Last time: • Application of Occam’s Razor: learning sparse disjunctions using few examples with (fairly) sparse disjunctions as hypotheses (Greedy Set Cover heuristic)

Today: • Proper vs Improper learning (improperly learning: 3-term DNF is computationally easy, but properly learning 3-term DNF is computationally hard)
• Start VC dimension (KV Chapter 3)

Reminder: midterm in-person (closed book/notes) next Thurs Oct 20

Questions?

\[ C = \text{all 3-term DNFs} \quad X = \{0, 1\}^n \]

Fact: if \( f = T_1 \lor T_2 \lor T_3 \) is a 3-term DNF, then \( f \) can be expr. as a 3-CNF.

Pf: De Morgan: \[ (a \land b) \lor (c \land d) = (a \lor c) \land (a \lor d) \land (b \lor c) \land (b \lor d) \]

Ex: \( \text{abcd vef vghi} \equiv (a \lor e \lor u \lor g \lor v \lor i) \land (a \lor e \lor h \lor i) \land (a \lor e \lor h \lor g) \land \ldots \)
So, to PAC learn 3-term DNFs, enough to give PAC learning alg for 3CNFs.

Thus: The class of 3CNFs is efficiently PAC learnable using the hyp. class $\mathcal{H}$ = all 3CNF formulas.

Proof:

Recall that class of all mon. conj. over $\{0,1\}^n$ is (properly) PAC learnable with $m = \frac{\log(1/\epsilon)}{\epsilon}(n + \log \frac{1}{\delta})$ ex.

Reduce to learning mon. conj.

There are $N = O(\epsilon^{-3})$ many poss. clauses of length 3 over $x_1, \ldots, x_n$ (for each $y_{17,25,61}$)

Introduce new vars. $y_1, \ldots, y_n$, for each $y_{17,25,61}$

Given an example $(x_1, \ldots, x_n) \in \{0,1\}^n$, in $O(n^3)$ time can create $y$-version $y \in \{0,1\}^n$

Target $c(x)$ is $c'(y)$, $c'$ some mon. conj.:

- orig dist $\mathcal{D}$ over $\{0,1\}^n$ $\Rightarrow$ $\mathcal{D}$ over $\{0,1\}^n$.
- Use conj. learning alg over $\{0,1\}^n$: we're simulating $EX(c', \mathcal{D})$ given $EX(c, \mathcal{D})$. 

$\Rightarrow \frac{1}{\epsilon}(N + \log \frac{1}{\delta})$ ex
get h(y), a mon conj over y₁,...,yₙ; rewrite each yᵢ as an x-clause, and get a 3CNF h(y). Efficient.

This was improper: 3-term target, 3-CNF h(y).

Proper learning 3-term ONF is hard:

Theorem: Suppose there's an efficient proper PAC learner for C = 3-term ONFs. Then there is an efficient randomized algo. to solve GRAPH 3-COLORABILITY, hence NP ⊆ RP.

3COL:
given a simple graph G, can we color its nodes using only 3 colors s.t. no edge is monochromatic? NP-complete; we don't believe there's a poly-time algo.
Pf: Key to pf: eff. mapping $f$ is computable.

$n$-node graph $G$\[\rightarrow f\] $\rightarrow \text{set } S = S^+ \cup S^-$

$S^+ = \text{coll of } n$-bit $\leftrightarrow \text{pos. ex.}$ strings, likewise $S^-$. $\leftrightarrow \text{neg. ex.}$

(note $|S| \leq \text{poly}(n)$).

Key prop. of $f$: $G$ is 3-colorable iff $S$ is consistent w/some 3-term ONF.

For now, assume we have

- such an $f$ (will do soon)
  - eff.
  - proper PAC learner $A$, for 3-term ONFs.

Here's an eff. rand. alg. for 3COL:

1. Input: $n$-node $G$  
2. Apply $f$ to $G$, get $S = S^+ \cup S^-$
3. (idea: use $S$ to run PAC learner)
   - Define $\mathcal{D} = \text{uniform dist. over } S$ (on except pt. in $S$)
   - Set $\epsilon = \frac{1}{2|S|}$, set $\delta = 0.01\quad \text{note } \epsilon = \frac{1}{\text{poly}(n)}$
   - Run $A$ on $\text{EX}(S, \mathcal{D})$ to acc $\epsilon$, conf $\delta$

$A$ eff $\Rightarrow$ this is $\text{poly(n)}$ time $(\approx \text{human!})$

Let $h = \text{hyp. it outputs. Note } h$ is a 3-term ONF.

