Direction des Relations Internationales (DRI)

Programme INRIA "Equipes Associées"

(*Demande de prolongation*)

I. DEFINITION

EQUIPE ASSOCIEE	*InToHyLo*
sélection	*2009*

Equipe-Projet INRIA : Traitement Automatique des Langues : Représentations, Inférences et Sémantique (TALARIS)	Organisme étranger partenaire : Grupo de Lógica y Computabilidad (GLyC)
Centre de recherche INRIA : INRIA Nancy, Grand Est Thème INRIA : SyM	Pays : Argentina

	Coordinateur français	Coordinateur étranger
Nom, prénom	ARECES, Carlos	FIGUEIRA, Santiago
Grade/statut	CR1 INRIA	Professor
Organisme d'appartenance *(précisez le département et/ou le laboratoire)*	TALARIS, LORIA	GLyC, Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires
Adresse postale	INRIA Nancy Grand Est TALARIS, Bat B 615, rue du Jardin Botanique 54600 Villers lès Nancy Cedex France	Departamento de Computación Facultad de Ciencias Exactas y Naturales. (C1428EGA) Pabellón I, Ciudad Universitaria, Buenos Aires, Argentina
URL	http://talaris.loria.fr http://talaris.loria.fr/~areces	http://www.glyc.dc.uba.ar http://www.glyc.dc.uba.ar/santiago
Téléphone	+33 (0)3 54 95 84 90	+54 11 4576 3390/6 int. 705
Télécopie	+33 (0)3 83 41 30 79	+54 11 4576 3359
Courriel	carlos.areces@loria.fr	sfigueir@dc.uba.ar

La proposition en bref

Titre de la thématique de collaboration *(en français et en anglais)* : *English:* Inference Services for Natural Language Processing *Francais*: Services d'Inférence pour le Traitement Automatique de la Langue Naturelle
Descriptif *(environ 10 lignes)* : The project main aim is to investigate inference methods for hybrid logics, to develop highly optimized inference tools based on these methods, and to use these tools in natural language applications. The teams Talaris and GLyC have been working in collaboration in this area since 2003. As a result of this collaboration they can be considered nowadays as leaders in the area of computational hybrid logic and automated inference for hybrid logics, and they are the developers of two advanced automated theorem provers (HyLoRes, based on resolution, and Htab, based on tableaux). Moreover, team Talaris is a specialist in the area of representation and inference in natural language processing. In the framework of the InToHyLo (Inference Tools for Hybrid Logics, http://led.loria.fr/dokuwiki/doku.php?id=intohylo_-_inria_equipes_associees) project we are investigating the complexity, and develop and implement optimized algorithms for new reasoning tasks (e.g., model generation and retrieval). We want to investigate also how to combine resolution and tableaux algorithms to allow our provers to collaborate (share partial results) while working on a given problem. Finally, we want to develop parallel inference algorithms to improve performance, and distributed testing to speed up developing.

II. BILAN 2009

Changements majeurs survenus concernant l'Equipe Associée (modifications des objectifs scientifiques, des chercheurs impliqués)

Rapport scientifique de l'année 2009

(joindre la page du programme de travail initialement prévu fin 2008 pour l'année 2009 ou insérer un lien vers cette page)

Description de l'activité scientifique de l'équipe associée et des résultats obtenus : publications, communications, organisation de colloques, formation, soutenances de thèse, valorisation économique, sociale, industrielle, enregistrement de logiciels, dépôt de brevets ... (1 à 2 pages)

Scientific Tasks for the Period 2009-2011

Besides continuing our ongoing research on inference in hybrid logics, the scientific tasks that we set to tackle during the three years the collaboration between the Talaris and GLyC are listed below. After each particular item we comment on the advances obtained during 2009.

- Currently our tools treat “one problem at a time”: if the same inference task is repeatedly submitted to the reasoner, it will be each time solved from scratch. We will integrate the different tools we have developed into a server/client architecture. The inference algorithm hosted in the

server will be able to reuse information from previous calls from the client;
We have redefined the input/output format of the HyLoRes and Htab provers so that now they can receive a background theory together with a lists of inference tasks to solve with respect to it. We have standarized the input/output of the provers so that they can now be integrated in a server/client architecture. This will be the main implementation challenge of 2010.

