

Speeding-up Method of Inductive Logic Programming System using Linguistic Bias

Hiroyuki KAMEDA ^{a,1}, Chiaki KUBOMURA ^b, Ryuichi SAITA ^c, and Shun HATTORI ^a

^a School of Computer Science, Tokyo University of Technology, Japan

^b School of Beauty and Health Science, Yamano College of Aesthetics, Japan

^c Creative Field, Nihon Kogakuin College of Hachioji, Japan

Abstract. Inductive logic programming (ILP) is a powerful machine learning method but time-consuming. This paper proposes a method to speed up ILP systems by using language bias (difference-list). The method was evaluated by simulating a simple English grammar acquisition. As a result, speed was about 20 times faster than the previous system.

Keywords. Inductive logic programming, speed-up method, linguistic bias, difference-list, Aleph ILP system, grammar acquisition.

Introduction

Natural language processing (NLP) is one of key techniques in a wide range of research fields including machine translation, advanced user interface, and even requirement engineering [1,2]. This paper proposes a new speed-up method for inductive logic programming (ILP) by using language-constraint condition (language bias). This method was applied to a simple English syntactic rule acquisition system to be evaluated with a small set of English text. Two kinds of evaluation experiments showed process time was improved roughly about 20 times faster than before.

1. Previous methods, their problems and our new speed-up method

One of the oldest ILP researches is Model Inference System. But the system did not handle with complex problems. On the other hand, ILP system Golem adopts rlgg (relatively Least General Generalization). Progol [3] and Aleph [4] adopts mode declaration as language bias [5]. But in naive grammar acquisition, the knowledge of syntactic rules are usually described in the difference list form (d-list form). We adopted this constraint to reduce search space of assumptions.

¹ Corresponding Author: Hiroyuki KAMEDA, 1404-1 Katakura, Hachioji Tokyo 192-0982 JAPAN, kameda@cs.teu.ac.jp .

2. Experiments

2.1. Experiment I

- Objective: To confirm effectiveness of d-list constraint as language bias.
- Method: To count the number of feasible, i.e., language biased assumption candidates, out of all general assumption candidates ILP system Progol generates.
- Experiment materials: 22 simple English sentences as positive examples.
- Result: The ratio of the number of all assumptions generated by Progol to that of feasible ones was about 43.5.
- Discussion: Only about 2% of assumption candidates were feasible ones because Progol cannot handle the d-list language bias directly. This suggests our proposed idea is fundamentally valid.

2.2. Experiment II

- Objective: To confirm effectiveness of our proposal in term of CPU time, when embedding it in ILP system Aleph.
- Method: ILP system with speed-up method and the normal ILP system were compared in terms of running CPU time.
- Experiment materials: Same in experiment I except that 7 negative examples were additionally adopted.
- Result: ILP system with speed-up algorithm ran about 23 times faster than the normal ILP system.
- Discussion: Our proposed method is really effective to reduce CPU time.

Conclusions

This paper proposed a new speed-up method and its fundamental validity was confirmed by two experiments.

References

- [1] S. Nienhuys-Cheng, et al. *Foundations of Inductive Logic Programming*. Springer; 1997.
- [2] S. Dzeroski. Inductive Logic Programming in a Nutshell. In: *Introduction to Statistical Relational Learning*. L. Getoor, et al.(Ed). MIT Press; 2007.
- [3] S. Muggleton. Inverse entailment and Progol, *New Generation Computing*,13:245–286;1995.
- [4] A. Srinivasan. The Aleph Manual (Version 4 and above). [Online] Available from: http://www.comlab.ox.ac.uk/activities/machinelearning/Aleph/aleph_toc.html [Accessed 23 April 2010].
- [5] C. Nedellec, et al. Declarative Bias in Inductive Logic Programming. In: *Advances in Inductive Logic Programming*. L. De Raedt ed. IOS Press;1996, p.82-103.