Last time: • Applic. of Occam's Razor $\mathcal{H}_n$ thm: learning sparse disj., using few examples, with hyp. that's a (pretty) sparse disj, via "greedy set cover" heuristic

• (start) proper vs improper learning:

– "can efficiently improperly learn 3-term DNF;

Today: – finish

– " can't eff. properly learn 3-term DNF (unless $NP \subseteq RP$).

– start VC dimension unit (KV Chap. 3)

Reminder: Midterm, in-person next Mon 10/23; closed-book, closed-note. (can cover stuff thru today)

Questions?

Let's finish: $\boxed{\text{can efficiently improperly learn 3-term DNF.}}$ over $\{0,1\}^n$

using 3-CNF hypotheses.

Recall Fact 1: N-var mon. conj. eff. learnable, w/ $\approx \frac{1}{\varepsilon}(N + \log\frac{1}{\delta})$ examples & runtime

Fact 2: any 3-term DNF is $\equiv$ to some 3-CNF.

Combine facts to get $\mathcal{D}$: DNF $T_1 \vee T_2 \vee T_3$, over x-lits. $\equiv$ to a 3CNF, also over x.

Reduce to learning N-var conj, $N = \Theta(n^3)$.

Target c (a 3-term DNF) is $\overset{F2}{\equiv}$ to some

3-CNF over $\{0,1\}^n$: $\wedge$ of $\underbrace{(\ell_1 \vee \ell_2 \vee \ell_3)}_{\text{lit over } x_1,\ldots,x_n}$

$\leq 8n^3$ poss length-3 clauses

Define $y_1, \ldots, y_{N=8n^3}$ new vars, one for each clause: e.g.

$\underline{y_{174}} = \underline{(x_{13} \vee \bar{x}_{23} \vee x_4)}$ 3-term DNF $\equiv$ 3-CNF

Explicitly convert each $(x, \underline{c(x)}) \sim EX(c, \mathcal{D})$ into

$(y, \underline{c'(y)})$ c' is a mon conj over the y_j's!

expanded $8n^3$-bit ver. of x

dist. acc. to some $\underline{\mathcal{D}' \text{ over } \{0,1\}^{8n^3}}$ no matter the dist.

(no prob!)

↳ PAC alg. succeeds

Run conj learning alg. over the y_j's (over length-3 clauses on orig x_i-lits.)

Get $h(y)$ our hyp, w.p. $1-\delta$ is ε-acc. w.r.t ϑ' → a 3-CNF over orig x_i-literals.

Acc of this under ϑ ≡

|| $h(y)$ under ϑ'.

So... non-properly, 3-term DNF eff. learnable!

||

✓ (using 3-CNF hyp's)

!! Thm: Sps exists an eff. proper PAC learning alg for 3-term DNF.

→ time poly$(n, \frac{1}{\varepsilon}, \frac{1}{\delta})$

Then there exists an eff. rand. alg. to solve GRAPH 3-COLORABILITY.

→ (poly(n)-time on n-node graphs)

w.p. 99%, alg. correctly announces whether or not G is properly 3-colorable.

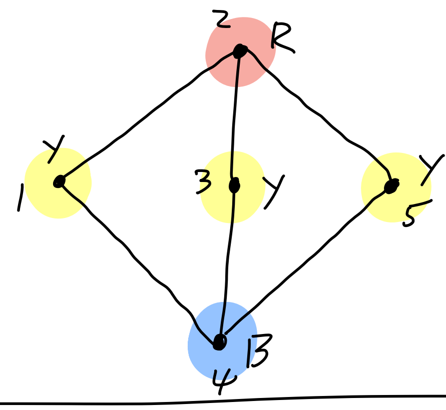
Prob. doesn't exist...

↳ G3COL is NP-complete! would mean $NP \subseteq RP$

G3COL: input: n-node graph G

output: a legit coloring of G's nodes using 3 colors R, Y, B

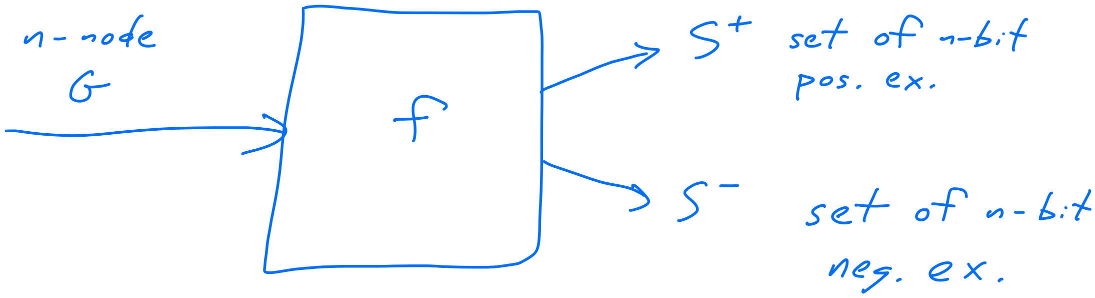
legit: no R-R edge,



no B-B edge,
no Y-Y edge.

NP-complete: almost certainly no poly(n) time (rand.) alg.

Key pf ingred: mapping f. $S = S^+ \cup S^-$



---

f is ① eff. computable (poly(n) time; so S is of poly(n) size)

② $\forall G$,

G is 3-colorable $\iff$ there exists some 3-term DNF lab. all ex in S^+ as $+$,
" " " S^- as $-$.

Assume for now we $\boxed{\text{have } f.}$

By assump. of thm, also have alg A an eff. proper PAC learner for 3-term DNFs.