- Develop optimizations that take advantage of this new architecture. In particular, different reasoning algorithms can collaborate and share results in a way that is transparent to the client;
We investigated the optimization called 'catching' and implemented it in Htab (this resulted in a master thesis at the Université Henri Poincare).
- Design, implement and test parallel versions of the inference algorithms to improve performance;
We have developed a parallel version of HyLoRes (this will result in a master thesis at the Universidad de Buenos Aires).
- Investigate ways to enhance the range of inference services provided (e.g., model building, retrieval);
We have investigated the inference task of model minimization (given a model M, obtain a model M' which corresponds to the quotient of M through the maximal L-bisimulation for a certain logic L). The complexity of these algorithms for a number of sub-Boolean modal logics was determined (this material will be submitted to an international conference). The relation of this problem with the problem of generation of referring expressions in natural language generation was investigated (this material will be submitted to an international conference).
- Verify the algorithms to ensure correctness and test them for efficiency;
The GridTest environment for testing provers in a computer grid was extended and used for testing of the provers developed by the teams. The testing platform was presented at the Description Logic Workshop.
- Finally, we want to extend our results, algorithms and tools to alternative semantics (e.g., topological and neighbourhood semantics) and to more expressive logics (e.g., first order hybrid logics).
Most of the theoretical production of the collaboration during 2009 was focused on Memory Logics. We have investigated their expressive power, defined sound and complete axiomatizations, investigated Craig interpolation and Beth definability, defined tableaux calculi, and proved decidability / undecidability results. These results were presented at WoLLIC, Tableaux, and LFCS. We also investigated the expressive power of modal logics with neighbourhood semantics and obtained complexity results. This was presented at IJCAI.

During 2009 we have decided to include an additional item to the list of scientific task to investigate during the collaboration between TALARIS and GLyC.

- Recent developments on the natural language processing systems being developed by TALARIS point out that the inference task of planning (i.e., given a specification of actions in terms of preconditions and effects, an initial state and a final state, compute a sequence of actions that permits the update of the initial state into the final state) is of crucial relevance.
Moreover, most current automated planners do not provide all inference services required. Of particular interest are: planning in terms of a background theory, requests for alternative plans, request for plans with intermediate check points, and request for incomplete plans in case of absence of complete plans.
We plan to investigate how to integrate these inference tasks into the InToHyLo framework.

Interchanges during 2009

- 2 week visit of Carlos Areces to GLyC [25 April - 10 May, 2009]
- 2 week visit of Santiago Figueira to Talaris [2 July - 16 July, 2009]
- First InToHyLo Meeting coordinated with a Loria Automated Deduction Day at Nancy, France [July 16th, 2009]
- 2 month visit of Daniel Gorin to Talaris [15 June - 13 August, 2009]

- 1 month visit of Sergio Mera to Talaris [4 July - 3 August, 2009]
  (Cofunded by MINCYT - INRIA/CNRS "Service d'inference pour des systemes de dialogue" project)
- 2 month visit of Daniel Koile to Talaris [10 July - 10 August, 2009]
  (Cofunded by INRIA Internship Programme and PARLO)
- 2 month visit of Facundo Carreiro to Talaris [10 July - 10 August, 2009]
  (Cofunded by INRIA Internship Programme, UBA and PARLO)

The following interchanges are programmed for the rest of 2009

- Inference Week at Buenos Aires, Argentina [December, 2009]
  (Cofunded by UBA)
- 2 month visit of Guillaume Hoffmann to GLyC [5 November 2009 - 4 January, 2010]
  (Cofunded by UBA)
- 2 week visit of Patrick Blackburn to GLyC [December, 2009]
- 1 month visit of Carlos Areces to GLyC [6 December 2009 - 6 January, 2010]
  (Cofunded by MINCYT - INRIA/CNRS "Service d'inference pour des systemes de dialogue" project)
- 1 month visit of Luciana Benotti to GLyC [6 December 2009 - 6 January, 2010]

Joint Publications during 2009

- Carlos Areces and Daniel Gorín. Coinductive Models and Normal Forms for Modal Logics. *Submitted to the Logic Journal of the IGPL.*

- Carlos Areces, Santiago Figueira, and Sergio Mera. Completeness results for memory logics. *Submitted to APAL*, 2009. Extended version of LFCS'09 paper of the same title.

