
Learning Talagrand DNF Formulas

Homin K. Lee

University of Texas at Austin
homin@cs.utexas.edu

Consider the following version of Talagrand’s probabilistic construction of a monotone function $f : \{0, 1\}^n \rightarrow \{0, 1\}$. Let f be an n -term monotone DNF formula where each term is selected independently and uniformly at random (with replacement) from the set of all $n^{\Omega(\log \log n)}$ possible terms of length $\log(n)$ over the first $\log^2(n)$ variables. Let us call such a DNF formula a *Talagrand DNF formula*. This is a scaled-down version of the construction by Talagrand (1996) which is defined over all n variables and has subexponentially many terms. I am interested in the following problem:

Does there exist a polynomial-time algorithm for learning the class of Talagrand DNF formulas over the uniform distribution on $\{0, 1\}^n$ given uniform random examples?

Ideally, I am looking for algorithms that will learn the class of *all* possible Talagrand DNF formulas in the worst-case. However, an average-case learning algorithm that succeeds with high probability over the choice of the Talagrand DNF formula as described above would be of significant interest as well.

Motivation: This problem is of course a special case of the corresponding learning problem for general polynomial-size DNF formulas. The problem of learning polynomial-size DNF formulas *without* queries has been open for almost twenty years (Valiant, 1984) and there has been no significant progress on the question until the last couple years.

Current Status: Recently, *random* DNF formulas were shown to be learnable in the following sequence of work (Jackson & Servedio, 2006; Jackson et al., To appear; Sellie, 2008; Sellie, 2009). The algorithms for learning random DNF formulas only work when the terms are well-separated, *i.e.*, when the terms share very few variables, which is clearly not the case for Talagrand DNF formulas.

Some Observations:

1. Unlike the class of all polynomial-size DNF formulas, the class of Talagrand DNF formulas as defined above do not suffer from the “junta” problem (Blum, 2003). We know exactly which variables are relevant.
2. The Talagrand functions are sensitive to noise as small as $1/\log(n)$, *i.e.*, $\Pr[f(x) \neq f(y)] \geq \Omega(1)$ where y is x with each bit flipped independently with probability $1/\log(n)$ (Mossel & O’Donnell, 2003). Thus, any variant of the “low-degree algorithm” (Linial et al., 1993) is unlikely to work for this problem.
3. Unlike for the case of all polynomial-size DNF formulas, there are no known reasons for ruling out statistical query (SQ) algorithms for this problem. (The algorithms for learning random DNF formulas cited above can all be couched as SQ algorithms.) Strong SQ lower bounds are known for depth-3 monotone formulas (Feldman et al., 2010), but there are no known strong SQ lower bounds for any subclass of monotone DNF formulas.

Rewards:

A hand shake: Demonstrate a uniform-distribution learning algorithm for Talagrand DNF formulas in the average-case.

A hand shake and a pat on the back: Demonstrate a uniform-distribution learning algorithm for Talagrand DNF formulas in the worst-case.

A surprised look: Prove a super-polynomial strong-SQ lower bound for the class of Talagrand DNF formulas.

References

- Blum, A. (2003). Learning a function of r relevant variables. *Proc. of the 16th Annual Conference on Computational Learning Theory (COLT)* (pp. 731–733). Springer-Verlag.
- Feldman, V., Lee, H. K., & Servedio, R. A. (2010). Lower bounds and hardness amplification for learning shallow monotone formulas. Manuscript.
- Jackson, J. C., Lee, H. K., Servedio, R. A., & Wan, A. (To appear). Learning random monotone DNF. *Discrete Applied Mathematics*. Prelim. ver. in *Proc. of RANDOM'08*.
- Jackson, J. C., & Servedio, R. A. (2006). On learning random DNF formulas under the uniform distribution. *Theory of Computing*, 2, 147–172.
- Linial, N., Mansour, Y., & Nisan, N. (1993). Constant depth circuits, Fourier transform, and learnability. *Journal of the ACM*, 40, 607–620. Prelim. ver. in *Proc. of FOCS'89*.
- Mossel, E., & O'Donnell, R. (2003). On the noise sensitivity of monotone functions. *Random Structures and Algorithms*, 23, 333–350.
- Sellie, L. (2008). Learning random monotone DNF under the uniform distribution. *Proc. of the 21th Annual Conference on Computational Learning Theory (COLT)* (pp. 181–192).
- Sellie, L. (2009). Exact learning of random dnf over the uniform distribution. *Proc. 41st Annual ACM Symposium on Theory of Computing (STOC)* (pp. 45–54).
- Talagrand, M. (1996). How much are increasing sets positively correlated? *Combinatorica*, 16, 243–258.
- Valiant, L. G. (1984). A theory of the learnable. *Communications of the ACM*, 27, 1134–1142. Prelim. ver. in *Proc. of STOC'84*.