *Path RDF*, *RDF*, *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 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 a single format or by reference to a single 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 alignments as representing the relationships between two ontologies on the semantic web. Ontology alignments express correspondences between 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 transformations (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 properties) 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.

2.2. Highlights of the year

Jérôme Euzenat, from INRIA project-team *Exmo*, and Pavel Shvaiko have written the first comprehensive reference book on Ontology matching: the art of finding relations between ontologies. The book presents currently available work in a uniform framework by providing a detailed account of matching techniques and matching systems in a systematic way from theoretical, practical and application perspectives. For more: http://book.ontologymatching.org.

3. Scientific Foundations

3.1. Knowledge representation semantics

We usually work with semantically defined knowledge representation languages (like description logics [31], 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. RDF is a knowledge representation language dedicated to the annotation of resources within the framework of the semantic web. OWL is designed for expressing ontologies: it offers the description of concepts and relations that can be used within RDF.

We consider a language L as a set of syntactically defined expressions (often inductively defined by applying constructors over other expressions). A representation ($o \subseteq L$) 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 expressions. An expression (δ) is then a consequence of a set of expressions (o) if it is satisfied by all of their 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 procedural 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), if any, 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 (γ) 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. A model of such a distributed system is a tuple of local models and an equalising function such that 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 ones 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 mean 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, these 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 content.

The vision of a "semantic web" [32] 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 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 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*.

In addition, *Exmo* also considers a more specific use of semantic web technologies in semantic peer-to-peer systems. Semantic peer-to-peer systems are made of a distributed network of independent peers which share local resources annotated semantically and locally. This means that each peer can use its own ontology for annotating resources and these ontologies have to be confronted before peers can communicate.

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 an 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 that a part of software development can be reduced to the 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 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, Chan Le Duc.

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

We have provided a first implementation for this API which can be used for producing transformations, rules or bridge axioms independently from the algorithm which produced the alignment. The proposed implementation features:

- a base implementation of the interfaces with all useful facilities;
- a library of sample matchers;
- 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 have developed on top of the *Alignment API* an Alignment server that can be used by remote clients for matching ontologies and for storing and sharing alignements [25]. The Alignement server can be accessed through HTML, web service and agent communication interfaces. It is developed as an extensible platform which allows to plug-in new interfaces.

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 Université du Québec à Montréal, see §8.3.1).

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 that we designed by extending SPARQL with regular expressions. CPSPARQL extends PSPARQL with constraints on path steps (see §6.2.1). We have implemented a PSPARQL query evaluator in Java. This evaluator can parse SPARQL, PSPARQL and CPSPARQL queries, parse RDF documents written in the 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 variable in different places of the query, i.e., to take into account the current mappings.

This evaluator passed 435 test cases out of the 440 in the W3C Data Access Working Group SPARQL first test base. The 5 missed tests are those that use the DESCRIBE result form which is not implemented.

The PSPARQL query evaluator is available for download as well as an online test servlet under http://psparql.inrialpes.fr.

6. New Results

The results in 2007 are mainly related to the use of our semantics of alignments in evaluation, modules, and alignment language (§6.1) and distributed systems (§6.2.2), the extension of the PSPARQL interpreter (§6.2.1), the introduction of semantic social networks (§6.2.3), experiments with our context management framework (§6.3) and the adaptation of the interactive structure in multimedia documents (§6.4).

6.1. Ontology matching and alignment

Pursuing the work on ontology matching and alignment [7] carried out in the framework of the *Knowledge web* network of excellence, we have developed 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 four years, we run the Ontology Alignment Evaluation Initiative which organises evaluation campaigns for assessing the degree of achievement of actual ontology matching algorithms [28]. This year's event has been held in Busan, Korea [6]. 18 different teams entered the evaluation which consisted in 7 different sets of tests. As the year before, the participation in this effort is thus increasing both in the number of participants and challenges.

