CS 381 – FALL 2019

Week 14.3, Friday, Nov 22

Homework 7 Due: November 26th at 11:59PM on Gradescope

Monday Office Hours (November 25th): ~~2:30-3:30PM~~ 5:30-6:30PM

SUBSET SUM

SUBSET SUM:

Instance: n integers $x_1, \dots, x_n$ and a separate integer k

Question: Does there exist a subset $S \subseteq \{1, \dots, n\}$ s.t. $\sum_{i \in S} x_i = k$

Example 1: (YES Instance)

Instance: $x_1 = 4, x_2 = 5, x_3 = 8, x_4 = 9$ and k=14

Witness: $S = \{2,4\} \rightarrow \sum_{i \in S} x_i = 5 + 9 = 14 = k$

Example 2: (NO Instance)

Instance: $x_1 = 4, x_2 = 5, x_3 = 8, x_4 = 9$ and k=15

FACT: SUBSET SUM is NP-Complete

3SAT $\leq_p$ SUBSET SUM

We give a polynomial reduction f mapping 3SAT Φ instances to SUBSET SUM instances.

Goal: Φ has satisfying assignment if and only if $(y_1, \ldots, y_{n'}$,k)=f(Φ) is satisfiable (i.e., there exist a subset $S \subseteq \{1, \ldots, n'\}$ s.t. $\sum_{i \in S} x_i = k$)

Reduction: Φ has n variables $x_1, \ldots, x_n$ and m clauses

m + n + 1 digits (m "clause" digits & n "variable" digits + padding)

(1) Set k = 3333…311…11110

(2) Let $S_i \subseteq \{1, \ldots, m\}$ be the clauses that contain the literal x_i for $i \leq n$

Example: Clauses m-1 and 1 contain x_i $(S_i = \{1, m-1\})$

$$y_i = 10^i + \sum_{j \in S_i} 10^{n+j}$$

$$y_i = 010\ldots0100\ldots010..000$$

(10^i)

$3SAT \leq_p$ SUBSET SUM

Reduction: Φ has n variables $x_1, \ldots, x_n$ and m clauses

m +n digits (m clause digits & n literal digits + padding)

(1) Set k = 3333…311…1110

(2) Let $S_i \subseteq \{1, \ldots, m\}$ be the clauses that contain the literal x_i for $i \leq n$

Example: Clauses m-1 and 1 contain x_i

$$y_i = 10^i + \sum_{j \in S_i} 10^{n+j} \qquad y_i = 010\ldots0100\ldots010..000$$

(10^i)

(3) Let $S_{i+n} \subseteq \{1, \ldots, m\}$ be the clauses that contain the literal $\bar{x}_i$ for $i \leq n$

Example: Clauses m and 2 contain $\bar{x}_i$ $(S_{i+n} = \{2, m\})$

$$y_{i+n} = 10^i + \sum_{j \in S_{i+n}} 10^{n+j} \qquad y_{i+n} = 100\ldots1000\ldots010..000$$

(10^i)

(4) For each clause C_i add two numbers $y_{i,1} = y_{i,2} = 10^{n+i} = 0\ldots010\ldots00\ldots00$

3SAT $\leq_p$ SUBSET SUM

Observation 1: No Overflows/Carries (even if we sum every number)

$$y_1 + \cdots + y_{2n} + y_{i,1} + \cdots + y_{m,1} + y_{i,2} \ldots + y_{m,2} = 5555\ldots522\ldots2220$$

Only y_i and y_{i+n} contain a 1 at digit i

(1) Set k = 3333…311…11110

(2) Let $S_i \subseteq \{1, \ldots, m\}$ be the clauses that contain the literal x_i for $i \leq n$

Example: Clauses m-1 and 1 contain x_i

$$y_i = 10^i + \sum_{j \in S_i} 10^{n+j} \qquad y_i = 010\ldots0100\ldots010..000$$

