Decision trees and disjoint covers

Balázs Szörényi and György Turán

We investigate the relation between two complexity measures used for a Boolean function: the decision tree size (DTS), which is the minimal number of leaves of a decision tree for the function, and the disjoint cover size (DCS), which is the minimal number of subcubes needed to cover the n-dimensional cube $\{0,1\}^n$, such that the subcubes are disjoint and the cover is consistent with the function (i. e., for each subcube, the function evaluates the same on each vertex of the subcube). Note that $\mathrm{DTS} \geq \mathrm{DCS}$, and that determining a disjoint cover for a function f is just the same as determining for f and for $\neg f$ a pair of disjunctive normal forms (DNFs) in which each two distinct terms conflict in at least one variable.

Our investigation is motivated by the paper of Jukna et al [2]. They have shown that there is superpolynomial gap between the DTS and the cover size of the Boolean functions, where the cover size (CS) is the same as the DCS without requiring the subcubes to be disjoint—note again that determining a cover for a function f is just the same as determining for f and for $\neg f$ a pair of DNFs. More specifically, they have presented a Boolean function for which $\mathrm{DTS} = \exp\left(\Omega\left((\log \mathrm{CS})^2\right)\right)$. Their result almost matches the upper bound, $\mathrm{DTS} = \exp\left(\mathrm{O}\left((\log \mathrm{CS})^2 \log \mathrm{n}\right)\right)$, which was proved by Ehrenfeucht and Haussler in [1] to hold for any Boolean function. A question raised by their result is whether one can prove a similar separation between DCS and DTS. The Fourier technique used in their result cannot be used for this purpose.

In our paper we show that there is a superpolynomial gap between the DCS and the DTS. More specifically we present a Boolean function for which $\mathrm{DTS} = \exp\left(\Omega\left((\log \mathrm{DCS})^\delta\right)\right)$ for any positive $\delta < \log_{\left(1+\sqrt{3}\right)} 3 \approx 1.093$. For this, of course, we have to develop a technique, different from the one used in [2], to lower bound the DTS. We also show that our technique gives “essentially” the same lower bound on the DTS of the example used by Jukna et al. in [2] as their method.

References

[1] A. Ehrenfeucht and D. Haussler. Learning decision trees from random examples. *Inf. Comput.*, 82(3):231–246, 1989.

[2] S. Jukna, A. Razborov, P. Savický, and I. Wegener. On P versus NP∩co-NP for decision trees and read-once branching programs. *Comput. Complex.*, 8(4):357–370, 1999.