- Carlos Areces, Daniel Gorin, Alejandra Lorenzo, and Mariano Perez Rodriguez. Testing Provers on a Grid – Framework Description. In *Proceedings of DL09*, 2009. [pdf]

- Carlos Areces, Diego Figueira, Daniel Gorin, and Sergio Mera. Tableaux and Model Checking for Memory Logics. In *Proceedings of Tableaux09*, 2009. [pdf]

- Carlos Areces and Diego Figueira. Which Semantics for Neighbourhood Semantics?. In *Proceedigns of IJCAI 09*, 2009. [pdf]

- Carlos Areces, Santiago Figueira, and Sergio Mera. Completeness results for memory logics. In Sergei Artemov and Anil Nerode, editors, *LFCS 09: Symposium on Logical Foundations of Computer Science*, volume 5407 of LNCS, Deerfield Beach, FL, USA, 2009. Springer. [pdf]

Moreover, two PhD theses (in co-tutelle between the Universidad of Buenos Aires and the Université Henri Poincare, and co-supervised by members of the EA) were completed during this period and will be defended in December:

- Daniel Gorín. Resolution-based automated reasoning techniques for hybrid logics. Co-supervised Carlos Areces, Patrick Blackburn and Verónica Becher.

- Sergio Mera. Modal Memory Logics. Co-supervised by Carlos Areces, Patrick Blackburn and Verónica Becher.

Rapport financier 2009

Avant de remplir les tableaux, consultez les règles au paragraphe "Financement" de la page d'accueil du programme.

1. Dépenses EA (*effectuées sur les crédits de l'Equipe Associée*)	Montant dépensé
Invitations des partenaires	5000
Missions INRIA	6750
Total	11750

The budget was mainly used to cover travel and accommodation expenses of researchers of the two teams. Long visits by young researchers were privileged when allocating funds (40 weeks vs. 10 weeks). The one and two month visits of Mera and Gorín, respectively, were influential to ensure termination in time of their PhD thesis. A small amount of the budget was used for the organization of the first InToHyLo Meeting (given that it was coordiated with the Loria Automated Deduction Day, most expenses were covered by special funds available for that event).

In addition, 3000 euros will be spent in the organization of the second InToHyLo Meeting during December 2009. The main expense will be to invite Prof. Gert Smolka to participate as speaker. Prof. Smolka recent work on inference for hybrid logics is very relevant to the research of the Equipe Associeé.

2. Dépenses externes *(effectuées sur des financements hors EA)*	Montant dépensé
Nom de l'organisme 1 (*): Universidad de Buenos Aires	
Invitations des partenaires	2000
Missions INRIA vers le partenaire	1050
Total	3050
Nom de l'organisme 2 (*): MINCyT / INRIA /CNRS	
Invitations des partenaires	2130
Missions INRIA vers	2450

le partenaire	
Total	4580
Nom de l'organisme 3 (*): PARLO - MISN-TALC	
Invitations des partenaires	5600
Missions INRIA vers le partenaire	0
Total	5600
Nom de l'organisme 4 (*):INRIA Internships	
Invitations des partenaires	1400
Missions INRIA vers le partenaire	0
Total	1400

(*) Ajouter ou supprimer des lignes au tableau ci-dessus de façon à faire figurer tous les organismes ayant contribué au financement de l'équipe associée

Total des financements externes dépensés	14630
Total des financements EA et externes dépensés	26380

Bilan des échanges effectués en 2009

1. Chercheurs Seniors

Nom	statut (1)	provenance	destination	objet (2)	durée (3)	Coût (si financement EA)	Coût (si financement externe)
Areces	CR	France	Argentina	visit	2 weeks	650	250
Figueira	Professor	Argentina	France	visit	2	1400	600

					weeks		
Blackburn	DR	France	Argentina	Visit + doctoral defense	2 weeks	2000	200
Areces	CR	France	Argentina	Visit + doctoral defense	4 weeks	0	2450

Total des durées	10 weeks

(1) DR / CR / professeur
(2) colloque, thèse, stage, visite....
(3) précisez l'unité (mois, semaine..)

