wimmics

Web-Instrumented Man-Machine Interactions, Communities, and Semantics.

socio-semantic networks: combining formal semantics and social semantics on the web

Summary: Wimmics[1] is a proposal for a joint research team between INRIA Sophia Antipolis - Méditerranée and I3S (CNRS and University of Nice – Sophia Antipolis). The research fields of this team are graph-oriented knowledge representation, reasoning and operationalization to model and support actors, actions and interactions in web-based epistemic communities.

Main research area: interaction, knowledge, communities, graphs, semantics, web

Main application area: supporting and fostering interactions in online communities

INRIA Field: Perception, Cognition, Interaction

INRIA Theme: Knowledge and Data Representation and Management

1 The research team

Head (and INRIA contact): Fabien Gandon

Vice Head (and I3S contact): Catherine Faron-Zucker

Researchers:

Michel Buffa, MdC (UNS)
Olivier Corby, CR1 (INRIA)
Catherine Faron-Zucker, MdC (UNS)
Fabien Gandon, CR1, HDR (INRIA)
Alain Giboin, CR1 (INRIA)
Nhan Le Thanh, Pr. (UNS)
Isabelle Mirbel, MdC, HDR (UNS)
Peter Sander, Pr. (UNS)

Assistants: Christine Foggia (INRIA) and Magali Richir (I3S)

PhD students:

Pavel Arapov, to start in October (EDSTIC-INRIA)
Adrien Basse, 3rd year (UGB-INRIA)
Franck Berthelon, 3nd year (UNS-EDSTIC)
Ahlem Bouchahda, 3rd year (UNS-SupCom Tunis)
Khalil Riad Bouzidi, 3rd year (UNS-CSTB)
Luca Costabello, 2nd year (INRIA-CORDI)
Corentin Follenfant, 2nd year (SAP)
Maxime Lefrançois, 2nd year (EDSTIC-INRIA)
Nguyen Thi Hoa Hue, 1st year (Vietnam-CROUS)
Nicolas Marie, 2nd year (Bell-ALU, INRIA)
Hasan Rakebul, 1st year (ANR-Kolflow)
Oumy Seye, 2nd year, (Rose Dieng allocation)
Imen Tayari, 3rd year (UNS-Sfax Tunisie)

Post-doc: Elena Cabrio (CORDIS), Serena Villata (ANR-Datalift)

Research engineers:

Julien Cojan (INRIA, Ministry of Culture)
Nicolas Delaforge (INRIA, ANR ISICIL)
Erwan Demairy (ADT)

External collaborators: Claude Frasson (University of Montreal), Alexandre Monnin (La Sorbonne, IRI Georges Pompidou)

[1] wimi is a variety of roses

2 Wimmics: "we mix" Edelweiss and Kewi

The main reason for closing down the Edelweiss INRIA team and starting the Wimmics team was the decease of Rose Dieng-Kuntz former leader of the Edelweiss team. While preparing the new team we decided to merge with the Kewi I3S research team working on closely related topics.

The research team Edelweiss (INRIA), previously known as Acacia, aims at offering models, methods and techniques for supporting knowledge management and collaboration in virtual communities interacting with information resources through the Web and using graph-based and ontology-based formalisms and algorithms. Latest research topics include: multilingual interaction with knowledge bases, mobile access to the web of data, rule-based semantic web formalisms, online presence and resource centric sociality, semantic web in business intelligence, automatic indexing of triple-stores.

3 **Research Challenges:** Analyzing, Modeling, Formalizing and Implementing Graph-based Social Semantic Web Applications for Communities

The web is no longer the simple documentary system built on a simple protocol (HTTP), a simple addressing scheme (URI) and a simple document formatting language (HTML). It grew to become a huge network of distributed data, applications and users where many flows of heterogeneous messages travel. *The web is an object of science: it is a very complex system that requires a multidisciplinary scientific approach.* Our institutes are more and more solicited to understand how these data and interactions can be processed, supported, controlled, exploited or improved. Relying on semantic web formalisms, we believe Wimmics can contribute to this understanding in two manners:

(1) by proposing a multidisciplinary approach to analyze and model the many aspects of these intertwined information systems, communities of users and their interactions;
(2) by formalizing and reasoning on these models to propose new analysis tools and indicators, and support new functionalities and better management.

