Class 12

Algorithmic LLL

Announcements

- HW5 due Friday
- HW6 out now!
- HW7 isn't due until after fall break! (Friday 12/2)
- No class on Tuesday: Democracy Day!

If you are eligible to vote, then \bigvee \bigcirc

Recap: Algorithmic LLL (for k-SAT)

- Given φ :
	- Choose a random assignment σ for each of the variables that appear in φ
	- While there is some clause C of φ that is not satisfied:
		- Update σ by randomly re-selecting the variables that appear in C.
	- Return σ

• **Theorem:**

- Suppose that each clause C in φ shares variables with at most $d + 1 = 2^{k-c}$ clauses (including C itself), for some constant c .
- Then φ is satisfiable and the algorithm above finds a satisfying assignment quickly.

 $\mathcal A$ is a collection of bad events determined by variables in V. Vbl(A) is the set of variables involved with $A \in \mathcal{A}$

Algorithmic LLL more generally

- Given V and \mathcal{A} :
	- Choose a random assignment σ_{ν} for each of the random variables $\nu \in V$
	- While there is some $A \in \mathcal{A}$ so that $A(\sigma) = 1$:
		- Choose (arbitrarily) an event A with $A(\sigma) = 1$.
		- Update σ by re-selecting $\{\sigma_v : v \in \text{Vbl}(A)\}\)$ randomly.
- Suppose that for all $A \in \mathcal{A}$:
	- $|\Gamma(A)| \leq d+1$
	- $Pr[A] \leq \frac{1}{e^{d}}$ $e(d+1)$
- Then this algorithm will find an assignment to the variables in V so that no event of $\mathcal A$ occurs with $O\left(\frac{\overline{a}}{d}\right)$ $d + 1$ re-randomizations.

Proof of Algorithmic LLL

- Add some print statements to our algorithm.
- If the algorithm runs for too long, it will be too good of a compression algorithm.

Questions? Algorithmic LLL, Quiz?

Q1: Applying alg. LLL

- $S_1, S_2, ..., S_M \subset X$ are sets of size $k < |X| = N$
- Each S_i intersects at most 10 other sets S_i
- Color points of X red or blue iid with prob ½.
- A_i is the event that S_i is monochromatic.
- \bullet | V | =
- \bullet d =
- For what k does alg. LLL apply?
- What is expected number of re-randomizations?

Q2. Changing the proof

- What if we print "trying to fix clause i_{ℓ} " instead of "trying to fix the l' th child"?
- Q2.1 How many bits get outputted in print statements?

• Q2.2. What would that prove?

Today: More practice with the Algorithmic LLL

- We saw the proof for k-SAT
- Today you'll prove it for set coloring!

The problem

- *n* points, $\{1, 2, ..., n\}$
- *m* sets, S_1 , S_2 , ..., $S_m \subseteq \{1, 2, ..., n\}$
- Each set has size k .
- Each set overlaps with no more than d other sets.
- Goal: color the n points red or blue so that none of the sets is monochromatic.

The problem

- *n* points, $\{1, 2, ..., n\}$
- *m* sets, S_1 , S_2 , ..., $S_m \subseteq \{1, 2, ..., n\}$
- Each set has size k .
- Each set overlaps with no more than d other sets.
- Goal: color the n points red or blue so that none of the sets is monochromatic.

Algorithmic LLL gives an algorithm to do this

• While not done:

- Pick a monochromatic set, S_i .
- Re-color all of the numbers in S_i , uniformly at random.
- But we didn't prove that this works.
	- We only proved it for k-SAT
- Goal of today:
	- Mimic the k-SAT argument to give an algorithm that provably works for nomonochromatic-coloring.

Quick recap of the proof idea for k-SAT

- We wrote the algorithm in a recursive way and added some print statements.
- From the print statements, you could figure out the random bits that went into the algorithm.
- If the algorithm runs for too long (too many re-randomizations), then we can compress the random bits!
- But that's impossible.

