### 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!



### Recap: Algorithmic LLL (for k-SAT)

#### • Given $\varphi$ :

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

#### • Theorem:

- Suppose that each clause C in  $\varphi$  shares variables with at most  $d + 1 = 2^{k-c}$  clauses (including C itself), for some constant c.
- Then  $\varphi$  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  $\sigma_v$  for each of the random variables  $v \in V$
  - While there is some  $A \in \mathcal{A}$  so that  $A(\sigma) = 1$ :
    - Choose (arbitrarily) an event A with  $A(\sigma) = 1$ .
    - Update  $\sigma$  by re-selecting  $\{\sigma_v : v \in Vbl(A)\}$  randomly.
- Suppose that for all  $A \in \mathcal{A}$ :
  - $|\Gamma(A)| \le d+1$
  - $\Pr[A] \le \frac{1}{e(d+1)}$
- Then this algorithm will find an assignment to the variables in V so that no event of  $\mathcal{A}$  occurs with  $O\left(\frac{|\mathcal{A}|}{d+1}\right)$  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

- $S_1, S_2, \dots, S_M \subset X$  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 1/2.
- $A_i$  is the event that  $S_i$  is monochromatic.
- |V|=
- 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  $i_{\ell}$ " instead of "trying to fix the  $\ell$ '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

- $n \text{ points}, \{1, 2, ..., n\}$
- *m* sets,  $S_1, S_2, \dots, S_m \subseteq \{1, 2, \dots, n\}$
- 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

- $n \text{ points}, \{1, 2, ..., n\}$
- *m* sets,  $S_1, S_2, \dots, S_m \subseteq \{1, 2, \dots, n\}$
- 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<sub>i</sub>.
- 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( $\varphi = C_1 \land C_2 \land \cdots \land C_m$ ):
  - Choose a random assignment  $\sigma$  for each of the variables that appear in  $\varphi$

Fixing

set *i*!

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

• **Fix**(*φ*, *i*, *σ*):

- Update  $\sigma$  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( $\varphi = C_1 \land C_2 \land \cdots \land C_m$ ):
  - Choose a random assignment  $\sigma$  for each of the variables that appear in  $\varphi$

Fixing

set *i*!

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

• **Fix**( $\varphi$ , *i*,  $\sigma$ ):

level.

- Update  $\sigma$  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  $C_i$
- For j = 1, ..., d + 1: Trying to fix All done • If  $C_{i_i}$  is violated: the *j*'th child with this After T re-randomizations,  $\checkmark$ •  $\sigma \leftarrow Fix(\varphi, i_i, \sigma)$ Return  $\sigma$ I give up. I've got  $\sigma$

### Our algorithm?

- FindSat( $S_1, S_2, ..., S_m$ ):
  - Choose a random coloring  $\sigma$  for each of numbers
  - For each *S<sub>i</sub>* that is monochromatic:
    - $\sigma \leftarrow \mathsf{Fix}(i, \sigma)$
  - Return  $\sigma$

• **Fix**(*i*, σ):

• Update  $\sigma$  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...

5 6 3 2 4 . . . 1 n Trying to fix the j'th child 4 5 6 3 2 n . . .

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...

Since I know 5 2 3 6 ... 4 1 n the recursion tree, I know that at this Trying to fix point "the *j*'th the j'th child child" means S<sub>4</sub>  $S_4$ 2 5 6 3 4 n 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...



#### Our algorithm

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

• **Fix**(*i*, *σ*):

- Update  $\sigma$  by re-randomizing every number in  $S_i$
- Let  $S_{i_1}, S_{i_2}, \dots S_{i_{d+1}}$  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<sub>4</sub>



• The print statements allow us to reconstruct the recursion tree.