10^i

(3) Let $S_{i+n} \subseteq \{1, \ldots, m\}$ be the clauses that contain the literal $\bar{x}_i$ for $i \leq n$

Example: Clauses m and 2 contain $\bar{x}_i$

$$y_{i+n} = 10^i + \sum_{j \in S_{i+n}} 10^{n+j} \qquad y_{i+n} = 100\ldots1000\ldots010..000$$

10^i

(4) For each clause C_i add two numbers $y_{i,1} = y_{i,2} = 10^{n+i} = 0\ldots010\ldots00\ldots00$

3SAT $\leq_p$ SUBSET SUM

Observation 1: No Overflows/Carries (even if we sum every number)

$$y_1 + \cdots + y_{2n} + y_{i,1} + \cdots + y_{m,1} + y_{i,2} \ldots + y_{m,2} = 5555\ldots522\ldots2220$$

At most five different y values have a 1 at digit n+j; all other y's have 0 at this digit
→ $y_{j,1} = y_{j,2} = 10^{n+j}$ (1 at digit n+j)
→ Clause C_j has 3 literals

(1) Set k = 3333…311…11110

(2) Let $S_i \subseteq \{1, \ldots, m\}$ be the clauses that contain the literal x_i for $i \leq n$

Example: Clauses m-1 and 1 contain x_i

$$y_i = 10^i + \sum_{j \in S_i} 10^{n+j} \qquad y_i = 010\ldots0100\ldots010..000$$

10^i

(3) Let $S_{i+n} \subseteq \{1, \ldots, m\}$ be the clauses that contain the literal $\bar{x}_i$ for $i \leq n$

Example: Clauses m and 2 contain $\bar{x}_i$

$$y_{i+n} = 10^i + \sum_{j \in S_{i+n}} 10^{n+j} \qquad y_{i+n} = 100\ldots1000\ldots010..000$$

10^i

(4) For each clause C_i add two numbers $y_{i,1} = y_{i,2} = 10^{n+i} = 0\ldots010\ldots00\ldots00$

$3SAT \leq_p$ SUBSET SUM

Observation 1: No Overflows/Carries (even if we sum every number)

$$y_1 + \cdots + y_{2n} + y_{1,1} + \cdots + y_{m,1} + y_{1,2} \ldots + y_{m,2} = 5555\ldots522\ldots2220$$

Suppose that there is a satisfying assignment e.g., $x_1=1, x_2=0, \ldots$

Build subset S as follows

1) If x_i is true then include y_i otherwise include y_{i+n}
2) Add 3-t_j items from the set $\{y_{i,1}, y_{i,2}\}$ where $t_j>0$ denotes number of true literals in clause C_j (Note that $3-t_j \leq 3$)

$$\sum_{y \in S} y = ???????11\ldots11110$$

1×10^i comes from y_i or y_{i+n}
(S includes exactly one of these values)

$$y_{i+n} = ?????????00\ldots010..000$$
$$y_i = ?????????00\ldots010..000$$

10^i

3SAT $\leq_p$ SUBSET SUM

Observation 1: No Overflows/Carries (even if we sum every number)

$$y_1 + \cdots + y_{2n} + y_{1,1} + \cdots + y_{m,1} + y_{1,2} \ldots + y_{m,2} = 5555\ldots522\ldots2220$$

Suppose that there is a satisfying assignment e.g., $x_1=1, x_2=0, \ldots$

Build subset S as follows

1) If x_i is true then include y_i otherwise include y_{i+n}
2) Add 3-t_j items from the set $\{y_{i,1}, y_{i,2}\}$ where t_j>0 denotes number of true literals in clause C_j (Note that 3-$t_j \leq 3$)

$$\sum_{y \in S} y = ???3???11\ldots11110$$

$1 \times 10^{n+j}$ appears in three y values from S

- y_i or y_{i+n} for t_j>1 true literals in clause C_j
- (3- t_j) from $y_{j,1} = y_{j,2} = 10^{n+j}$