Group work!

- Give a proof!
	- What is the same between the k-SAT proof and this proof?
	- What needs to change?

Outline:

- We wrote the algorithm in a recursive way and added some print statements.
- From the print statements, you could figure out the random bits that went into the algorithm.
- If the algorithm runs for too long (too many re-randomizations), then we can compress the random bits!
- But that's impossible.

For inspiration, here was the k-SAT algorithm Your job: adapt to set-coloring!

- **FindSat**($\varphi = C_1 \wedge C_2 \wedge \cdots \wedge C_m$):
	- Choose a random assignment σ for each of the variables that appear in φ

Fixing

set *i*!

- For each clause C_i in φ that is not satisfied:
	- $\sigma \leftarrow \text{Fix}(\varphi, i, \sigma)$
- Return σ

• **Fix** (φ, i, σ) :

- Update σ by re-randomizing every variable that appears in the clause C_i
- Let C_{i_1} , C_{i_2} , ... $C_{i_{d+1}}$ be the clauses that share variables with C_i

What needs to change?

- **FindSat**($\varphi = C_1 \wedge C_2 \wedge \cdots \wedge C_m$):
	- Choose a random assignment σ for each of the variables that appear in φ

Fixing

set *i*!

- For each clause C_i in φ that is not satisfied:
	- $\sigma \leftarrow \text{Fix}(\varphi, i, \sigma)$
- Return σ

• **Fix** (φ, i, σ) :

- Update σ by re-randomizing every variable that appears in the clause C_i
- Let C_{i_1} , C_{i_2} , ... $C_{i_{d+1}}$ be the clauses that share variables with C_i

Our algorithm?

- **FindSat** $(S_1, S_2, ..., S_m)$:
	- Choose a random coloring σ for each of numbers
	- For each S_i that is monochromatic:
		- $\sigma \leftarrow$ **Fix**(*i*, σ)
	- Return σ

• **Fix** (i, σ) :

• Update σ by re-randomizing every number in S_i

Fixing set

 $i!$

To do the proof

- We need to count the number of random bits that go in in the first T re-randomizations.
- We need to count the number of bits of print statements that come out in the first T re-randomizations.
- We need to argue that we can recover the random bits that go in from the print statements that come out.

To do the proof

- We need to count the number of random bits that go in in the first T re-randomizations.
- We need to count the number of bits of print statements that come out in the first T re-randomizations.
- We need to argue that we can recover the random bits that go in from the print statements that come out.

- The print statements allow us to reconstruct the recursion tree.
- Then…

Say we know the coloring AFTER we re-randomized to fix the j'th child. (We know the final assignment since it was printed out, and we're working backwards.)

- The print statements allow us to reconstruct the recursion tree.
- Then…

Say we know the coloring AFTER we re-randomized to fix the j'th child. (We know the final assignment since it was printed out, and we're working backwards.)

- The print statements allow us to reconstruct the recursion tree.
- Then…

 $\begin{array}{|c|c|c|c|c|}\hline \end{array}$ $\begin{array}{|c|c|c|c|}\hline \end{array}$ $\begin{array}{|c|c|c|}\hline \end{array}$ $\begin{array}{|c$ 1 2 3 4 5 6 … n Trying to fix the j' th child

Say we know the coloring AFTER we re-randomized to fix the j'th child. (We know the final assignment since it was printed out, and we're working backwards.)