We have also pursued our investigations on generalising precision and recall by developing new measures [14] based on the new semantics of alignments (see §3.2). These measures go as far as possible in taking into account the actual meaning of ontologies and alignments.

6.1.2. Expressive alignment language

Participants: Jérôme Euzenat [Contact], Antoine Zimmermann.

The Alignment format [1] designed by *Exmo* is widely used. However, it is only able to express simple alignments between ontologies. In order to offer a format which is both expressive and independent from concrete ontology languages we developed the Expressive Alignment Language [26].

The high expressivity of the language allows the expression of complex alignments even if the ontology languages are not themselves expressive. The language independence guarantees that we can define expressive alignments between any languages and provides a declarative definition of the alignments which will be usable in various manners (ontology merging, data translation, etc.).

We defined for this language an abstract syntax (used for describing the semantics), an exchange syntax (in RDF/XML) and a more readable surface syntax. We provided a model theoretic semantics for the language which rely upon the semantics of aligned ontologies while remaining independent from their details. We also provided support for this language by combining and extending the API for the Alignment Mapping Language developed at Innsbruck and our Alignment API (see §5.1).

This work is carried out in cooperation with Innsbruck Universität (François Scharffe) in the context of the *Knowledge web* project (see §8.2.1).

6.1.3. Ontology modules

Participants: Jérôme Euzenat, Antoine Zimmermann [Contact].

The goal of the semantic web is to share knowledge. In this context, knowledge is expressed in interlinked chunks rather than large monolitic ontologies. Ontologies can be assembled from ontology modules like programme modules in software engineering.

We have designed a model of modules which combines an interface and an ontology implementation, in which a module can import other modules through alignments with their interface [15]. This is a very natural way to do since alignments can be used to adjust the components in the ontologies. We have provided the semantics of such modules which is a combination of ontology semantics and our own alignment semantics.

This work is carried out in cooperation with Frederico Freitas (see §8.3.2) and in the framework of the NeOn project (see §8.2.2).

6.2. Distributed semantic systems

Dealing with the semantic web, we are interested in distributed semantic systems, i.e., sets of distributed ontologies that have to work together. One way for these systems to interact consists of exchanging queries and answers. For that reason, we pay particular attention to query systems.

6.2.1. Constrained Path RDF as a query language for RDF

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

Though RDF itself can be used as a query language for an RDF knowledge base (using RDF semantic consequence), the need for added expressivity in queries has led to the definition of the SPARQL query language. SPARQL queries are defined on top of graph patterns that are basically RDF (and more precisely GRDF) graphs.

To be able to characterise paths of arbitrary length in a query, e.g. "does there exist a trip from town A to town B using only trains and buses?", we have already proposed the PRDF (for Path RDF) language effectively mixing RDF reasoning with database-inspired regular paths [23][10]. However, PRDF graphs do not allow expressing constraints on the nodes, e.g. "Moreover, one of the correspondences must provide a wireless connection.". To express these constraints, we propose an extension of PRDF, called CPRDF (for Constrained Path RDF [24]).

For these two extensions of RDF, we have provided an abstract syntax and an extension of RDF semantics. We characterise query answering (the query is a PRDF or a CPRDF graph, the knowledge base is an RDF graph) as a particular case of PRDF or CPRDF entailment that can be computed using some kind of graph homomorphism. Query answering thus remains an NP-hard problem in all these languages. Finally, we use these PRDF or CPRDF graphs as graph patterns in SPARQL, defining the PSPARQL and CPSPARQL extensions of that query language.

6.2.2. Reasoning with distributed semantic systems

Participants: Faisal Al-Khateeb [Contact], Antoine Zimmermann.

