CPSC 421: Introduction to Theory of Computing
Assignment #6, due Monday November 12th by 11:59pm (midnight), via Gradescope

[1] 1. **LATEX BONUS!** You get 1 bonus mark if the homework is typeset using Latex.

[10] 2. Suppose that

- $A \subseteq \Sigma^*$ is NP-complete,
- $B \subseteq \Sigma^*$ is in P,
- $A \cap B = \emptyset$, and
- $A \cup B \neq \Sigma^*$. That is, there exists a string $z \in \Sigma^* \setminus (A \cup B)$ that is known to you.

Prove that $A \cup B$ is NP-complete.

[10] 3. In class (and in Sipser Theorem 7.32) a reduction f is described showing that $3SAT \leq_P CLIQUE$. In class we described the independent set problem:

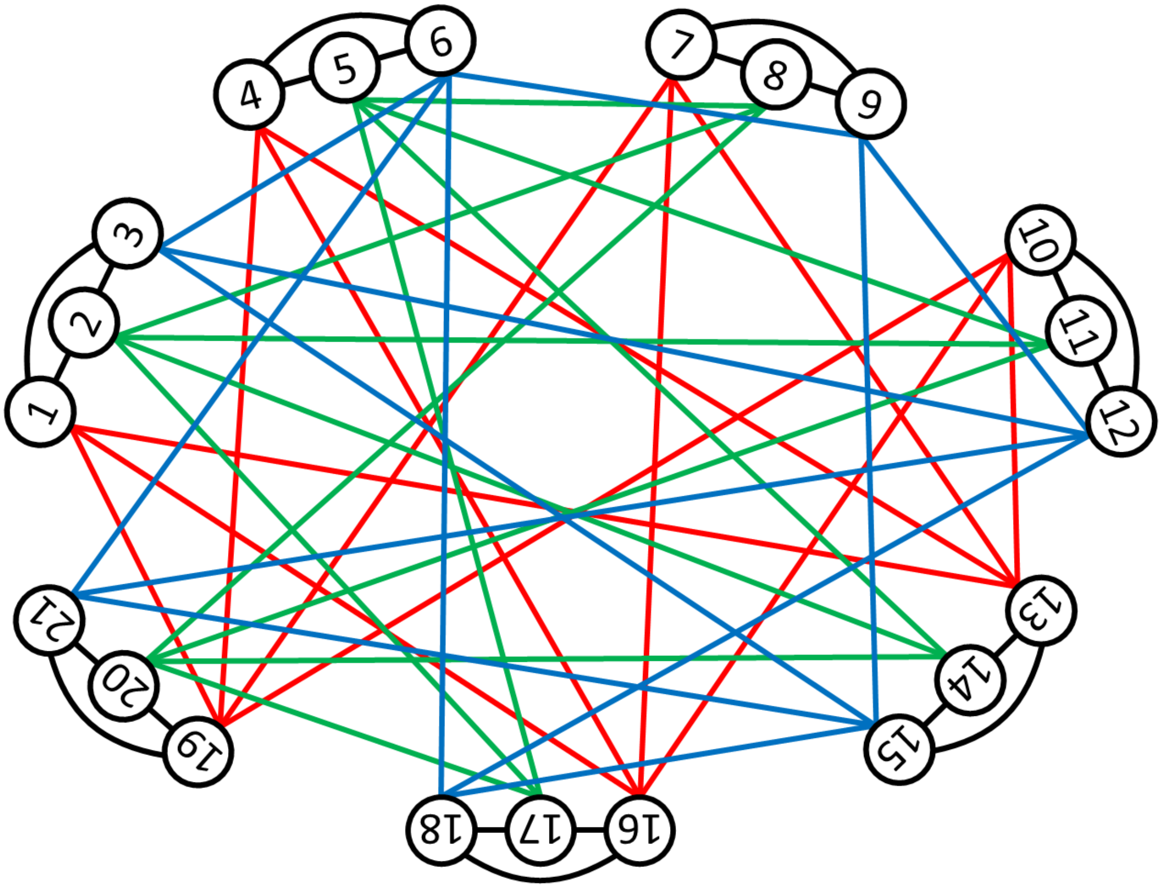
$$INDSET = \{ \langle G, k \rangle \ : \ G \text{ has an independent set of size } k \}.$$

