Section 7

Prof. Boaz Barak

1 Polynomial Time Reductions

Recall how we have been using reductions to prove the uncomputability of functions. We now extend that idea to incorporate notions of time complexity.

1.1 Definition

Let $F, G : \{0,1\}^* \to \{0,1\}$. We say that F reduces to G, denoted by $F \leq_p G$ if there is a polynomial-time computable R such that for every $x \in \{0,1\}^*$,

$$F(x) = G(R(x))$$

This is equivalent to saying that F is "easier" than G, as we can efficiently compute F so long as we have a way to efficiently compute G. Remember that "efficient" implies polynomial time.

Exercise: Show that for every $F, G, H : \{0,1\}^* \to \{0,1\}$, if $F \leq_p G$ and $G \leq_p H$, then $F \leq_p H$.

Solution: By definition, we have polynomial time computable functions R_1, R_2 such that $F(x) = G(R_1(x))$ and $G(x) = H(R_2(x))$. Then $F(x) = H(R_2(R_1(x)))$. But $R_1, R_2 \in$ P, and P is closed under composition, so $R_2(R_1(x))$ is polynomial computable.

1.2 Practice

Problem 1: The vertex cover of size k over a graph is a set of k vertices such that each edge in the graph has an endpoint in the set. $\mathsf{VERTEX-COVER}(G, k)$ returns 1 if there is a vertex cover of size at most k in the graph G, and 0 otherwise. Prove that $\mathsf{3SAT} \leq_p \mathsf{VERTEX-COVER}$.

Solution: (Dumbbells and Triangles). We seek to create a graph representing the 3SAT equation such that it's satisfiable if and only if there is a vertex cover of size $n + 2m$, where n is the number of literals in the 3SAT equation and m is the number of clauses.

Consider the following construction of such a graph. For each variable x in the formula, add 2 nodes x and $\bar{x}$ and connect them with an edge, kind of like a dumbbell. For each clause x_i, x_j, x_k, add a new node that represents each of the literals in that clause, and connect those new nodes together in a triangle so that each literal is connected to the others in the same clause. Then, connect each of the new literal nodes to the original variable nodes.

We posit that this graph has a vertex cover of size $n + 2m$ if and only if the 3SAT expression it represents is satisfiable. We see first that if there is a satisfiable solution, we can construct a vertex cover from it. We do so by selecting the n variable nodes (1/2 of each pair of the dumbbell nodes). This will cover each of the edges connecting the variable nodes, and 1 out of each of the three nodes connecting the variable nodes with the literal nodes (the cross edges between dumbbells and triangles). Now, to cover the remaining cross edges and the edges between literal nodes in each of the triangles, you will have to select up to 2 nodes for each of the clauses, which is up to $2m$ overall. Thus, there will be a vertex cover of at most $n + 2m$ nodes if the 3SAT expression is satisfiable.

To show the other direction, we see that given a vertex cover of size $n + 2m$, we must have an assignment that satisfies the expression. n of the vertices must be used to cover the variable nodes, since there are n edges there and at least one of their endpoints must be covered. Then, the remaining $2m$ vertices must cover the edges within each triangle, as well as the cross edges between the triangles and the variable nodes, and so we have a satisfying assignment.

Finally, we need to show that the transformation from the expression to the graph takes polynomial time. There are exactly $2n + 3m$ vertices, and there are n edges to connect the variable nodes, $3m$ cross edges, and $3m$ nodes in all of the triangles, so that gives us both $O(n+m)$ vertices and edges. Thus, we have a polynomial time transformation.

