ERC Advanced Grant 2014

Research proposal [Part B1]

Alternative Web Ontology Language

AWOL

Name of the Principal Investigator (PI): HASSAN AÏT-KACI
Name of the PI's host institution for the project: UNIVERSITÉ CLAUDE BERNARD LYON 1
Proposal duration in months: 60 MONTHS (5 YEARS)

Proposal summary The "*Semantic Web*" is a collaborative research effort promoted by the World-Wide Web Consortium with aim to add meaning to data linked over the Internet. Little that was planned has been achieved. For it to succeed, we must redesign it from scratch! Our proposal offers to design an experimental net-aware knowledge representation and reasoning software system in this context. What we propose departs from the current state of the art with original ideas from constraint-based reasoning adapted for distributed architectures. Contrary to standard approaches, we offer to use Constraint Logic Programming (CLP) and Functional Programming (FP) to encode and process semantic information in the form of constrained graphs interconnected on the Internet. Specifically, we intend to extend such graphs with nodes spanning scoped Internet clouds. Such graphs can be readily encoded using the Resource Description Framework (RDF) and organized in networks of distributed so-called "triplestores." The *AWOL* project will consist of four interwoven tracks: 1. **specification track**—definition of a network-aware knowledge representation and reasoning system; 2. **implementation track**—development of a proof-of-concept prototype implementing the specification on a massively scalable architecture and processing engine capable of managing distributed knowledge and data; 3. **experimentation track**—realistic use case scenarios and contexts; experiment with intelligent communication protocols capable of handling incremental Semantic Web computing; 4. **demonstration track**—successive versions of the *AWOL* system running over a cloud. Anywhere decidability of reasoning is not guaranteed, it will systematically be opted to choose the largest possible decidable sublogic, even if at the cost of incompleteness. Efficiency challenges can also be accommodated whenever incremental reasoning may be ensured by "lazy" (*i.e.*, only if/when needed) processing of information as it materializes, whether synchronously or not.

Section a: Extended Synopsis of the Scientific Proposal

What is being proposed in this document:

- touches on one of the most ambitious Computer Science research objectives of the past couple of decades; *viz.*, Sir Tim Berners-Lee's dream of a "*Semantic Web*;"
- offers to use a formal, practical, tested, effective, and efficient formalism for representing knowledge, and automated reasoning therefrom, that has so far been neglected by mainstream research in the field; and,
- takes on the challenge of demonstrating the practical effectiveness of this formalism for the Semantic Web, as being on a par with, and often more *efficient* than, many of the tools and techniques thus far proposed.

My essential scientific motivation in submitting the AWOL project proposal is the concern that mainstream research has so far failed to deliver what was envisaged. Quoting Wikipedia:

> *The original 2001 Scientific American article by Berners-Lee, Hendler, and Lassila described an expected evolution of the existing Web to a Semantic Web, [18] but this has yet to happen. In 2006, Berners-Lee and colleagues stated that: "This simple idea . . . remains largely unrealized." [25].*

And, one might add, such is still the case as years keep ticking by [5, 7]. My opinion as to the essential cause for this state of affairs is the unsatisfactoriness of the mainstream Semantic Web formalism—*viz.*, Description Logic (DL)—and technology—*viz.*, the "Web Ontology Language" (OWL) adopted by the W3C as an *exclusive* standard for Semantic-Web knowledge representation and reasoning.[1]

One important question emerges as a consequence of this; *viz.*, "***Should the Semantic Web be abandoned as a wild-goose chase?***" *I, personally, most emphatically do not think so!* Rather, what should be revised is the limiting of the means toward it to a single logical formalism and technology, while neglecting potential alternatives.[2] Putting all the SW's eggs in the same (OWL) basket simply makes no sense [5]. Also, and concommitantly, we should readjust our expectations, taking into account the pragmatics of realistic Internet processing.

Much in Constraint Logic Programming (CLP) technology has been understood and put to efficient use [24, 22, 23]. I contend that this know-how that can help the realization of a significantly more effective and efficient than current mainstream DL-based technology, yet still remarkably expressive.

Indeed, offering a formalism for the Semantic Web, however expressive it may be, is not sufficient. It is as important that such a formalism be as well operationally *efficient*; in particular, in the pragmatics of the Internet—*viz.*, how can knowledge representation and reasoning be made *scalable* in a *distributed* environment. In this proposal, I am offering to put to use such know-how, which I pioneered, in the formal design and *efficient implementation* of constraint-based formal knowledge-representation and automated-reasoning systems [1, 13]. Thus, what is novel in what I am proposing is to exploit implementation techniques developed for CLP for constraint processing, but adapt it to fit specifically the Semantic Web's context.

> The AWOL project proposes to leverage the formal understanding of *Constraint Logic Programming* and the efficient practical techniques that have been developed for it, in the context of the Semantic Web.

There are several reasons for expecting this heretofore neglected combination of fields to be beneficial to the Semantic Web in terms of efficiency. First, in the same way Constraint Processing benefitted pure Logic Programming by factoring out efficient specific reasoning from general-purpose reasoning, it should do so as well for general-purpose SW reasoning. Second, efficient *compilation techniques* such as Warren's Abstract Machine [1] can be used [12, 16]. Finally, taxonomic reasoning, which is central to ontology processing, can be made drastically more performant on actual knowledge bases of enormous size—as we have recently demonstrated [9, 11]. However, challenges in achieving high efficiency of these techniques in the specific area of the Semantic Web are also non trivial. The most important issues concern adapting them to large *distributed* information.

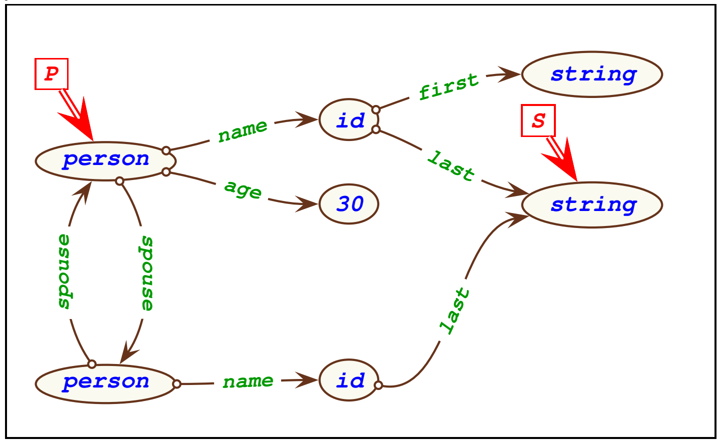
> Our novel key technical insight is to ***view data structures as constraints***.

