

Zugzwang | Logic and Artificial Intelligence

Team

- Francisco Coelho
- Salvador Abreu (PI)
- Bruno Dinis (External Collaborator)

Research Questions

How to extend probability annotations on an ASP program to a distribution over the possible observations? In general, the association of probabilities to some facts is not enough to uniquely define a probability of stable models. This lack of information must be carefully expressed and handled, to avoid biased results.

Once made explicit, how to use such distribution, together with an empiric distribution from a dataset, to do general probabilistic tasks such as the estimation of a marginal or of the joint probability? Having a probability on a set of observations, including the stable models, might not cover all the sample space. Again, information and structure entailed from the ASP program must guide unbiased extensions to the sample space.

How to use distribution measures to score ASP programs? We view ASP programs as formalized knowledge about an observable system. Some of those programs will be "better" than others, as determined by a quantitative, objective, measure, rooted on well-known functions such as the Kullback–Leibler divergence.

This last question leads to the application of evolutionary algorithms to inductive logic programming and to a bridge with common machine learning methods, where a model is scored by a dataset. Other important questions lie behind the scope of this project. For example, how to deal with latent, unobserved, variables?

Novelty/Impact

ASP has some key advantages over Prolog, most of which result from ASP being a truly declarative language and modern APS systems, such as CLASP, apply efficient optimisation techniques. Other systems, like ILASP, learn ASP programs with normal rules, choice rules, and hard and weak constraints. Moreover, ASP can use recent important advances concerning SAT solvers to ILP tasks. However, the assumption that the knowledge base (BK) is perfect, leaving no room for uncertainty, poses here a major limitation.

One approach to overcome this restriction on logic programs is Statistical Relationship Learning (StarAI), that extends the BK with probabilities in order to setup a distribution representation. The formal setting for this approach rests on Sato's Distribution Semantics and frames systems such as Problog and PRISM. But these systems are oriented towards Prolog-like programs and semantic and leave out ASP program learning.

One important ongoing research question is the precise semantic of an ASP program annotated with probabilities. Sato's semantic specify an unique probability distribution over Prolog-like program's atoms, but this uniqueness fails for ASP programs. Efforts to address this problem are either based on Credal sets or on selecting one probability over others, such as the P-log, the LP^{MLN} or the PrASP languages.

The novelty of this research is to define the semantic of an ASP program + probability annotated facts + observations from the following process: (a) parametrise the uncertainty on stable models and annotations, (b) setup a partition of the sample space around on the stable models and (c) use observations to estimate the value of the uncertainty parameters of step (a). If successful, it would have an impact on how logic programs express and are used to deal with real-world problems, where both uncertainty and formal KB are required ingredients. Also, when used in ILP problems the resulting models are logic programs, much simpler to understand than numeric models.

Expected Results / Demonstrators

Our aim is to develop a set of software tools to apply and evaluate the theoretical results on well-known, and relevant, problems. The initial target framework is the Potassco suite, that provides a Python API to state-of-the-art grounding (CLASP) and solving (CLINGO) tools, as well as ASP-Core-2 Language support.

The researchers will do the theoretical study and supervise one to three undergraduate, master or Ph.D. students on the implementation tasks. Theoretical work includes formalization and assessment of methods; The students will implement the tools required to experiment, explore and test those methods. The degree of complexity of the implementation tasks must be adapted to the individual competences and interests of the students.

The theoretical work is to be reported in one or two papers with intermediate results and a final comprehensive paper for a conference.

Relationship of activity w/ NOVA LINCS strategic program and Research Group work plan

This project involves two members from the NOVA LINCS "Knowledge-Based Systems" research group, from the Universidade de Évora pole and a member of the CIMA center from the Universidade de Évora.

Part of the proposed research and software is currently being developed.

Resources required and justification

Students are expected to benefit from a BIC or a BI scholarship. Depending on the candidates the respective durations can range from 3 to 12 months and the respective amounts from 486.12€ to 1144.64€ per month.

Considering one undergraduate student with a 3 month scholarship:

- $972.24 \text{ €} = 3 \text{ months} \times 486.12 \text{ €/month}$;
- TOTAL: 972.24 €