3SAT $\leq_p$ SUBSET SUM

Observation 1: No Overflows/Carries (even if we sum every number)

$$y_1 + \cdots + y_{2n} + y_{1,1} + \cdots + y_{m,1} + y_{1,2} \ldots + y_{m,2} = 5555\ldots522\ldots2220$$

Suppose that there is a satisfying assignment e.g., $x_1=1, x_2=0, \ldots$

Build subset S as follows

1) If x_i is true then include y_i otherwise include y_{i+n}
2) Add $3-t_j$ items from the set $\{y_{i,1}, y_{i,2}\}$ where $t_j>0$ denotes number of true literals in clause C_j (Note that $3-t_j \leq 3$)

$$\sum_{y \in S} y = 333333311\ldots1110$$

Each clause contains $t_j>0$ true literals

$1 \times 10^{n+j}$ appears in three y values from S

- y_i or y_{i+n} for $t_j>1$ true literals in clause C_j
- $(3-t_j)$ from $y_{j,1} = y_{j,2} = 10^{n+j}$

3SAT $\leq_p$ SUBSET SUM

Suppose S is a valid subset sum solution s.t.

$$\sum_{y \in S} y = 33333331 1\ldots 11110$$

We define a truth assignment as follows:

Set x_i=true if y_i is included S and x_i=false if y_{i+n} is included S

Observe that S must contain exactly one of y_i and y_{i+n} (only numbers with a 1 at digit i) → consistent/complete truth assignment

Claim: Each clause C_j has a true literal.

Proof: (sketch) Consider clause $C_j = x_i \vee \bar{x}_j \vee x_k$

→ only 5 numbers contain a 1 at digit n+j y_i, y_{j+n}, y_k , and $y_{j,1} = y_{j,2} = 10^{n+j}$

↑

S must include at least one of these numbers to ensure that digit n+j is set to 3

→ at least one of the literals x_i, $\bar{x}_j$ or x_k is true

SUBSET SUM in NP

SUBSET SUM:

Instance: n integers $x_1, \ldots, x_n$ and a separate integer k

Question: Does there exist a subset $S \subseteq \{1, \ldots, n\}$ s.t. $\sum_{i \in S} x_i = k$

Witness: S

Certifier: C(($x_1, \ldots, x_n$,k),S)