2. Juniors

Nom	statut (1)	provenance	destination	objet (2)	durée (3)	Coût (si financement EA)	Coût (si financement externe)
Gorín	PhD	Argentina	France	PhD thesis	8 weeks	3600	400
Mera	PhD	Argentina	France	PhD thesis	4 weeks	0	2130
Koile	intern	Argentina	France	Internship	8 weeks	0	4000
Carreiro	intern	Argentina	France	Internship	8 weeks	0	4000
Benotti	PhD	France	Argentina	PhD thesis	4 weeks	1500	200
Hoffmann	PhD	France	Argentina	PhD thesis	8 weeks	2600	400

Total des durées	40 weeks

(1) post-doc / doctorant / stagiaire
(2) colloque, thèse, stage, visite....
(3) précisez l'unité (mois, semaine..)

III. PREVISIONS 2010

Programme de travail

Description du programme scientifique de travail pour l'année 2010

During 2010 we plan to continue, on one hand, the work on the tools being developed in the framework of the collaboration. In particular, the standardization work carried out during 2009 (on the input/output format of HyLoRes and Htab) provides the first step to the new envisioned architecture as inferences servers. The main implementation goal of next year is to finally achieve this model.

On the theoretical side, we will continue the investigation of algorithms and complexity of different inference tasks for modal and hybrid logics. Much of the work of 2009 was focused on the study of the family of languages called Memory Logics. Most of the explored languages are expressively weaker than the hybrid logic containing the downarrow binder, but they turn to still be undecidable. One of the remaining tasks is to look for possible decidable fragments of these languages. We are also interested in exploring the following interesting phenomena: memory logics can be seen as truly *dynamic* logic, in the sense that the semantic clauses of the memory operators provide instructions of how to modify the model during the evaluation of a particular formula. Despite this, it is possible to define a translation of this languages into a fragment of classical first order logic, where models are *static* and they do not change during the evaluation of the model. The phenomenon is similar to the well know duality between the *internal* perspective of standard modal formulas, and the *external* behaviour of their first-order translation. We plan to investigate this issue by providing a 'van Benthem' style characterization of first-order formulas equivalent to the translation of a formula of memory logic.

Two main themes of collaboration for 2010 will be model minimization and planning in the framework of modal and hybrid logics. Both have great potential in NLP applications but have received almost no attention beside a few theoretical results (there are no tools available that can provide these services). Model minimization can be obtained by computing the coarsest auto-bisimulation of the input model for the desired logical language. During the next year we want to develop efficient algorithms for computing bisimulations for various modal languages. Moreover, model minimization is also related to the NLP problem of generating referring expressions (a topic that has recently being investigated by Carlos Areces). During 2010 we plan to use these new ideas to develop and test new algorithms for the generation of referring expressions. Finally, TALARIS has been developing dialogue systems that use planning as a core inference service, but most current automated planners do not provide all inference services required. Of particular interest are: planning in terms of a background theory, requests for alternative plans, request for plans with intermediate check points, and request for incomplete plans in case of absence of complete plans. We plan to investigate how to integrate these inference tasks into the InToHyLo framework.

The points mentioned above integrate well with the themes of the doctoral thesis of Guillaume Hoffmann and Laura Perez in TALARIS, and the master thesis of Facundo Carreiro in GLyC.

Programme d'échanges avec budget prévisionnel

1. Echanges

- Décrivez les échanges prévus dans les deux sens : invitations de chercheurs de votre partenaire et missions INRIA vers votre partenaire ;
- Précisez s'il s'agit de chercheurs confirmés ou de juniors *(stagiaires, doctorants, post-doctorants) ;*
- Motivez, si possible, les raisons scientifiques *(travail commun, workshop,..)* et précisez la durée prévue ;
- Indiquez les étudiants impliqués dans la proposition. Donnez une estimation de leur nombre, pour chaque partenaire, et précisez si des thèses en cotutelle sont prévues ;

▪ Résumez ensuite ces informations dans les tableaux 1 et 2 ci-dessous en faisant une estimation budgétaire :

