15-251: Great Theoretical Ideas In Computer Science

Recitation 15

- Our final is 8:30 - 11:30 am on Monday, May 8 in GHC 4401 (Rashid). Look on Piazza for information about review sessions, practice problems, and office hours.
- Look out for FCEs and feedback forms! Help us make 251 a better experience in the future.

All The Stuff From February You Forgot About

(a) Our ancient DFA interpreter Chris is getting too old to run, so he only accepts DFA jobs if he can get a circuit to do the work for him.

CHRIS $= \{ \langle D, n, s \rangle$: D is a DFA with $\Sigma = \{0, 1\}$, $n, s \in \mathbb{N}$, and
$\exists$ circuit of size at most s accepting the same inputs of length n as D $\}$

Prove that CHRIS is decidable.

(b) Anna is from Canada, the 51st US state. Feeling some Canadian pride, she comes up with the following language:

ANNA $= \{ \langle M \rangle$: M is a TM, and $\nexists$ DFA with at most 50 states which decides L(M) $\}$

Prove that ANNA is undecidable.

(c) Ji An is Number One, so he is disgusted by any decimal digits other than 1. He creates a set called JIAN containing all real numbers that have decimal representations consisting entirely of ones (1.11 and 111.111... are examples of numbers in JIAN). Show that JIAN is countable.

(d) The set of classes that Andrew has been a TA for at CMU is uncountable. He wants his set to be uncountable too!

$$\text{ANDREW} = \{a_1 a_2 a_3 \ldots \in \{\text{A,B}\}^\infty \mid \forall n \geq 1 \text{ the number of A's in the string } a_1 \ldots a_{3n} \text{ is twice the number of B's } \}$$

Prove that ANDREW is uncountable.

(e) Joohyun is perfect, so he wants the lengths of strings in his language to be perfect squares.

JOOHYUN $\subseteq \{1\}^* = \{1^n \mid \sqrt{n} \in \mathbb{N}\}$. Show that JOOHYUN is not regular.

(f) Calvin likes graphs, so he chooses his language:

CALVIN $= \{ \langle G, k \rangle \mid G$ contains a clique of size k but not an independent set of size k $\}$

Calvin thinks that CALVIN is not decidable in polytime, so he sketches the following proof:

- AFSOC CALVIN is polytime decidable, and show that IND-SET Cook reduces to CALVIN
- define IND-SET_DECIDER($\langle G, k \rangle$):
  - construct G' by making a copy of G
  - add to G' an additional component of $k+1$ completely connected nodes
  - return $\neg$CALVIN_DECIDER($\langle G', k+1 \rangle$)
- we know that IND-SET cannot be decided in polytime, so we have a contradiction

Is Calvin's reduction correct? What about the rest of his proof?

Ramsey Theory

Carolyn's cat and Apoorva's dog are playing a game with graphs. They start with an empty graph on n vertices (one with no edges). On the cat's turn, she adds a red edge to the graph. On the dog's turn, he adds a blue edge to the graph (a single pair of vertices cannot be joined by both a red and a blue edge). The cat wins if she creates a red clique of size k in the graph. The dog wins if he creates a blue clique of size k in the graph. The game is a draw if the graph is complete (all possible edges have been added) and it has neither a red clique nor a blue clique of size k. Prove that if n is large enough, this game cannot end in a draw (no matter how poorly the pets play). First show the following stronger claim:

For every $i, k \in \mathbb{N}$, $\exists\, n \in \mathbb{N}$ such that every graph on n vertices contains either an independent set of size i or a clique of size k.

(HINT: Prove this by induction. Consider choosing an aribitrary vertex v and separating the other vertices into two sets based on whether or not they are adjacent to v)

Chernoff Bounds

Josh and Woo are playing a series of high stakes poker games. The loser of each game must pay the winner \$100. Josh has been practicing Poker all semester, so he has a $\frac{2}{3}$ chance of winning each game. He reasons that he can't lose money as long as he plays enough games. Prove that the chance that Josh loses money overall decreases exponentially in the number of games n that the TAs play.

Here are some steps to help:

(a) Let A be the number of games won by Woo. Using Markov's inequality, show that for every $t \in \mathbb{R}^+$, $\Pr\{A > \frac{n}{2}\} \leq \frac{\mathbf{E}[e^{tA}]}{e^{0.5tn}}$

(b) Show that $\mathbf{E}[e^{tA}] \leq e^{\frac{n}{3}(e^t-1)}$ (You'll want to write A as a sum of Bernoulli random variables, and you'll need to use the "useful inequality" from the randomized algorithms lecture)

(c) Let $t = \ln(\frac{3}{2})$ (found using the magic of calculus). Show that $\Pr\{A > \frac{n}{2}\} \leq e^{-\frac{n}{30}}$ (You may use the fact that $\ln(\frac{3}{2}) > \frac{2}{5}$)