Aditya Shirode, Rhythm Shah, Shrenuj Gandhi

4. [6 points]

The **Set Splitting (SSP)** problem is defined as follows: Given a collection of subsets C of a finite set S, does there exist a partition of S into subsets S' and $S - S'$ such that no subset in C is entirely inside S' or entirely inside $S - S'$? Prove that SSP is NP-Complete. Hint: Use 3-Sat.

An example of set splitting is

$$C = \{\{u, v, x\}, \{w, x, y\}, \{u, w, y\}, \{u, x, z\}, \{y, z\}, \{v, x\}\}$$

The choice of $S' = \{w, x, z\}$ will work. All but two of the sets in C contain x but not both w and z, and are therefore "split" by S'. The remaining sets are $\{u, w, y\}$, which contains w but not x or z, and $\{y, z\}$, which contains z but not w or x. What if we added the set S' to C? Then our certificate would no longer work. Is there another one?

[Bonus problem (3 points):] Prove that a polynomial algorithm for SSP implies a polynomial algorithm for the certificate version of SSP, i.e., instead of an output of yes or no, the output is a set S' that satisfies the condition or an indicator that no such set exists.

Answer:

To prove that set splitting is NP complete, we need to prove the following -

(1) Prove that set Splitting $\in NP$
Given a choice of S', it is easy to verify in polynomial time that no set in c contains all the elements n S' or in $S - S'$.
In other words, if elements in S' were colored Red and elements in $S - S'$ were colored $Green$, we can verify in polynomial time that no set in c is monochromatic.

(2) Prove that some problem (say $3 - CNF - SAT$) in NPC is polynomial time reducible to *SET-SPLITTING* i.e. $3 - CNF - SAT \leq_p$ SET-SPLITTING
(a) Mapping f that converts instance of $3 - CNF - SAT$ into instance of *SET-SPLITTING*
Let ϕ be the $3 - CNF - SAT$ formula with the variable set V.
Create the *SET-SPLITTING* instance

S as -
- One element for each variable $x_i \in V$
- One element for each variable $\neg x_i, x_i \in V$
- A new element f that is not related to any variable

C as -
- $S_{x_i} = \{x_i, \overline{x}_i\}$ for each variable $x_i \in V$
- $S_{c_i} = \{X, Y, Z, F\}$ for each clause of the form $c_i = X \vee Y \vee Z$

So the reduction is $f(\phi) = (S, C)$, where S and C are as described above.
This reduction takes place in polynomial time.

(b) Prove that certificate for $3 - CNF - SAT$ implies certificate for $SET - SPLITTING$.
i.e. if $3 - CNF - SAT$ formula was satisfiable, then the $SET - SPLITTING$ instance is solvable

as well.

Consider the following assignment:

- Assign element f to *red*
- For each variable x_i that is assigned $False$ (in the $3-CNF-SAT$ problem), color the elements x_i and $\overline{x}_i$ (in set S_{x_i} of C) *green* and *red* respectively.
- For each variable x_i that is assigned $True$ (in the $3-CNF-SAT$ problem), color the elements x_i and $\overline{x}_i$ (in set S_{x_i} of C) *red* and *green* respectively.

As long as the assignment was satisfying, this coloring makes no set monochromatic, which implies that the $SET-SPLITTING$ instance can be solved.

(c) Prove that the certificate for $SET-SPLITTING$ implies certificate for $3-CNF-SAT$ i.e. $SET-SPLITTING$ instance is satisfiable,
then $3-CNF-SAT$ is also solvable.

Consider the following assignment to the variable of ϕ.

- For each variable x_i, assign *True* (in $3-CNF-SAT$) if its color in $SET-SPLITTING$ instance differs from that of element f.
- For each variable x_i, assign *False* (in $3-CNF-SAT$) if its color in $SET-SPLITTING$ instance is same as that of element f.

Then each clause c in ϕ is satisfied, because the set S_{c_i} has at least one element that is colored differently than f, thus satisfying the $3-CNF-SAT$ instance.

BONUS

Let *SSP-DEC(C, S')* be the polynomial time decision algorithm for *SET-SPLITTING* problem.

SSP-DEC(C, S') returns *True* if no subset in c is entirely inside S' or entirely inside $S - S'$ and returns *False* otherwise.

Lets define function *SET-SHRINK(C,x)* such that it replaces x with special symbol # in all the sets occurring in C.
Moreover it returns *True* if after the replacement all sets in C have cardinality greater than 1 and returns *False* otherwise. (Also retaining all occurences of x).

The following is the certificate algorithm -

```
SSP-CERT(C,S):

        for each element x in S, do:
                if SET-SHRINK(C,x) = True, then:
                        S' ← S' ∪ x
                if SSP-DEC(C,S') = True, then:
                        return S'

        return False
end
```

This algorithm ouptuts the set S' that satisfies the condition or indicates that no such set exists.

Let $n = |S|$ and $m = |C|$, and $D(m, n)$ be the time taken by decision algorithm.
Then the runtime of the above algorithm is $\mathrm{O}(nm+nD(m, n))$ which is polynomial in size of m and n.

Thus if the decision algorithm of *SSP* is polynomial,
then the corresponding certificate version is also polynomial.