[1] See for example [21, 26].

[2] So far, Semantic Web research has essentially been assimilated to Description Logic research. This unfortunately confuses one potential *means* (DL) with its *end* (SW), when this end ought to be an objective independent of any specific approach.

This insight is what the "*Order-Sorted Feature*" ($\mathcal{OSF}$) constraint formalism exploits [3]. It allows formulating ontologies where concepts are seen as constraints describing distributed graph structures subject to relational and functional properties. In this way, it is possible to specify knowledge declaratively while enabling operationally enforceable conceptual properties [3]. In other words, it makes it possible to define ontological concepts organized into inheritance taxonomies and subject to efficiently enforceable constraints.

The $\mathcal{OSF}$ formalism essentially proposes to see (and represent) ***everything*** as a labelled graph: nodes (called *tags*) and arrows (called *features*) between nodes. The nodes are labelled with partially-ordered sorts (denoting values, sets, and sets of sets) where the order relation is set containment; the arrows are labelled with functional attribute names. In the box to the right is an example of an $\mathcal{OSF}$ graph.




```
P : person(name ⇒ id(first ⇒ string,
                     last ⇒ S : string),
           age ⇒ 30,
           spouse ⇒ person(name ⇒ id( ⇒ S),
                           spouse ⇒ P)).
```

Syntactically, $\mathcal{OSF}$ graphs may be viewed as $\mathcal{OSF}$ terms expressing constraint structures that are made exclusively using three entities (denoted syntactically as identifiers obeying various conventions): (1) *tags*; (2) *sorts*; and (3) *features*.

For example, the $\mathcal{OSF}$ term on the right is a possible syntax for the $\mathcal{OSF}$ graph above. Clearly, $\mathcal{OSF}$ terms are in 1–1 correspondence with $\mathcal{OSF}$ graphs, up to permutation of subterm order.

There is much to gain to see such graphs as constraints that can be expressed using syntax generalizing that of First-Order Terms used in CLP [8, 3]. Much of this was demonstrated in practice in the CLP language $\mathcal{LIFE}$ [2, 13].[3] Tags generalize the so-called "*logical variables*" of (Constraint) Logic Programmming. They are (scoped) addresses referring to type and data structures—types constraining the syntactic and semantic nature of data. As the structures to which they refer are refined by constraints that define them, all constraint-enforcing events are synchronized through tags that may appear in constraints. Indeed, any constraint in need of further information from a tag's binding will suspend as long as this information is missing. Binding and refining a tag's sorted structure triggers all constraints waiting for it to be awakened. Thus does constraint-driven reasoning become an automated event-driven process managing the concurrent orchestration of verifying codependent structures [15].

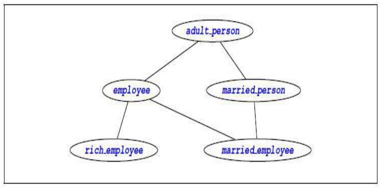
Sorts generalize the notion of type and value. As in the Scala programming language, *all* types are subtypes of *any*, the type of all types—*i.e.*, built-in "unboxed" types, such as characterm, integer, floating-poing or double-precision numbers, *etc.*, are too: unboxing is automatic and off-the-shoulders of the user. I will be referring to Scala because it is to be the implementation language used in the $\mathcal{AWOL}$ project (see Part B2). Features denote functional attributes of any node they originate from and point to a node sorted either by (1) a sort—including the least-constraining sort `Any`, the maximal all-encompassing element ($\top$) of the taxonomic sort ordering; or, (2) a value (a minimal sort); or, (3) an aggregate sort or value (of which: a **set-of** sort, or a set of values).

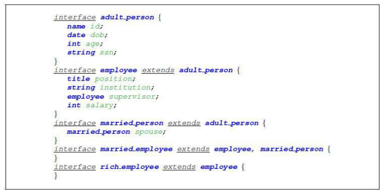
Basic $\mathcal{OSF}$ term matching and unification can be extended to more expressive systems [4], such as capable of handling: non-lattice sort taxonomies; disjunction; negation; partial features; extensional sorts (*i.e.*, value-denoting sorts); relational features (*a.k.a.*, "roles" in DL—*i.e.*, powerset-valued features), and more generally any monoidal aggregates (collections *à la* monoid comprehensions [20]);[4] regular-expression feature paths; and, (possibly recursive) sort definitions (*a.k.a.*, "$\mathcal{OSF}$ theories"). Most importantly, all this can be done by rule-driven constraint normalization processes preserving incrementality and independent of order of application—a key property of the system to be made concurrent. In addition, arbitrary definable functions and predicates can in turn be used in the definitions of concepts as constrained sorts in a taxonomy and implemented efficiently using well-established compilation techniques from Logic Programming (LP) and Functional Programming (FP).[5]

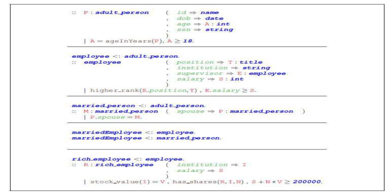
[3] $\mathcal{LIFE}$ stand for (programming with) $\mathcal{L}$ogic, $\mathcal{I}$nheritance, $\mathcal{F}$unctions, and $\mathcal{E}$quations.

[4] Scala's `for` construct is a (restricted) form of a monoid comprehension.

[5] ***N.B.*: Please click on each of the three following thumbnails for larger views illustrating the preceding statements.**







These capabilities will be realized in the design and implementation of "𝓗𝓞𝓞𝓣," a concept-oriented distributed constraint logic language to be developed and used in the 𝒜𝒲𝒪ℒ project.[6] The plan is to realize an abstract machine *à la* **Warren Abstract Machine** [1, 12, 16] in Scala and compile 𝓗𝓞𝓞𝓣 constructs into it.