3.1 Analyzing and Modeling Communities Interactions through Social Semantic Web Applications: interacting with dynamic semantic systems of the web.

The web is a worldwide system never sleeping. In all its dimensions the complexity of the web is growing. The dynamics of the system and its ever-changing contexts make it very difficult to interact with it. The flexibility and extensibility of semantics-based systems is more and more used to tackle the diversity of web resources and applications through metadata describing web resources. But the growing use of semantic representations also augments the complexity of the web and makes it more difficult to interact with.

How do we improve our interactions with such an information system that keeps getting more and more complex? We propose to rely on cognitive studies to build models of the system, the user and the interactions in order to support and improve these interactions. We intend:

- to revisit requirement models and reorganize them in RDFS to support mutual understanding and interoperability between requirements expressed with these heterogeneous models.
- to apply argumentation theory to requirement engineering to improve participant awareness and support decision-making.
- to adapt the "Personas" user modeling technique to include and/or develop relational and emotional aspects appropriate to Web applications
- to pursue ontology-based modeling of users and communities combining generalist schemas (e.g. FOAF) and extensions to capture additional aspects (e.g. schemas of emotional states)

- to adapt to the design of such systems the incremental formalization approach originally introduced in the context of CSCW and HCI communities

Semantics in knowledge representation and in computer science in general are fixed (e.g. semantics based on first order logics) while semantics in social context are renegotiated all the time (e.g. semantics based on "natural logics").

How do we reconcile and integrate the formalized stable semantics of computer science and the negotiable social interactions? We propose to rely on social studies to build models of the communities, their vocabularies, activities and protocols in order to identify where and when formal semantics is useful. We intend:

- to further develop our method for elaborating collective personas, and to compare it to the related “collaboration personas” method.
- to rely on and adapt participatory sketching and prototyping to support the design of interfaces for visualizing and manipulating representations of collectives.
- to consider mixed representations containing social semantic representations (e.g. folksonomies) and formal semantic representations (e.g. ontologies) and study the operations that allow us to couple them and exchange knowledge between them
- to consider compatible linguistic approaches to cover the cases where users interact with a knowledge base using different languages or jargons

When users interact with the web they can use a variety of devices (e.g. mobile phone), of modalities (e.g. vocal interaction), of languages (e.g. Chinese) and be in various contexts (e.g. in the bus).

How do we reconcile local contexts of users and global characteristics of the world-wide virtual machine and information systems that the web has become? We propose to rely on knowledge representation methodologies and theories (e.g. ontologies) to build models of the contexts, devices and mediums ensuring the effectiveness, quality and precision of the information delivered, to provide proofs and explanations of the processes applied and ultimately foster acceptation and trust from the users. We intend:

- to develop ontology based and rule based models that can be merged with RDF based representations of resources, users and systems to integrate their context while reasoning on them.
- to revisit reasoning operations on social semantic web applications to extend them by taking into account the context of their use, of their resources and users
- to study modularity through named graphs, graph annotation and SPARQL extensions to take context into account when managing accesses to a triple store
- to empirically study how the user and the machine articulate and explain their reasoning by the joint use of techniques for analyzing human behaviors and techniques for analyzing machine operations

3.2 Formalizing Models and Implementing Social Semantic Web Applications: calculating on heterogeneous joined typed graphs of the web.

The models identified in the previous section need to be formalized in order to automate their analysis and processing when supporting web applications.

What kind of formalism is the best suited for such models? We defend that the network nature of linked data, social communities and service compositions on the web and the large variety of types of links that compose them call for typed graphs as formalized in languages like conceptual graphs or RDF/S. We intend to build on our experience with such formalisms to identify, propose and characterize fragments of typed graph formalisms best suited for each type of model identified before. We intend:

- to specify the required characteristics of such formalisms and systematically evaluate their effectiveness in implementing these abstract graph models in real applications.

- to work on abstraction of knowledge representation models following conceptual graphs and semantic networks approaches.
- to extend our model and update our abstract machine to dissociate the semantics of the languages and to parameterize graph operators.

