Best Available Copy
L.G. Valiant
Harvard University
Cambridge, MA 02138

Work was centered on quantifiable aspects of machine learning, particularly efficient algorithms for learning new classes of representations, and proofs of limitations. We list here the publications arising from the last four years of the work, omitting earlier work.


We shall mention five highlights from the above papers:

(a) NP-completeness results: In [42] it was shown that some very restricted representation classes, such as 2-term dnf, are difficult to learn, if the learner is forced to express the hypothesis in the restricted class, but becomes tractable if the learner is allowed a more expressive representation in learning. The technique has been applied by others since to other representations, such as neural nets.

(b) Representation independent hardness results: In [20] it was shown that general Boolean formulas and finite automata are hard to learn in the pac model (assuming some cryptographic conjectures) whatever representation is chosen by the learner. This says, roughly, that a black box containing an unknown formula can behave essentially as a random function, about which little can be found out by feasible experimentation. Some remarkable extensions of these results have been developed recently by others.

(c) Algorithms for learning several new subclasses of decision trees and disjunctive normal form were discovered by Hancock and described in his thesis
and several papers. A powerful application of the Fourier spectrum method was applied by Kushilevitz and Mansour to learn decision trees under the uniform distribution with membership queries [27].

(d) The concept of weak learning was introduced [20, 25]. The result of Schapire while at MIT and others subsequently showed that any weak learning algorithm in the pac model could be "boosted" to higher accuracy by a very general procedure. Recent experiments at AT&T suggest that this boosting method is applicable to natural data, such as images of digits.

(e) A monograph based on the neuroidal model introduced in [44] is near completion. The model is an attempt at a basis for giving a computational account of some of the most basic tasks of memorization and learning. The tasks considered are restricted to "random access" tasks, which are defined to be those that potentially involve any part of memory. In this context the model aims to capture the relevant properties of real neurons in cortex. The algorithms developed attempt to be consistent with the quantitative parameters of human performance, such as processing time and memory capacity.