This proposal in context My strongest motivation for proposing the 𝒜𝒲𝒪ℒ project is the work I have done within 𝒞ℰ𝒟𝒜ℛ project.[7] The preliminary results obtained, and experience acquired, in this project have opened exciting perspectives that are strongly encouraging: namely, (1) astonishingly fast performance of implementation of taxonomic reasoning [10, 17, 11], and (2) deep understanding through experimentation with the current state of the art in scalable distributed triplestore processing [14], and the on-going design and implementation of our own distributed file system [19].[8]

Hence, this justifies my taking the bold move to submit this proposal even though it is in clear rupture with the state of the art in Semantic Web reasoning, of which we exposed the serious drawbacks regarding scalability and distribution. Still, the issues to be faced and solved by the 𝒜𝒲𝒪ℒ project represent daunting challenges. Nevertheless, I proffer that its feasibility is technically coherent and my plan reasonably cogent. This is further debated next.

Project feasibility I next put what 𝒜𝒲𝒪ℒ is proposing through a set of legitimate questions one may be prompted to ask reading this document, and briefly address each in turn. More details will be provided in Part B2.

Q: **Most extant ontological specifications are already formulated using OWL; how can 𝒜𝒲𝒪ℒ be reasonably expected to process or use them in any way?**

A: This is simply a compiling issue. OWL (and any successive version thereof) has a syntax and a semantics. So will 𝓗𝓞𝓞𝓣, the ontological programming language we intend to develop. It should be a simple matter of compiling (various versions of) OWL into (appropriate versions of) 𝓗𝓞𝓞𝓣. In fact, this is the way we have proceeded in the 𝒞ℰ𝒟𝒜ℛ project, compiling DL-based OWL constructs into semantically equivalent 𝒪𝒮ℱ-based constructs. This is how we could do our initial comparative performance analysis to demonstrate the higher efficiency of our taxonomic reasoning techniques [9, 11].

Q: **How can you expect to achieve in 5 years and with about 10 staff members (postdocs and grad students) what hundreds of highly qualified researchers and developers could not in almost 15 years?**

A: The 𝒜𝒲𝒪ℒ project is only proposing making a *contribution* toward an adequately working Semantic Web. However, it offers to resolve effectively the current failure of state of the art by using a radically different way for representing distributed knowledge and reasoning in a scalable incremental fashion. The team I am proposing to put together to meet this daring challenge will consist of, besides myself, two highly qualified post-docs, one wizard engineer, four selectively chosen PhD students, and four MSc students of similar excellence. In addition, I have planned collaborations in the form of student co-supervision with world-renown researchers at SRI International, the University of Louvain, the Technical University of Madrid, and the University of Minnesota. (See details in B2.)

[6] 𝓗𝓞𝓞𝓣 stands for 𝓗ierarchical 𝓞ntologies with 𝓞bjects and 𝓣ypes. More details are given in Part B2.

[7] http://cedar.liris.cnrs.fr/

[8] There are two post-𝒞ℰ𝒟𝒜ℛ project proposals (excluding 𝒜𝒲𝒪ℒ's) that I have submitted to this day: (1) ℒivℰℳ𝒰𝒮ℐ𝒞—ℒiving ℰnvironment ℳonitoring 𝒰se 𝒮cenarios with ℐntelligent 𝒞ontrol. This was submitted to a French regional research program; *Programme Avenir Lyon St-Etienne* (PALSE 2014). This is to provide a credible use case in the area on living-environment monotoring. (2) ℬℒℐ𝒩𝒦—ℬig ℒinked ℐnformation as 𝒩etworked 𝒦nowledge. This was sumbitted to the French national research program: ANR @TRaction 2014, whose objective is attracting high-level international researchers to France. It is expected (although not required) to provide an advanced concurrent distributed environment for 𝒜𝒲𝒪ℒ. See Appendix Section, Page 11.

Q: **Alternative? Fine. But why this one rather than another?**

A: It is by no means affirmed that ours is an *exclusive* approach—most emphatically not! There are other formalisms that could provide such possible alternatives—Rudolf Wille's Formal Concept Analysis (FCA), John Sowa's Conceptual Graphs (CG), to name a couple. They could, and should, be considered and investigated as such, although not incompatible, alternatives to the mainstream in the same manner I am proposing *OSF* constraint logic to be. Each (including DL and *OSF* logics) can constitute a part of a whole in the pursuit of the Semantic Web. I believe that success of the Semantic Web will come from a federated effort toward a common objective, rather than the single exclusive means that is currently being *de facto* imposed toward this end.

Q: **Why choose Scala as *AWOL*'s development and *HOOT*'s compilation target language?**

A: This is answered in more technical details in Part B2. Suffice it here to say that I have been a fervent Java developer since it appeared twenty years ago. So I can really appreciate the many improvements Scala has brought to Java while staying fully compatible at the JVM level.[9] In fact, for the past 15-or-so years, I have built for myself comprehensive language-design Java libraries and tools: the Jacc system. With hindsight, many of these were motivated by the need to palliate Java's shortcomings, most of which are now supported by Scala [6]. The fact that Scala is Java-compatible makes it possible for me to keep using my tool set, which has proven quite effective for language rapid-prototyping (just as for our *CEDAR* language design track). Although not strictly needed, my intention is to convert most of it from Java to Scala (using Jacc to bootstrap the conversion process).

Q: ***AWOL* relies on *BLINK* for cloud deployment.[10] What if *BLINK* does not get funded?**

A: Ideally, the *AWOL* project would greatly benefit from exploiting the results of the *BLINK* project I have recently submitted to the French *Agence Nationale de la Recherche*. As such, it would provide our ideal vision for a scalable distributed architecture—the one for which we have laid the ground work in the *CEDAR* project. This is because it is meant to be specifically designed for the nature and structure of the *HOOT* language. I can say this with assurance since what I have proposed with *BLINK* also stems from our in-depth investigation and experimentation with the best existing distributed file systems for triplestores in the course of the *CEDAR* project. Be that as it may, however, *any* extant distributed file system could also be used, if only with less convenient an interface, and perhaps also less optimal performance.

