IN PARTNERSHIP WITH:
CNRS

INRA

Université des sciences et techniques du Languedoc (Montpellier 2)

Activity Report 2012

Project-Team GRAPHIK

GRAPHs for Inferences and Knowledge representation

IN COLLABORATION WITH: Laboratoire d'informatique, de robotique et de microélectronique de Montpellier (LIRMM)

RESEARCH CENTER
Sophia Antipolis - Méditerranée

THEME
Knowledge and Data Representation and Management

Table of contents

Project-Team GRAPHIK

Keywords: Artificial Intelligence, Knowledge Representation, Reasoning, Data Management, Rule-based Languages

Creation of the Project-Team: January 01, 2010 .

1. Members

Research Scientists

Jean-François Baget [Junior Researcher, Inria]
Rallou Thomopoulos [Junior Researcher, INRA, on leave until July 2012]

Faculty Members

Michel Chein [Emeritus Professor, Université Montpellier 2]
Madalina Croitoru [Associate Professor, Université Montpellier 2]
Jérôme Fortin [Associate Professor, Université Montpellier 2]
Souhila Kaci [Professor, Université Montpellier 2, HdR]
Michel Leclère [Associate Professor, Université Montpellier 2, on sabbatical leave from Feb. 2012]
Marie-Laure Mugnier [Team leader, Professor, Université Montpellier 2, HdR]

External Collaborator

Patrice Buche [Research Engineer, INRA, HdR]

Engineers

Alain Gutierrez [Engineer, CNRS, 50% member of GraphIK]
Luc Menut [Engineer, INRA, 50% member of GraphIK]

PhD Students

Akila Ghersedine [2012 – INA Cifre Grant]
Léa Guizol [2011 – co-funded with ABES]
Mélanie König [2011 – Ministry Grant]
Bruno Paiva Lima Da Silva [2010 – Ministry Grant]
Tjitze Rienstra [2010 – funded by University of Luxembourg]
Michaël Thomazo [2010 – Grant from École Normale Superieure de Cachan]

Administrative Assistant

Annie Aliaga [Inria, shared with other teams]

2. Overall Objectives

2.1. Logic and Graph-based KRR

The main research domain of GraphIK is Knowledge Representation and Reasoning (KRR), which studies paradigms and formalisms for representing knowledge and reasoning on these representations. We follow a logic-oriented approach of this domain: the different kinds of knowledge have a logical semantics and reasoning mechanisms correspond to inferences in this logic. However, in the field of logic-based KRR, we distinguish ourselves by using graphs and hypergraphs (in the graph-theoretic sense) as basic objects. Indeed, we view labelled graphs as an *abstract representation* of knowledge that can be expressed in many KRR languages (different kinds of conceptual graphs —historically our main focus—, the Semantic Web language RDFS, expressive rules equivalent to the so-called tuple-generating-dependencies in databases, some description logics dedicated to query answering, etc.). For these languages, reasoning can be based on the structure of objects (thus on graph-theoretic notions), with homomorphism as a core notion, while being sound and complete with respect to entailment in the associated logical fragments. An important issue is to study *trade-offs* between the expressivity of languages and the computational tractability of (sound and complete) reasoning in these languages.

2.2. From Theory to Applications, and Vice-versa

We study logic- and graph-based KRR formalisms from three perspectives:

- theoretical (structural properties, expressiveness, translations between languages, problem complexity, algorithm design),
- software (developing tools to implement theoretical results),
- applications (which also feed back into theoretical work).

2.3. Main Challenges

GraphIK focuses on some of the main challenges in KRR:

- ontological query answering, i.e., query answering taking an ontology into account, and able to process large datasets;
- reasoning with rule-based languages;
- dealing with heterogeneous and hybrid knowledge bases (*i.e.* composed of several modules that have their own formalism and reasoning mechanisms);
- reasoning with "imperfect knowledge" (*i.e.* vague, uncertain, partially inconsistent, multi-viewpoints and/or with multi-granularity).

2.4. Scientific Axes

GraphIK has three main scientific directions:

1. **decidability, complexity and algorithms** for problems in languages corresponding to first-order logic fragments;
2. the addition of expressive and **non-classical features** (to the first-order logic languages studied in the first axis) with a good expressivity/efficiency trade-off;
3. the integration of theoretical tools to **real knowledge-based systems**.

From an applicative viewpoint, two themes are privileged for the next years:

- knowledge representation for agronomy, the final objective being a knowledge-based system to aid decision-making for the quality control in food processing.
- data integration and quality improvement, specifically for document metadata bases.

2.5. Highlights of the Year

- Organization of ECAI 2012 (European Conference on Artificial Intelligence), one of the major conferences in AI, together with the Coconut team at LIRMM (see Sect. 9.1).
- Several keynote talks at international conferences and workshops: RuleML 2012, Effost 2012, Datalog 2.0 2012, MPREF 2012 (see Sect. 9.1).
- Three new ANR projects: ASPIQ, PAGODA and Qualinca, the latter being coordinated by GraphIK (see Sect. 8.1). The three projects tackle different aspects of ontology-based data management, with our rule-based framework as the kernel formalism.

BEST PAPER AWARD :
[31] **A Sound and Complete Backward Chaining Algorithm for Existential Rules in RR'2012: International Conference on Web Reasoning and Rule Systems**. M. KÖNIG, M. LECLÈRE, M.-L. MUGNIER, M. THOMAZO.

3. Scientific Foundations

3.1. Logic-based Knowledge Representation and Reasoning

We follow the mainstream *logical* approach to the KRR domain. First-order logic (FOL) is the reference logic in KRR and most formalisms in this area can be translated into fragments (i.e., particular subsets) of FOL. A large part of research in this domain can be seen as studying the *trade-off* between the expressivity of languages and the complexity of (sound and complete) reasoning in these languages. The fundamental problem in KRR languages is entailment checking: is a given piece of knowledge entailed by other pieces of knowledge, for instance from a knowledge base (KB)? Another important problem is *consistency* checking: is a set of knowledge pieces (for instance the knowledge base itself) consistent, i.e., is it sure that nothing absurd can be entailed from it? The *query answering* problem is a topical problem (see Sect. 3.3). It asks for the set of answers to a query in the KB. In the special case of boolean queries (i.e., queries with a yes/no answer), it can be recast as entailment checking.

3.2. Graph-based Knowledge Representation and Reasoning

Besides logical foundations, we are interested in KRR formalisms that comply, or aim at complying with the following requirements: to have good *computational* properties and to allow users of knowledge-based systems to have a maximal *understanding and control* over each step of the knowledge base building process and use.

These two requirements are the core motivations for our specific approach to KRR, which is based on labelled *graphs*. Indeed, we view labelled graphs as an *abstract representation* of knowledge that can be expressed in many KRR languages (different kinds of conceptual graphs —historically our main focus—, the Semantic Web language RDF (Resource Description Framework), its extension RDFS (RDF Schema) expressive rules equivalent to the so-called tuple-generating-dependencies in databases, some description logics dedicated to query answering, etc.). For these languages, reasoning can be based on the structure of objects, thus based on graph-theoretic notions, while staying logically founded.

More precisely, our basic objects are labelled graphs (or hypergraphs) representing entities and relationships between these entities. These graphs have a natural translation in first-order logic. Our basic reasoning tool is graph homomorphism. The fundamental property is that graph homomorphism is sound and complete with respect to logical entailment i.e. given two (labelled) graphs G and H, there is a homomorphism from G to H *if and only if* the formula assigned to G is entailed by the formula assigned to H. In other words, logical reasonings on these graphs can be performed by graph mechanisms. These knowledge constructs and the associated reasoning mechanisms can be extended (to represent rules for instance) while keeping this fundamental correspondence between graphs and logics.

3.3. Ontological Query Answering

Querying knowledge bases is a central problem in knowledge representation and in database theory. A knowledge base (KB) is classically composed of a terminological part (metadata, ontology) and an assertional part (facts, data). Queries are supposed to be at least as expressive as the basic queries in databases, i.e., conjunctive queries, which can be seen as existentially closed conjunctions of atoms or as labelled graphs. The challenge is to define good trade-offs between the expressivity of the ontological language and the complexity of querying data in presence of ontological knowledge. Classical ontogical languages, typically description logics, were not designed for efficient querying. On the other hand, database languages were able to process complex queries on huge databases, but without taking the ontology into account. There is thus a need for new languages and mechanisms, able to cope with the ever growing size of knowledge bases in the Semantic Web or in scientific domains.

