POSTDOCTORAL PROPOSAL IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Pushing the computational frontiers of reasoning with logic, probabilities and preferences

Advisor (s): Simon.de-Givry@inra.fr; http://www.inra.fr/mia/T/degivry, Thomas.Schiex@inra.fr; http://www.inra.fr/mia/T/schiex/

Net salary: Negotiable with a minimum of 2600€ per month with some teaching (64 hours per year on average).

Duration: from one to four years

DESCRIPTION

Constraint programming (CP [Rossi et al, 2006]) is an a*utomated reasoning* technology with tight connections with propositional logic. It offers a problem modeling and solving framework where the set of solutions of a complex (NP-hard) problem is described by discrete variables, connected by constraints (Boolean functions).Together with propositional satisfiability, it is one of the automated reasoning approaches of AI, where problems are solved *exactly* to provide rigorous solutions to hardware or software testing, verification, system configuration, scheduling or planning problems.

Discrete Stochastic Graphical Models (GMs [Koller et al, 2009]) define a m*achine learning* technology where a probability mass function is described by discrete variables, connected by potentials (numerical functions). GMs can be learned from data. The NP-hard problem of identifying a Maximum a Posteriori (MAP) labelling is often solved *approximately* to tackle several problems in Image [Kappes et al, 2013] and Natural Language Processing [Bilmes, 2004], among others.

The Cost Function Network framework [Cooper et al, 2010] with its associated exact solvers such as toulbar2, merge Constraint Programming and Stochastic Graphical Models processing. By solving the so-called Weighted Constraint Satisfaction problem, it becomes possible to simultaneously reason on logical information and gradual, possibly Machine Learned, information described as local numerical functions.

To process the available information, solvers rely on exact Branch and Bound algorithms that combine Best and Depth-First search [Allouche et al, 2015] and where pruning follows from a variety of mechanisms that can either simplify the problem at hand, provide primal and dual solutions and lower bounds. Multi-threading and parallel solving offer new opportunities to organize these various mechanisms differently in time, to exploit problem decompositions, to apply stronger primal/dual reasoning and use Machine learning to guide search or decide which mechanism to activate based on the current solving and/or a collection of instances of the same problem.

The short-term aim of the PostDoc will be to enhance these algorithms with parallel solving capabilities on multi-core, cluster, and cloud architectures following the ongoing trend followed by CP, SAT, and also Integer Linear Programming solvers [Gent et al, 2018 ; Hamadi et al, 2018]. A first master-worker parallel version of the algorithm underlying toulbar2 has been implemented. The PostDoc will continue to adapt the parallel algorithm to exploit tree decompositions [Jégou et al, 2017]. He or she will explore new machine learning rules for better load balancing and reduced workload of the master process. The resulting algorithm will be integrated into a metaheuristic [Ouali et al, 2017] for improved global anytime lower and upper bounds.

Experiments will be performed on large collections of real problem instances, many of which are not known to be currently solvable. This includes applications to current exciting problems in Computational Protein Design

(CPD) [Allouche et al, 2014 ; Pan et al, 2016], in collaboration with molecular modellers and biochemists, in the context of the ongoing development of dedicated CPD software with applications in bioenergies, health and nanotechnologies.

The position is specifically open to highly creative researchers that may quickly want to develop and explore their own ideas. As such, we expect that the PostDoc will be increasingly capable of injecting their own ideas in the project, in interaction with all the members of the project team as well as external collaborators, and contribute to the supervision of PhD students.

Candidate profile

The PostDoc is at the intersection of CP, SAT, integer programming, metaheuristics, and distributed computing. The ideal candidate should ideally be familiar with CP or SAT algorithms. He or she may also benefit from background knowledge in the weighted variants of SAT/CP, Integer Linear Programming, or Stochastic Graphical Models processing (eg. In Computer Vision). Some experience in the design and implementation of multi-threaded/distributed code is a nice plus. Good programming abilities (in C++ ideally) will be required. Additional knowledge in bioinformatics, biochemistry, and molecular modelling would be a plus in the context of Computational Protein Design applications.

The monthly net salary is about 2 700 euros ; starting date : autumn 2019

References

• D Allouche, I André, S Barbe, J Davies, S de Givry, G Katsirelos, B O'Sullivan, S Prestwich, T Schiex, S Traoré. Computational protein design as an optimization problem. In Journal of Artificial Intelligence 212, 59-79, 2014.

• D Allouche, S. de Givry, G. Katsirelos, T. Schiex, M. Zytnicki. Anytime Hybrid Best-First Search with Tree Decomposition for Weighted CSP. In Proc. of CP-15, pp 12-28, Cork, Ireland, 2015.

• J Bilmes. Graphical Models and Automatic Speech Recognition. Mathematical Foundations of Speech and Language Processing pp 191-245, 2004.

• M Cooper, S. de Givry, M. Sanchez, T. Schiex, M. Zytnicky, T. Werner. Soft Arc-consistency revisited. In Journal of Artificial Intelligence, 174(7-8), pp 449-478, 2010.

• I Gent, I Miguel, P. Nightingale, C. McCreesh, P. Prosser, N. Moore, and C. Unsworth. A review of literature on parallel constraint solving. In Journal of Theory and Practice of Logic Programming, 18(5-6), 725-758, 2018.

• Y. Hamadi and Lakhdar Sais. Handbook of Parallel Constraint Reasoning. Springer, 2018.

• P Jégou, H Kanso, and C Terrioux. Adaptive and Opportunistic Exploitation of Tree-Decompositions for Weighted CSPs. In Proc. of ICTAI 2017, pp 366-373, Boston, MA, USA, 2017.

• J Kappes, M Speth, G Reinelt, and C Schnorr. Towards efficient and exact MAP-inference for large scale discrete computer vision problems via combinatorial optimization. In Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2013.

• D Koller, N Friedman. Probabilistic graphical models: principles and techniques. The MIT Press, 2009.

• A Ouali, D Allouche, S de Givry, S Loudni, Y Lebbah, F Eckhardt, and L Loukil. Iterative Decomposition Guided Variable Neighborhood Search for Graphical Model Energy Minimization. In Proc. of UAI-17, pp 550-559, Sydney, Australia, 2017.

• Y Pan, Y Dong, J Zhou, M Hallen, B Donald, J Zeng, and W Xu. cOSPREY: A Cloud-Based Distributed Algorithm for Large-Scale Computational Protein Design. In Journal of Computational Biology, 23(9), 2016.

• F Rossi, P Van Beek, and T Walsh. Handbook of constraint programming. Elsevier, 2006.

APPLICATION PROCEDURE

Formal applications should include a detailed cv, a motivation letter, samples of published research by the candidate and reference letters.
Applications should be sent by email to: advisor emails
More information about ANITI: https://aniti.univ-toulouse.fr