Lecture 11

Lance Fortnow(notes for 02/01/08 scribed by Sanchit Misra)

February 4, 2008

1 3-SAT

3-SAT = { ϕ | ϕ is a satisfiable boolean formula and ϕ is a CNF with 3 literals in each clause}
Claim: 3-SAT is NP-Complete.
Proof: Clearly, 3-SAT $\in$ NP (Guess and check satisfying assignment).
Now we will prove that CNF-SAT $\leq_m^p$ 3-SAT

- $\forall$ clauses of the form x, introduce two new variables y and z and create the following clauses: $(x \bigvee y \bigvee z) \bigwedge (x \bigvee \overline{y} \bigvee z) \bigwedge (x \bigvee y \bigvee \overline{z}) \bigwedge (x \bigvee \overline{y} \bigvee \overline{z})$
- $\forall$ clauses of the form $x_1 \bigvee x_2$, introduce a new variable z (this z is different from the z above) and create the following clauses: $(x_1 \bigvee x_2 \bigvee z) \bigwedge (x_1 \bigvee x_2 \bigvee \overline{z})$
- $\forall$ clauses of the form $x_1 \bigvee x_2 \bigvee ... \bigvee x_k (k > 3)$, introduce a new variable z and create the following clauses: $(x_1 \bigvee x_2 \bigvee ... \bigvee x_{k-2} \bigvee z) \bigwedge (\overline{z} \bigvee x_{k-1} \bigvee x_k)$
  This gives us $k-1$ literals in first term and 3 literals in second term.
  This is repeated till all the clauses have 3 literals.

Hence 3-SAT is NP-Complete

2 Vertex Cover

G is an undirected graph G(V,E), vertex cover is a set of vertices which cover all the edges. Vertex Cover problem can be defined as:
VC = $\{(G, k) \mid \exists V' \subseteq V \ni |V'| = k$ and $\forall (u, v) \in E$, either $u \in V'$ or $v \in V'\}$
Claim: VC is NP-Complete

Proof: VC $\in$ NP (Guess a subset of V and verify that the size is k and it covers all the edges)
3-SAT $\leq_m^p$ VC

Let ϕ be a 3-CNF with n variables and m clauses.
Create graph G as follows:
2 vertices for each variable x and $\overline{x}$.Join them with an edge
3 vertices corresponding to c_1, c_2 and c_3 for each clause of the form $c_1 \bigvee c_2 \bigvee c_3$. Join all three of them. For each vertex corresponding to a clause, add an edge from the vertex to the corresponding vertex for the variable. Take $k = n + 2m$. Figure 1 shows an example.