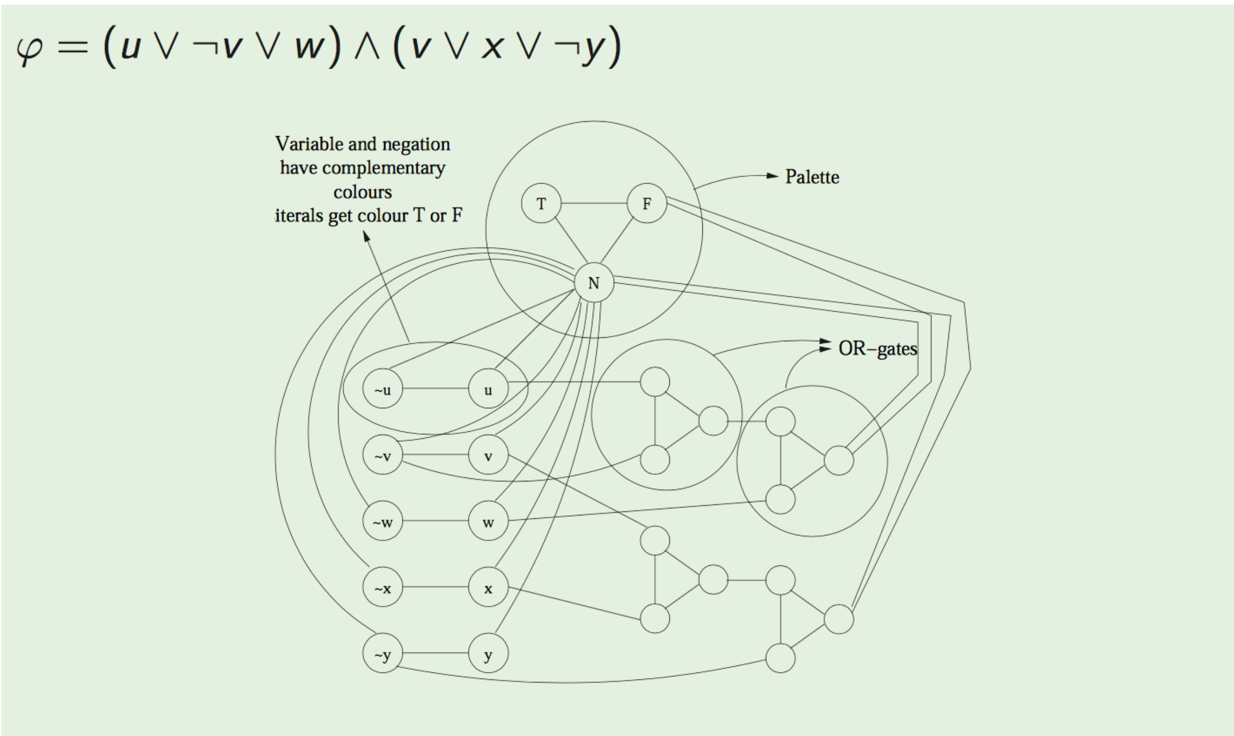
Problem 2: A "half cover" of size k over a graph is a set of k vertices such that at least half the edges of the graph have an endpoint in the set. Prove that VERTEX-COVER $\leq_p$ HALF COVER

Solution: Let the cover size in question be k. Given a graph G with vertices V and edges E, we create G' by augmenting G with $|E|$ disjoint vertex pairs and $|E|$ edges (one for each pair). Now we run HALF COVER on (G', k). If there is a k cover of graph G, then there is a k half cover of graph G', since the k cover of graph G itself suffices as there are $2|E|$ edges total and $|E|$ of them have endpoints in the cover. Conversely, if there is a k half cover on graph G', then we know that $|E| - m$ of the edges are not covered for some non-negative m. But we also know that the half cover covers at least $|E|$ edges, so it must cover at least m of the edges of the pairs unique to G'. Uncover m of these dumbbells and cover the m uncovered edges of E instead. This does not change the size of the half cover, but now all the edges of E are covered, so we have a vertex cover for G.

The transformation is polynomial time because it is bounded linearly by E, the number of edges in the original graph.

Problem 3: A 3 coloring of a graph is an assignment of a color $\in \{R, B, G\}$ to every vertex of the graph such that no two vertices connected by an edge are assigned the same color. Prove that 3SAT $\leq_p$ 3COLOR.

Solution: Construct a graph as follows. Create triangle with nodes True, False, and Neutral. For each variable, create a dumbbell (x and $\neg x$) with both ends also connected to Neutral. For each clause, create an "OR gadget" with edges connecting to the corresponding variables and an output node that connects to False and Neutral. See diagram for example.



Problem 4 Let SINGLE2SAT take in a 2CNF formula, and returns 1 if there is exactly one solution, and 0 otherwise. Similarly, let DOUBLE2SAT take in a 2CNF formula, and returns 1 if there is exactly two solutions, and 0 otherwise. Show that SINGLE2SAT $\leq_p$ DOUBLE2SAT.

Solution: Add an extra clause $(y \vee \neg y)$, where y is a new variable not present in the original 2CNF.