Query answering in expressive knowledge bases is a difficult task and becomes more tedious when knowledge is described in different knowledge bases. We have considered query answering over a system in Distributed Description Logics, such that each ontology is a description logic knowledge base and alignments are made of directed relations called bridge rules. More precisely, we have defined the distributed answers of a given query expressed in terms of one ontology (called the target ontology) in the system [9]. These answers may contain individuals from different ontologies. Our definition of distributed answers is well founded since it covers local query answers. To compute these answers, we have provided an algorithm that reduces the problem of distributed query answering to local query answering. This algorithm extends the target ontology with knowledge deduced using bridge rules and local axioms of other ontologies, then locally evaluates the query over the extended ontology alone. Although this algorithm has been proved correct, it is not complete with highly expressive logic. However, the completeness of the algorithm for fragments of less expressive description logics is still open. There are two main defects for this algorithm: First, it is much time and resource consuming; Second, it requires total access to the foreign knowledge.

To overcome these problems, we provided another approach that consists in broadcasting a set of queries to each foreign ontology [11]. Each query in the set is obtained by translating terms of the initial query using bridge rules. Despite the incompleteness of this approach, it provides highly parallelised distributed query evaluation.

In order to effectively reason on distributed systems of ontologies and alignments, the abstract semantics proposed in §3.2 has been specifically formalised for particular logics and alignments. More precisely, we described a new kind of distributed logics, namely Integrated Distributed Description Logics [22], where ontologies are represented as description logic knowledge bases, and alignments assert cross-ontology concept/role subsumptions or disjunctions, or cross-ontology instance membership. In particular, this formalism is adapted for reasoning with OWL ontologies aligned by automatic ontology matching tools.

6.2.3. Semantic social networks

Participant: Jérôme Euzenat [Contact].

Social networks are simply the graph between people along social relations (usually denoting that they know each others). There has been much work on social network analysis for finding central people in a network or connecting efficiently an individual to another.

We introduced the notion of semantic social networks in order to describe networks embedding not only relations between people, but also the ontologies that people use. These ontologies can be used, for instance, in order to annotate resources such as documents, pictures, etc. We proposed an organisation for semantic social networks in three layers: social layer, ontology layer and concept layer. Each layer features a network based on different relations [17].

People in the social network are related to the ontologies they use, and ontologies are related to the concepts they use and they define. However, it would be useful to be able to infer relations between people from the relations between concepts and ontologies. This has the advantage of providing potential proximity relations for people who do not even know each others. Such techniques can be useful for instance, for finding people to which it will be easier to forward a query or group of homogeneous people who will be more prone to design a consensus ontology [18]. We proposed some tentative propagation rules as well as measures for computing network analysis.

6.3. Dynamic aspects of alignments

We apply the results obtained on alignments in various contexts where semantic web technologies and alignments 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 management component in charge of maintaining its local context. It can communicate with other context management 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 right producer for the information it needs. Context management 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.

We have developed a library to build the distributed context management system. It provides support for most operations of context management, i.e., searching, broadcasting and updating, context information.

In 2007, we have developed the Alignment server [18] and, in particular, the JADE plug-in for communicating with agents which is used in our distributed context management system. The Alignement server allows the context information manager component to find the correspondences between various ontologies with which it is confronted and thus to match application needs in terms of context information with the information provided by the other devices.

We have built a complete easily deployable ambient home environment. Our infrastructure manages context information flows from sensors and web services to pervasive application and a dynamic service composition infrastructure. We demonstrated it through several applications composed of a set of potentially interchangeable sensors and actuators. These applications are combined to present an integrated scenario which shows how an ambient home environment can improve the experience of a typical Grenoble resident and helps him to organise his leisure. This environment was showned at the Ubicomp 2007 conference.

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

6.4. Semantic adaptation of multimedia documents

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

When a multimedia document is played on platforms with limited resources, e.g., a mobile phone that can only display one image at a time 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 contributing in various areas:

Firstly, when considering the hypermedia dimension of multimedia documents, we have shown that it is possible to identify several loosely dependent sub-specifications. In order to adapt them and let the profile changing between each user interaction, we proposed an incremental adaptation approach. It adapts document sub-specifications step by step according to user interactions.

