ELSEVIER
Available online at www.sciencedirect.com
SCIENCE DIRECT®
Theoretical Computer Science 350 (2006) 1–2
Theoretical Computer Science
www.elsevier.com/locate/tcs

Foreword

This special issue of *Theoretical Computer Science* is dedicated to the Thirteenth International Conference on Algorithmic Learning Theory (ALT 2002) held at the Mövenpick Hotel in Lübeck, Germany, November 24–26, 2002. It contains ten articles that were among the best in the conference.[1] The authors of these papers have been invited by the Special Issue Editors to submit completed versions of their work for this Special Issue. Once received, these papers underwent the usual refereeing process of *Theoretical Computer Science*.

The ALT series is focusing on all areas related to algorithmic learning theory such as statistical learning, learning via queries, identification of formal languages, learning logical formulae, information extraction, inductive inference, inductive logic programming, on-line learning as well as to specific algorithmic approaches, e.g., margin-based algorithms, MDL estimation. The diversity of approaches to learning is also reflected in this Special Issue.

The first three papers are concerned with computationally efficient statistical learning of Boolean functions, a field initiated in theoretical computer science by Valiant who introduced the PAC learning model. In this model examples of the target Boolean function are drawn from an unknown probability distribution, and the learner is required to approximate the target function to within a given accuracy. The paper by Amano and Maruoka deals with the important class of monotone Boolean functions. Their first result shows that a single variable function has the minimum correlation with the majority function among all fair monotone functions. This theorem had been conjectured by Blum, Burch and Langford in 1998 and improves the performance guarantee of the best-known learning algorithm for monotone Boolean functions under the uniform distribution. Their second result goes part of the way to establish an optimal algorithm for learning monotone Boolean functions.

The paper by Servedio studies the PAC learnability of another important class of Boolean functions: those representable by constant-depth, polynomial-size circuits of unbounded fan-in (the ACC class). A first result relates the efficient learnability of ACC to that of a different class: the embedded midbit functions. Several positive and negative results on the efficient learnability of this latter class, providing important insights on the learnability of ACC, are shown for the PAC model and some of its restrictions.

The third paper is by Bshouty and Burroughs. The main theme is the proof of impossibility results for efficient learning of several subclasses of Boolean functions in an extension of PAC learning called co-agnostic learning. The key proof technique is the construction of reductions between learning problems and well-studied approximability problems for which there exist known hardness results.

Large-margin learning algorithms, like support vector machines, boosting and their variants, build on results from statistical learning theory to generate linear hypotheses with good generalization abilities. Forster and Simon study fundamental questions about the use of linear-threshold hypotheses. Statistical learning theory ensures that a good generalization error is achieved whenever the sample can be embedded in a space where some linear hypothesis can be found that separates positive examples from negative ones with large margin. The paper proves bounds relating the spectrum of a matrix derived from the sample to the margin which can be obtained via the embedding technique.

The paper by Köbler and Lindner deals with a model where the learning algorithm is allowed to ask certain queries to a teacher. The learning performance is measured by the number and type of queries needed by the learner to identify the target concept. They use the general dimension to characterize the complexity of learning using queries. Their main

[1] The conference proceedings, including preliminary versions of these papers, appeared as Lecture Notes in Artificial Intelligence, Vol. 2533, Springer, Berlin, 2002.

0304-3975/$ - see front matter
doi:10.1016/j.tcs.2005.10.011

result shows that any class learnable with a polynomial number of polynomial-size superset and subset queries can be learned in polynomial time with such queries and the help of an oracle in Σ_3^p. Their main result also implies that DNF formulas can be properly learned in polynomial time with subset and superset queries with the help of an oracle from Σ_3^p.

The paper by Suzuki, Shoudai, Uchida, and Miyahara belongs to the area of information extraction. This is a rapidly growing applied research area whose goal is the design of systems able to extract meaningful information from structured, semi-structured, and unstructured data sources. Important subareas of information extraction are database mining, web mining, and text analysis. In their work, algorithms for the efficient learning of ordered tree structures are considered.

Reidenbach solves a long standing open problem, i.e., the learnability of E-pattern languages from positive data. Classically, this paper also belongs to the area of inductive inference. On the other hand, learning algorithms for pattern languages have found many interesting applications. A pattern is simply a string over a terminal alphabet and variables. The language generated by a pattern is obtained by substituting the variables by strings over the terminal alphabet. For the case that only non-empty strings are allowed, the learnability of pattern languages from positive data has been established roughly 25 years ago. If empty substitutions are allowed we obtain the class of E-pattern languages. Their learnability from positive data had remained open. As far as the whole class of E-pattern languages is concerned, Reidenbach shows that they are *not learnable* from positive data.

The work by Martin, Sharma, and Stephan goes towards providing a unified view over logic, learning, and topology. The outcome is a notion of parameterized logic, based on topology, that makes use of learning-theoretic notions for its interpretation and application. The framework is also seen to provide a nice interpolation between applications based on induction and deduction.

Hayashi presents to a learning theory audience a new approach to the foundations of mathematics called “Limit-Computable Mathematics (LCM).” This approach is based on the notion of functions computable in the limit. In his exposition Hayashi aims at indicating the motivations and the basics of LCM, as well as at indicating its relationships to learning theory and to reverse mathematics. LCM can be seen as an extension of intuitionistic mathematics, in which the concept of “construction” in the classical Brouwer–Heyting–Kolmogorov interpretation of first-order formulas is interpreted as “function computable in the limit.”

The last paper in this special issue makes a contribution to a problem fundamental in various types of inference, e.g., abductive inference, inductive inference, machine learning, and machine discovery, i.e., it addresses the famous hypothesis finding problem from background knowledge and examples. Fronhöfer and Yamamoto propose a new approach to this problem which is based on the definition of a minimal residue hypothesis. In this way, the authors shed new light on the question what does it mean for a hypothesis to be appropriate. This paper has also some interesting connections to relevance logic which are used to show a tradeoff between “logical smallness” and “syntactic smallness.”

We would like to express our immense gratitude to the referees for their fine reports and their efficient work, and to all the members of the program committee of ALT 2002 for selecting the papers. Our special thanks go to all authors for submitting their papers and for all their efforts to improve and to polish their articles. Unfortunately, this process caused some unusual delay and we would like to apologize for it and thank all authors, who provided their work in time, for their patience.

Moreover, we are particularly thankful to Giorgio Ausiello for providing the opportunity to compile this special issue.

Milano, Lübeck and Sapporo, May 29, 2005,

Special Issue Editors
Nicolò Cesa-Bianchi
Rüdiger Reischuk
Thomas Zeugmann
Hokkaido University, Graduate School of Information Science and Technology,
Division of Computer Science, Kita 14-jo, Nishi 9-chome, Sapporo, 060-0814, Japan
E-mail address: thomas@ist.hokudai.ac.jp (T. Zeugmann)