This problem is related to two other problems identified as fundamental in KRR:

- *Query-answering with incomplete information.* Incomplete information means that it might be unknown whether a given assertion is true or false. Databases classically make the so-called closed-world assumption: every fact that cannot be retrieved or inferred from the base is assumed to be false. Knowledge bases classically make the open-world assumption: if something cannot be inferred from the base, and neither can its negation, then its truth status is unknown. The need of coping with incomplete information is a distinctive feature of querying knowledge bases with respect to querying classical databases (however, as explained above, this distinction tends to disappear). The presence of incomplete information makes the query answering task much more difficult.
- *Reasoning with rules.* Researching types of rules and adequate manners to process them is a mainstream topic in the Semantic Web, and, more generally a crucial issue for knowledge-based systems. For several years, we have been studying some rules, both in their logical and their graph form, which are syntactically very simple but also very expressive. These rules can be seen as an abstraction of ontological knowledge expressed in the main languages used in the context of KB querying. See Sect. 6.1 for details on the results obtained.

A problem generalising the above described problems, and particularly relevant in the context of multiple data/metadata sources, is *querying hybrid knowledge bases*. In a hybrid knowledge base, each component may have its own formalism and its own reasoning mechanisms. There may be a common ontology shared by all components, or each component may have its own ontology, with mappings being defined among the ontologies. The question is what kind of interactions between these components and/or what limitations on the languages preserve the decidability of basic problems and if so, a "reasonable"complexity. Note that there are strong connections with data integration in databases.

3.4. Imperfect Information and Priorities

While classical FOL is the kernel of many KRR languages, to solve real-world problems we often need to consider features that cannot be expressed purely (or not naturally) in classical logic. The logic- and graph-based formalisms used for previous points have thus to be extended with such features. The following requirements have been identified from scenarios in decision making in the agronomy domain (see Sect. 4.2):

1. to cope with vague and uncertain information and preferences in queries;
2. to cope with multi-granularity knowledge;
3. to take into account different and potentially conflicting viewpoints ;
4. to integrate decision notions (priorities, gravity, risk, benefit);
5. to integrate argumentation-based reasoning.

Although the solutions we will develop need to be validated on the applications that motivated them, we also want them to be sufficiently generic to be applied in other contexts. One angle of attack (but not the only possible one) consists in increasing the expressivity of our core languages, while trying to preserve their essential combinatorial properties, so that algorithmic optimizations can be transferred to these extensions. To achieve that goal, our main research directions are: non-monotonic reasonings (see ANR project ASPIQ in Sect. 8.1), as well as argumentation and preferences (see Sect. 6.2).

4. Application Domains

4.1. Introduction

We currently focus on two application domains: knowledge representation in agronomy, more precisely applied to the quality in agri-food chains, and metadata management, in particular for bibliographic metadata.

The choice of the agronomy domain is motivated both by the local context of GraphIK (UMR IATE) and by its adequation to our research themes. Indeed, the agri-food domain seems to be particularly well-adapted to artificial intelligence techniques: there are no mathematical models available to solve the problems related to the quality of agrifood chains, which need to be stated at a more conceptual level; solving these problems requires an integrated approach that takes into account expert knowledge, which is typically symbolic, as well as numeric data, vague or uncertain information, multi-granularity knowledge, multiple and potentially conflicting viewpoints and actors.

The second area, metadata management, is not strictly speaking an application domain, but rather a cross-cutting axis. Indeed, metadata can be used to describe data in various areas (including for instance scientific publications in agronomy). We have a long experience in this domain, and we currently focus on document metadata.

4.2. Agronomy

Quality control within agri-food chains, but also non-food chains relies on numerous criteria (environmental, economical, functional, sanitary quality, etc.). The objectives of quality are based on several actors. The current structure of chains is questioned as for system perenniality, protection of the environment, cost and energy. In all cases, the following questions have to be taken into account:

1. the actors' viewpoints are divergent, hence it is necessary to define reasoning mechanisms able to model and take into account the balance between viewpoints, and the risks and benefits they imply;
2. the successive steps involved in a chain, impacting the quality of end products, have limiting factors. Their improvement is a complex objective that has no simple solution;
3. data from literature are dispersed and scattered, which makes their use difficult.

These questions highlight the need for an integrated approach of agri-food chains, respectively with symbolic reasoning mechanisms, reverse engineering methods, and knowledge organization and modelling.

Our general objective is the conception of a decision support tool for the actors of an agri-food chain, in presence of contradictory viewpoints and priorities, including the concepts of gravity and certainty of a risk or a benefit. The first step is to build a knowledge-based system able to represent the different kinds of knowledge needed, and provided with consistency checking, querying and symbolic simulation mechanisms, which will allow to refine and validate the modelling.

Our results in Sect. 6.1 and Sect. 6.2 can be seen as theoretical requirements towards this objective.

4.3. Document Metadata

Semantic metadata, in particular semantic annotations for multimedia documents, are at the core of the applications we are working on for several years. In the applications we developed in the previous years, mainly with INA (National Institute of Audiovisual) and FMSH (Fondation Maison des Sciences de l'Homme), we have built tools aimed at helping the manual construction of semantic annotations. In these projects, manual construction was unavoidable because semantically rich annotations, not obtainable by automatic processes, had to be built. In our current project Qualinca (see Sect. 8.1), the semantic metadata considered consists of information present in bibliographic databases and authority notices (which respectively describe documents and so-called authorities, such as authors typically). The challenge is not to build these metadata, which have been built by human specialists and already exist, but, for instance, to check their validity, to link or to merge different metadata bases.

5. Software

5.1. Cogui

Participants: Alain Gutierrez, Michel Leclère, Marie-Laure Mugnier, Michel Chein, Madalina Croitoru.

Cogui (http://www.lirmm.fr/cogui) is a tool for building and verifying knowledge bases. It is a freeware written in Java (version 1.2, 2005–2010 GPL Licence). Currently, it supports Conceptual Graphs and import/export in RDFS.

Here are the major evolutions of the version delivered this year:

- Cogui now allows import/export in the Datalog+/- language thanks to a new Datalog+/- parser (see Sect. 5.4).
- Scripted rules were introduced. It is a new type of object that allows users to attach a script to a traditional rule in order to modify or control its behavior.
- A new interface ensures connectivity to a NoSQLdatabase (MongoDB).
- Large graphs can now be stored. In a near future, we will be able to perform queries on data too big to fit in central memory (see Sect. 5.2).

5.2. Alaska

Participants: Bruno Paiva Lima Da Silva, Jean-François Baget, Madalina Croitoru.

Alaska (http://alaska.bplsilva.com/) is a java library dedicated to the storage and querying of large knowledge bases. It intends to be the foundation layer of our OBDA (Ontology Based Data Access) software developments. It has been built, first as part of a Master's thesis, and now of the PhD of Bruno Paiva Lima da Silva [34].

In Alaska, facts and queries are defined via a generic interface that favors a logical view of these objects. Implementations of this interface allow for the storage of facts w.r.t. different storage paradigms and systems (e.g., relational databases *MySQL* and *Sqlite*; triple stores *Sesame* and graph databases *Neo4J*, *DEX*, *HyperGraphDB* and *OrientDB*). For the time being, we can store 10^7 to 10^8 atoms. In the same way, logical queries can be evaluated through different methods, be it the native querying mechanism of the database used (*e.g. SPARQL* or *SQL*), or specifically designed algorithms (from a simple backtrack to a full constraint solver based upon *Choco* for hard problem instances). Note that all these methods provide the same answer set to queries.

This library already allows for testing our OBDA algorithms on large instances (it is already used by other PhD students for their experiments), and will soon be ready to be distributed to a broader audience. Our generic approach will ease this dissemination to different research domains.

5.3. Kiabora

Participants: Swan Rocher [first year master internship], Michel LeclÃ¨re, Marie-Laure Mugnier.

http://www2.lirmm.fr/~mugnier/graphik/kiabora/index.html

Kiabora is a tool dedicated to the analysis of a set of existential rules. It can check if this set belongs to a known *decidable* class of rules, either directly or by means of its *Graph of Rule Dependencies (GRD)*. Kiabora analyzes the properties of the strongly connected components in the GRD, which allows to determine properties of the rule set with respect to decidability as well as the kind of paradigm (forward or backward chaining) ensuring decidability.

Besides, Kiabora also provides format conversion and rule decomposition services. It is written in Java.