Secondly, we have extended our adaptation approach with the capability to suppress multimedia objects [19]. For example, a profile may indicate that only a few multimedia objects are allowed in a presentation. When multimedia objects are removed, we forced the adapted document to satisfy properties such as presentation contiguity. Moreover, we presented a method for summarising multimedia documents using a relevance degree for each multimedia object.

Thirdly, we have considered relation graphs containing mixed quantitative and qualitative relations. To efficiently check the satisfiability of this kind of temporal constraint network, we have to deal with the infinite domains of variables, which can generated an infinite number of candidates during backtrack algorithm. For solving this problem, we rely upon finite partitions of domains using bi-intervals (intervals of intervals) [12]. We have implemented sound and complete backtrack and forward checking algorithms and shown that bi-intervals, used in a hybrid algorithm which also instantiates constraints, improve our backtrack efficiency.

Finally, we have considered media adaptation [20]. For that purpose, we propose to adapt media items by replacing incompatible media items by others found on the web. The adapted media items must convey the same message as the original ones, while satisfying the target profile. We have presented a possible architecture to implement this and we have shown that search engines can already do it to a limited extent. Nonetheless, some results are unsatisfactory because media annotations lack semantics, are partial and are heterogeneous. Hence, we have proposed to use semantic web technologies, such as RDF descriptions, ontologies, ontology merging and matching, in order to select better alternatives, thus improving this adaptation framework.

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 (also known as Orange labs). 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 (LIG) and Jean-Marc Petit (LIRIS) [5].

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

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.

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], Chan Le Duc, 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.

8.3. 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 together a matching algorithm called OLA. The new version of OLA has obtained very good results at the OAEI-2007 evaluation campaign [13].

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 (see §6.1.3).

Frederico visited *Exmo* in September 2007.

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 and member of the steering committee of the LMO conference series.
- 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 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 2007 issues of the conferences "International Semantic Web Conference (ISWC)", "European Semantic Web Conference (ESWC)", "(US) National conference on AI (AAAI)",

9.3. Conferences, meetings and tutorial organisation

- Faisal Al-Khateeb, Sébastien Laborie and Antoine Zimmermann have been members of the organising committee of the 2007 AFIA platform, Grenoble, 2007.
- Jérôme Euzenat has organised (with Paolo Bouquet, Chiara Ghidini, Deborah L. McGuinness, Valeria de Paiva, Luciano Serafini, Pavel Shvaiko, and Holger Wache) the "Context and ontologies: representation and reasoning" workshop of the 6th Context, Rokskilde (DK), 2007 [4].
- Jérôme Euzenat has organised (with Pavel Shvaiko, Bin He and Fausto Giunchiglia) the "Ontology matching" workshop of the 6th ISWC, Busan (KR), 2007 [6].
- Jérôme Euzenat has organised (with Jean-Marc Petit and Marie-Christine Rousset) the "Passage à l'échelle des techniques de découverte de correspondances" workshop of the 7th EGC, Namur (BE), 2007 [5].
- Jérôme Euzenat has organised (with Laurent Vercouter, Yves Demazeau and François Jacquenet) the "Intelligence artificielle et web intelligence" workshop of the 4th AFIA platform, Grenoble (FR), 2007.
- Jérôme Euzenat has organised (with many other colleagues) the Ontology Alignment Evaluation Initiative 2007 at the "Ontology matching" workshop of the 6th ISWC, Busan (KR), 2007 [16].