Check whether $h$ is indeed cons. w/$S$ $\Rightarrow h$ is eff. eval. so
If \( Y \), output "G is 3-colorable": this step poly time

Works!

- \( S \) s G is 3-colorable. By prop. of \( f \),
  \( \exists \) 3-term \( \text{ONF} \) \( c \) cons. w/ \( S \).
  So our simul. \( \text{EX}(S, \theta) \) is legit.
  So A's hyp \( h \) is, w/ 99% prob, e-acc.
  \( \theta \) puts prob. at \( \frac{1}{151} \) on every pt, \( + e = \frac{1}{2151} \): \( h \)
got every pt in \( S \) right. So w. p. \( \geq 0.99 \) \( h \) says
  "G is 3-col."

- \( S \) s G not 3-col. By prop. of \( f \),
  \( \exists \) no 3-term \( \text{ONF} \) \( c \) cons. w/ \( S \).
  Since \( h \) is a 3-term \( \text{ONF} \), can't be cons. w/ \( S \).
  So for sure \( h \) says "not 3-col."

Remains to see \( f \) & verify its key prop.
Here's \( f \): effi. mapping \( f \).
computable

\[
\text{set } S = S^+ \cup S^-
\]
\( S^+ = \text{coll of } n \)-bit
\( \leq \text{pos. ex. } \) strings,
likewise \( S^- \).
\( \leq \text{neg ex. } \)
Input: $G = (V, E)$, $V = \{1, \ldots, n\}$.

$S^+$ has $n$ ex: if $i$ has 0 in pos $j$, then there is an edge $(i, j)$.

$S^-$ has 1E/ex: edge $(i, j) \rightarrow \text{ex w/ 0 in i + j}$.

**Example:**

$n = 5$

$G = \begin{array}{cccc}
1 & 2 & 3 & 4 \\
2 & 3 & 4 & 5 \\
3 & 4 & 5 & 1 \\
4 & 5 & 1 & 2 \\
5 & 1 & 2 & 3 \\
\end{array}$

$S^+ = \{01111,
10111,
11011,
11101,
11110\}$

$S^- = \{00111,
01110,
10011,
10101,
11010,
11100\}$

**To show:**

Key prop. of $G$: $G$ is 3-colorable iff $S$ is consistent w/ some 3-term DNF.

**Claim:** 5ps $G$ is 3-colorable. Then there is a 3-term DNF cons. w/ $S^+, S^-$. 

**PF:** Consider legal 3-coloring $R, B, Y$ of graph $G$. 
Let \( T_R = \text{conj. cont. } x_i \text{ for all non-red } v_i \text{ in } G \)
\( T_B = \text{" " " " " blue " " " "} \)
\( T_Y = \text{" " " " " yellow " "} \)

Equiv., \( T_R \) missing the red \( i \)'s

\[
\begin{align*}
E x: & \quad T_R = x_1 x_3 x_4 x_5 \\
& \quad T_B = x_1 x_2 x_3 x_4 \\
& \quad T_Y = x_2 x_5.
\end{align*}
\]

Consider ex. in \( S^+ \): \( \underline{01111} \) (yellow vert) sat. by \( T_Y \):
vert. \( i \)'s color's term will sat. \( i \)th pos ex.
b/c it's missing the vert. set to 0.
the DNF labels \( S^+ \) pos \( i \)

Consider ex in \( S^- \): \( \underline{00111} \)
For \((i,j)\) ex. in \( S^- \) to sat. \( T_R, T_R \) would need
to be missing both \( x_i \) \& \( x_j \), but that'd mean
\( i,j \) both \( R \), but impossible b/c coloring is valid.

Claim 2: If there's a 3-term DNF cons. \( S^+, S^- \), then \( G \) is 3-colorable.

\[\text{Pf:} \quad \text{Let } T_R \cup T_B \cup T_Y \text{ be 3-term DNF cons. } S^+ \cup S^- .\]

Construct col. by col. each \( v \) in \( G \)
\[
i \in [n] \quad R \text{ if } i \text{th } S^+ \text{ ex. sat. } T_.
\]
(break ties arb.)

Remains to show: coloring valid, i.e. given edge (i,j), can’t make both endpoints R. Consider $i=1, j=2$:

Suppose both sat. $T_R$. This means $T_R$ doesn’t have $x_i, \bar{x}_i, x_j, \bar{x}_j$; but then if $T_R$ makes $+$, it’d make pos. too $\Rightarrow$ Contrad.