- The print statements allow us to reconstruct the recursion tree.
- Then…

- The print statements allow us to reconstruct the recursion tree.
- Then…

Our algorithm

- **FindSat** $(S_1, S_2, ..., S_m)$:
	- Choose a random coloring σ for each of numbers
	- For each S_i that is monochromatic:
	- $\sigma \leftarrow$ **Fix**(*i*, σ) • Return σ Fixing set $i!$

• **Fix** (i, σ) :

- Update σ by re-randomizing every number in S_i
- Let S_{i_1} , S_{i_2} , ... $S_{i_{d+1}}$ be the sets that intersect S_i

- The print statements allow us to reconstruct the recursion tree.
- Then…

Since I know the recursion tree, I know that at this point "the j' th child" means S_4

- The print statements allow us to reconstruct the recursion tree.
- Then…

Since I know the recursion tree, I know that at this point "the j' th child" means S_4

To do the proof

We need to count the number of random bits that go in in the first T re-randomizations.

We need to count the number of bits of print statements that come out in the first T re-randomizations.

We need to argue that we can recover the random bits that go in from the print statements that come out.

Our algorithm

- **FindSat** $(S_1, S_2, ..., S_m)$:
	- Choose a random coloring σ for each of numbers
	- For each S_i that is monochromatic:
		- $\sigma \leftarrow$ **Fix**(*i*, σ)
	- Return σ

Random bits in:

 $n + k$. T original o l'éluits
per re-randomization

• **Fix** (i, σ) :

- Update σ by re-randomizing every number in S_i
- Let S_{i_1} , S_{i_2} , ... $S_{i_{d+1}}$ be the sets that intersect S_i
- For $j = 1, ..., d + 1$:

Fixing

set i!

Our algorithm

- **FindSat** $(S_1, S_2, ..., S_m)$:
	- Choose a random coloring σ for each of numbers
	- For each S_i that is monochromatic:
		- $\sigma \leftarrow$ **Fix**(*i*, σ)
	- Return σ
- **Fix** (i, σ) :
	- Update σ by re-randomizing every number in S_i
	- Let S_{i_1} , S_{i_2} , ... $S_{i_{d+1}}$ be the sets that intersect S_i
	- For $j = 1, ..., d + 1$:

Fixing

set i!

Bits out:

 $\leq m\left[\log(m)+C\right]$

 $+ n + C$

"I give up. 6."

" Fixing clause i"

 $+ 7$ [$log(d+1) + 1 + C$]

"tying to fix jth child

call Fix Ttimes

Also

"all

done"

$$
\text{Want: } \mathsf{N} + \mathsf{k} \top \gg \mathsf{m} \Big(\log \mathsf{m} + C \Big) + \top \Big(\log(d+1) + 1 + C \Big) + \mathsf{n} + C
$$

$$
\begin{array}{ll}\n\text{Want:} & \mathsf{N} + \mathsf{k} \top \Rightarrow \mathsf{m} \big(\log \mathsf{m} + C \big) + \top \big(\log(\mathsf{d} + \mathsf{l}) + \mathsf{l} + C \big) + \mathsf{n} + C \\
\text{Aka:} & \mathsf{m} \big(\log \mathsf{m} + C \big) \ll \top \big(\mathsf{k} - \log(\mathsf{d} + \mathsf{l}) + \mathsf{l} + C \big) \\
\end{array}
$$

$$
\text{Want: } \mathsf{R} + \mathsf{R} \top \gg \mathsf{m} \Big(\log \mathsf{m} + C \Big) + \top \Big(\log(\mathsf{d} + \mathsf{l}) + \mathsf{l} + C \Big) + \mathsf{n} + C
$$
\n
$$
\text{Aka: } \mathsf{m} \Big(\log \mathsf{m} + C \Big) \ll \top \Big(\mathsf{R} - \log(\mathsf{d} + \mathsf{l}) + \mathsf{l} + C \Big)
$$

Provided that
$$
k \ge log(d+1) + 100000
$$
, this happens for $T = poly(m)$.

What happens if there are $t > 2$ colors?