Each type of network of the web is not an isolated island. Networks interact with each other: the networks of communities influence the message flows, their subjects and types, the semantic links between terms interact with the links between sites and vice-versa, etc.

How do we analyze these typed graph structures and their interactions? We believe that type-based inference algorithms (e.g. conceptual graph projection, inference rules) and type-parameterized operators (e.g. parameterized betweeness centrality) provide declarative formalisms to flexibly define operations to monitor, filter, query, mine, validate, protect, etc. these imbricated graph structures taking into account constraints spanning several types of network at once. We intend:

- to explore graph representations and algorithms to propose methods and tools to analyze the states of a collective and the relations within it that are mediated by a social semantic web application.
- to continue specify the required characteristics of each operator encountered in our scenarios in terms of graph manipulations and systematically evaluate their effectiveness in implementing them in an abstract graph machine.
- to extend our abstract graph model to cover as many features as possible of the upcoming SPARQL 1.1 language and the RIF rule format.
- to extend the graph operators of our abstract model to integrate approximation, clustering and analysis operations
- to revisit classical structural metrics and adapt their definition to go beyond the pure structural calculation and take into account the types in the graphs
- to study spreading algorithm and extend them to work on typed graphs in particular to propose type-based propagation functions.

These graphs are not available in a single central repository but distributed over many different sources. Some sub-graphs are small and local (e.g. a users' profile on a device), some are huge and hosted on clusters (e.g. Wikipedia), some are largely stable (e.g. thesaurus of Latin), some change several times per second (e.g. social network statuses), etc.

How do we support different graph life-cycles, calculations and characteristics in a coherent and understandable way? We believe that moving to graphs languages with open-world logics, temporal aspects, distributed and loosely coupled algorithms and model-driven programming relying on higher abstractions (e.g. formal ontologies) provides an adequate theoretical framework to allow at the same time the specification and operationalization of the models and algorithms and the opening of these black boxes to be able to explain, document, prove and trace results for the users. We intend:

- to develop an RDF/S model to enable a declarative representation of workflows of operations and to develop an engine able to interpret these representations and execute the workflows.
- to combine this model with standards like SPARQL 1.1 to support heterogeneous operations on possibly heterogeneous and distributed data.
- to study the handling of graphs with possible different semantics (coming from distributed heterogeneous sources) relying on an abstract graph model and an abstract graph virtual machine.
- to support temporal reasoning approaches to include temporal context and change patterns to identify trends, mine temporal propagation to build oriented networks, track behavioral patterns to qualify actors and communities

- to port and extend previous work on automated explanation in expert systems to allow our abstract virtual machine to explain its results and failures

The research team Kewi (I3S), is interested in knowledge engineering techniques and web-based applications to capture, extract analyze, organize, store and share knowledge. Latest research topics include: semantic web and graph-based knowledge modeling; semantic security and access control; affective computing and emotion detection; requirement engineering and collaborative work.

The two teams have been collaborating for now more than seven years and, as shown by their descriptions, exhibit both common interests (web, semantic web, graph-based formalisms, ontologies, etc.) and complementarities (interaction design/ affective computing, triple-stores, rules / access control, etc.). Merging the two teams really is a natural acknowledgement of an ongoing collaboration and numerous co-publications. Moreover this merge will create a group including full-time researchers, assistant professors and professors that could naturally cover all the activities of a research team.

4 Some selected peer-reviewed publications

From books and book chapters:

- Rose Dieng, Olivier Corby, Fabien Gandon, Alain Giboin, Joanna Golebiowska, Nada Matta, Myriam Ribière, Knowledge management: Méthodes et outils pour la gestion des connaissances, 3ème édition, DUNOD, 2005
- Alain Giboin. Correspondances : cadre dialogique pour analyser et assister la coordination entre production et compréhension de documents, in: Production, compréhension et usages des écrits techniques au travail, Toulouse, France, D. Alamargot, P. Terrier, J. Cellier (editors), Octarès, 2005, p. 225-243.
- Fabien Gandon, Ontologies in Computer Science, Book Chapter in Ontology Theory, Management and Design: Advanced Tools and Models, Ed. Faiez Gargouri, Wassim Jaziri, Pages 1-26, ISBN10: 1615208593.
- Alain Giboin (2004). Construction de référentiels communs dans le travail coopératif. In J. Hoc & F. Darses (Eds.), Psychologie ergonomique : tendances actuelles (pp. 119-140). Paris: Presses Universitaires de France.

