Learning Grammars by Unification

Wojciech Buszkowski
Faculty of Mathematics and Computer Science
Adam Mickiewicz University in Poznań

Early results in formal learning theory used recursion-theoretic tools; learning procedures were understood as Turing machines computing some (partial) recursive functions. These algorithms often involve a process of running all grammars from the given class to find the first one compatible with the data and satisfying additional conditions. Clearly algorithms of that kind are strongly inefficient. A survey of these methods can be found in (Jain et al. 1999).

We are interested here with different algorithms, which restrict the search space of possible grammars to those which can be computed from the given data. Such algorithms have been designed for various classes of finite-state automata (Angluin 1978, 1982, Yokomori 1995) and context-free grammars (Sakakibara 1990, 1992, Clark 2010, 2011).

We focus on the unification-based procedures, first proposed for type grammars (categorial grammars) in (Buszkowski 1987, Buszkowski and Penn 1990, also related to van Benthem 1987) as a formalization of a natural approach of a linguist who designs a type grammar for a language fragment. In (Kanazawa 1996, 1998) these algorithms were further extended and studied in connection with the Gold-style theory of learning as identification in the limit. Negative data were considered in (Marciniec 1994, Dziemidowicz 2002). Further developments are due to (Foret 2001): unification modulo conjoinability, (Marciniec 1997, 2011): unification of infinite sets of types, (Costa Florêncio 2003, Costa Florêncio and Fernau 2010): complexity results.

In (Buszkowski 2007) it has been shown that the unification framework yields certain procedures of approximate learning: given a class of grammars $\mathcal{G}$ and a finite sample set D, one computes the intersection (resp. union) of all languages from $L(\mathcal{G})$ which contain D, called the lower (resp. upper) approximation of the searched language with respect to D. A description of a concept by means its lower and upper approximation resembles the rough-set approach of (Pawlak 1982, 1991), though we do not use indiscernibility relations.

In this talk I will discuss some basic properties of approximate learning by unification, extending the method beyond type grammars, e.g. for context-free grammars and finite automata (for the latter, it works similarly as earlier methods of forming quotient-sets from the set of states). I also show that a learning method, computing lower approximation, can be treated as an effective example of belief-revision learning methods, considered in (Gierasimczuk 2010).