9.4. Invited conferences and other talks

- XML transformation composition: early attempts and current state: Seminar Xerox research centre europe, Meylan (FR), March 1st, 2007 (Jérôme Euzenat)
- Réseaux dynamiques d'ontologies hétérogènes personnelles: Séminaire de prospective en modélisation des processus perceptifs et cognitifs, Saint-Dié des Vosges (FR), October 26th, 2007 (Jérôme Euzenat)
- Dynamic networks of heterogeneous personal ontologies: Web intelligence project meeting, Grenoble (FR), November 5th, 2007 (Jérôme Euzenat)

- The semantic web: from document annotation to ambient knowledge: Seminar Yeungnam university, Gyeongsan (KR), November 16th, 2007 (Jérôme Euzenat)

9.5. Teaching

- Co-ordination of the Web 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 (Jérôme Euzenat, 3h): European Summer School on Ontology Engineering and the Semantic Web, Cercedilla (ES), resp. John Domingue
- Semantic interoperability tutorial (Jérôme Euzenat with Natasha Noy, 1h): ISWC invited tutorial, Busan (KR), resp. John Domingue.

9.6. Miscellaneous

- Development and maintainance of the web site (http://jedai.afia-france.org) for the Journal électronique d'intelligence artificielle (JEDAI).

10. Bibliography

Major publications by the team in recent years

[1] JÉRÔME EUZENAT. *An API for ontology alignment*, in "Proc. 3rd 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] PAOLO BOUQUET, JÉRÔME EUZENAT, CHIARA GHIDINI, DEBORAH MCGUINNESS, VALERIA DE PAIVA, LUCIANO SERAFINI, PAVEL SHVAIKO, HOLGER WACHE (editors).*Proc. Context 2007 workshop on Context and ontologies: representation and reasoning (C&O:RR), Roskilde (DK)*, 2007, http://ceur-ws.org/Vol-298/.

[5] JÉRÔME EUZENAT, JEAN-MARC PETIT, MARIE-CHRISTINE ROUSSET (editors).*Actes atelier EGC 2007 sur Passage à l'échelle des techniques de découverte de correspondances (DECOR), Namur (BE)*, 2007, ftp://ftp.inrialpes.fr/pub/exmo/reports/EGC2007-decor-ws.pdf.

[6] PAVEL SHVAIKO, JÉRÔME EUZENAT, FAUSTO GIUNCHIGLIA, BIN HE (editors).*Proc. 2nd ISWC 2007 international workshop on ontology matching (OM), Busan (KR)*, 2007, http://ceur-ws.org/Vol-304/.

[7] JÉRÔME EUZENAT, PAVEL SHVAIKO. *Ontology matching*, Springer-Verlag, Heidelberg (DE), 2007, http://book.ontologymatching.org.

Articles in refereed journals and book chapters

[8] PAVEL SHVAIKO, JÉRÔME EUZENAT (editors).*Special issue on Ontology matching*, in "International journal of semantic web and information systems", IGI publishing, Hershey (PA US), vol. 3, n^{o} 2, p. 1-122, 2007.

Publications in Conferences and Workshops

[9] FAISAL ALKHATEEB, ANTOINE ZIMMERMANN. *Query answering in distributed description logics*, in "Proc. conference on New technologies, mobility and security (NTMS), Paris (FR)", HOUDA LABIOD, MOHAMAD BADRA (editors), Springer Verlag, Heildelberg (DE), 2007, p. 523-534, ftp://ftp.inrialpes.fr/pub/exmo/publications/alkhateeb2007a.pdf.

[10] FAISAL ALKHATEEB. *Une extension de RDF avec des expressions régulières*, in "Actes 8e rencontres nationales sur jeunes chercheurs en inteligence artificielle (RJCIA), Grenoble (FR)", p. 1-14, 2007.

[11] FAISAL ALKHATEEB, ANTOINE ZIMMERMANN. *Répondre à des requêtes dans un système distribué à base de connaissances*, in "Actes atelier sur Intelligence artificielle et web intelligence (IAWI), Grenoble (FR)", no pagination, 2007, http://www.emse.fr/~vercouter/iawi/16-alkhateeb.pdf.