1. ESTIMATION DES DÉPENSES EN MISSIONS INRIA VERS LE PARTENAIRE	Nombre de personnes	Coût estimé
Chercheurs confirmés	2	6000
Post-doctorants	0	0
Doctorants	2	5000
Stagiaires	0	0
Autre (précisez) :	0	0
Total	5	11000

The planned interchanges for 2010 are the following. Carlos Areces (CR1 from TALARIS) will visit GLyC for a total period of 6 weeks (probably in two trips), combining one of the trips with a 2 week visit of Patrick Blackburn (DR, director of TALARIS). The first visit of Areces will be early during next year (March/April) to work on final versions of the submitted journal articles. During the joint visit of Areces and Blackburn, a post-graduate seminar will be organized at the Universidad de Buenos Aires.

Gullaumme Hoffmann (PhD from TALARIS) will visit GLyC for 8 weeks later in 2010, to work on the development of the HyLoRes and Htab provers, in particular with respect to the new client/server architecture. Finally Laura Perez (PhD from TALARIS) will visit GLyC for 4 weeks at the end of the year, to work on the use of our tools for generation.

2. ESTIMATION DES DÉPENSES EN INVITATIONS DES PARTENAIRES	Nombre de personnes	Coût estimé
Chercheurs confirmés	1	2500
Post-doctorants	2	6000
Doctorants	0	0
Stagiaires	1	3500
Autre (précisez) :	0	0
Total	4	12000

Santiago Figueira (Professor from GLyC) will visit TALARIS for 2 weeks in mid 2010 to work on model minimization and bisimulations for hybrid logics. The visit will overlap with a 4 week visit of Sergio

Mera (Postdoc from GLyC) who is working in similar topics. The 4 week visit of Daniel Gorín will follow the 8 week visit of Hoffmann to GLyC. This will give us three full months where member of TALARIS and GLyC will be able to work together in the development of the InToHyLo tools.

A 2 month internship will be cofunded by the PARLO project and the INRIA Internship programm to work on the parallel resolution algorithm of HyLoRes.

2. Cofinancement

Cette coopération bénéficie-t-elle déjà d'un soutien financier de la part de l'INRIA, de l'organisme étranger partenaire ou d'un organisme tiers *(projet européen, NSF, ...) ?*
Indiquez ces éléments et donnez les montants associés.

The PARLO project funded by the MISC-TALC (CPER) continues during 2010 and will help fund the visit of an intern for 8 weeks (approx. 2000 euros). We will also apply for a two month INRIA Internship position to combine with these funds (approx. 1500 euros). GLyC is applying to an UBACyT project (Science and Technology Projects funded by the Universidad de Buenos Aires) for three years (total budget approx. 4500 euros), part of these funds will be used to fund the collaboration with TALARIS.

3. Demande budgétaire

Indiquez, dans le tableau ci-dessous, le coût global estimé de la proposition et le budget demandé à la DRI dans le cadre de cette Equipe Associée.
(*maximum 20 K€ pour une prolongation en 2e année et 10 K€ pour une 3e année*).

Commentaires	Montant
A. Coût global de la proposition (*total des tableaux 1 et 2 : invitations, missions, ...*)	23000
B. Cofinancements utilisés (*financements autres que Equipe Associée*)	4000
Financement "Équipe Associée" demandé (A.-B.) *(maximum 20 K€ pour une 2e année et 10 K€ pour une 3e année)*	19000

Remarques ou observations :

The collaboration during this first year of the EA project has been extremely successful, with four joint articles accepted in international conferences. Moreover two other articles have been submitted to international journals. Two other articles (a journal and a conference article) are being finished at the time of writing of this report. The collaboration was also the framework for the finalization of two PhD thesis in cotutelle, and a master thesis.
The implementation of the provers HyLoRes and Htab has also advanced during this year, having now a standardized input/output format which will ease the transition to the client/server architecture. HyLoRes has now a “parallel mode” which takes advantage of multiple CPUs. The GridTest distributed testing environment has been improved and extensively used.

The main topics of collaboration for 2010 will be to explore two inference tasks: model minimization and planning in the framework of modal and hybrid logics. Both have great potential in applications but have

received almost no attention beside a few theoretical results (there are no tools available that can provide these services). The funds provided by a second year of the EA will be instrumental to the achievement of these goals and to the consolidation of the collaboration between the two research groups.