

**Deadline for applications:** 1 May 2019

**Start date:** September/October 2019

**Funding Duration:** 36 months

**Title:** Belief merging for improving ontologies and open-domain knowledge exploitation

**Keywords:** Belief merging, inductive reasoning, vector space embeddings, natural language processing.

**Supervisors:**

Zied Bouraoui (<http://www.cril.fr/~bouraoui/>),

Sebastien Konieczny (<http://www.cril.fr/~konieczny>) and

Ivan Varzinczak (<https://www.ijv.ovh>).

The CRIL (CNRS & Artois University) invite applications for a fully funded 3 years Ph.D. scholarship. It will focus on the development of suitable merging methods for improving ontologies and open-domain knowledge exploitation

**Background:** Structured knowledge, available for instance in formal ontologies or knowledge graphs, plays an increasingly important role in domains such as information retrieval, natural language processing and machine learning. Such applications differ from the settings that have traditionally been considered in the field of knowledge representation, in that they require knowledge bases with a wide coverage, even if that means accepting some inaccuracies. This has motivated a large body of recent work on methods for automatically extending knowledge bases (e.g. [1]). One possible strategy consists in defining automated methods for finding plausible rules and facts that can be added to the knowledge base (e.g. [2, 3, 4, 5]). Another strategy consists in gathering information provided by several available sources. However, this often leads to inconsistencies and ambiguities. The goal of merging, as it is classically studied [6], is to define operators taking as input the information provided by the different sources and produce as output a base that is consistent and that synthesises the best of the sources while dealing with possible inaccuracies.

**Aim:** The general goal of the work to be carried out during the Ph.D. project is the delivery of flexible and suitable methods for merging open-domain knowledge. The first step will consist in studying classical belief merging operators and comparing them in terms of how fruitful their results can be for merging open-domain data. The second step will consist in devising new merging techniques that use vector-space embeddings [7, 8, 9, 10] in order to deal with problems such as ambiguities and typicalities [11]. Several requirements should be fulfilled by these methods. For instance, the results of merging should be supported by intuitive explanations, as this will allow us to determine the plausibility of the resulting knowledge. The scalability of sound and complete merging methods is likely to be limited. We will hence work on identifying tractable fragments, i.e., syntactic restrictions under which polynomial-time sound and complete reasoning is possible. Finally, the models that will be proposed will not only be used for merging knowledge bases, but also for dealing with inconsistency.

**Selection criteria:**

- Master Degree or equivalent in computer science or other relevant disciplines.
- Strong background in formal logic, machine learning and linear algebra.
- Excellent programming skills in python, Matlab and Java.
- Excellent organisational skills with a proven ability to work independently, to prioritise work and meet deadlines within the framework of an agreed programme.

**Application Procedure:**

- Step 1 - Pre-application documents : (1) CV, (2) academic transcripts, and (3) 1-page personal statement outlining motivation, interest and eligibility for the post. Application should be emailed to Zied Bouraoui ([zied.bouraoui@cril.fr](mailto:zied.bouraoui@cril.fr)), Sebastien Konieczny ([konieczny@cril.fr](mailto:konieczny@cril.fr)) and Ivan Varzinczak ([varzinczak@cril.fr](mailto:varzinczak@cril.fr)).
- Step 2 - Job Interview with selected candidates.

**Deadline for application : 1 May 2019**

**Interviews : Throughout May 2019**

**More information:** For more details about the position, please contact Zied Bouraoui ([zied.bouraoui@cril.fr](mailto:zied.bouraoui@cril.fr)), Sebastien Konieczny ([konieczny@cril.fr](mailto:konieczny@cril.fr)) and Ivan Varzinczak ([varzinczak@cril.fr](mailto:varzinczak@cril.fr)).

## References

- [1] Jay Pujara, Danqi Chen, Bhavana Dalvi, and Sameer Singh Tim Rocktäschel, editors. *Proceedings of the 6th Workshop on Automated Knowledge Base Construction*, 2017.
- [2] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko. Translating embeddings for modeling multi-relational data. In *Proc. NIPS*, pages 2787–2795, 2013.
- [3] Z. Bouraoui, S. Jameel, and S. Schockaert. Inductive reasoning about ontologies using conceptual spaces. In *Proceedings of AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA.*, pages 4364–4370, 2017.
- [4] Tim Rocktäschel and Sebastian Riedel. End-to-end differentiable proving. In *Proc. NIPS*, pages 3791–3803, 2017.
- [5] Z. Bouraoui and S. Schockaert. Automated rule base completion as bayesian concept induction. In *Proceedings of AAAI Conference on Artificial Intelligence*, 2019.
- [6] Sébastien Konieczny and Ramón Pino Pérez. Merging information under constraints: A logical framework. *J. Log. Comput.*, 12(5):773–808, 2002.
- [7] Tomas Mikolov, Wen-tau Yih, and Geoffrey Zweig. Linguistic regularities in continuous space word representations. In *Proc. NAACL-HLT*, pages 746–751, 2013.
- [8] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. Glove: Global vectors for word representation. In *EMNLP*, pages 1532–1543, 2014.
- [9] Shoaib Jameel, Zied Bouraoui, and Steven Schockaert. Member: Max-margin based embeddings for entity retrieval. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, Shinjuku, Tokyo, Japan, August 7-11, 2017*, pages 783–792, 2017.
- [10] Shoaib Jameel, Zied Bouraoui, and Steven Schockaert. Unsupervised learning of distributional relation vectors. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers*, pages 23–33, 2018.
- [11] Ivan Varzinczak. A note on a description logic of concept and role typicality for defeasible reasoning over ontologies. *Logica Universalis*, 12(3-4):297–325, 2018.