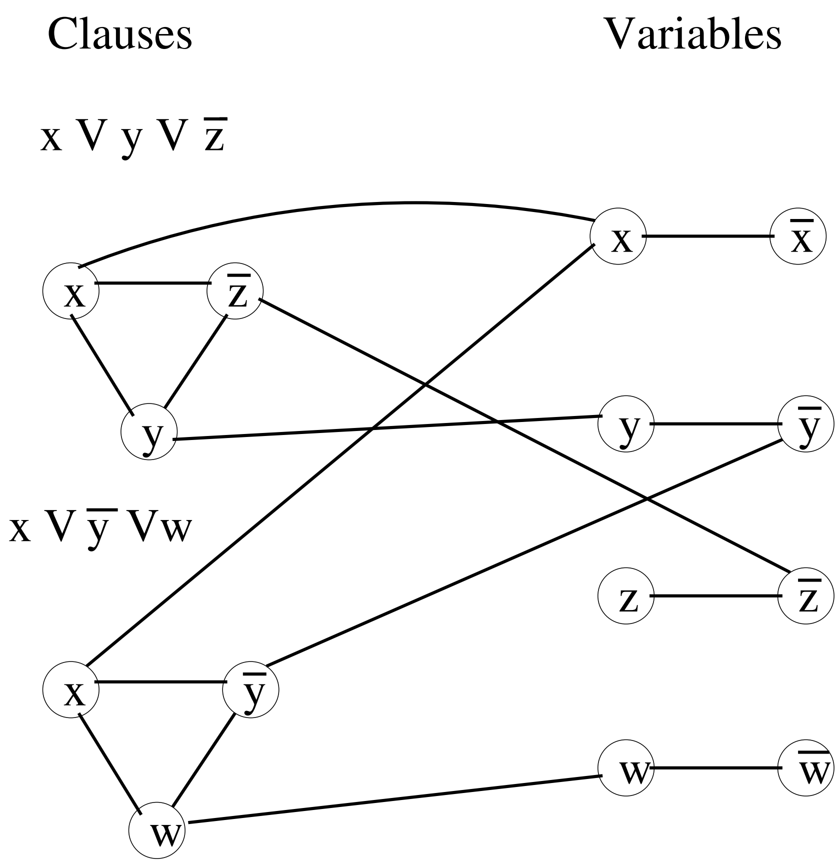

Figure 1: Reduction of 3-SAT to vertex cover

To prove that this reduction works:

- Suppose the formula is satisfiable. Take the corresponding assignment of variables. If a particular variable, say x, is assigned to TRUE, then include the vertex corresponding to variable x in the vertex cover and for all the clauses in which x is used, include the other two vertices of the clause (vertices which are not marked x). Similarly if x is assigned to FALSE, include vertex corresponding to $\overline{x}$ and in the clauses which use $\overline{x}$, include the other two vertices. This way the vertex cover size $= n + 2m$. For example, in the example given by Figure 1, a possible assignemt is:

  $x : TRUE$, $y : FALSE$, $z : TRUE$ and $w : FALSE$

  So we include vertices x, $\overline{y}$, z, $\overline{w}$ from variables and $\overline{z}$, y, w, $\overline{y}$ from clauses.

- Suppose there exists a vertex cover of size n+2m. Since every triangle should have atleast two nodes and each variable pair should have atleast one node in the vertex cover, we have a solution to the 3-SAT problem.

3 Other NP-Complete Problems

- Independent Set, IS $= \{(G, m) \mid I \ni I \subseteq V,\ |I| = m$ and $\forall u, v \in I, (u, v) \notin E\}$

  Clearly V' is a vertex cover iff $I = V - V'$ is an independent set. So, VC $\leq_m^p$ IS. Reduction function being $F(G, k) = (G, n - k)$

- Clique Set, CS $= \{(G, k) \mid V' \subseteq V \ni |V'| = k$ and $\forall u, v \in V', (u, v) \in E\}$

  This is same as finding independent set of the complement of the graph. Hence, IS $\leq_m^p$ CS. Reduction function: $F(G, k) = (\overline{G}, k)$

- 3-coloring:

  $\{G \mid \exists c : V \rightarrow \{R, G, B\} \ni \forall (u, v) \in E, c(u) \neq c(v)\}$

- Subset Sum:

  $\{(x_1, x_2, ..., x_k, y) \mid$ Is there $S \subseteq \{1, ..., k\} \ni \sum_{i \in S} X_i = y\}$

  Other similar problems are Bin Packing and Knapsack

- Hamiltonian Cycle:

  $\{G \mid \text{There is a } \sigma \text{ permutation on } \{1, ..., n\} \ni \forall i < n, (\sigma(i), \sigma(i+1)) \in E$ and $(\sigma(n), \sigma(1)) \in E\}$

  Eulerian cycle is a similar problem where one has to visit every edge exactly once but it is in P owing to the following theorem:

  A graph is eulerian iff G is connected and every node has even degree.

- TSP (Travelling Salesperson Problem):

  Euclidean TSP is a variant of TSP, in which there are n points in a plane and we are given the Euclidean distance between them. Can we visit every node and do it in total m units distance?

  Hamiltonian cycle $\leq_m^p$ Euclidean TSP

  But we don't know if "$TSP \in NP$?" since we need to compare the Euclidean distance with m; i.e.;

  $\sum_i \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} < m$

  Some of the square-roots can be very long and different values of m require different precisions. Hence, so far we don't know whether we can compute it in polynimial time.