Doctopic: Algorithms, automata, complexity and games ARTICLE IN PRESS TCS:10876

Theoretical Computer Science ••• (••••) •••–•••

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

A map of update constraints in inductive inference

Timo Kötzing [a,*], Raphaela Palenta [b]

[a] Friedrich-Schiller-Universität Jena, Germany
[b] Technische Universität Müchen, Germany

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Inductive inference
Language learning
Update constraints

ABSTRACT

We investigate how different learning restrictions reduce learning power and how the different restrictions relate to one another. We give a complete map for nine different restrictions both for the cases of complete information learning and set-driven learning. This completes the picture for these well-studied *delayable* learning restrictions. A further insight is gained by different characterizations of *conservative* learning in terms of variants of *cautious* learning.

Our analyses greatly benefit from general theorems we give, for example showing that learners which have to obey only delayable restrictions can always be assumed total.

1. Introduction

This paper is set in the framework of *inductive inference*, a branch of (algorithmic) learning theory. This branch analyzes the problem of algorithmically learning a description for a formal language (a computably enumerable subset of the set of natural numbers) when presented successively all and only the elements of that language. For example, a learner h might be presented more and more even numbers. After each new number, h outputs a description for a language as its conjecture. The learner h might decide to output a program for the set of all multiples of 4, as long as all numbers presented are divisible by 4. Later, when h sees an even number not divisible by 4, it might change this guess to a program for the set of all multiples of 2.

Many criteria for deciding whether a learner h is *successful* on a language L have been proposed in the literature. Gold, in his seminal paper [9], gave a first, simple learning criterion, **TxtGEx**-*learning*,[1] where a learner is *successful* iff, on every *text* for L (listing of all and only the elements of L) it eventually stops changing its conjectures, and its final conjecture is a correct description for the input sequence. Trivially, each single, describable language L has a suitable constant function as a **TxtGEx**-learner (this learner constantly outputs a description for L). Thus, we are interested in analyzing for which *classes of languages* $\mathcal{L}$ there is a *single learner* h learning *each* member of $\mathcal{L}$. This framework is also sometimes known as *language learning in the limit* and has been studied extensively, using a wide range of learning criteria similar to **TxtGEx**-learning (see, for example, the textbook [11]).

A wealth of learning criteria can be derived from **TxtGEx**-learning by adding restrictions on the intermediate conjectures and how they should relate to each other and the data. For example, one could require that a conjecture which is consistent with the data must not be changed; this is known as *conservative* learning and known to restrict what classes of languages

* Corresponding author.
E-mail address: timo.koetzing@hpi.de (T. Kötzing).

[1] **Txt** stands for learning from a *text* of positive examples; **G** stands for Gold, who introduced this model, and is used to indicate full-information learning; **Ex** stands for *explanatory*.

http://dx.doi.org/10.1016/j.tcs.2016.07.028

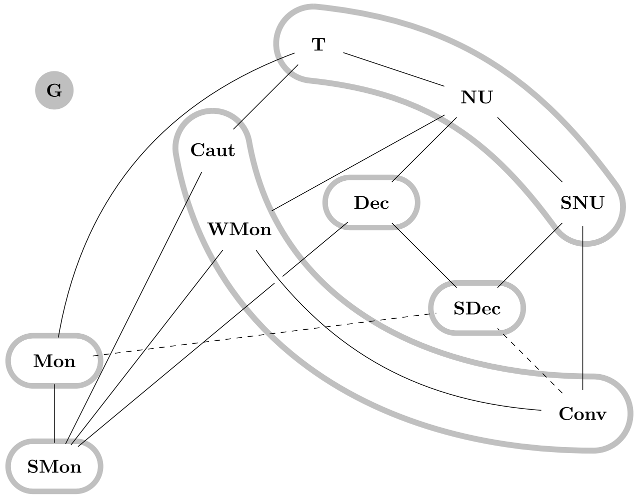

Fig. 1. Relation of criteria.

can be learned ([1], we use **Conv** to denote the restriction of conservative learning). Additionally to conservative learning, the following learning restrictions are considered in this paper (see Section 2.1 for a formal definition of learning criteria including these learning restrictions).

In *cautious* learning (**Caut**, [18]) the learner is not allowed to ever give a conjecture for a strict subset of a previously conjectured set. In *non-U-shaped* learning (**NU**, [3]) a learner may never *semantically* abandon a correct conjecture; in *strongly non-U-shaped* learning (**SNU**, [7]) not even syntactic changes are allowed after giving a correct conjecture.

In *decisive* learning (**Dec**, [18]), a learner may never (semantically) return to a *semantically* abandoned conjecture; in *strongly decisive* learning (**SDec**, [14]) the learner may not even (semantically) return to *syntactically* abandoned conjectures. Finally, a number of monotonicity requirements are studied [10,24,17]: in *strongly monotone* learning (**SMon**) the conjectured sets may only grow; in *monotone* learning (**Mon**) only incorrect data may be removed; and in *weakly monotone* learning (**WMon**) the conjectured set may only grow while it is consistent.

The main question is now whether and how these different restrictions reduce learning power. For example, non-U-shaped learning is known not to restrict the learning power [3], and the same for strongly non-U-shaped learning [7]; on the other hand, decisive learning *is* restrictive [3]. The relations of the different monotone learning restriction were given in [17]. Conservativeness is long known to restrict learning power [1], but also known to be equivalent to weakly monotone learning [16,12].

Cautious learning was shown to be a restriction but not when added to conservativeness in [18,19], similarly the relationship between decisive and conservative learning was given. In Exercise 4.5.4B of [19] it is claimed (without proof) that cautious learners cannot be made conservative; we claim the opposite in Theorem 4.4.

This list of previously known results leaves a number of relations between the learning criteria open, even when adding trivial inclusion results (we call an inclusion trivial iff it follows straight from the definition of the restriction without considering the learning model, for example strongly decisive learning is included in decisive learning; formally, trivial inclusion is inclusion on the level of learning restrictions as predicates, see Section 2.1). With this paper we now give the complete picture of these learning restrictions. The result is shown as a map in Fig. 1. A solid black line indicates a trivial inclusion (the lower criterion is included in the higher); a dashed black line indicates an inclusion which is not trivial. A gray box around criteria indicates equality of (learning of) these criteria.

A different way of depicting the same results is given in Fig. 2 (where solid lines indicate inclusion). Results involving monotone learning can be found in Section 7, results on the particularly difficult relations of decisive learning in Section 5, all others in Section 4.

For the important restriction of conservative learning we give the characterization of being equivalent to cautious learning. Furthermore, we show that even two weak versions of cautiousness are equivalent to conservative learning. Recall that cautiousness forbids to return to a strict subset of a previously conjectured set. If we now weaken this restriction to forbid to return to *finite* subsets of a previously conjectured set we get a restriction still equivalent to conservative learning. If we forbid to go down to a correct conjecture, effectively forbidding to ever conjecture a superset of the target language, we also obtain a restriction equivalent to conservative learning. On the other hand, if we weaken it so as to only forbid going to *infinite* subsets of previously conjectured sets, we obtain a restriction equivalent to no restriction. These results can be found in Section 4.

In *set-driven* learning [23] the learner does not get the full information about what data has been presented in what order and multiplicity; instead, the learner only gets the set of data presented so far. For this learning model it is known that, surprisingly, conservative learning is no restriction [16]! We complete the picture for set driven learning by showing that