5.4. DLGP

Participants: Jean-FranÃ§ois Baget, Michel LeclÃ¨re, Marie-Laure Mugnier, Alain Gutierrez, Swan Rocher [first year master internship], ClÃ©ment Sipieter [first year master internship].

http://www2.lirmm.fr/~mugnier/graphik/kiabora/downloads/datalog-plus_en.pdf

DLGP (for Datalog Plus) is a textual exchange format at once human-friendly, concise and easy to parse. This format can be seen as an extension of the commonly used format for plain Datalog. A file may contain four kinds of knowledge elements: facts, existential rules, negative constraints and conjunctive queries. This format will allow us to easily exchange data and ontologies with groups working on the equivalent Datalog+/- formalism, developped in Oxford.

A DLGP parser is now available.

For this section, participants are listed in alphabetical order.

6. New Results

6.1. Ontology-Based Query Answering with Existential Rules

Participants: Jean-François Baget, Mélanie König, Michel Leclère, Marie-Laure Mugnier, Michaël Thomazo.

Note that for this section, as well as all sections in New Results, participants are given in alphabetical order.

In collaboration with: Sebastian Rudolph (Karlsruhe Institute of Technology)

We have pursued the work on the existential rule framework in the context of Ontology-Based Query Answering. See the 2011 activity report for details on this framework also known as Datalog+/-. The ontology-based query answering problem consists of querying data while taking into account inferences enabled by an ontology (described by existential rules in our case).

From 2009 to 2011, we mainly investigated decidability and complexity issues. In 2012, while still interested in deepening decidability and complexity results, we tackled the next step: algorithms. Our aim is to develop algorithms with good theoretical properties (at least they should run in "the good worst-case complexity class") and with good performance in practice. There are two main ways of processing rules, namely forward chaining and backward chaining. In forward chaining, rules are applied to enrich the initial facts and query answering is solved by evaluating the query against the "saturated" facts (as in a classical database system). When it is finite, the backward chaining process can be divided into two steps: first, the query is rewritten into a first-order query (typically a union of conjunctive queries) using the rules; then the rewritten query is evaluated against the initial facts (again, as in a classical database system).

6.1.1. Forward Chaining Algorithms

Considering the expressive class of greedy bounded-treewidth set of rules (in short *gbts*), which we defined in 2011, we have designed a query answering algorithm which has several advantages over 2011 algorithm, while staying optimal with respect to worst-case combined and data complexities.

1. It is much more implementable (previous algorithm was using an oracle).
2. It is generic in the sense that it works for any class of rules that fulfills the gbts property, but it can also be easily specialized for specific gbts subclasses whith lower complexities, such as frontier-guarded or guarded rules, in such a way that it runs in the good complexity class.
3. It allows for separation between offline and online processing steps: the knowledge base can be compiled independently from queries, which are evaluated against the compiled form.

One of the lightweight description logics used for ontology-based query answering is $\mathcal{EL}$. We designed a subclass of existential rules that covers $\mathcal{EL}$ with the same complexity of reasoning, while allowing for any predicate arity and some cycles on variables. We also added complex role inclusions like transitivity and right/left identity rules to enhance expressivity, while staying polynomial in data complexity and generalizing existing results.

- *Results published in [36], [37] and [32] (invited conference). See also our research report [49] for a longer version.*
- *A journal version extending the papers at IJCAI 2011 and KR 2012 is in preparation, to be submitted to a major artificial intelligence journal.*

6.1.2. Backward Chaining Algorithms

We consider query rewriting techniques that output a union of conjunctive queries, which we see as a set of conjunctive queries. More specifically, only the most general elements of this set need to be kept in the output. We first proved that all sound and complete query rewriting algorithms necessarily produce the same result (up to redundancy) when restricted to their most general elements. It follows that comparing existing algorithms with respect to the size of the produced query is pointless.

Existing query rewriting algorithms accept only specific classes of existential rules (mainly corresponding to the translation of some lightweight description logics). We designed an algorithm that accept as input any set of existential rules and stops if this set of rules fufills so-called *fus* property (meaning that the set of most general rewritings of any initial conjunctive query is finite). This algorithm has been implemented and first experimentations have been led on rule bases obtained by translating description logic bases.

- *Results published in [31] (best paper price)*

6.1.3. Querying Optimization (Work in Progress)

Our current work aims at improving previous algorithms, in particular: the online querying step in the gbts algorithm; the query rewriting algorithm, by avoiding generating several times equivalent rewritings; for specific subclasses, query rewriting into a set of so-called semi-conjunctive queries instead of conjunctive queries, which reduces the size of the output query.

6.2. Reasoning with Imperfect Information and Priorities

Participants: Madalina Croitoru, Jérôme Fortin, Souhila Kaci, Tjitze Rienstra, Rallou Thomopoulos.

In collaboration with: Joël Abecassis (IATE/INRA), Patrice Buche (IATE/INRA), Nir Oren (Univ. of Aberdeen, Scotland), Leon van der Torre (University of Luxembourg) and Nouredine Tamani (post-doc IATE).

This year, we mainly investigated decision support based on argumentations systems and preferences, either in relation with application needs in agronomy or on more fundamental aspects.

6.2.1. Argumentation for Decision Making in Agronomy

Historically, scientific investigations in this axis are guided by applications of our partners in agronomy (IATE laboratory). Part of the work has consisted of analyzing the proposed applications and the techniques they require in order to select appropriate applications with respect to our team project.

In the context of the EcoBioCap project (see Sect. 8.2), the different stakeholders have expressed conflicting preferences for the packaging quality. However, when discussing with domain experts they have raised the need for a tool which allows them to highlight a conflict and see the reasons behind it. In order to achieve this goal two steps were taken. First we have instantiated a popular logical argumentation framework (ASPIC+) with a simple preference logic. This allowed the different experts to express arguments about their preferences. We can then extract maximal consistent subsets of preferences by the means of extensions.

- *This work was performed in collaboration with the University of Aberdeen (Dr. Nir Oren) and the results were published and presented at the COMMA conference [24].*

Second, a negotiation phase was introduced to the previously described system in order for the domain experts to refine and extend their preferences. This tool was the aim of the master thesis of Patricio Mosse.

- *This work was published and presented at the Effost conference [23], based upon Patricio Mosse's Master Thesis [48]. A detailed journal article reporting on the two steps is under preparation and will be submitted beginning 2013.*

Let us mention additional results related to the applications in agronomy on semi-automatic data extraction from web data (tables) [39], [40], [41], data reliability, and the representation and flexible querying of imprecise data with fuzzy sets [42], [15]. These investigations are complementary to the above mentioned results on argumentation and generally relate to other aspects in the same applicative projects.

6.2.2. Conditional Acceptance Functions

Dung-style abstract argumentation theory centers on argumentation frameworks and acceptance functions. The latter take as input a framework and return sets of labelings. A labeling assigns "in", "out" or "undecided" to each arguments. Arguments having "in" assignment are acceptable arguments. This methodology however assumes full awareness of the arguments relevant to the evaluation. There are two reasons why this is not satisfactory. Firstly, full awareness is, in general, not a realistic assumption. Second, frameworks have explanatory power, which allows us to reason abductively or counterfactually, but this is lost under the usual semantics. To recover this aspect, we generalized conventional acceptance, and we present the concept of a conditional acceptance function which copes with the dynamics of argumentation frameworks.

- *Results published in [28].*

6.2.3. Foundational Aspects of Preferences

Preferences are the backbone of various fields as they naturally arise and play an important role in many real-life decisions. Preferences are fundamental in scientific research frameworks as well as applications. One of the main problems an individual faces when expressing her preferences lies in the number of variables (or attributes or criteria) that she takes into account to evaluate the different outcomes. Indeed, the number of outcomes increases exponentially with the number of variables. Moreover, due to their cognitive limitation, individuals are generally not willing to compare all possible pairs of outcomes or evaluate them individually. These facts have an unfortunate consequence that any preference representation language that is based on the direct assessment of individual preferences over the complete set of outcomes is simply infeasible.
Fortunately, individuals can abstract their preferences. More specifically, instead of providing preferences over outcomes (by pairwise comparison or individual evaluation), they generally express preferences over partial descriptions of outcomes. Often such statements take the form of qualitative comparative preference statements e.g., "I like London more than Paris" and "prefer tea to coffee". Conditional logics aim at representing such partial descriptions of individual preferences which we refer to as comparative preference statements. They use different completion principles in order to compute a preference relation induced by a set of preference statements. In particular they use various more or less strong semantics to interpret comparative preference statements. So far the main objective in artificial intelligence has been to rank-order the set of outcomes given a set of comparative preference statements and one or several semantics. We addressed this problem from a different angle. We considered a set of postulates studied in preference logics and non-monotonic reasoning which formalize intuition one may have regarding the behavior of preference statements. We analyzed the behavior of the different semantics w.r.t. these postulates. Our analysis gives a complete picture of the behavior of our (five) semantics.