Q: ***LIFE* is a major reference in this proposal—if it was such a clever design, why hasn't it endured as a CLP language?**

A: For three reasons: (1) it was being developed at DEC's Paris Research Lab when Digital collapsed in 1994, and the company diseappeared; (2) LP went in relative disfavor in the late 90's due to the Japanese 5th generation failure;[11] and, (3) its principal designer (myself) undertook a personal career change at the end of the 90's away from research and closer to Industry.

High risk/high gain? What is proposed in the *AWOL* project is admittedly a difficult challenge. In spite of my feeling confident in its feasibility, I agree that there is a real risk of falling short of the effciency I expect to contribute toward Semantic Web processing. Indeed, while CLP technology has been shown to boost efficiency in reasoning and problem-solving, this has been the case in *confined* systems, not systems *distributed* over the Internet. This novel aspect in the proposed research presents several open issues such as how to deal efficiently with communication latency, incremental information processing, *etc.*, in addition to pure reasoning challenges. Facing (and resolving) such issues is what makes this project highly non-trivial. However, there is never any loss at pursuing challenging goals, even if they turn out to be unattainable—*e.g.*, JFK's Moon Technology.

On the other hand, as I hope to have illustrated, the challenges the *AWOL* project is taking on, if met, would contribute to breaking a few momentous locks in the pursuit toward making the Semantic Web a reality. This, methinks, would arguably outweigh the risks acknowledged above.

[9] In a way, Scala is to Java what I would like *HOOT* to be to OWL.

[10] See footnote 8 on previous page.

[11] This was a failure at the hardware level, though not at the software level. But the former unrealistic pursuits made the latter fall in disrepute as well. Today, most schools barely teach LP.

References

[1] AÏT-KACI, H. *Warren's Abstract Machine: A Tutorial Reconstruction*. The MIT Press, Cambridge, MA, USA, 1991. (online)

[2] AÏT-KACI, H. An introduction to LIFE—Programming with Logic, Inheritance, Functions, and Equations. In *Proceedings of the International Symposium on Logic Programming* (October 1993), D. Miller, Ed., MIT Press, pp. 52–68. (online)

[3] AÏT-KACI, H. Data models as constraint systems—a key to the Semantic Web. *Constraint Processing Letters 1* (November 2007), 33–88. (online)

[4] AÏT-KACI, H. Objects as constraints—a formalism of order-sorted featured structures. Lecture notes, 2nd International School on Rewriting (ISR 2007), Nancy, France, July 2007. (online)

[5] AÏT-KACI, H. Children's magic won't deliver the Semantic Web. Letter to the Editor, Communications of the ACM, March 2009. (online)

[6] AÏT-KACI, H. An abstract, reusable, and extensible programming language design architecture. In *Essays dedicated to Peter Buneman—In Search of Elegance in the Theory and Practice of Computation* (Edinburgh, UK, October 27–28 2013), V. Tannen, L. Wong, L. Libkin, W. Fan, W.-C. Tan, and M. Fourman, Eds., Lecture Notes in Computer Science Volume 8000, Springer-Verlag, pp. 112–166. (online)

[7] AÏT-KACI, H. Is it possible to make the Semantic Web a reality? Keynote presentation at the 3rd edition of the international workshop on Innovation and New Trends in Information Systems (INTIS 2013), Tangiers, Morocco, November 22–23, 2013. (online)

[8] AÏT-KACI, H. Reasoning and the Semantic Web. Invited presentation at the Ontolog Forum's Panel on Ontology, Rules, and Logic Programming for Reasoning and Applications (RulesReasoningLP) mini-series of virtual panel sessions, November 21, 2013. (online)

[9] AÏT-KACI, H., AND AMIR, S. Classifying and querying very large taxonomies—a comparative study to the best of our knowledge. CEDAR Technical Report Number 2, *CEDAR* Project, LIRIS, *Département d'Informatique, Université Claude Bernard Lyon 1*, Villeurbanne, France, May 2013. (online)

[10] AÏT-KACI, H., AND AMIR, S. Classifying and querying very large taxonomies with bit-vector encoding. *(submitted)* (2013). (online)

[11] AÏT-KACI, H., AND AMIR, S. Design and implementation of an efficient semantic web reasoner. CEDAR Technical Report Number 12, *CEDAR* Project, LIRIS, *Département d'Informatique, Université Claude Bernard Lyon 1*, Villeurbanne, France, October 2014. (online)

[12] AÏT-KACI, H., AND DI COSMO, R. Compiling order-sorted feature term unification. PRL Technical Note 7, Digital Paris Research Lab, Rueil-Malmaison, France, December 1993. (online)

[13] AÏT-KACI, H., DUMANT, B., MEYER, R., PODELSKI, A., AND VAN ROY, P. The Wild LIFE handbook (prepublication edition, 1994). (online)

[14] AÏT-KACI, H., HACID, M.-S., HAQUE, R., AND FOURURE, D. Experiments with scalable triplestores. CEDAR Technical Report Number 5, *CEDAR* Project, LIRIS, *Département d'Informatique, Université Claude Bernard Lyon 1*, Villeurbanne, France, October 2013. (online)

[15] AÏT-KACI, H., AND PODELSKI, A. Functions as passive constraints in LIFE. *ACM Transactions on Programming Languages and Systems 16*, 4 (July 1994), 1279–1318. (online)

[16] AÏT-KACI, H., AND VORBECK, M. Compiling order-sorted feature unification. Unpublished Technical Report, February 1996. (online)

[17] AMIR, S., AND AÏT-KACI, H. *CEDAR*: a fast taxonomic reasoner based on lattice operationssystem demonstration. In *Proceedings of the Posters & Demonstrations Track of the 12th International Semantic Web Conference* (Sydney, Autralia, October 2013), E. Blomqvist and T. Groza, Eds., CEUR Workshop Proceedings, pp. 9–12. (online)

[18] BERNERS-LEE, T., HENDLER, J., AND LASSILA, O. The Semantic Web. *Scientific American* (May 13, 2001). (online)