Our algorithm

- **FindSat** $(S_1, S_2, ..., S_m)$:
	- Choose a random coloring σ for each of numbers
	- For each S_i that is monochromatic:
	- $\sigma \leftarrow$ **Fix**(*i*, σ) • Return σ Fixing set *i*!
- **Fix** (i, σ) :
	- Update σ by re-randomizing every number in S_i
	- Let S_{i_1} , S_{i_2} , ... $S_{i_{d+1}}$ be the sets that intersect S_i

- The print statements allow us to reconstruct the recursion tree.
- Then…

Since I know the recursion tree, I know that at this point "the j' th child" means S_4

To do the proof

We need to count the number of random bits that go in in the first T re-randomizations.

We need to count the number of bits of print statements that come out in the first T re-randomizations.

We need to argue that we can recover the random bits that go in from the print statements that come out.

Our algorithm

- **FindSat** $(S_1, S_2, ..., S_m)$:
	- Choose a random coloring σ for each of numbers
	- For each S_i that is monochromatic:
		- $\sigma \leftarrow \text{Fix}(i, \sigma)$
	- Return σ

Fixing set i!

Random bits in:

• **Fix** (i, σ) :

- Update σ by re-randomizing every number in S_i
- Let S_{i_1} , S_{i_2} , ... $S_{i_{d+1}}$ be the sets that intersect S_i
- For $j = 1, ..., d + 1$:

Our algorithm

- **FindSat** $(S_1, S_2, ..., S_m)$:
	- Choose a random coloring σ for each of numbers
	- For each S_i that is monochromatic:
		- $\sigma \leftarrow$ **Fix**(*i*, σ)
	- Return σ
- **Fix** (i, σ) :
	- Update σ by re-randomizing every number in S_i
	- Let S_{i_1} , S_{i_2} , ... $S_{i_{d+1}}$ be the sets that intersect S_i
	- For $j = 1, ..., d + 1$:

Fixing

set i!

(or blue, **or purple, or….**

as appropriate)

 $\leq m\left\{\log(m)+C\right\}$ Bits out: "Fixing clause i" $+ \top \left[\log(d+1) + \chi + C\right]$ "tying to fix jth child call Fix Ttimes Also "all done" $+ n + C$ "I give up. 6."

got σ

Want:
$$
n+k
$$
 $\overline{p}e^{kt}$ $m(pgm + C) + \overline{p}(\log(dt)) + N + C) + n + C$

$$
\begin{array}{ll}\n\text{Want:} & \mathsf{N} + \mathsf{k} \, \mathbb{R}^{\mathsf{ce}(t)} & \text{m} \big(\, \log \mathsf{m} + \mathsf{C} \, \big) + \mathsf{T} \big(\, \log(\mathsf{d} + \mathsf{l}) \, \mathsf{+} \mathsf{N} + \mathsf{C} \, \big) + \mathsf{n} + \mathsf{C} \\
\text{Ata:} & \mathsf{m} \big(\log \mathsf{m} + \mathsf{C} \, \big) \ll \mathsf{T} \big(\, \mathsf{k}^{\mathsf{ce}(t)} \log(\mathsf{d} + \mathsf{l}) \, \mathsf{+} \mathsf{N} + \mathsf{C} \, \big) \\
\end{array}
$$

Want:
$$
1 + k \overline{\log b}
$$
 and $\log m + C$ + $\overline{\log (d+1) + X + C}$ + $n + C$

\nAta: $m(\log m + C) \ll T(\log^{o(e)}\log(d+1) + X + C)$

\nProvided that $k = \frac{\log(d+1) + \log(t)}{\log(t)} + \log q = \frac{\log(d+1)}{\log(t)} + 10000$

\nThis happens for $T = \text{poly}(m)$

Conclusion

As long as $k \geq \frac{\log(d+1)}{\log(d)}$ $log(t)$ + 10000, we can find a good coloring with $poly(m)$ re-randomizations!

