Learning Talagrand DNF Formulas

Homin K. Lee
UT-Austin
Disjunctive Normal Form:
OR of AND of literals

Size is the number of AND gates (terms).
PAC Learning DNF Formulas

A is a PAC-learner for poly(n)-size DNF if ∀f in the class given uniform random examples (x,f(x)) w.h.p. outputs h s.t.
\[ \Pr[h(x) = f(x)] \geq 1 - \varepsilon \]

Best alg takes time \( n^{O(\log n/\varepsilon)} \) [V84]

[V90]
Juntas

Boolean funcs that depend on $\leq k$ vars.

Best alg takes time $n^{0.7k}$ \[\text{[MOS03]}\]

Learning DNF $\Rightarrow$ Learning $O(\log n)$ Juntas

\[\text{[B03]}\]
Parity with Noise

\[ S = \{x_1, x_5, x_8, x_9\} \]

\[ \chi_S(x) = 1 \text{ if odd } \# \text{ of vars in } S \text{ are set to } 1. \]

\[ \chi_S(x) \oplus \eta, \ \eta = 1 \text{ w.p. } p \]

Best alg takes time \( 2^{O(n/\log n)} \) \[\text{[BKW00]}\]

Learning PWN, \( |S| = O(\log n) \implies \) Learning DNF \[\text{[FGKP06]}\]
An SQ-oracle given $g$, outputs a good estimate to $E[g(x,f(x))]$

SQ-learners for DNF take $n^{\omega(1)}$ queries [K93]

Almost all PAC-learning algs are SQ algs!
Monotone DNF

Monotone: no negations on the literals

\[ x_1x_2x_4x_6 \lor x_1x_2x_5 \lor x_1x_2x_3 \]

A Theory of the Learnable

advantageous. The question as to whether monotone
DNF expressions can be learned from EXAMPLES alone
is open. A positive answer would be especially signifi-
No Excuses!

Monotone juntas are easy.
MDNF can’t compute parity.
No SQ lower bounds.
No consequences!
Known Results

• Poly(n)-size read-k MDNF. [HM91]
• Size-2^{\sqrt{\log(n)}} MDNF [S01]
• Random poly(n)-size MDNF [S08, JLSW08]
  – Pick t terms uniformly from all terms of size log(t)
  – Relies on terms not overlapping too much

Pretty pitiful.
Setting a Goal

- Read-o(1)
- Size $\Omega(n)$
- Overlapping terms
Talagrand DNF

Pick $n$ terms from set of all terms of length $\log(n)$ defined over first $\log^2(n)$ variables. [T96]

- Size $n$, read-$o(1)$.
- Know all relevant variables.
- Lots of overlap.
Talagrand DNF

Pick $n$ terms from set of all terms of length $\log(n)$ defined over first $\log^2(n)$ variables. 

[T96]

- $f$ is sensitive to low noise
  \[ \Pr[f(x) \neq f(y)] = \Omega(1) \]
  $y = x$ with each bit flipped with prob $1/\log(n)$

- $f$ has high “surface area” $\Omega(\sqrt{\log(n)})$
Prizes

• PAC-learn Talagrand DNFs w.h.p. over the choice of DNF.
• PAC-learn Talagrand DNFs in the worst case.
• Prove that Talagrand DNFs require $n^{\omega(1)}$ SQ-queries [FLS10].