From Journals:

- Olivier Corby, Rose Dieng-Kuntz, Catherine Faron-Zucker, Fabien Gandon. Searching the Semantic Web: Approximate Query Processing based on Ontologies, IEEE Intelligent Systems Journal, January/February 2006 (Vol. 21, No. 1).
- Michel Buffa, Fabien Gandon, Guillaume Ereteo, Peter Sander and Catherine Faron, SweetWiki: A semantic wiki, Special Issue of the Journal of Web Semantics on Semantic Web and Web 2.0, Volume 6, Issue 1, February 2008 , Edited by Mark Greaves and Peter Mika, Elsevier, Pages 84-97
- Isabelle Mirbel, Ralyte, J. Situational method engineering: combining assembly-based and roadmap-driven approaches. Requirement Engineering Journal, 11(1), 2006, pp. 58--78.
- A. Yurchyshyna, C. Faron-Zucker, N. Le Thanh, and A. Zarli. Knowledge Capitalisation and Organisation for Conformance Checking Model in Construction. International Journal of Knowledge Engineering and Soft Data Paradigms (IJKESDP), 2(1):15–32, 2010.
- Isabelle Mirbel, Pierre Crescenzo, Improving collaboration in the neuroscientist community. International Journal of Web Portals. To appear 2011.
- Amel Yessad, Catherine Faron-Zucker, Rose Dieng-Kuntz, M. Tayeb (2011), Ontology-based Semantic Relatedness for Detecting the Relevance of Learning Resources, In Interactive Learning Environments Journal 19(1), Special issue "Semantic Technologies for Multimedia-enhanced Learning Environments", pp. 63--80.

From international conferences:

- Fabien Gandon, Olivier Corby, Alain Giboin, Nicolas Gronnier, Cecile Guigard, Graph-based inferences in a Semantic Web Server for the Cartography of Competencies in a Telecom Valley, ISWC, Lecture Notes in Computer Science, Galway, 2005
- Olivier Corby and Catherine Faron-Zucker, The KGRAM Abstract Machine for Knowledge Graph Querying, IEEE/WIC/ACM International Conference, September 2010, Toronto, Canada.
- Olivier Corby, Web, Graphs & Semantics, Proc. of the 16th International Conference on Conceptual Structures (ICCS'2008), Invited talk, July 2008 Toulouse
- Olivier Corby, Rose Dieng-Kuntz, Catherine Faron-Zucker.Querying the Semantic Web with the Corese search engine. Proc. of the 16th European Conference on Artificial Intelligence (ECAI PAIS'2004), Valencia, 22-27 August 2004, IOS Press, p. 705-709.
- Guillaume Ereteo, Fabien Gandon, and Michel Buffa, SemTagP: Semantic Community Detection in Folksonomies, IEEE/WIC/ACM International Conference on Web Intelligence, WI 2011
- Guillaume Erétéo, Michel Buffa, Fabien Gandon, and Olivier Corby. Analysis of a Real Online Social Network using Semantic Web Frameworks. In Proc. International Semantic Web Conference, ISWC'09, Washington, USA, October 2009.
- Amira Tifous, Adil El Ghali, Alain Giboin, Rose Dieng-Kuntz. O'CoP, an Ontology for Communities of Practice, in "Networked Knowledge - Networked Media. Integrating Knowledge Management, New Media Technology and Semantic Systems", T. Pellegrini, S. Auer, K. Tochtermann, S. Schaffert (editors), Studies in Computational Intelligence, vol. 221, Springer, 2009, p. 155-169.
- Freddy Limpens, Fabien Gandon and Michel Buffa, Helping Online Communities to Semantically Enrich Folksonomies, Web Science Conference, April, 2010, Raleigh, NC, USA.