In the last decade, AI researchers have pointed out the existence of two types of information: positive information and negative information. This distinction has also been asserted in cognitive psychology. Distinguishing between these two types of information may be useful in both knowledge and preference representation. In the first case, one distinguishes between situations which are not impossible because they are not ruled out by the available knowledge, and what is possible for sure. In the second case, one distinguishes between what is not rejected and what is really desired. Besides it has been shown that possibility theory is a convenient tool to model and distinguish between these two types of information. Knowledge/Preference representation languages have also been extended to cope with this particular kind of information. Nevertheless despite solid theoretical advances in this topic, the crucial question of "which reading (negative or positive) one should have" remains a real bottleneck. We focused on comparative statements and presented a set of postulates describing different situations one may encounter. We provided a representation theorem describing which sets of postulates are satisfied by which kind of information (negative or positive) and conversely. One can then decide which reading to apply depending on which postulates she privileges.

- *Results published in [29] and [30].*

6.2.4. Argumentation for Inconsistency-Tolerant Query Answering (Work in Progress)

Argumentation allows to encode by the means of extensions maximal subsets of the knowledge base which are consistent (given the logic chosen). We are currently investigating the link between different argumentation extensions and the notion of a maximal repair as introduced by [51], [50] in the context of the positive existential subset of first order logic we are mainly working with. We are then interested in comparing the semantics proposed in the literature for query answering with inconsistent knowledge bases and argumentation reasoning paradigms. This study has been performed jointly with the University of Luxembourg during a research visit during end of November. We plan to submit our results at a conference beginning January.

6.3. Semantic Data Integration

Participants: Michel Chein, Madalina Croitoru, Léa Guizol, Michel Leclère, Rallou Thomopoulos.

It often happens that different references (i.e. data descriptions), possibly coming from heterogeneous data sources, concern the same real world entity. In such cases, it is necessary: (i) to detect whether different data descriptions really refer to the same real world entity and (ii) to fuse them into a unique representation. Since the seminal paper [52], this issue has been been studied under various names:"record linking", "entity resolution","reference resolution", "de-duplication", "object identification", "data reconciliation", etc., mostly in databases (cf. the bibliography by William E. Winckler [1]). It has become one of the major challenges in the Web of Data, where the objective is to link data published on the web and to process them as a single distributed database. Most entity resolution methods are based on classification techniques; Fatiha Saïs, Nathalie Pernelle and Marie-Christine Rousset proposed the first logical approach [53]. Many experiments on public data are underway, in France (cf. DataLift[2] and ISIDORE[3] projects) or internationally (e.g., VIAF project[4] led by OCLC[5], whose aim is to interconnect authority files coming from 18 national organizations).

Three years ago, we began a collaboration with ABES (National Bibliographic Agency for Universities, which takes part in the VIAF project). The aim of this collaboration is to enable the publication of ABES metadata based on the Web of Data and to provide an identification service dedicated to bibliographic notices. ABES bibliographic bases, and more generally document metadata bases, appear to be a privileged application domain for the representation and reasoning formalisms developed by the team. This work has an interdisciplinary dimension, as it also requires experts in the Library and Information Science domain. We think that a logical approach is able to provide a generic solution for entity resolution in document metedata bases, even though it is generally admitted in Library and Information Science that *"there is no single paradigmatic author name disambiguation task—each bibliographic database, each digital library, and each collection of publications, has its own unique set of problems and issues"* [54].

6.3.1. Implementation of an Entity Identification Service

Last year, we have developed a method and a prototype to perform entity resolution between on one hand the authors of a new bibliographic notice, and, on the other the domain experts of an authority catalog (and namely the Sudoc catalogue from the ABES agency). The prototype providing this service has been implemented on top of Cogui and experiments have been led in the context of the SudocAd project (jointly conducted by ABES and GraphIK). This work has been continued this year on the following issues as part of the Qualinca project:

- *generalizing the developed method with the aim to define a generic combined (numerical/logical) framework for entity resolution. This work is reported in the research report [44] that we plan to submit to a conference in January.*
- *Defining evaluation measures of the quality of an entity resolution tool. This work is still on-going.*

[1] http://www.hcp.med.harvard.edu/statistics/survey-soft/docs/WinklerReclinkRef.pdf
[2] DataLift, http://datalift.org/
[3] ISIDORE, http://www.rechercheisidore.fr/
[4] The Virtual International Authority File, http://www.oclc.org/research/activities/viaf/
[5] Online Computer Library Center, http://www.oclc.org

6.3.2. Quality of Document Catalogs

The SudocAd project showed the feasability and pertinence of a mixed approach for data interlinking problems. It also showed the immediate necessity of taking into account the existence of human errors already present in document catalogues. This led us to propose Qualinca, an ANR Contint project, accepted beginning 2012 and started in April 2012. The partners include two major actors in the document catalogues field: ABES and INA, as well as three academic research groups.

In this context we currently investigate a formal approach to the notion of a "key" in the web of data. Our immediate objective is to define the notion of a discovered key used then in order to evaluate the quality of data inter linking of a meta data catalogue.

We also study the methodology of linking error detection and fixing based on a partitioning (clustering) method on authors of bibliographic records. This study is part of the PhD thesis of LÃ©a Guizol (jointly funded by GraphIK and ABES). The above mentioned methodology is based on a set of criteria which will allow us to cluster "similar" authors together. Each criterion represents a point of view on the author: name, publication time span, publication domain etc. The first challenge consists of defining for each of such view points the respective criteria. The second challenge is to propose an aggregation semantics of such criteria which is well adapted for the problem at hand.

- *The methodology of using such clustering techniques for this problem has been published in [25]. A certain number of criteria have already been implemented and different partitioning semantics proposed. We are currently evaluating these on the ABES data.*

6.3.3. Multi Agent Knowledge Allocation

The assumption behind semantic data integration and querying is that different agents accessing the integrated data repository will have equal interest in the querying results. This is not always true in a data sensitive scenario where the knowledge provider might want to allocate the query answers to the agents based on their valuations. Furthermore, the agents might want some information exclusively (and thus offer a valuation that allows it) while others might want it shared. To this end we have proposed a new mechanism of allocation of query answers inspired from combinatorial auctions. We have defined the newly introduced scenario of Multi Agent Knowledge Allocation and proposed a graph based method, inspired on network flows, for solving it.

- *These results were published in [26] and [35]. We are currently investigating the mechanism design aspects of such valuations in collaboration with the University of Athens (Dr. Iannis Vetsikas).*

7. Bilateral Contracts and Grants with Industry

7.1. ABES

Participants: Michel LeclÃ¨re, Michel Chein, Madalina Croitoru, LÃ©a Guizol.

Collaboration with ABES. Funding of half a PhD grant (Léa Guizol, started in October 2011). See Sect. 6.3.

7.2. CTFC

Participants: Patrice Buche, JÃ©rÃ´me Fortin, Awa Diattara.

We have initiated a national collaboration with the technical center of Comptois' cheese (CTFC : Centre Technique des Fromages Comtois). The objective of this collaboration is to design and test a platform for expert knowledge management. This will allow us to validate the integration of our theoretical tools on a new real-world application and strengthen GraphIK's involvement in agronomy applications. A master degree internship in collaboration with CTFC is done by Awa Diattara (University Gaston Berger of Saint-Louis, SÃ©nÃ©gal).

7.3. INA

Participants: Michel LeclÃ¨re, Michel Chein, Marie-Laure Mugnier, Akila Ghersedine.

Funding of a PhD CIFRE-grant (Akila Ghersedine, started in May 2012). The objective of the collaboration is to propose automatic (or semi-automatic) technics for enriching authorities. An authority is a record that describes a named entity used in document metadata (e.g. a person, a domainÂ). The elaboration of a solution requires addressing different problems: extraction of knowledge from textual metadata, entity resolution which is the core problematic of the Akila Ghersedine's thesis subject, and authority fusion.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

8.1.1.1. ASPIQ

Participants: Jean-FranÃ§ois Baget, JÃ©rÃ´me Fortin, Marie-Laure Mugnier, Michel LeclÃ¨re.

