Theoretical Computer Science 410 (2009) 1747–1748

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Preface

This special issue contains expanded versions of papers that appeared in preliminary form in the proceedings of the 18th International Conference on Algorithmic Learning Theory (ALT 2007), which was held in Sendai, Japan, during October 1–4, 2007. *Algorithmic Learning Theory* is a conference series which is dedicated to the theoretical study of the algorithmic aspects of learning. The best papers of the conference ALT 2007 were invited for this special issue and after a thorough reviewing process, most of them qualified for this Special Issue on Algorithmic Learning Theory of Theoretical Computer Science. In the following, a short introduction is given to each of these papers.

Markus Maier, Matthias Hein, and Ulrike von Luxburg study a scenario in which a learning algorithm receives a sample of points from an unknown distribution which contains a number of distinct clusters. The goal in this setting is to construct a "neighborhood graph" from the sample, such that the connected component structure of the graph mirrors the cluster ancestry of the sample points. They prove bounds on the performance of the k-nearest neighbor algorithm for this problem and also give some supporting experimental results. This paper received the E. Mark Gold award at the conference for the most outstanding paper co-authored by a student.

Kevin Chang also considers an unsupervised learning scenario, but one in which a learner is given access to a sequence of samples drawn from a mixture of uniform distributions over rectangles in d-dimensional Euclidean space. He gives a streaming algorithm which makes only a small number of passes over such a sequence, uses a small amount of memory, and constructs a high-accuracy (in terms of statistical distance) hypothesis density function for the mixture. A notable feature of the algorithm is that it can handle samples from the mixture that are presented in any arbitrary order. This result extends earlier work of Chang and Kannan which dealt with mixtures of uniform distributions over rectangles in one or two dimensions.

V.V. V'yugin studies defensive forecasting in the setting of online prediction of the binary label associated with each instance in a sequence of instances. In this line of work no assumption is made that there exists a hidden function dictating the labels, and in contrast with other work in online learning there is no comparison class or "best expert" that is compared with. One well-studied parameter of algorithms in this setting is the calibration error, which roughly speaking measures the extent to which the forecasts are accurate on average. In his paper V'yugin establishes a tradeoff between the calibration error and the "coarseness" of any prediction strategy by showing that if the coarseness is small then the calibration error cannot also be too small. This negative result comes close to matching the bounds given in previous work by Kakade and Foster on a particular forecasting system.

Sanjay Jain, Frank Stephan and Nan Ye study some basic questions about how hypothesis spaces connect to the class of languages being learned in Gold-style models. Building on work by Angluin, Lange and Zeugmann, their paper introduces a comprehensive unified approach to studying learning languages in the limit relative to different hypothesis spaces. Their work distinguishes between four different types of learning as they relate to hypothesis spaces, and gives results for vacillatory and behaviorally correct learning. They further show that every behaviorally correct learnable class has a *prudent* learner, i.e. a learner using a hypothesis space such that it learns every set in the hypothesis space.

Ryo Yoshinaka studies learning languages in the limit from positive data. His paper addresses the question of what precisely is meant by the notion of efficient language learning in the limit; despite the clear intuitive importance of such a notion, there is no single accepted definition. The discussion focuses particularly on learning very simple grammars and minimal simple grammars from positive data, giving both positive and negative results on efficient learnability under various notions.

M.M. Hassan Mahmud analyzes transfer learning from the perspective of Kolmogorov complexity. The goal in transfer learning is to solve new learning problems more efficiently by leveraging information that was gained in solving previous related learning problems. One challenge in this area is to clearly define the notion of "relatedness" between tasks in a rigorous yet useful way. Mahmud shows that if tasks are related in a particular precise sense, then joint learning is indeed faster than separate learning. This work strengthens previous work by Bennett, Gacs, Li, Vitanyi and Zurek.

doi:10.1016/j.tcs.2009.01.008

Kilho Shin and Tetsuji Kuboyama give a sufficient condition under which it is ensured that new candidate kernels constructed in a particular way from known positive semidefinite kernels will themselves be positive semidefinite. Developing new kernel functions, and selecting the most appropriate kernels for particular learning tasks, is an active area of research. One difficulty in constructing kernel functions is in ensuring that they obey the necessary condition of positive semidefiniteness; thus it is useful to have an automatic way to ensure that candidate kernels satisfy this property. The work of Shin and Kuboyama gives new insights into several kernel functions that have been studied recently such as principal-angle kernels, determinant kernels, and codon-improved kernels.

John Case and Samuel Moelius III study iterative learning. This is a variant of the Gold-style learning model described above in which each of a learner's output conjectures may depend only on the learner's current conjecture and on the current input element. Case and Moelius analyze two extensions of this iterative model which incorporate parallelism in different ways. Roughly speaking, one of their results shows that running several distinct instantiations of a single learner in parallel can actually increase the power of iterative learners. This provides an interesting contrast with many standard settings where allowing parallelism only provides an efficiency improvement. Another result deals with a "collective" learner which is composed of a collection of communicating individual learners that run in parallel.

Jean-Yves Audibert, Rémi Munos and Csaba Szepesvári deal with the stochastic multi-armed bandit setting. They study an Upper Confidence Bound algorithm that takes into account the empirical variance of the different arms. They give an upper bound on the expected regret of the algorithm, and also analyze the concentration of the regret; this risk analysis is of interest since it is clearly useful to know how likely the algorithm is to have regret much higher than its expected value. The risk analysis reveals some unexpected tradeoffs between logarithmic expected regret and concentration of regret.

Last but not least, Vitaly Feldman and Shrenik Shah analyze two previously studied variants of Angluin's exact learning model that make learning more challenging: learning from equivalence and incomplete membership queries, and learning with random persistent classification noise in membership queries. They show that under cryptographic assumptions about the computational complexity of solving various problems the former oracle is strictly stronger than the latter, by demonstrating a concept class that is polynomial-time learnable from the former oracle but is not polynomial-time learnable from the latter oracle. They also resolve an open question of Bshouty and Eiron by showing that the incomplete membership query oracle is strictly weaker than a standard perfect membership query oracle under cryptographic assumptions.

We wish to thank the referees for their thorough and thoughtful reviews, Thomas Zeugmann for his help and, most of all, the authors for contributing their papers.

Marcus Hutter *
RSISE, Australian National University and SML,
NICTA Canberra ACT 0200, Australia
E-mail address: marcus.hutter@anu.edu.au.

Rocco A. Servedio
Department of Computer Science,
Columbia University, New York,
NY 10027, USA
E-mail address: rocco@cs.columbia.edu.

* Corresponding editor.