We described a reduction g showing that $CLIQUE \leq_P INDSET$. The definition was:

g : input $w \in \Sigma^*$
  If w not of the form $\langle G, k \rangle$, output w
  Otherwise
    Compute $\bar{G}$, the complement of G
    Output $\langle \bar{G}, k \rangle$

The combined reduction $g \circ f$ (which applies f, then g) shows that $3SAT \leq_P INDSET$.

The following picture shows the graph H obtained by starting from some Boolean formula ϕ (in 3CNF form) then applying $g \circ f$. (The reduction does not actually put colours on the edges or numbers on the vertices. I added the colours to make the picture easier on your eyes. I added the numbers to part (c) easier to answer.)



[1] a. Actually, the reduction $g \circ f$ produces a string of the form $\langle H, k \rangle$. What is the value of k produced by this reduction?

[6] b. What is a Boolean formula ϕ that would produce this graph H? Is this formula satisfiable?

[3] c. What is the maximum size of an independent set in the graph H? Find a maximum cardinality independent set (using the numbers on the vertices to indicate which vertices you have chosen).

[15] 4. The "Interval Depth" problem is as follows. Given a set of n intervals on the real line, we would like to determine the largest subset of these intervals that contain a common point. (Each interval is of the form $[x, y]$ where $x, y \in \mathbb{R}$ and $x < y$.)

We may write the Interval Depth problem as a language $INTDEPTH$, which contains strings of the form $\langle k, x_1, y_1, \ldots, x_m, y_m \rangle$, where $x_i < y_i$, and *there exist* k intervals containing a common point.

[7] a. Describe a polynomial-time reduction from $INTDEPTH$ to $CLIQUE$.

Hint: You may use Helly's theorem without proof.

[5] b. Describe and analyze a polynomial-time algorithm for $INTDEPTH$.

[3] c. Why don't these two results imply that $P = NP$?

[5] 5. This question is about Sipser's proof of Theorem 7.37 (the Cook-Levin theorem). (Beware: other proofs that you might find in other books or online resources might be different. In particular, they might not use the notion of a "configuration".)

Each row of the tableau is supposed to be a "configuration" (defined on Sipser, page 168). How does the formula ϕ ensure that each row (i) contains *at least one* state q_i, and (ii) does not contain *two or more* states q_i.

[10] 6. We showed in class that $CLIQUE$ is NP-complete. (Sipser Theorem 7.32.) So, $CLIQUE \in P$ if and only if $P = NP$. Consider instead the problem

$$MAXCLIQUE \;=\; \{\, \langle G, k \rangle : \text{ the maximum clique in } G \text{ has exactly } k \text{ vertices} \,\}.$$

The $MAXCLIQUE$ problem is *not believed* to be in NP.

[5] a. Use a polynomial-time mapping reduction (as in Sipser Definition 7.29) to show that $MAXCLIQUE$ is NP-hard.

Hint: Think carefully about Lecture 23.

[5] b. Prove that if $P = NP$ then $MAXCLIQUE \in P$.

[2] 7. **OPTIONAL BONUS QUESTION:**

Let us say that a boolean formula is a "four-occurrence CNF formula" if it is in conjunctive normal form and every variable appears at most four times. Define

$$CNF_4 \;=\; \{\, \langle \phi \rangle : \phi \text{ is a satisfiable, four-occurrence CNF formula} \,\}.$$

It is known that CNF_4 is NP-complete.

Let us say that a boolean formula is a "four-occurrence 4CNF formula" if it is in conjunctive normal form, every variable appears at most four times, and every clause contains exactly four literals (no repetitions). Define

$$4CNF_4 \;=\; \{\, \langle \phi \rangle : \phi \text{ is a satisfiable, four-occurrence 4CNF formula} \,\}.$$

Prove that $4CNF_4$ is in P.