[12] JEAN-FRANÇOIS BAGET, SÉBASTIEN LABORIE. *Bi-intervals for backtracking on temporal constraint networks*, in "Proc. 14th international conference on temporal representation and reasoning (TIME), Alicante (SP)", VALENTIN GORANKO, SEAN WANG (editors), IEEE Computer Society, Los Alamitos (CA), 2007, p. 163-168.

[13] JEAN-FRANÇOIS DJOUFAK-KENGUE, JÉRÔME EUZENAT, PETKO VALTCHEV. *OLA in the OAEI 2007 evaluation contest*, in "Proc. 2nd ISWC 2007 workshop on ontology matching (OM), Busan (KR)", p. 188-195, 2007, http://ceur-ws.org/Vol-304/paper16.pdf.

[14] JÉRÔME EUZENAT. *Semantic precision and recall for ontology alignment evaluation*, in "Proc. 20th International Joint Conference on Artificial Intelligence (IJCAI), Hyderabad (IN)", p. 248-253, AAAI Press, Menlo Park (CA US), 2007, http://www.ijcai.org/papers07/Papers/IJCAI07-054.pdf.

[15] JÉRÔME EUZENAT, ANTOINE ZIMMERMANN, FREDERICO FREITAS. *Alignment-based modules for encapsulating ontologies*, in "Proc. 2nd workshop on Modular ontologies (WoMO), Whistler (BC CA)", p. 32-45, 2007, ftp://ftp.inrialpes.fr/pub/exmo/publications/euzenat2007f.pdf.

[16] JÉRÔME EUZENAT, ANTOINE ISAAC, CHRISTIAN MEILICKE, PAVEL SHVAIKO, HEINER STUCKENSCHMIDT, ONDREJ SVAB, VOJTECH SVATEK, WILLEM ROBERT VAN HAGE, MIKALAI YATSKEVICH. *Results of the Ontology Alignment Evaluation Initiative 2007*, in "Proc. 2nd ISWC 2007 international workshop on ontology matching (OM), Busan (KR)", p. 96-132, 2007, http://ceur-ws.org/Vol-304/paper9.pdf.

[17] JASON JUNG, JÉRÔME EUZENAT. *Towards semantic social networks*, in "The semantic web: research and applications", ENRICO FRANCONI, MICHAEL KIFER, WOLFGANG MAY (editors), in "Lecture notes in computer science", Springer Verlag, Heildelberg (DE), vol. 4273, 2007, p. 267-280, ftp://ftp.inrialpes.fr/pub/exmo/publications/jung2007a.pdf.

[18] JASON JUNG, ANTOINE ZIMMERMANN, JÉRÔME EUZENAT. *Concept-based query transformation based on semantic centrality in semantic peer-to-peer environment*, in "Advances in data and web management", GUOZHU DONG, XUEMIN LIN, WEI WANG, YUN YANG, JEFFREY XU YU (editors), in "Lecture notes in computer science", Springer Verlag, Heildelberg (DE), vol. 4505, 2007, p. 622-629, ftp://ftp.inrialpes.fr/pub/exmo/publications/jung2007b.pdf.

[19] SÉBASTIEN LABORIE, JÉRÔME EUZENAT, NABIL LAYAÏDA. *Multimedia document summarization based on a semantic adaptation framework*, in "Proc. 1st international workshop on Semantically aware document processing and indexing (SADPI), Montpellier (FR)", HENRI BETAILLE, JEAN-YVES DELORT, PETER KING, MARIE-LAURE MUGNIER, JOCELYNE NANARD, MARC NANARD (editors), ACM Press, 2007, p. 87-94.

[20] SÉBASTIEN LABORIE, ANTOINE ZIMMERMANN. *A framework for media adaptation using the web and the semantic web*, in "Proc. 2nd international workshop on semantic media adaptation and personalization (SMAP), London (UK)", PHIVOS MYLONAS, MANOLIS WALLACE, MARIOS ANGELIDES (editors), IEEE Computer Society, Los Alamitos (CA US), 2007, p. 32-37.