ASPIQ (ASP technologIes for Querying large scale multisource heterogeneous web information), is an ANR white program that started in Oct. 2012. The project coordinator is Odile Papini (LSIS), and it involves partners from CRIL and LERIA.

The main objective of this project is to propose:

- extensions of standard ASP for representing OWL2 tractable sublanguages;
- new operations for merging conflicting information in this extended ASP;
- the identification of subclasses of this extended ASP allowing for efficient query answering mechanisms;
- an implementation of a prototype reasoning system.

8.1.1.2. Pagoda

Participants: Jean-FranÃ§ois Baget, Marie-Laure Mugnier, MÃ©lanie König, Michaël Thomazo.

Pagoda is an ANR JCJC (young researchers) that will begin in Jan. 2013. The project coordinator is Meghyn Bienvenu (LRI), and it involves partners from the EPI LEO, the LIG, and the Anatomy Laboratory of Grenoble.

The primary aim of this project is to help address challenges brought by scalability and the handling of data inconsistencies by developing novel OBDA query answering algorithms and practical methods for handling inconsistent data.

8.1.1.3. Qualinca

Participants: Michel LeclÃ¨re, Michel Chein, Madalina Croitoru, LÃ©a Guizol, Akila Ghersedine, Rallou Thomopoulos, Marie-Laure Mugnier.

Qualinca is an ANR Contint project that started in Apr. 2012 and will end in Sept. 2015. The project coordinator is Michel LeclÃ¨re (GraphIK), and it involves partners from LRI, LIG, ABES and INA.

The main objective is to elaborate mechanisms allowing to:

- evaluate the quality of an existing documents base;
- maintain a given level of quality by controlling updating operations;
- increase the quality of a given base;
- develop generic methods that take into account the quality of a given base (for instance for searching documents or interconnecting bases).

8.1.2. Competitivity Clusters

We are taking part in the Laboratory of Excellence ("labex") *NUMEV* (Digital and Hardware Solutions, Modelling for the Environment and Life Sciences), led by University of Montpellier 2 in partnership with CNRS, University of Montpellier 1 and Inria. This project aims at developping information and communication technologies for environmental and life sciences. We are participating to one of the four axis, namely "Scientific Data: processing, integration and security".

8.2. European Initiatives

8.2.1. FP7 Projects

8.2.1.1. EcoBioCap

Participants: Patrice Buche, Madalina Croitoru, JÃ©rÃ´me Fortin, Patricio Mosse.

FP7-KBEE, March 2011–March2015. Led by INRA (and scientifically managed by Montpellier IATE laboratory). Sixteen partners among which Cork University (Ireland), CSIC (Spain), Roma University La Sapienza (Italy), SIK (Sweden). The objective of EcoBioCAP is to “provide the EU food industry with customizable, ecoefficient, biodegradable packaging solutions with direct benefits both for the environment and EU consumers in terms of food quality and safety”. GraphIK is involved in this project via its common members with IATE-KRR team. The budget is managed by IATE team. This project will feed Axis 2.

- *See Sect. 6.2 (argumentation for decision making in agronomy) for the results obtained this year.*

8.2.2. Collaborations with Major European Organizations

Leon van der Torre: University of Luxembourg, Computer Science and Communications Research Unit (Luxembourg)

Souhila Kaci collaborates with Leon van der Torre on argumentation aspects. They co-supervise a PhD student (Tjitze Rienstra).

Sebastian Rudolph: University of Karlsruhe, AIFB (Germany)

Jean-FranÃ§ois Baget, Marie-Laure Mugnier and Michaël Thomazo collaborate with Sebastian Rudolph on the study of complexity classes for fragments of existential rules. This successful work has already led to major publications (see Sect. 6.1).

Srdjan Vesic: University of Luxmbourg, Individual and Collective Reasoning research group (Luxembourg)

Madalina Croitoru collaborates with Srdjan Vesic on the link between inconsistency tolerant reasoning and argumentation.

Nir Oren: University of Aberdeen, Department of Computing Science (United Kingdom)

Madalina Croitoru and JÃ©rÃ´me Fortin collaborate with Nir Oren on argumentation and preference logics.

Ioannis A. Vetsikas: University of Athens, IIT (Greece)

Madalina Croitoru collaborates with Ioannis A. Vetsikas on mechanism design aspects of multi-agent knowledge allocation.

8.3. International Research Visitors

8.3.1. Visits of International Scientists

- Feb. 2012: Leon van der Torre (Pr., University of Luxembourg), collaboration on argumentation systems (2 days)
- Mar. 2012: Meghyn Bienvenu (CR CNRS, IASI/LEO), collaboration on Ontology-Based Data Access (5 days)

- Apr. 2012: Karima Sedki (MdC AgroCampus-Rennes, IRISA), Seminar on "Reasoning with preferences and deciding under uncertainty"
- May. 2012: Safa Yahi (MdC University of Marseille, LSIS), Seminar on "Management of inconsistency with justified argumentative inference"
- Sept. 2012: Bernard Moulin (UniversitÃ© Laval, Canada), collaboration on argumentation and dynamic systems (1 month)
- Oct. 2012: Jean-FranÃ§ois Condotta (CRIL), collaboration on representation and treatment of inconsistencies (2 days)
- Nov. 2012: Frank van Harmelen (Freie Univ. Amsterdam), seminar on "Reasoning over very, VERY large knowledge bases: towards a web-scale knowledge base of a 100 million facts and beyond"

8.3.1.1. Internships

Patricio Mosse (6 months)

Subject: Argumentation based preference aggregation (cf Ecobiocap in Sect. 8.2)

Institution: University of Buenos Aires (Argentina)

Awa Diattara (6 months)

Subject: Default rules for an agronomy application (cf CTFC in Sect. 7.2)

Institution: University Gaston Berger of Saint-Louis, SÃ©nÃ©gal

8.3.2. Visits to International Teams

- Madalina Croitoru and JÃ©rÃ´me Fortin. Visit to the Department of Computer Science (University of Aberdeen). 5 days in January 2012.
- Souhila Kaci. Visit to Leon van der Torre (University of Luxembourg). January 2012.
- Michaël Thomazo. Visit to KIT (Karlsruhe Institute of Technology). 6 weeks in June/July 2012. Collaboration with Sebastian Rudolph on Ontology-Based Data Access. Funded by DAAD http://paris.daad.de/bourses_de_recherche_de_courte_duree.html.
- Madalina Croitoru. Visit to Srdjan Vesic (University of Luxembourg). 1 week in November 2012. Collaboration on the link between maximal repairs and argumentation extensions
- Jean-FranÃ§ois Baget, Marie-Laure Mugnier and Michaël Thomazo. Visit to the Information Systems Group (University of Oxford). 3 days in December 2012. Scientific Exchanges on Existential Rules and Datalog+/-.

9. Dissemination

9.1. Scientific Animation

9.1.1. Organization of Conferences/Workshops

ECAI 2012 (European Conference on Artificial Intelligence), held in August 2012 in Montpellier, gathered more than 750 participants. This major conference in artificial intelligence was locally organized by the LIRMM, mainly by the teams Coconut and GraphIK. http://www2.lirmm.fr/ecai2012/

GKR@IJCAI 2013 (International workshop on graph structures for knowledge representation and reasoning) will be held for the third time in conjunction with IJCAI. This workshop was initiated by Madalina Croitoru in 2009. The GKR post proceedings of 2011 were published in a Springer volume [43].

MPREF'13 (Multidisciplinary Workshop on Advances in Preference Handling): will be held in conjunction with IJCAI'13 and will be co-organized by Souhila Kaci. This workshop promotes the broadened scope of preference handling and continues a series of events on preference handling at AAAI-02, Dagstuhl in 2004, IJCAI-05, ECAI-06, VLDB-07, AAAI-08, ADT-09, ECAI-2010 and ECAI-2012. Since 2008, this series of workshops is organized by the multidisciplinary working group on Advances in Preference Handling, which is affiliated to the Association of European Operational Research Societies EURO.

Souhila Kaci will also co-organize the "Uncertain Reasoning" Track at FLAIRS'13.