Since I know the recursion tree, I know that at this point "the j'th child" means S<sub>4</sub>

#### 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  $\sigma$  for each of numbers
  - For each S<sub>i</sub> that is monochromatic:
    - $\sigma \leftarrow \mathbf{Fix}(i, \sigma)$
  - Return  $\sigma$

### $n + k \cdot T$ original or C k bits per re-randomization

Random bits in:

• Fix $(i, \sigma)$ :

- Update  $\sigma$  by re-randomizing every number in  $S_i$
- Let  $S_{i_1}, S_{i_2}, \dots S_{i_{d+1}}$  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  $\sigma$  for each of numbers
  - For each *S<sub>i</sub>* that is monochromatic:
  - $\sigma \leftarrow Fix(i, \sigma)$ • Return  $\sigma$
- **Fix**(*i*, σ):
  - Update  $\sigma$  by re-randomizing every number in  $S_i$
  - Let  $S_{i_1}, S_{i_2}, \dots S_{i_{d+1}}$  be the sets that intersect  $S_i$
  - For j = 1, ..., d + 1:



Fixing

set *i*!

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

+ n+C

"I give up. o."

"Fixing clause i

+ T[log(d+1)+1+C]

"trying blix jth child because it was red"

- call Fix Ttimes

Also

"all

done"

Bits out:





Want: 
$$n+kT \gg m(\log m+C) + T(\log(d+1)+1+C)+n+C$$

Want: 
$$n+kT \gg m(\log m+C) + T(\log(cl+1)+l+C)+n+C$$
  
Aka:  $m(\log m+C) \ll T(k-\log(d+1)+l+C)$ 

Want: 
$$n+kT \gg m(\log m+C) + T(\log(d+1)+1+C)+n+C$$
  
Aka:  $m(\log m+C) \ll T(k-\log(d+1)+1+C)$ 

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

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

#### Our algorithm

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

• **Fix**(*i*, σ):

- Update  $\sigma$  by re-randomizing every number in  $S_i$
- Let  $S_{i_1}, S_{i_2}, \dots S_{i_{d+1}}$  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<sub>4</sub>

#### 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  $\sigma$  for each of numbers
  - For each *S<sub>i</sub>* that is monochromatic:
    - $\sigma \leftarrow \mathbf{Fix}(i, \sigma)$
  - Return  $\sigma$

Fixing set *i*! Random bits in:



• **Fix**(*i*, σ):

- Update  $\sigma$  by re-randomizing every number in  $S_i$
- Let  $S_{i_1}, S_{i_2}, \dots S_{i_{d+1}}$  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  $\sigma$  for each of numbers

Fixing

set *i*!

- For each *S<sub>i</sub>* that is monochromatic:
- $\sigma \leftarrow Fix(i, \sigma)$ • Return  $\sigma$
- **Fix**(*i*, σ):
  - Update  $\sigma$  by re-randomizing every number in  $S_i$
  - Let  $S_{i_1}, S_{i_2}, \dots S_{i_{d+1}}$  be the sets that intersect  $S_i$
  - For j = 1, ..., d + 1:
- All done with this level. •  $\sigma \leftarrow \mathbf{Fix}(i_j, \sigma)$ 
  - Return  $\sigma$



...because it was all red! (or blue, or purple, or.... as appropriate)

Trying to fix

the *j*'th child

l give up. l've got  $\sigma$ 





Want: 
$$n + k \xrightarrow{\log(t)} m(\log(t) + C) + T(\log(d+1) + C) + n + C$$

aka, then we'd get a contradiction and conclude that there must be < T re-randomizations.

