Machine Learning, 5, 1–37 (1997)

On Restricted-Focus-of-Attention Learnability of Boolean Functions

ANDREAS BIRKENDORF birkendo@ls2.informatik.uni-dortmund.de

Universität Dortmund, Fachbereich Informatik, D-44221 Dortmund, Germany.

ELI DICHTERMAN eli@cdam.lse.ac.uk

Department of Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, UK. And, Department of Computer Science, Royal Holloway University of London, Egham, Surrey TW20 0EX, UK.

JEFFREY JACKSON jackson@mathcs.duq.edu

Math and Computer Science Department, Duquesne University, 600 Forbes Avenue, Pittsburgh, PA 15282, USA.

NORBERT KLASNER klasner@ls2.informatik.uni-dortmund.de

Universität Dortmund, Fachbereich Informatik, D-44221 Dortmund, Germany.

HANS ULRICH SIMON simon@ls2.informatik.uni-dortmund.de

Universität Dortmund, Fachbereich Informatik, D-44221 Dortmund, Germany.

Received August 20, 1997; Revised August 20, 1997

Editor: Dana Ron

Abstract. In the k-Restricted-Focus-of-Attention (k-RFA) model, only k of the n attributes of each example are revealed to the learner, although the set of visible attributes in each example is determined by the learner. While the k-RFA model is a natural extension of the PAC model, there are also significant differences. For example, it was previously known that learnability in this model is not characterized by the VC-dimension and that many PAC learning algorithms are not applicable in the k-RFA setting.

In this paper we further explore the relationship between the PAC and k-RFA models, with several interesting results. First, we develop an information-theoretic characterization of k-RFA learnability upon which we build a general tool for proving hardness results. We then apply this and other new techniques for studying RFA learning to two particularly expressive function classes, k-decision-lists (k-DL) and k-TOP, the class of thresholds of parity functions in which each parity function takes at most k inputs. Among other results, we prove a hardness result for k-RFA learnability of k-DL, $k \leq n-2$. In sharp contrast, an $(n-1)$-RFA algorithm for learning $(n-1)$-DL is presented. Similarly, we prove that 1-DL is learnable if and only if at least half of the inputs are visible in each instance. In addition, we show that there is a uniform-distribution k-RFA learning algorithm for the class of k-DL. For k-TOP we show weak learnability by a k-RFA algorithm (with efficient time and sample complexity for constant k) and strong uniform-distribution k-RFA learnability of k-TOP with efficient sample complexity for constant k. Finally, by combining some of our k-DL and k-TOP results, we show that, unlike the PAC model, weak learning does *not* imply strong learning in the k-RFA model.

Keywords: Restricted Focus of Attention, PAC-Learning, Learning Algorithms, Boolean Function Classes, Decision Lists, Threshold of Parities, Fourier Transform