[19] CHEN, M., HAQUE, R., AND HACID, M.-S. CedTMart—a triplestore for storing and querying blinked data. CEDAR Technical Report Number 2, *CEDAR* Project, LIRIS, *Département d'Informatique, Université Claude Bernard Lyon 1*, Villeurbanne, France, July 2013. (online)

[20] FEGARAS, L., AND MAIER, D. Optimizing object queries using an effective calculus. *ACM Transactions on Database Systems 25* (1998), 25–4. (online)

[21] GOODWIN, J. Experiences of using OWL at the ordnance survey. In *Proceedings of the OWLED'05 Workshop on OWL: Experiences and Directions* (Galway, Ireland, November 11–12, 2005), B. C. Grau, I. Horrocks, B. Parsia, and P. Patel-Schneider, Eds., CEUR Workshop Proceedings Vol. 188. (online)

[22] JAFFAR, J., AND LASSEZ, J.-L. Constraint logic programming. In *Proceedings of the 14th ACM Symposium on Principles of Programming Languages (POPL'87)* (New York, NY, USA, 1987), Association for Computing Machinery, pp. 111–119. (online)

[23] JAFFAR, J., AND MAHER, M. J. Constraint logic programming—a survey. *The Journal of Logic Programming 19* (July 1994), 503–581. (online)

[24] SARASWAT, V., AND RINARD, M. Concurrent constraint programming. In *Proceedings of the 17th Annual ACM Symposium on Principles of Programming Language (POPL'90)* (New York, NY, USA, 1990), F. E. Allen, Ed., Association for Computing Machinery, pp. 232–245. (online)

[25] SHADBOLT, N., HALL, W., AND BERNERS-LEE, T. The Semantic Web revisited. *IEEE Intelligent Systems* (2006), 96–101. (online)

[26] SRINIVAS, K. OWL reasoning in the real world: Searching for Godot. In *Proceedings of the 22nd International Workshop on Description Logics (DL 2009)* (Oxford, United Kingdom, July 27–30 2009), B. C. Grau, I. Horrocks, B. Motik, and U. Sattler, Eds., CEUR Workshop Proceedings. Invited lecture (online)

Section b: Candidate's Curriculum Vitae

PERSONAL INFORMATION

Family name, First name: **Aït-Kaci, Hassan**
Researcher unique identifier(s): LinkedIn site: **https://www.linkedin.com/in/hak2007**
Nationality: **France**
Date of birth: **June 14, 1954**
URL for web site: **hassan-ait-kaci.net**

EDUCATION

1990 Habilitation, Informatique, Université Pierre et Marie Curie (Paris 7), Paris, France—supervisor: **Maurice Nivat**; committee: **Jean Gallier**, **Jean-Jacques Lévy**, **Patrick Baudelaire**, **Guy Cousineau**, **Irène Guessarian**, **Gérard Huet**

1984 PhD, Computer and Information Science, University of Pennsylvania, Philadelphia, PA, USA—supervisor: **Peter Buneman**; committee: **Alex Borgida**, **Timothy Finin**, **Jean Gallier**, **David MacQueen**, **Fernando Pereira**

1982 MSc, *ibid.*

CURRENT POSITION

2013 – 2015 ANR Chair of Excellence, Computer Science, LIRIS, Université Claude Bernard (Lyon 1), Villeurbanne, France

PREVIOUS POSITIONS

2009 – 2012 Senior Technical Staff Member, Application and Integration Middleware, IBM, Burnaby, BC, Canada

2003 – 2009 Distinguished Scientist, ILOG, Gentilly, France

2000 – 2003 Extended Sabbatical Visitor, *ibid.*

1994 – 2001 Tenured Full Professor of Computing Science, Simon Fraser University, Burnaby, BC, Canada

1988 – 1993 Senior Member of Research Staff, Digital Paris Research Lab, Rueil-Malmaison, France

1991 – 1992 Adjunct Professor of Computer Science, Université de Paris 7 and École Normale Supérieure, Paris, France

1984 – 1988 Member of Research Staff, Microelectronics and Computer Technology Corporation, Austin, TX, USA

1984 – 1988 Adjunct Assistant Professor, Department of Computer Sciences, University of Texas, *ibid.*

FELLOWSHIPS AND AWARDS

2013 – 2015 PI, ANR Chair of Excellence, *CEDAR* Project (Constraint Event-Driven Automated Reasoning) LIRIS, *Université Claude Bernard Lyon 1*, Villeurbanne, France

2004 – 2008 Co-PI, MANIFICO Project (*Métacompilation Non-Intrusive du Filtrage par Contraintes*), *Réseau National sur la Technologie des Langages, Ministère de la Recherche et Ministère de l'Économie, des Finances et de l'Industrie*; joint industrial research project with LORIA, Nancy, France; INRIA, Rocquencourt, France; ILOG, Gentilly, France

1998 – 2000 PI, NSERC Research Grant, Investigation of a Formalism for Indefinite Computation and Approximation, School of Computing Science, Simon Fraser University, Burnaby, BC, Canada

1994 – 1999 PI, MPR Teltech NSERC Industrial Chair in Intelligent Software Systems, *ibid.*

1994 – 1995 Fellow of the Advanced Systems Institute of British Columbia, Vancouver, BC, Canada

1993 – 1994 Co-PI, ESPRIT Basic Research Action CONTESSA (Constraints and Extended Support for Storage and Access); consortium gathering 8 european academic and industrial research centers; Digital Paris Research Lab, Rueil-Malmaison, France

1992 – 1994 Co-PI,: ESPRIT Basic Research Action ACCLAIM Project (Advancing Concurrent Constraint LAnguages Implementation and Methodology); consortium gathering 9 european academic and industrial research centers;*ibid.*

SUPERVISION OF GRADUATE STUDENTS AND POSTDOCTORAL FELLOWS

2013 – 2015 2 Postdocs, 4 MSc's, Université Claude Bernard Lyon 1, France; 1 State Engineer; Institute National d'Informatique, Algeria

1996 – 1999 3 Postdocs, 2 PhD's, 1 MSc's, 2 Engineering projects, 2 visiting researchers; Simon Fraser University, Canada

1990 – 1996 1 PhD (*Université de Paris 7, France*), 2 MSc (*École Normale Supérieure*, Paris, France), 1 Research Engineer (*École des Mines*, Paris, France)

TEACHING ACTIVITIES

1996 – 2000 SFU, School of CS: Introduction to Internet Programming in Java and its Friends, MS Windows Application Programming with C++, Principles of Programming Languages, Symbolic Computing (undergraduate); Seminar on advanced Internet application design (graduate).

