CSE547 Class 21

Jeremy Buhler

April 12, 2017

1 SUBSET-SUM

Are things that are *not* SAT-like NP-hard?

- Let's talk about the subset-sum problem.
- **Input**: a set S of positive integers, and a target t
- **Problem**: does S contain subset S' whose members add to exactly t?
- **Ex**: $S = \{2, 5, 3, 6, 9\}$: true for $(S, 14)$, but not for $(S, 4)$.

Well, that doesn't seem hard, does it?

- **Lemma**: SUBSET-SUM is NP-complete.
- **Pf**: First, will show it's in NP.
- Given instance (S, t), certificate is subset S'.
- Surely, $|S'| \leq |S|$.
- Moreover, can check in time $O(|S||S'|)$ that all elts of S' are in S, and that their sum is t.

So far, so good. But what about NP-hardness?

- Will reduce from 3-SAT (i.e. prove 3-SAT $\leq_p$ SUBSET-SUM)!
- Given a 3-CNF formula ϕ, will construct an instance (S, t) of SUBSET-SUM.
- Will show that S has subset with sum t iff ϕ is satisfiable.
- WLOG, will assume ϕ does not contain any clause with both literals x and $\neg x$.
- (Any such clause is trivially satisfiable, so transducer function f can delete it.)

Yikes! What's the construction?

- Suppose ϕ has n variables and m clauses.
- Set S will contain base-10 integers, each with $n + m$ digits.

- Label the digit positions $p_1 \ldots p_n, q_1 \ldots q_m$.
- For variable x_i, define integers v_i and $\bar{v}_i$ as follows:
  1. Digit p_i of both v_i and $\bar{v}_i$ is 1.
  2. All other digits p_j, $j \neq i$, are 0 for both v_i and $\bar{v}_i$.
  3. Digit q_k of v_i is 1 iff x_i appears in clause C_k of ϕ.
  4. Digit q_k of $\bar{v}_i$ is 1 iff $\neg x_i$ appears in clause C_k of ϕ.
- For each clause C_k, define integers y_k and z_k as follows:
  1. All digits p_j of both y_k and z_k are zero.
  2. Digit q_k of y_k is 1; digit q_k of z_k is 2.
  3. All other digits q_ℓ, $\ell \neq k$, of both y_k and z_k are zero.
- Finally, define target t as follows:
  1. Every digit p_j of t is 1.
  2. Every digit q_k of t is 4.
- **Example**:

That's really funky (but at least it's polynomial-time). Why does it work?

- *Observation 0*: S really is a set – v_i and $\bar{v}_i$ differ because x_i and $\neg x_i$ do not occur together in any clause, and all other numbers are pairwise different.
- *Observation 1*: if we add up any subset of values in S, no digit position causes a carry.
- (No position p_j exceeds 2, and no position q_k exceeds 6.)
- Hence, we can consider each position separately.
- **Claim 1**: If ϕ is satisfiable, (S, t) is solvable.
- **Pf**: let A be a satisfying assignment to ϕ.
- Construct S' as follows.
- If A makes variable x_i true, add v_i to S'; otherwise, add $\bar{v}_i$.

- This alone guarantees that posn p_i sums to 1 for $1 \leq i \leq n$.
- For clause C_k, let a_k be the sum in posn q_k of all values in S' so far.
- Note that a_k must be at least 1 (since ϕ is satisifed) and at most 3.
  - If a_k is 3, add y_k to S'.
  - If a_k is 2, add z_k to S'.
  - If a_k is 1, add both y_k and z_k to S'.

  This ensures that position q_k sums to 4 over S'.
- Conclude that total of all values in S' is exactly t.

Halfway there...

- **Claim 2**: If (S, t) is solvable, ϕ is satisfiable.
- **Pf**: Let S' be a valid solution to (S, t).
- Construct assignment A for ϕ as follows:
  - If $v_i \in S'$, set x_i true.
  - If $\bar{v}_i \in S'$, set x_i false.
- Note first that every valid solution to (S, t) includes exactly one of v_i and $\bar{v}_i$.
- (Only way to make sum in posn p_i equal to 1.)
- Hence, A assigns every x_i a unique truth value.
- Suppose clause C_k contains literals ℓ_1, ℓ_2, and ℓ_3.
- **Notn**: for a literal ℓ, let $v(\ell)$ be the value in S corresponding to ℓ.
- Every valid solution to (S, t) must contain at least one of the values $v(\ell_1)$, $v(\ell_2)$, $v(\ell_3)$.
- (We cannot make 4 in posn q_k using only y_k and z_k.)
- But then A makes at least one of the these three literals true!
- Conclude that A satisfies every clause of ϕ. QED

2 CLIQUE

Here's another problem for which 3-SAT helps us out.

- Remember the CLIQUE problem?
- Given a graph G, does G contain a clique of size at least k?
- **Lemma**: CLIQUE is NP-complete.

- **Pf**: we argued earlier that CLIQUE is in NP.
- Will now show that 3-SAT $\leq_p$ CLIQUE.
- Given 3-CNF formula ϕ with m clauses, such that clause C_k contains literals ℓ_k^1, ℓ_k^2, and ℓ_k^3.
- Construct a graph G as follows.
- For each literal ℓ_k^r, create a vertex v_k^r.
- Add an edge between each pair of vertices v_k^r and v_j^s for which
  1. $j \neq k$ (literals are in distinct clauses)
  2. $v_k^r \neq \neg v_j^s$ (literals are not logical opposites)
- Let (G, m) be the instance of CLIQUE corresponding to ϕ.
- (Clearly, constructible in time $O(|\phi|^2)$.)
- **Example**:

On to the proof!

- **Claim 1**: If ϕ is satisfiable, then G has a clique of size m.
- **Pf**: Let A be a satisfying assignment for ϕ.
- Every clause C_k has at least one true literal; designate one such literal ℓ_k^* per clause.
- Consider the subgraph G' of G containing verts v_k^* for all k.
- G' contains m vertices.
- Moreover, every pair ℓ_k^* and ℓ_j^* are in different clauses, and no pair can be logically contradictory.
- Hence, every such pair is joined by an edge.
- Conclude that G' is a clique of size m!
- **Claim 2**: If G has a clique of size m, then ϕ is satisfiable.
- **Pf**: Let G' be a clique of size m in G.

- Construct assignment A for ϕ as follows.
- If G' contains vertex v_k^r, assign value to var of literal ℓ_k^r so as to make this literal true.
- No two verts in G' correspond to contradictory literals (since each pair is connected by an edge), so this assignment is logically consistent.
- (For any unassigned vars of ϕ, set them arbitrarily in A.)
- No two verts of G' correspond to literals from same clause.
- Hence, A makes one literal from each of m clauses true.
- Conclude that A satisfies every clause in ϕ. QED