9.1.2. Editorial Boards

ICCS (International Conference on Conceptual Structures)
RIA (Revue Francophone d'Intelligence Artificielle)

9.1.3. Program Committees

International: IJCAI 2013 (senior PC and PC), KR 2012 (Principles of Knowledge Representation and Reasoning), RR 2012-2013 (Reasoning the Web), ECAI 2012 (European Conference on Artificial Intelligence), Datalog 2.0 2012, SGAI 2012-2013 (SGAI International Conference on Artificial Intelligence), FUZZ-IEEE 2012 (International Conference on Fuzzy Systems), Effost 2012 (European Federation of Food Science and Technology), MPREF'12 (Multidisciplinary Workshop on Advances in Preference Handling, in conjunction with ECAI), WL4AI'12 (Weighted Logics for AI, in conjunction with ECAI), FLAIRS'12 (Florida Artificial Intelligence Research Society), ICAART 2012-2013 (International Conference on Agents and Artificial Intelligence), Web Science 2012-2013

National: JIAF 2012 (JournÃ©es d'Intelligence Artificielle Fondamentale), RFIA 2012 (Reconnaissance des Formes et Intelligence Artificielle), IC 2012-2013 (IngÃ©nierie des Connaissances), LFA 2012 (rencontres francophones sur la Logique Floue et ses Applications).

9.1.4. Invited Talks

Keynote Talks at International Conferences

- RuleML: Ontology-Based Query Answering with Existential Rules, Marie-Laure Mugnier, Montpellier, August 2012. http://dbis.informatik.tu-cottbus.de/ruleml2012/
- Effost 2012: Food science and knowledge engineering: a challenging encounter, Patrice Buche, November 2012. http://www.effostconference.com/index.html

Invited Talks at International Workshops

- Workshop Datalog 2.0: Existential Rules: A Graph-based View, Marie-Laure Mugnier, Vienne, September 2012. http://www.dbai.tuwien.ac.at/datalog2.0/
- MPREF'12: What do we prefer? From philosophical stimulus to Artificial Intelligence solutions, Souhila Kaci, Montpellier, August 2012.

Invited Seminars

- Jean-FranÃ§ois Baget and Marie-Laure Mugnier. Existential Rules: Decidability and Algorithms, Department of Computer Science, University of Oxford, December 2012.
- Madalina Croitoru and JÃ©rÃ´me Fortin. Graph based Knowledge Representation and Reasoning, Department of Computer Science, University of Aberdeen, January 2012.

9.1.5. Scientific Advisory Boards

- ABES (National Bibliographic Agency for Universities) Scientific Advisory Board, Michel Chein (since its creation in 2010)
- Advisory Board of the Center of Excellence in Semantic Technologies (MIMOS, Malaysia), Marie-Laure Mugnier (since its creation in 2008)
- Scientific board of INRA-CEPIA department (CaractÃ©risation et Elaboration des Produits Issus de l'Agriculture – Agricultural Products Engineering), Marie-Laure Mugnier (since Septembre 2011)

9.1.6. Expertise Tasks:

Experts for ANR, INRA and Inria (project proposal reviewing); reviewers for Artificial Intelligence Journal, JAIR (Journal of Artificial Intelligence Research), IEEE Transactions on Fuzzy Systems, Journal of Visual Languages & Computing, European Journal of Operational Research, ...

9.1.7. Local Collective Tasks:

LIRMM Scientific Council (Jean-François Baget), Vice-chair of "Expert Pool" section 27–Computer Science (Michel LeclÃ¨re), member of Expert Pool Section 27 (Marie-laure Mugnier), LIRMM Laboratory Concil (Marie-Laure Mugnier until Oct. 2012, Souhila Kaci from Oct. 2012).

9.1.8. Participation to the W3C RDF Working Group

(Jean-François Baget) The mission of the RDF Working Group, part of the Semantic Web Activity, is to update the 2004 version of the Resource Description Framework (RDF) Recommendation. The scope of work is to extend RDF to include some of the features that the community has identified as both desirable and important for interoperability based on experience with the 2004 version of the standard, but without having a negative effect on existing deployment efforts. http://www.w3.org/2011/01/rdf-wg-charter

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

The next table details the number of lecture hours as well as the number of module responsibilities for each team member.

Name	Position	2011/12	Cursus (*)	Module Resp. (per year)
J.-F. Baget	Research Scientist	40	M (UM2)	1
M. Croitoru	Assistant Prof.	96	L (IUT) and M	2
M. Chein	Emeritus Prof.	0		
J. Fortin	Assistant Prof	192	Polytech	2
S. Kaci	Professor	198	L and M(UM2)	3
M. LeclÃ¨re	Assistant Prof.	96	L and M (UM2)	1
M. -L. Mugnier	Professor	192	L and M (UM2)	4
R. Thomopoulos	Research Scientist	0		no
L. Guizol	PhD	64	L	no
M. König	PhD	64	L	no
B. Paiva Lima	PhD	64	L (UM2)	no
M. Thomazo	PhD	64	L and M1 (UM2)	no

(*) L =Licence, M = Master (M1 = first year, M2 = second year), UM2 = Univ. Montpellier 2 (Sciences), IUT = Institute of Technology of UM2 (Licence Cursus), Polytech = Engineering School of UM2, UM3 = Univ. Montpellier 3 (Art and Humanities)

Globally, the team ensures the courses in logics (propositional logic and first-order logic in L, logics for Artificial Intelligence in M2) at the Montpellier 2 University, as well as the Master courses in Artificial Intelligence, Decision Support, Knowledge Representation and Knowledge Engineering. We are also responsible of modules in Web Technologies (Professional L at IUT) and Databases (L). Michel LeclÃ¨re, on sabbatical leave since Feb. 2012, and Madalina Croitoru, on Inria delegation, ensured each a half-service.

We have some specific responsibilities in the Computer Science Master:

- Michel LeclÃ¨re is co-responsible of the master speciality DECOL (about 20 students) started in September 2011.
- Marie-Laure Mugnier is co-responsible of the Computer Science Master started in September 2011 (about 240 students), which gathers six specialties.

9.2.2. Supervision

9.2.2.1. PhDs

No PhD was defended in 2012.

PhD in progress : MichaÃ«l Thomazo, Querying knowledge bases: decidability, complexity and algorithms, Sept. 2010, Marie-Laure Mugnier and Jean-FranÃ§ois Baget

PhD in progress : Bruno Paiva Lima da Silva, Comparing Storage Systems for Large knowledge bases, Oct. 2010, Jean-FranÃ§ois Baget and Madalina Croitoru

PhD in progress : Tjitze Rienstra, Dynamic argumentation systems, Oct. 2010, Souhila Kaci and Leon van der Torre (University of Luxembourg)

PhD in progress : MÃ©lanie KÃ¶nig, Algorithms for querying large knowledge bases, Oct. 2011, Michel LeclÃ¨re and Marie-Laure Mugnier

PhD in progress : LÃ©a Guizol, Entity identification in metadata bases, Oct. 2011, Michel LeclÃ¨re and Madalina Croitoru

PhD in Progress: Akila Ghersedine, Creating and enriching lists of authorities from heterogeneous resources, Apr. 2012, Michel LeclÃ¨re and Marie-Laure Mugnier

9.2.2.2. Internships

- Joris Lamare (ENS Cachan, 6 weeks)
- Swan Rocher (Master 1 UM2, 2 months + 3 months TER)
- ClÃ©ment Sipieter (Master 1 UM2, 2 months)
- Patricio Mosse (Master 2, 6 months)
- Awa Diattara (Master 2, 6 months)

9.2.3. Juries

- Patrice Buche
  - Reviewer (PhD): Interrogation personnalisÃ©e des systÃ¨mes d'information dÃ©diÃ©s au transport : une approche bipolaire floue, Nouredine Tamani, IRISA, ENSSAT Lannion, April 2012
  - Reviewer (PhD): Approches hybrides pour la recherche sémantique de l'information : Intgration des bases de connaissances et des ressources semi-structurées. Yassine Mrabet, University of Paris-Sud, July 2012
- Madalina Croitoru
  - Reviewer (PhD): Decision Support Systems for Brain Tumor Diagnosis: Classification and Evaluation Approaches, Javier Vicente Robledo, University of Valencia, October 2012
- Marie-Laure Mugnier
  - President of jury (PhD): Structuration de dÃ©bats en ligne Ã l'aide d'annotations socio-sÃ©mantiques, Antoine Seilles, University of Montpellier 2, April 2012
  - Reviewer (PhD): Extraction de motifs de graphes pour caractÃ©riser des sources de donnÃ©es RDF, Adrien Basse, University of Nice, December 2012
  - President of jury (PhD): Enhancing Ontology Matching by Using Machine Learning, Graph Matching and Information Retrieval Techniques, DuyHoa Ngo, University of Montpellier 2, December 2012.
  - Jury member(PhD): Static analysis of semantic web queries, Melisachew Wudage Chekol, University of Grenoble, December 2012.
- Michel Chein

– Jury Member (HDR): Knowledge Representation meets DataBases for the sake of ontology-based data management, FranÃ§ois GoasdouÃ©, University Paris-Sud, June 2012.