[21] LOREDANA LAERA, IAN BLACOE, VALENTINA TAMMA, TERRY PAYNE, JÉRÔME EUZENAT, TREVOR BENCH-CAPON. *Argumentation over Ontology Correspondences in MAS*, in "Proc. 6th International conference on Autonomous Agents and Multiagent Systems (AAMAS), Honolulu (HA US)", p. 1285-1292, 2007, http://www.aamas-conference.org/Proceedings/aamas07/html/AAMAS07_0321_4301ade6dd0b32710792554358d9b28e.xml.

[22] ANTOINE ZIMMERMANN. *Integrated distributed description logics*, in "Proc. 20th International workshop on description logic (DL), Brixen/Bressanone (IT)", DIEGO CALVANESE, ENRICO FRANCONI, VOLKER HAARSLEV, DOMENICO LEMBO, BORIS MOTIK, SERGIO TESSARIS, ANNI-YASMIN TURHAN (editors), Bolzano University Press, Bolzano (IT), 2007, p. 507-514, ftp://ftp.inrialpes.fr/pub/exmo/publications/zimmermann2007a.pdf.

Internal Reports

[23] FAISAL ALKHATEEB, JEAN-FRANÇOIS BAGET, JÉRÔME EUZENAT. *RDF with regular expressions*, Research report, n° RR-6191, INRIA Rhône-Alpes, Grenoble (FR), 2007, https://hal.inria.fr/inria-00144922.

[24] FAISAL ALKHATEEB, JEAN-FRANÇOIS BAGET, JÉRÔME EUZENAT. *Constrained regular expressions in SPARQL*, Research report, n° RR-6360, INRIA Rhône-Alpes, Grenoble (FR), 2007, https://hal.inria.fr/inria-00188287/en.

[25] JÉRÔME EUZENAT, ANTOINE ZIMMERMANN, MARTA SABOU, MATHIEU D'AQUIN. *Matching ontologies for context*, Deliverable, n° 3.3.1, NeOn, 2007, ftp://ftp.inrialpes.fr/pub/exmo/reports/neon-331.pdf.

[26] JÉRÔME EUZENAT, FRANÇOIS SCHARFFE, ANTOINE ZIMMERMANN. *Expressive alignment language and implementation*, Deliverable, n° 2.2.10, Knowledge web, 2007, ftp://ftp.inrialpes.fr/pub/exmo/reports/kweb-2210.pdf.

[27] FRANÇOIS SCHARFFE, JÉRÔME EUZENAT, CHAN LE DUC, PAVEL SHVAIKO. *Analysis of knowledge transformation and merging techniques and implementations*, Deliverable, n° 2.2.7, Knowledge web, 2007, ftp://ftp.inrialpes.fr/pub/exmo/reports/kweb-227.pdf.

[28] PAVEL SHVAIKO, JÉRÔME EUZENAT, HEINER STUCKENSCHMIDT, MALGORZATA MOCHOL, FAUSTO GIUNCHIGLIA, MIKALAI YATSKEVICH, PAOLO AVESANI, WILLEM ROBERT VAN HAGE, ONDREJ SVAB, VOJTECH SVATEK. *Description of alignment evaluation and benchmarking results*, Deliverable, n^{o} 2.2.9, Knowledge web, 2007, ftp://ftp.inrialpes.fr/pub/exmo/reports/kweb-229.pdf.

Miscellaneous

[29] FRANÇOIS SCHARFFE, JÉRÔME EUZENAT, YING DING, DIETER FENSEL. *Correspondence patterns for ontology mediation*, in "Proc. ISWC 2007 Poster session , Busan (KR)", p. 89-90, 2007.

[30] PAVEL SHVAIKO, JÉRÔME EUZENAT. *Guest editorial preface on the special issue on Ontology matching*, in "International journal of semantic web and information systems", IGI publishing, Hershey (PA US), vol. 3, n^{o} 2, p. i-iii, 2007.

References in notes

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

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