Here's our alg. to decide G3COL:

- input graph to alg: G

The alg:

1) Apply f to G to get $S = S^+ \cup S^-$

2) Let $\mathcal{D}'$ be dist. over lab. ex. which is uniform over S.

   Let $\varepsilon = \frac{1}{2|S|}$. Let $\delta = 0.01$

   Run A with examples from $\mathcal{D}'$

   (note: we toss coins to sim. draws from $\mathcal{D}'$)

   A outputs a 3-term DNF h (b/c it's proper).

3) Check whether h agrees with every lab. ex. in S.

   If Y, output "G is 3-col."

   If N, " "G is not 3-col."

Eff? 1) poly(n) time (f)

Y 2) also poly(n) time: $\varepsilon = \frac{1}{\text{poly}(n)}$ b/c

$|S| \le \mathrm{poly}(n)$, $\delta =$ const, A eff.

3) poly(n) time.

Correct?

$\rightarrow$ Sps G is 3-colorable.

$\Longrightarrow$ S is cons. w/ some 3-term DNF (f).

so our sim of $EX(c, \mathcal{D})$ for A was valid :)

so w.p. $\ge 99\%$, h output by A is a 3-term DNF w/ error $\le \varepsilon = \frac{1}{2|S|}$.

$\rightarrow$ this h is perfect on S. (even 1 error would mean error rate $\ge \frac{1}{|S|}$)

so in 3) we say "G is 3-col."

$\rightarrow$ Sps G is not 3-colorable.

$\Longrightarrow$ S is not consist. w/ any 3-term DNF (f guarantee)

(b/c A is proper!)

So the 3-term DNF hyp h is surely not cons. w/ S, so in step 3 we surely say "G not 3-col."

Remains to give f + argue it has claimed properties.

Here's f:

On input $G = (V, E)$ where

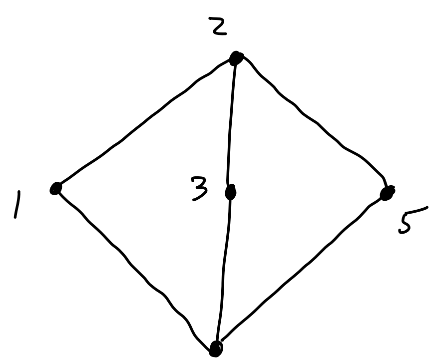
$V = \{1, \dots, \underline{n}\}$, $E =$ bunch of $\{i,j\}$ pairs, (m many)

output $f(G)$ is (S^+, S^-) where

- $S^+ = n$ examples, (n-bit; in $\{0,1\}^n$) (vertices) i^{th} one is $\overbrace{11\dots1 \underset{\uparrow \text{pos. } i}{0} 1\dots1}^{n \text{ bits}}$
- $S^- = m$ examples, (n-bit; each ex is in $\{0,1\}^n$) corr. to edges;

edge $\{i,j\} \Longleftrightarrow \underbrace{1\dots1\overset{i}{\downarrow}{0}1\dots1\overset{j}{\downarrow}{0}1\dots1}_{n \text{ bits}}$

Ex:

$G =$ 

$|| f$

$S^+ = $ 01111 +
10111 +
11011 +
11101 +
11110 +

$n =$ 5 pos. examples

$S^- = $ 00111 −
01101 −
10011 −
10110 −
11001 −
11100 −

$m = 6$ neg ex.

---

f easy to compute. ✓

To show:

L1 If G is 3-colorable, then ∃ 3-term DNF cons. w/ S^+ + S^-

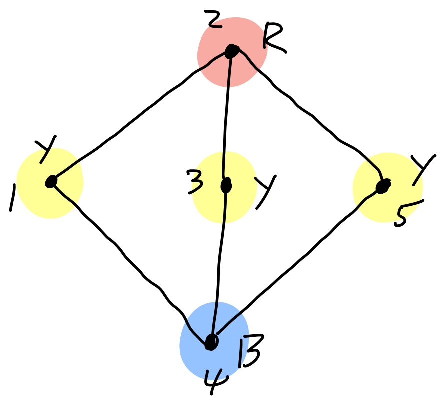
+ L2 If ∃ 3-term DNF cons. w/ S^+ + S^-, then G is 3-colorable.

Pf of L1: Let R, B, Y be some legal 3-coloring of G.

Let T_R = conj containing x_i for every non-red vtx i

T_Y = " " x_i " " non-yellow "

$T_B =$ " " " " " non-blue vtx i.



$T_R = x_1 x_3 x_4 x_5$

$T_Y = x_2 x_4$

$T_B = x_1 x_2 x_3 x_5$.

[$\equiv$ view: T_R is missing only red x_i's, etc.]

• Each ex in S^+ is labeled $+$ by $T_R \vee T_Y \vee T_B$

consider i^{th} pos ex. $11\ldots 1\overset{i}{0}1\ldots 1$

Say it's R in col. Then T_R is missing x_i

so ex sat. T_R

• Each ex. in S^- is lab. $-$ by $T_R \vee T_Y \vee T_B$:

consider $\{i,j\}$ ex $1\ldots 1\overset{i}{0}1\ldots 1\overset{j}{0}1\ldots 1$

Does this sat. T_R? No - would mean T_R missing x_i & x_j, ie i,j both R, but can't be since we had a legit col.

Same arg gives that T_B, T_Y also not sat by this ex.

so this is lab – by $T_R \vee T_Y \vee T_B$.

L1

Next time: L2

start unit on sample cxity of PAC learning.