9.3. Popularization

- JÃ©rÃ´me Fortin gave a talk about argumentation during the CIAG (Carrefours de l'Innovation Agronomique) in Clermont-Ferrand http://www.inra.fr/ciag/colloques_alimentation/cereales. This presentation led to an article published in *Innovations Agronomiques* [19].
- Patrice Buche gave a talk about argumentation to JTIC (Congress on Cereals) http://www.jtic.eu/, http://www4.inra.fr/cepia/Restez-informes/Evenements/Les-videos-et-diaporamas-des-conferences-passees/JTIC-2012 in Reims. [38].

10. Bibliography

Major publications by the team in recent years

[1] F. ALKHATEEB, J.-F. BAGET, J. EUZENAT. *Extending SPARQL with regular expression patterns (for querying RDF)*, in "Journal of Web Semantics", 2009, vol. 7, n° 2, p. 57-73.

[2] J.-F. BAGET, J. FORTIN. *Default conceptual graph rules, atomic negation and Tic-Tac-Toe*, in "ICCS'10: 18th International Conference on Conceptual Structures - Conceptual Structures: From Information to Intelligence", Malaisie, M. CROITORU, S. FERRÉ, D. LUKOSE (editors), LNAI, Springer, July 2010, p. 42-55, ICCS'10 Best Paper Award, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00537338/en/.

[3] J.-F. BAGET, M. LECLÈRE, M.-L. MUGNIER, E. SALVAT. *Extending Decidable Cases for Rules with Existential Variables*, in "Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence (IJCAI 2009), Pasadena, California, USA,", 2009, p. 677-682.

[4] J.-F. BAGET, M. LECLÈRE, M.-L. MUGNIER, E. SALVAT. *On Rules with Existential Variables: Walking the Decidability Line*, in "Artificial Intelligence", March 2011, vol. 175, n° 9-10, p. 1620-1654 [*DOI :* 10.1016/J.ARTINT.2011.03.002], http://hal.inria.fr/lirmm-00587012/en.

[5] J.-F. BAGET, M.-L. MUGNIER, S. RUDOLPH, M. THOMAZO. *Walking the Complexity Lines for Generalized Guarded Existential Rules*, in "IJCAI'11: International Joint Conference on Artificial Intelligence", Barcelona, Spain, T. WALSH (editor), AAAI Press, July 2011, p. 712-717, http://hal.inria.fr/lirmm-00618081/en.

[6] M. CHEIN, M.-L. MUGNIER. *Graph-based Knowledge Representation and Reasoning—Computational Foundations of Conceptual Graphs*, Advanced Information and Knowledge Processing, Springer, 2009.

[7] S. DESTERCKE, P. BUCHE, V. GUILLARD. *A Flexible Bipolar Querying Approach with Imprecise Data and Guaranteed Results*, in "Fuzzy Sets and Systems", April 2011, vol. 169, n° 1, p. 51-54 [*DOI :* 10.1016/J.FSS.2010.12.014], http://hal-lirmm.ccsd.cnrs.fr/lirmm-00611940/en/.

[8] J. FORTIN, A. KASPERSKI, P. ZIELINSKI. *Some methods for evaluating the optimality of elements in matroids with ill-known weights*, in "Fuzzy Sets and Systems", 2009, vol. 160, n° 10, p. 1341-1354.

[9] A. HADJALI, S. KACI, H. PRADE. *Database Preference Queries–A Possibilistic Logic Approach with Symbolic Priorities*, in "Annals of Mathematics and Artificial Intelligence", 2012, vol. 63, n° 3-4, p. 357-383 [*DOI :* 10.1007/S10472-012-9279-9], http://hal-lirmm.ccsd.cnrs.fr/lirmm-00680458.

[10] S. KACI. *Working with Preferences: Less Is More*, Cognitive Technologies, Springer, 2011, p. 11-193.

[11] M. KÖNIG, M. LECLÈRE, M.-L. MUGNIER, M. THOMAZO. *A Sound and Complete Backward Chaining Algorithm for Existential Rules*, in "RR'2012: International Conference on Web Reasoning and Rule Systems", Australie, Springer, 2012, p. 122-138, RR'2012 Best Paper Award, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00764341.

[12] M.-L. MUGNIER, G. SIMONET, M. THOMAZO. *On the Complexity of Entailment in Existential Conjunctive First Order Logic with Atomic Negation*, in "Information and Computation", March 2012, vol. 215, p. 8-31 [*DOI :* 10.1016/J.IC.2012.03.001], http://hal.inria.fr/lirmm-00681914.

[13] M. THOMAZO, J.-F. BAGET, M.-L. MUGNIER, S. RUDOLPH. *A Generic Querying Algorithm for Greedy Sets of Existential Rules*, in "KR'12: International Conference on Principles of Knowledge Representation and Reasoning", Italie, 2012, p. 096-106, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00763518.

Publications of the year

Articles in International Peer-Reviewed Journals

[14] M. CROITORU, N. OREN, S. MILES, M. LUCK. *Graphical norms via conceptual graphs*, in "Knowledge-Based Systems", May 2012, vol. 29, p. 31-43 [*DOI :* 10.1016/J.KNOSYS.2011.06.025], http://hal-lirmm.ccsd.cnrs.fr/lirmm-00763678.

[15] S. DESTERCKE, P. BUCHE, B. CHARNOMORDIC. *Evaluating Data Reliability: an Evidential Answer with Application to a Web-Enabled data Warehouse*, in "IEEE Transactions on Knowledge and Data Engineering", January 2013, vol. 25, n° 1, p. 92-105, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00611954.

[16] A. HADJALI, S. KACI, H. PRADE. *Database Preference Queries–A Possibilistic Logic Approach with Symbolic Priorities*, in "Annals of Mathematics and Artificial Intelligence", 2012, vol. 63, n° 3-4, p. 357-383 [*DOI :* 10.1007/S10472-012-9279-9], http://hal-lirmm.ccsd.cnrs.fr/lirmm-00680458.

[17] M.-L. MUGNIER, G. SIMONET, M. THOMAZO. *On the Complexity of Entailment in Existential Conjunctive First Order Logic with Atomic Negation*, in "Information and Computation", March 2012, vol. 215, p. 8-31 [*DOI :* 10.1016/J.IC.2012.03.001], http://hal.inria.fr/lirmm-00681914.

Articles in National Peer-Reviewed Journals

[18] G. DELLA VALLE, N. PERROT, C. BAUDRIT, P. BUCHE, K. KANSOU, A. NDIAYE, R. THOMOPOULOS. *L'ingénierie des Connaissances, un outil pour intégrer les connaissances sur les procédés de la chaîne alimentaire et aider à la décision.*, in "Innovations Agronomiques", January 2012, vol. 19, p. 107-116, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00764369.

[19] J. FORTIN, R. THOMOPOULOS, J.-R. BOURGUET, J. ABÉCASSIS, P. BUCHE. *Un modèle d'argumentation pour l'aide à la décision dans les filières céréalières*, in "Innovations Agronomiques", March 2012, vol. 19, p. 117-126, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00764992.

International Conferences with Proceedings

[20] P. BUCHE. *Food science and knowledge engineering: a challenging encounter*, in "EFFoST'2012: European Federation of Food Science and Technology", France, Elsevier, November 2012, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00764987.

[21] P. BUCHE, S. DERVAUX, J. DIBIE-BARTHÉLEMY, L. IBANESCU, L. SOLER. *Semi-Automatic Loading of a Microbial Risk in Food Database Thanks to a Domain Ontology*, in "EFFoST'12 Annual Meeting", France, November 2012, 1, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00764998.

[22] P. BUCHE, V. GUILLARD, C. GUILLAUME, N. GONTARD, L. MENUT. *Decision Support System for designing biodegradable packaging for fresh produce: a knowledge engineering approach.*, in "EMB'2012", Italie, March 2012, 1, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00764997.

[23] M. CROITORU, J. FORTIN, P. MOSSE, P. BUCHE, V. GUILLARD, C. GUILLAUME. *Bio-Sourced and Biodegradable Packaging design Using Argumentation to Aggregate Stakeholder Preferences*, in "EFFoST'12 Annual Meeting", France, November 2012, 1, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00765000.

