

Zugzwang, Stochastic Adventures in Inductive Logic Specifications

Abstract. A major limitation of logical representations is the implicit assumption that the Background Knowledge (BK) is perfect. This assumption is problematic if data is noisy, which is often the case. Here we aim to explore how ASP specifications with probabilistic facts can lead to characterizations of the joint distribution of the specification's atoms.

Introduction and Motivation

Answer Set Programming (ASP) is a logic programming paradigm based on the Stable Model semantics of Normal Logic Programs (NP) that can be implemented using the latest advances in SAT solving technology. ASP is a truly declarative language that supports language constructs such as disjunction in the head of a clause, choice rules, and hard and weak constraints.

The Distribution Semantics (DS) is a key approach to extend logical representations with probabilistic reasoning. Probabilistic Facts (PF) are the most basic stochastic DS primitive and they take the form of logical facts labelled with a probability p ; Each probabilistic fact represents a boolean random variable that is true with probability p and false with probability $1 - p$.

Crucially, a joint distribution of atoms derived from an ASP specification can be used to *quantitatively measure the performance of that specification* given data observed from the system it is intended to describe. Then, given competing specifications to describe a certain system, these performance measures can be applied in various optimization techniques in order to obtain one that best describes the target system.

The following example illustrates a critical problem with stochastic semantics of ASP specifications. The specification

```
a::0.3.  
b ; c :- a.
```

has three stable models, $\{-a\}$, $\{a, b\}$ and $\{a, c\}$. While it is quite natural to set $P(-a) = 0.7$, there is no further information to assign values to $P(a, b)$ and $P(a, c)$. At best, we can use a parameter α such that

$$\begin{aligned}P(a, b) &= 0.3\alpha, \\P(a, c) &= 0.3(1 - \alpha).\end{aligned}$$

In summary, if an ASP specification is intended to describe some system and that system can be observed then:

1. The observations can be used to estimate the value of the parameters (such as α above and others entailed from the laws of probability).
2. Once probabilities are set for each stable model, we can proceed to define a joint distribution of the atoms.
3. If that specification is only one of many possible candidates then the joint distributions together with the observations can be used to score and sort the specifications.

Work Plan

A team of two (**or three?**) researchers and a graduate student, working over six months with adequate resources, should be able to produce an intermediate progress report for a workshop and a final comprehensive paper for a conference, describing:

- The formalization of the methods outlined above including the parameter estimation from observations and the joint distribution extending the probabilities of the stable models.
- Application and evaluation of this approach to well-known problems, using available software tools, such as
 - Problems: **Assim, de momento, não me lembro de nenhum!**
 - Software tools: `s (casp)`, Potassco suit, etc.

References

1. Victor Verreert, Vincent Derkinderen, Pedro Zuidberg Dos Martires, Luc De Raedt, Inference and Learning with Model Uncertainty in Probabilistic Logic Programs (2022)
2. Andrew Cropper, Sebastijan Dumancic, Richard Evans, Stephen H. Muggleton, Inductive logic programming at 30 (2021)
3. Fabio Gagliardi Cozman, Denis Deratani Mauá, The joy of Probabilistic Answer Set Programming: Semantics - complexity, expressivity, inference (2020)
4. Fabrizio Riguzzi, Foundations of Probabilistic Logic Programming Languages, Semantics, Inference and Learning. Rivers Publishers (2018)
5. Martin Gebser, Roland Kaminski, Benjamin Kaufmann, and Torsten Schaub, Answer Set Solving in Practice, Morgan & Claypool Publishers (2013)