On the Hardness of Approximating the Minimum Consistent OBDD Problem

Kouichi Hirata\textsuperscript{1} Shinichi Shimozono\textsuperscript{2} Ayumi Shinohara\textsuperscript{3}

\textsuperscript{1} Department of Artificial Intelligence
\textsuperscript{2} Department of Control Engineering and Science
Kyushu Institute of Technology,
Kawazu 680-4, Iizuka, 820 JAPAN
E-mail: hirata@ai.kyutech.ac.jp, sin@ces.kyutech.ac.jp
\textsuperscript{3} Research Institute of Fundamental Information Science
Kyushu University, Fukuoka 812 JAPAN
E-mail: ayumi@rifis.kyushu-u.ac.jp

Abstract. Ordered binary decision diagrams (OBDDs, for short) represent Boolean functions as directed acyclic graphs. The minimum consistent OBDD problem is, given an incomplete truth table of a function, to find the smallest OBDD that is consistent with the truth table with respect to a fixed order of variables. We show that this problem is NP-hard, and prove that there is a constant $\epsilon > 0$ such that no polynomial time algorithm can approximate the minimum consistent OBDD within the ratio $n^\epsilon$ unless P=NP, where $n$ is the number of variables. This result suggests that OBDDs are unlikely to be polynomial time learnable in PAC-learning model.

1 Introduction

For a class of representations of languages, the minimum consistent problem is to find a representation that is as small size as possible and is consistent with given positive and negative examples. The computational complexity of the problem closely relates to the efficiency of learning with the target class. Roughly speaking, the polynomial-time learnability of a class is equivalent to the existence of a polynomial-time algorithm which can produce a consistent representation whose size is not too large compared to the smallest one. Therefore, for the minimum consistent problem that is intractable, a polynomial-time algorithm that can find an approximately small representation is undoubtedly important.

One of the minimum consistent problems that are most frequently tackled is that for deterministic finite automata (DFAs for short). The problem was first shown to be NP-hard [1, 12]. This negative result was enhanced by Li and Vazirani [14]: they showed that the minimum consistent DFA cannot be
approximated within the ratio $\frac{9}{8}$ in polynomial time, unless P=NP. Pitt and Warmuth [17] improved to the ratio $opt^k$, where $opt$ is the minimum number of states and $k$ is any positive integer.

By applying recent results on non-approximabilities of combinatorial optimization problems, Hancock et al. [13] investigated the minimum consistent problem for decision lists and decision trees. They showed that decision lists cannot be approximated in polynomial time within a factor of $n^c$ for some constant $c > 0$, unless P=NP. They also showed that decision trees cannot be approximated in polynomial time within a factor of $n^c$ for any $c > 0$ unless NP is included in DTIME[$2^{poly \log n}$].

This paper deals with the minimum consistent problem for ordered binary decision diagrams (OBDDs for short). OBDDs succinctly represent many useful Boolean functions, such as symmetric functions and threshold functions [8], as directed acyclic graphs. For a fixed order of variables, OBDDs can be regarded as acyclic DFAs whose sizes are measured by the number of only the branching states [11]. It is known that the problem of finding the optimal order that realizes the minimum size OBDD is intractable even for complete truth tables [7]. To concentrate on the complexity to deal with incomplete truth tables, we fix the order of variables in the minimum consistent OBDD problem. Our result claims that the problem cannot be approximated within the ratio $n^\epsilon$ for some $\epsilon > 0$ in polynomial time unless P = NP. This is in sharp contrast to the well-known fact that the minimum size OBDD representing the function can be computed in linear time from the complete truth table [19].

The remaining part of this paper is organized as follows. First, we show that the minimum consistent OBDD problem is NP-hard, even the number of either positive examples or negative examples is only one. Then, we show that there is a constant $\epsilon > 0$ such that no polynomial time algorithm can approximate the minimum consistent OBDD within the ratio $n^\epsilon$ unless P=NP. Our results suggest that any efficient algorithm for learning OBDDs would have to produce very large hypotheses.

2 Minimum Consistent OBDD Problem

We first give definitions and notations for OBDDs. Then we introduce the minimum consistent problem of OBDD as a combinatorial optimization problem, and show that the problem is NP-hard.

---

4 Some results are known for similar problems. See Section 6.