[24] M. CROITORU, J. FORTIN, N. OREN. *Arguing with Preferences in EcoBioCap*, in "COMMA'2012: Fourth International Conference on Computational Models of Argument", Autriche, Frontiers in Artificial Intelligence and Applications, IOS Press, September 2012, vol. 245, p. 51-58 [*DOI :* 10.3233/978-1-61499-111-3-51], http://hal-lirmm.ccsd.cnrs.fr/lirmm-00763656.

[25] M. CROITORU, L. GUIZOL, M. LECLÈRE. *On Link Validity in Bibliographic Knowledge Bases*, in "IPMU'2012: 14th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems", Italie, Springer, July 2012, vol. 297, p. 380-389 [*DOI :* 10.1007/978-3-642-31709-5_39], http://hal-lirmm.ccsd.cnrs.fr/lirmm-00763664.

[26] M. CROITORU, S. RUDOLPH. *Exclusivity-based allocation of knowledge*, in "AAMAS'12: 11th International Conference on Autonomous Agents and Multiagent Systems", Espagne, ACM, June 2012, vol. 3, p. 1249-1250, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00763649.

[27] J. FORTIN, R. THOMOPOULOS, J.-R. BOURGUET, M.-L. MUGNIER. *Supporting Argumentation Systems by Graph Representation and Computation*, in "Graph Structures for Knowledge Representation and Reasoning", Barcelona, Spain, France, LNAI, Springer-Verlag Berlin Heidelberg, 2012, p. 119-136 [*DOI :* 10.1007/978-3-642-29449-5_5], http://hal-lirmm.ccsd.cnrs.fr/lirmm-00768328.

[28] S. KACI, R. BOOTH, T. RIENSTRA, L. VAN DER TORRE. *Conditional Acceptance Functions*, in "COMMA'12: Computational Models of Argument", Autriche, I. PRESS (editor), September 2012, p. 470-477, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00762955.

[29] S. KACI. *A Postulate-Based Analysis of Comparative Preference Statements*, in "FLAIRS'12: 25th International Florida Artificial Intelligence Research Society Conference", États-Unis, May 2012, p. 543-548, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00762958.

[30] S. KACI. *Characterization of Positive and Negative Information in Comparative Preference*, in "ECAI'12: 20th European Conference on Artificial Intelligence", France, I. PRESS (editor), August 2012, p. 450-455, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00762956.

[31] *Best Paper*
M. KÖNIG, M. LECLÈRE, M.-L. MUGNIER, M. THOMAZO. *A Sound and Complete Backward Chaining Algorithm for Existential Rules*, in "RR'2012: International Conference on Web Reasoning and Rule Systems", Australie, Springer, 2012, p. 122-138, RR'2012 Best Paper Award, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00764341.

[32] M.-L. MUGNIER. *Existential Rules: A Graph-Based View*, in "2nd International Workshop on Datalog 2.0", Autriche, P. BARCELÓ, R. PICHLER (editors), LNCS, Springer, September 2012, p. 21-26, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00763474.

[33] M.-L. MUGNIER. *Ontology-Based Query Answering with Existential Rules*, in "RuleML'2012: 6th International Symposium on Rules: Research Based and Industry Focused", France, A. BIKAKIS, A. GIURCA (editors), LNCS, Springer, August 2012, 16, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00763475.

[34] B. PAIVA LIMA DA SILVA, J.-F. BAGET, M. CROITORU. *A Generic Platform for Ontological Query Answering*, in "AI'2012: 32nd SGAI International Conference on Artificial Intelligence", Royaume-Uni, Springer, December 2012, p. 151-164, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00763634.

[35] S. RUDOLPH, M. CROITORU. *Multi-Agent Knowledge Allocation*, in "AI'2012: 32nd SGAI International Conference on Artificial Intelligence", Royaume-Uni, Springer, December 2012, p. 165-178, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00763642.

[36] M. THOMAZO, J.-F. BAGET, M.-L. MUGNIER, S. RUDOLPH. *A Generic Querying Algorithm for Greedy Sets of Existential Rules*, in "KR'12: International Conference on Principles of Knowledge Representation and Reasoning", Italie, 2012, p. 096-106, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00763518.

[37] M. THOMAZO. *From EL to Tractable Existential Rules with Complex Role Inclusions*, in "DL'12: International Workshop on Description Logics", Italie, 2012, p. 563-573, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00763586.

National Conferences with Proceeding

[38] P. BUCHE. *Argumentation et décision dans les filières céréalières*, in "JTIC'2012", France, October 2012, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00764993.

[39] A. GHERSEDINE, P. BUCHE, J. DIBIE-BARTHÉLEMY, N. HERNANDEZ, M. KAMEL. *Extraction de relations n-aires interphrastiques guidée par une RTO*, in "CORIA'2012", France, January 2012, p. 179-190, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00764371.

[40] L. LEBRAS, P. BUCHE, J. DIBIE-BARTHÉLEMY, M. ROCHE. *Enrichissement d'une RTO par l'ajout de termes spécialisés*, in "EGC'13: 13ème Conférence Francophone sur l'Extraction et la Gestion des Connaissances", France, January 2013, 2, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00764373.

[41] R. TOUHAMI, P. BUCHE, J. DIBIE-BARTHÉLEMY, L. IBANESCU. *Evolution d'une RTO dédiée à la représentation de relations n-aires.*, in "EGC'2013", France, January 2013, 6, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00764984.

Scientific Books (or Scientific Book chapters)

[42] P. BUCHE, S. DESTERCKE, V. GUILLARD, O. HAEMMERLÉ, R. THOMOPOULOS. *Flexible bipolar querying of uncertain data using an ontology*, in "Springer Series " Studies in Computational intelligence "", Springer, June 2013, 25, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00764985.

[43] M. CROITORU, S. RUDOLPH, N. WILSON, J. HOWSE, O. CORBY. *Graph Structures for Knowledge Representation and Reasoning*, Lecture Notes in Computer Science, Springer, July 2012, vol. 7205, p. 1-77, Second International Workshop, GKR 2011, Barcelona, Spain, July 16, 2011. Revised Selected Papers, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00763674.

Research Reports

[44] M. CHEIN, M. LECLÈRE, Y. NICOLAS. *SudocAD: A Knowledge-Based System for Object Identification*, LIRMM, December 2012, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00765100.

[45] M. CROITORU, S. RUDOLPH. *Towards Multi-Agent Knowledge Allocation*, LIRMM, January 2012, 18, http://hal.inria.fr/lirmm-00657422.

[46] L. GUIZOL, M. CROITORU, M. LECLÈRE. *On Link Validity and Entity Resolution*, LIRMM, January 2012, 17, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00647284.

[47] M. KÖNIG, M. LECLÈRE, M.-L. MUGNIER, M. THOMAZO. *A Sound and Complete Backward Chaining Algorithm for Existential Rules*, LIRMM, 2012, http://hal.inria.fr/lirmm-00713182.

[48] P. MOSSE. *Argumentation based Preference Aggregation*, LIRMM, September 2012, http://hal-lirmm.ccsd.cnrs.fr/lirmm-00735214.

[49] M. THOMAZO, J.-F. BAGET, M.-L. MUGNIER, S. RUDOLPH. *A Generic Querying Algorithm for Greedy Sets of Existential Rules*, LIRMM, 2012, http://hal.inria.fr/lirmm-00675560.

References in notes

[50] M. BIENVENU. *On the Complexity of Consistent Query Answering in the Presence of Simple Ontologies*, in "AAAI", J. HOFFMANN, B. SELMAN (editors), AAAI Press, 2012.

[51] D. LEMBO, M. LENZERINI, R. ROSATI, M. RUZZI, D. F. SAVO. *Inconsistency-Tolerant Semantics for Description Logics*, in "Proceedings of the Fourth International Conference on Web Reasoning and Rule Systems (RR 2010)", 2010.

[52] H. NEWCOMBE, J. KENNEDY, S. AXFORD, A. JAMES. *Automatic Linkage of Vital Records*, in "Science", 1959, vol. 130, p. 954-959.

[53] F. SAÏS, N. PERNELLE, M.-C. ROUSSET. *L2R: A Logical Method for Reference Reconciliation*, in "AAAI", AAAI Press, 2007, p. 329-334.

[54] N. R. SMALHEISER, V. I. TORVIK. *Author name disambiguation*, in "Annual Review of Information Science and Technology", 2009, vol. 43.