1991 – 1992 École Normale Supérieure, Paris, France: Graduate seminar on Multi-Paradigm and Constraint-Based Programming.

1984 – 1988 University of Texas at Austin, CS Department: Discrete Mathematics for Computer Science, Principles of Programming Languages (undergraduate); Introduction to Artificial Intelligence, Seminar on Symbolic Computation (graduate).

1982 – 1983 University of Pennsylvania, CIS Department: Database Theory, Compiler Design, Theory of Computation (undergraduate).

ORGANISATION OF SCIENTIFIC MEETINGS

2013 Big Data Forum, LIRIS, Université Claude Bernard Lyon 1, France—co-organized by the *CEDAR* and PetaSky projects; *ca.* 20 participants.

2012 RuleML at ECAI Challenge and Doctoral Consortium, International Symposium on Rules, Montpellier, France—co-organized with Yuh-Jong Hu, Grzegorz J. Nalepa, Monica Palmirani, and Dumitru Roman; *ca.* 40 participants.

1987 International Conference on Resolution of Equations in Algebraic Structures (CREAS), Lakeway TX, USA; co-organized with Maurice Nivat of LITP, Université de Paris 7, France—co-sponsored by MCC and INRIA; *ca.* 80 participants.

INSTITUTIONAL RESPONSIBILITIES

2013 Member of INRIA's juries for its *concours de Directeur de Recherche 2ème classe* (DR2), June 2013, INRIA, Paris; and *concours de Chargé de Recherche 2ème classe* (CR2), May 2013, INRIA, Rhône-Alpes, France

2003 Member of INRIA's jury for its *concours de Chargé de Recherche 2ème classe* (CR2), May/June 2003, INRIA, France

2002 Member of INRIA's Evaluation committee for Software Engineering and Symbolic Computing, October 2002, INRIA, France

COMMISSIONS OF TRUST

2004 – 2009 ILOG's principal representative to the "Rule Interchange Format" (RIF) W3C Working Group

1994 – present Editorial board member, Methods of Logic in Computer Science, Ralph Wilkerson and Rick L. Smith, eds., Ablex Publishing Corporation, Norwood, NJ, USA

1988 – 2004 Referee and/or external examinator for several PhD theses and Research Habilitations in France; non-exhaustively: at Université de Nancy/LORIA (**Pierre-Étienne Moreau**); at École Polytechnique (**Gilles Dowek**, the late **Alain Deutsch**); at Université de Paris-Sud Orsay (**François-Xavier Josset**); at Université Joseph Fourier Grenoble, (**Cécile Capponi**, **Rachid Echahed**, **Jerôme Euzenat**, **Yves Lepage**); at Université Claude Bernard Lyon (**Mohand-Saïd Hacid**); *etc.*, as well as member in several INRIA review committees, including the *Conseil Scientifique d'Évaluation du Laboratoire d'Informatique de l'École Polytechnique* (LIX), Palaiseau, France, under the late **Gilles Kahn**'s chairmanship.

FORMER MENTORS I had the great fortune to collaborate with **Maurice Nivat**, an international figure largely and highly regarded as the father of Computer Science research in France, and a major influencer for making Computer Science a *formal science*, applied especially to establishing rigorous program correctness. We organized an international Colloquium on Resolution of Equations in Algebraic Structures gathering the best world research figures in formal automated reasoning in 1987, and in 1989, we co-edited a two-volume book at Academic Press. He was also my habilitation supervisor. I also owe much to **Peter Buneman**, my PhD supervisor at Penn. He taught me all I know in Computer Science and made me understand what scientific research is all about.

FORMER MENTEES I managed several reputed international research teams in high-grade institutions, both in Academia and Industry. Several of my mentees have now become internationally recognized research figures. **At MCC: Pat Lincoln**, now Director of Stanford Research Institute's Computer Science Lab; **at MCC and PRL**: **Gert Smolka**, now Professor at Saarland University; **at PRL: Andreas Podelski**, now professor at the University of Freiburg; **Peter Van Roy**, now Professor of Computer Science at Université Catholique de Louvain; **Roberto Di Cosmo**, now Professor of Computer Science at University of Paris Diderot; **Jacques Garrigue**, now Associate Professor at Nagoya University; **at SFU: Yutaka Sasaki**, now Professor of Advanced Science and Technology at Toyota Technological Institute; **Denys Duchier**, now professor at University of Orléans. Please refer to **my LinkedIn site** for more, as well as public endorsements of former collaborators.[12]

[12] **https://www.linkedin.com/in/hak2007**

Section c: Ten years track-record

Hassan Aït-Kaci[13] has 30+ years of solid experience in advanced research in programming systems for intelligent software design and implementation. He is widely recognized as a pioneer in formalizing inheritance in knowledge representation using Constraint Logic Programming. He has acquired this experience both in first-class industrial and academic internationally renowned research institutions, in the United States, Canada, and France. He has also acquired an excellent hands-on experience in Industry on transfer of research into software product features (2000–2012 at ILOG, acquired by IBM in 2009).