Corollary 3. Let V be a finite set of independent random variables. Let A be a finite set of events determined by the random variables in V. If for all $A \in \mathcal{A}, |\Gamma(A)| \leq d+1$, and $\Pr[A] \leq \frac{1}{e(d+1)}$, then Algorithm 2 will find an assignment to the variables V such that no event of A oscurs. Additionally, the expected number of "re-randomizations" performed by the algorithm is bounded by $O(|A|/(d+d))$ 1).

Corollary 3. Let V be a finite set of independent random variables. Let A be a finite set of events determined by the random variables in V. If for all $A \in \mathcal{A}, |\Gamma(A)| \leq d+1$, and $\Pr[A] \leq \frac{1}{e(d+1)}$, then Algorithm 2 will find an assignment to the variables V such that no event of A oscurs. Additionally, the expected number of "re-randomizations" performed by the algorithm is bounded by $O(|A|/(d+d))$ 1).

$$
A_i = \{S_i \text{ is monodmonadic}\}
$$

 $\mathbb{P}\{A_i\} = \frac{t}{t^k} \text{ for t colors.}$

Corollary 3. Let V be a finite set of independent random variables. Let A be a finite set of events determined by the random variables in V. If for all $A \in \mathcal{A}$, $|\Gamma(A)| \leq d+1$, and $\Pr[A] \leq \frac{1}{e(d+1)}$, then Algorithm 2 will find an assignment to the variables V such that no event of A decurs. Additionally, the expected number of "re-randomizations" performed by the algorithm is bounded by $O(|A|/(d+d))$ 1).

$$
A_i = \{S_i \text{ is monodmonadic}\}
$$

 $\mathbb{P}\{A_i\} = \frac{t}{t^k} \text{ for t colors.}$

Need: $\mathbb{P}\{A_i\} \leq \frac{1}{e(d+1)}$

Corollary 3. Let V be a finite set of independent random variables. Let A be a finite set of events determined by the random variables in V. If for all $A \in \mathcal{A}, |\Gamma(A)| \leq d+1$, and $\Pr[A] \leq \frac{1}{e(d+1)}$, then Algorithm 2 will find an assignment to the variables V such that no event of A oscurs. Additionally, the expected number of "re-randomizations" performed by the algorithm is bounded by $O(|A|/(d+d))$ 1).

$$
A_i = \{S_i \text{ is monodmonadic}\}
$$

 $\mathbb{P}\{A_i\} = \frac{t}{t^k} \text{ for t colors.}$

Need: $P\{A_i\} \leq \frac{1}{e(d+1)}$ $t^{-(k-1)} \leq \frac{1}{e(dt)}$ $(h-1) log(t) \ge 1 + log(d+1)$

Corollary 3. Let V be a finite set of independent random variables. Let A be a finite set of events determined by the random variables in V. If for all $A \in \mathcal{A}, |\Gamma(A)| \leq d+1$, and $\Pr[A] \leq \frac{1}{e(d+1)}$, then Algorithm 2 will find an assignment to the variables V such that no event of A oscurs. Additionally, the expected number of "re-randomizations" performed by the algorithm is bounded by $O(|A|/(d+d))$ 1).

$$
A_i = \{S_i \text{ is monodmonadic}\}
$$

 $\mathbb{P}\{A_i\} = \frac{t}{t^k} \text{ for t colors.}$

$$
Need: \mathbb{P} \{ A_i \} \le \frac{1}{e(d+1)}
$$
\n
$$
t^{-(k-1)} \le \frac{1}{e(d+1)}
$$
\n
$$
(\frac{1}{k-1}) \log(k) \ge 1 + \log(d+1)
$$
\n
$$
\frac{1}{k} \ge \frac{\log(d+1)}{\log(k)} + \text{[constant]}
$$
\nSame thing

Conclusions

- As long as $k \geq \frac{\log(d+1)}{\log(d)}$ $log(t)$ + 10000, we can find a good coloring with $poly(m)$ re-randomizations!
- You now have some idea of how you might adapt this proof to deal with other examples (aka, ones with $Pr[A_i] \leq p$ for a general $p)....$

