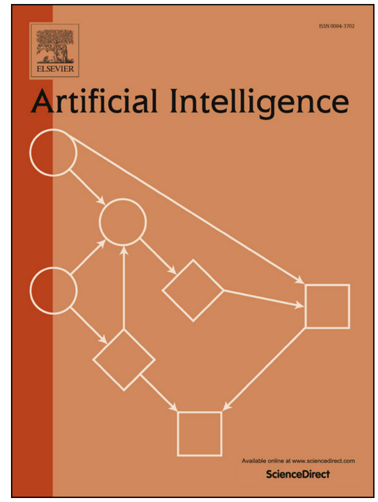
Accepted Manuscript



The complexity and generality of learning answer set programs

Mark Law, Alessandra Russo, Krysia Broda

PII: S0004-3702(18)30105-X
DOI: https://doi.org/10.1016/j.artint.2018.03.005
Reference: ARTINT 3063

To appear in: *Artificial Intelligence*

Received date: 2 September 2016
Revised date: 20 February 2018
Accepted date: 15 March 2018

Please cite this article in press as: M. Law et al., The complexity and generality of learning answer set programs, *Artif. Intell.* (2018), https://doi.org/10.1016/j.artint.2018.03.005

Available online at www.sciencedirect.com

ScienceDirect

Artificial Intelligence 00 (2018) 1–44

Artificial Intelligence

The Complexity and Generality of Learning Answer Set Programs

Mark Law*, Alessandra Russo, Krysia Broda

Department of Computing, Imperial College London, SW7 2AZ

Abstract

Traditionally most of the work in the field of Inductive Logic Programming (ILP) has addressed the problem of learning Prolog programs. On the other hand, Answer Set Programming is increasingly being used as a powerful language for knowledge representation and reasoning, and is also gaining increasing attention in industry. Consequently, the research activity in ILP has widened to the area of Answer Set Programming, witnessing the proposal of several new learning frameworks that have extended ILP to learning answer set programs. In this paper, we investigate the theoretical properties of these existing frameworks for learning programs under the answer set semantics. Specifically, we present a detailed analysis of the computational complexity of each of these frameworks with respect to the two decision problems of deciding whether a hypothesis is a solution of a learning task and deciding whether a learning task has any solutions. We introduce a new notion of *generality* of a learning framework, which enables us to define a framework to be more general than another in terms of being able to *distinguish* one ASP hypothesis solution from a set of incorrect ASP programs. Based on this notion, we formally prove a generality relation over the set of existing frameworks for learning programs under answer set semantics. In particular, we show that our recently proposed framework, *Context-dependent Learning from Ordered Answer Sets*, is more general than brave induction, induction of stable models, and cautious induction, and maintains the same complexity as cautious induction, which has the highest complexity of these frameworks.

Keywords: Non-monotonic Logic-based Learning, Answer Set Programming, Complexity of Non-monotonic Learning

1. Introduction

Over the last two decades there has been a growing interest in Inductive Logic Programming (ILP) [1], where the goal is to learn a logic program called a *hypothesis*, which together with a given background knowledge base, explains a set of examples. The main advantage that ILP has over traditional statistical machine learning approaches is that the learned hypotheses can be easily expressed in plain English and explained to a human user, so facilitating a closer interaction between humans and machines. Traditional ILP frameworks have focused on learning definite logic programs [1, 2, 3, 4, 5, 6] and normal logic programs [7, 8]. On the other hand, Answer Set Programming [9] is a powerful language for knowledge representation and reasoning. ASP is closely related to other declarative paradigms such as SAT, SMT and Constraint Programming, which have each been used for inductive reasoning [10, 11, 12]. Compared with these other paradigms, due to its non-monotonicity, ASP is particularly suited for common-sense reasoning [13, 14, 15]. Because of its expressiveness and efficient solving, ASP is also increasingly gaining attention in industry [16]; for example, in decision support systems [17], in e-tourism [18] and in product configuration [19].

*Corresponding Author:
Email address: mark.law09@imperial.ac.uk
Address: Mark Law, Department of Computing, Huxley Building, 180 Queen's Gate, Imperial College London, London, SW7 2AZ