Detailed introduction Hassan Aït-Kaci is currently occupying an ANR Chair of Excellence, at the LIRIS in the *Université Claude Bernard Lyon 1* (UCBL): the *CEDAR* Project ("Constraint Event-Driven Automated Reasoning").[14] This project started on January 15, 2013, and will end on January 14, 2015. The project's plan is an experimental exploration of how Semantic Web technology can live up to full-scale distribution.[15]

Before January 2013, Hassan Aït-Kaci held the position of Senior Member of Technical Staff at IBM Canada.[16] He became so after IBM's acquisition of ILOG in February 2009, the French INRIA spin-off that made its business success and fame in the technology of constraint-processing and business rules. Dr. Aït-Kaci's interests and contributions have been in automated reasoning, knowledge representation, declarative programming, and language processing. In these areas, he has been a fervent advocate of constraint-based computing as the versatile key to essential locks that all these subjects have in common, and that we are facing in the pursuit of making the Semantic Web an intelligent reality. He had joined ILOG in 2000, as a Distinguished Scientist, originally on leave from Simon Fraser University (SFU), where he was senior NSERC Research Chair in Intelligent Systems, a tenured full professor in the SFU School of Computing Science since 1994. Before that, he was Project Leader at Digital Equipment's Paris Research Lab, where he led the Paradise project developing the $\mathcal{OSF}$-constraint programming language $\mathcal{LIFE}$ (an acronym for "$\mathcal{L}$ogic, $\mathcal{I}$nheritance, $\mathcal{F}$unctions, and $\mathcal{E}$quations"). Before joining Digital in 1989, he was a member of technical staff at the Microelectronics and Computer technology Corporation (MCC), in Austin, TX (USA), in **Bob Boyer's** Intelligent Architecture group, part of MCC's AI Program headed by the late **Woody Bledsoe**. There, he headed a research team that designed and prototyped the original version of $\mathcal{LIFE}$.

Hassan Aït-Kaci's scientific production reflects the fact that his career has been balanced between Academic Research Centers (at the start of his professional career), and Industry (later in his professional career). Accordingly, **his publication record is stronger earlier in his career**. From 2000 until his recent return to Academia in January 2013 at *Université Claude Bernard Lyon 1*, France, to lead the ANR Chair-of-Excellence project $\mathcal{CEDAR}$, Hassan Aït-Kaci worked in Industry (ILOG, IBM). Most of his work there was proprietary software-product oriented (specification and development)—ILOG's Optimization Programming Language (OPL) and the JRules Business Rule Management System. This explains the relatively scant publication record in the past 10 years. However, it is hereby proposed that (1) judging the PI's scientific notoriety and credibility strictly over the past 10 years only is arguably moot in that it misrepresents his earlier research record, whose specific contributions are, as a matter of fact, crucially *technically* relevant to the proposed project; and, (2) the essential technology upon which this proposal will be resting stems from work he performed earlier than this relatively short time span for a senior scientist of his generation. Indeed, most of the technical results to be put to use he developed from the late 1980's to the late 1990's before taking a career break from pure research to industrial R&D (*viz.*, from 2000 until 2012 at ILOG,[17] then IBM). Nevertheless, it should also be noted that **his tenure in Industry has given him a unique and quite relevant experience in actual high-quality market-grade software development**. More so, *this was at ILOG, the best internationally recognized constraint-solving tools vending company in the market for the past 15 years*. This fact was clearly established by IBM's decision to acquire it in 2009.[18]

[13] http://www.hassan-ait-kaci.net/
[14] http://cedar.liris.cnrs.fr
[15] http://cedar.liris.cnrs.fr/documents.html
[16] Level "10" (*i.e.*, top of IBM's scale, under "Distinguished Engineer," and "IBM Fellow").
[17] http://www-01.ibm.com/software/info/ilog/
[18] http://www-03.ibm.com/press/us/en/pressrelease/26403.wss

Refereed journals

- "Order-Sorted Feature Theory Unification," (with Andreas Podelski and Seth Copen Goldstein) *Journal of Logic Programming*, 30(2), pp. 99–124, February 1997. [**JLP rank: 100**]
- "Label-Selective λ-Calculus—Syntax and Confluence," (with Jacques Garrigue) *Theoretical Computer Science* 151, pp. 353–383, 1995. [**TCS rank: 28**]
- "Functions as Passive Constraints in $\mathcal{LIFE}$," (with Andreas Podelski) *ACM Transactions on Programming Languages and Systems*, 16(4), pp. 1279–1318, July 1994. [**TCS rank: 9**]
- "A Feature Constraint System for Logic Programming with Entailment," (with Andreas Podelski and Gert Smolka) *Theoretical Computer Science*, 122, pp. 263–283. 1994. [**TCS rank: 28**]
- "Towards a Meaning of $\mathcal{LIFE}$," (with Andreas Podelski) *Journal of Logic Programming*, 16(3-4), pp. 195–234. 1993. (**343 citations**) [**JLP rank: 100**]
- "Inheritance Hierarchies: Semantics and Unification," (with Gert Smolka), *Journal of Symbolic Computation*, 7, pp. 343–370, 1989. [**JSC rank: 19**]
- "Efficient Implementation of Lattice Operations," (with Robert Boyer, Patrick Lincoln and Roger Nasr) *ACM Transactions on Programming Languages and Systems*, 11(1), pp. 115–146, January 1989. (**279 citations**) [**TCS rank: 9**]
- "An Algebraic-Semantics Approach to the Effective Resolution of Type Equations," *Theoretical Computer Science*, 45, pp. 293–351, 1986. [**TCS rank: 28**]
- "LOGIN: A Logic-Programming Language with Built-In Inheritance," (with Roger Nasr), *Journal of Logic Programming* 3, pp. 185–215,1986. (**579 citations**) [**JLP rank: 100**]

Books

- "*Warren's Abstract Machine—A Tutorial Reconstruction*," MIT Press, 1991. (**678 citations**)