Want: 
$$n + k \operatorname{Tog}^{(t)} m(\log m + C) + T(\log(cl+1) + n + C) + n + C$$
  
Aka:  $m(\log m + C) \ll T(k^{\log(t)}\log(d+1) + n + C)$ 

laa/+

Want: 
$$n + k T^{og(c)} m(\log m + C) + T(\log(cl+1) + N + C) + n + C$$
  
Aka:  $m(\log m + C) \ll T(k^{\log(c)}\log(d+1) + N + C)$   
Provided that  $k = \frac{\log(d+1) + \log(t)}{\log(t)} + 9999 = \frac{\log(d+1)}{\log(t)} + 10000$   
This happens for  $T = \operatorname{poly}(m)$ 

#### Conclusion

As long as  $k \ge \frac{\log(d+1)}{\log(t)} + 10000$ , 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  $A \in A$ ,  $|\Gamma(A)| \le d+1$ , and  $\Pr[A] \le \frac{1}{e(d+1)}$ , then Algorithm 2 will find an assignment to the variables V such that no event of A occurs. Additionally, the expected number of "re-randomizations" performed by the algorithm is bounded by  $O(|\mathcal{A}|/(d+1))$ .

**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  $A \in A$ ,  $|\Gamma(A)| \le d+1$ , and  $\Pr[A] \le \frac{1}{e(d+1)}$ , then Algorithm 2 will find an assignment to the variables V such that no event of A occurs. Additionally, the expected number of "re-randomizations" performed by the algorithm is bounded by  $O(|\mathcal{A}|/(d+1))$ .

$$A_{i} = \left\{ \begin{array}{l} S_{i} \text{ is monochromabic} \right\} \\ P\left\{A_{i}\right\} = \frac{t}{t^{k}} \quad \text{for $t$ colors.} \end{array}$$

**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  $A \in A$ ,  $|\Gamma(A)| \le d+1$ , and  $\Pr[A] \le \frac{1}{e(d+1)}$ , then Algorithm 2 will find an assignment to the variables V such that no event of A occurs. Additionally, the expected number of "re-randomizations" performed by the algorithm is bounded by  $O(|\mathcal{A}|/(d+1))$ .

$$A_{i} = \left\{ \begin{array}{l} S_{i} \text{ is monodomonabic} \end{array} \right\}$$
$$P\left\{A_{i}\right\} = \frac{t}{t^{R}} \quad \text{for t colors.}$$

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  $A \in A$ ,  $|\Gamma(A)| \le d+1$ , and  $\Pr[A] \le \frac{1}{e(d+1)}$ , then Algorithm 2 will find an assignment to the variables V such that no event of A occurs. Additionally, the expected number of "re-randomizations" performed by the algorithm is bounded by  $O(|\mathcal{A}|/(d+1))$ .

$$A_{i} = \left\{ \begin{array}{l} S_{i} \text{ is monodronnabic} \right\} \\ P\left\{A_{i}\right\} = \frac{t}{t^{k}} \quad \text{for $t$ colors.} \end{array}$$

Need:  $\{P_{k}^{2}\} \leq \frac{1}{e(d+1)}$  $t^{-(k-1)} \leq \frac{1}{e(d+1)}$  $(f_{k-1}) \log(t) \geq 1 + \log(d+1)$ 

**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  $A \in A$ ,  $|\Gamma(A)| \le d+1$ , and  $\Pr[A] \le \frac{1}{e(d+1)}$ , then Algorithm 2 will find an assignment to the variables V such that no event of A occurs. Additionally, the expected number of "re-randomizations" performed by the algorithm is bounded by  $O(|\mathcal{A}|/(d+1))$ .

$$A_{i} = \left\{ \begin{array}{l} S_{i} \text{ is monochromatic} \right\} \\ P\left\{A_{i}\right\} = \frac{t}{t^{k}} \quad \text{for t colors.} \end{array}$$

Need:  $P\{A_i\} \leq \frac{1}{e(d+1)}$  $t^{-(k-1)} \leq \frac{1}{e(d+1)}$  $(f_{k-1}) \log(t) \geq 1 + \log(d+1)$  $f_k \geq \frac{\log(d+1)}{\log(t)} + [constant]$ Same thing!

#### Conclusions

- As long as  $k \ge \frac{\log(d+1)}{\log(t)} + 10000$ , 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  $Pr[A_i] \le p$  for a general p)....