Refereed conferences

- "$\mathcal{CEDAR}$: a Fast Taxonomic Reasoner Based on Lattice OperationsSystem Demonstration," (with Samir Amir) in *Proceedings of the Posters & Demonstrations Track of the 12th International Semantic Web Conference (ISWC'13)* pp. 9–12. October, 2013. [**ISWC rank: A+**]
- "An Axiomatic Approach to Feature Term Generalization," (with Yutaka Sasaki), in *Proceedings of the European Conference on Machine Learning (ECML'01*, Freiburg, Germany, September 2001. [**ECML rank: A**]
- "A Database Interface for Complex Objects," (with Marcel Holsheimer and Rolf de By), in *Proceedings of the 11th International Conference on Logic Programming (ICLP'94)*, (Genoa, Italy), June 13–17, 1994. [**ICLP rank: A**]
- "The Typed Polymorphic Label-Selective Lambda-Calculus,"[19] (with Jacques Garrigue), in *Proceedings of the 21st Annual ACM Symposium on Principles of Programming Languages (POPL'94)*, Portland, Oregon, pp. 35–47. January, 1994. [**POPL rank: A+**]
- "An Introduction to $\mathcal{LIFE}$—Programming with Logic, Inheritance, Functions, and Equations," in *Proceedings of the 10th International Logic Programming Symposium*, Vancouver, BC, Canada, pp. 1–17, October 1993. [ILPS is now **ALP rank: B**]
- "Entailment and Disentailment of Order-Sorted Feature Constraints," (with Andreas Podelski), in *Proceedings of the 4th International Conference on Logic Programming and Automated Reasoning (LPAR'93)*, Saint Petersburg, Russia, Andrei Voronkov, ed., Lecture Notes in A.I. 698, Springer-Verlag, pp. 1–18, 1993. [**LPAR rank: A**]
- "BABEL: A Base for an Experimental Library," with Roger Nasr and Jungyun Seo) *Proceedings of the ACM SIGIR International Conference on Information Retrieval*, Grenoble, France, June 1988. [**SIGIR rank: A+**]
- "Logic Programming and Inheritance,"(with Roger Nasr), in *Proceedings of the 13th ACM Symposium on Principles of Programming Languages (POPL'86)*, Saint-Petersburg, FL, pp. 219–228, January 1986. [**POPL rank: A+**]

Software development

- "Jacc: Just another compiler compiler (Java software tool—LALR(k) compiler generator)," Java package suites.
- "CEDAR Taxonomic Reasoner Evaluation Tool V.1.0," (with Samir Amir).
- "Wild LIFE—A C Implementation of $\mathcal{LIFE}$ (Logic, Inheritance, Functions, Equations)," (with Bruno Dumant, Richard Meyer, Andreas Podelski, Peter Van Roy (V.1.0), and Denys Duchier (V.1.02)).

[19] `http://hassan-ait-kaci.net/pdf/labsel-tcs94.pdf`

Appendix: All ongoing and submitted grants and funding of the PI (Funding ID)

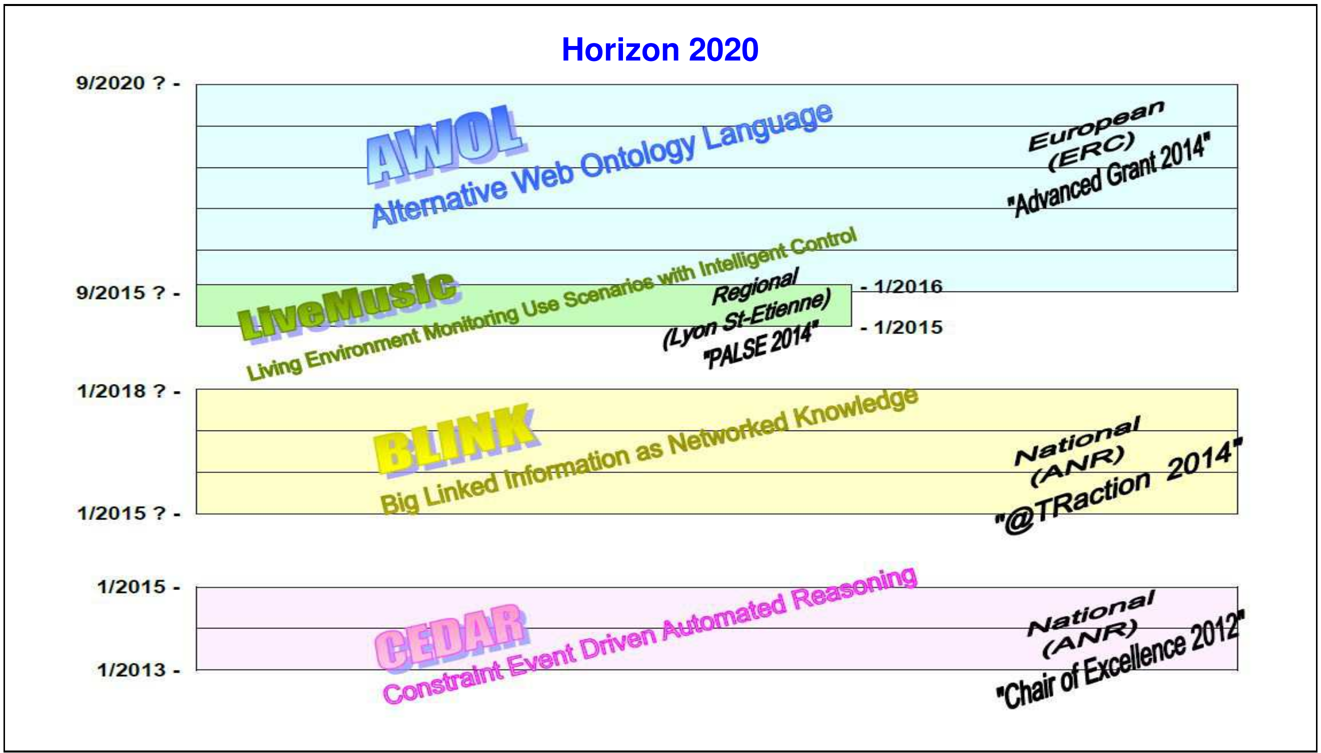
Ongoing Grants

Project Title	***Funding source***	***Amount (€)***	***Period***	***Role of the PI***	***Relation to current ERC proposal***
CEDAR	Agence Nationale de la Recherche	540,996	1/2013 – 1/2015	Chair of Excellence	(See diagram below)

Applications

Project Title	***Funding source***	***Amount (€)***	***Period***	***Role of the PI***	***Relation to current ERC proposal***
BLINK	Agence Nationale de la Recherche	decision pending	1/2015 – 1/2018	PI	(See diagram below)
LivEMUSIC	Programmme Avenir Lyon St-Etienne	150,000	1/2015 – 1/2016	PI	(See diagram below)

N.B.: We received notice that the LivEMUSIC proposal has been granted 150,000 € for a one-year project; implementation protocol (funds repartition, project's final schedule, *etc.*) yet to be determined.



AWOL Project's Grander Context: Relation with the PI's Ongoing and Planned Projects