if $\sum_{i \in S} x_i = k$ output ACCEPT; otherwise output REJECT

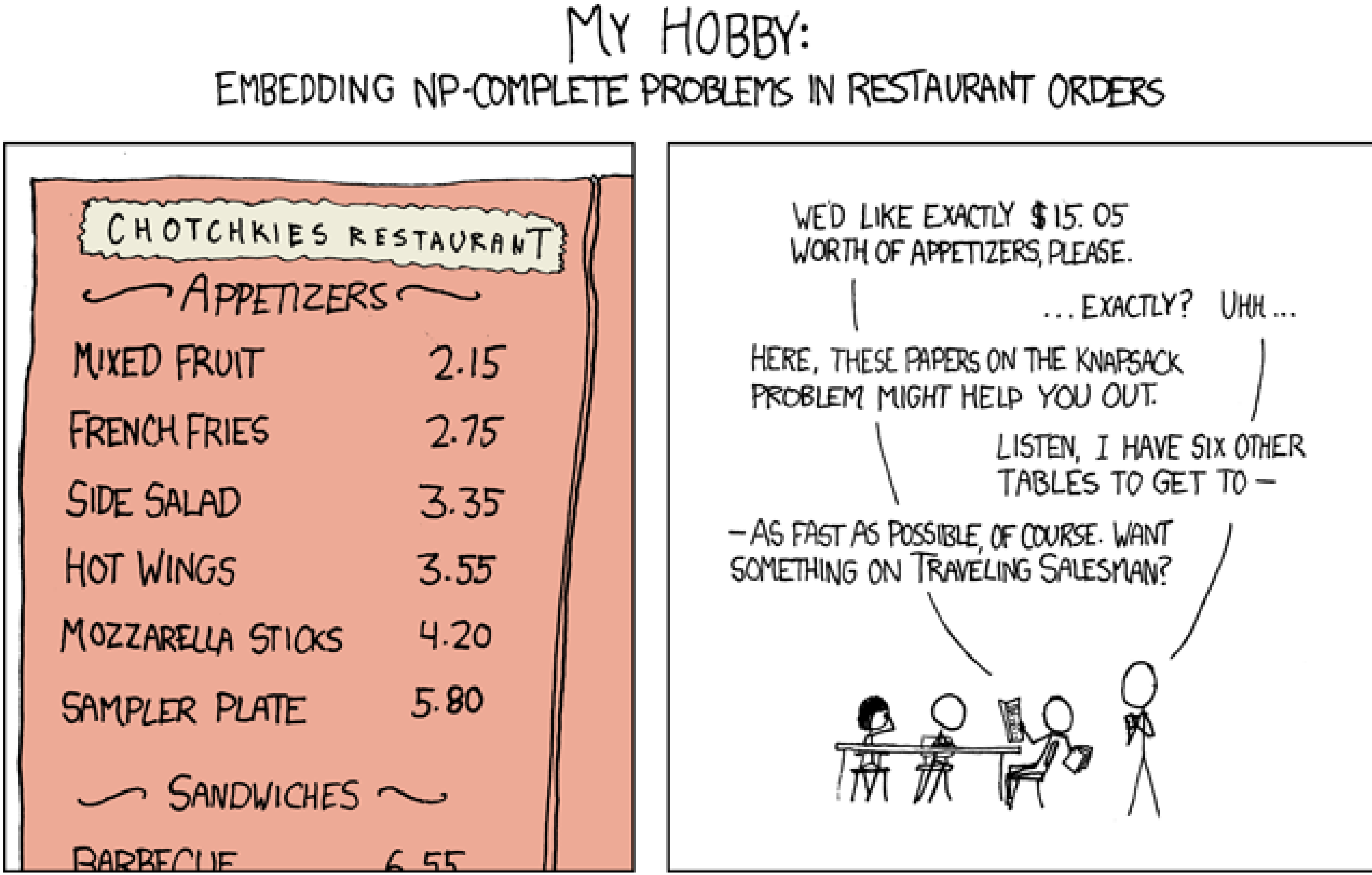
Example: (YES Instance)

Instance: $x_1 = 4, x_2 = 5, x_3 = 8, x_4 = 9$ and k=14

Witness: $S = \{2,4\} \rightarrow \sum_{i \in S} x_i = 5 + 9 = 14 = k$

Obs 1: Every YES instance ($x_1, \ldots, x_n$,k) contains a valid witness S

Obs 2: If I=($x_1, \ldots, x_n$,k) is a NO instance then C(I,w)=REJECT for all ``witnesses" w.



Randall Munro: `http://xkcd.com/c287.html`

8.9 co-NP and the Asymmetry of NP

Asymmetry of NP

Asymmetry of NP. We only need to have short proofs of yes instances.

Ex 1. SAT vs. TAUTOLOGY.

- Can prove a CNF formula is satisfiable by giving such an assignment.
- How could we prove that a formula is not satisfiable?

Ex 2. HAM-CYCLE vs. NO-HAM-CYCLE.

- Can prove a graph is Hamiltonian by giving such a Hamiltonian cycle.
- How could we prove that a graph is not Hamiltonian?

Remark. SAT is NP-complete and SAT $\equiv_P$ TAUTOLOGY, but how do we classify TAUTOLOGY?

↑ not even known to be in NP

↑ Φ is TAUTOLOGY if and only $\neg\Phi$ is a unsatisfiable

NP and co-NP

NP. Decision problems for which there is a poly-time certifier.

Ex. SAT, HAM-CYCLE, COMPOSITES.

Def. Given a decision problem X, its complement $\overline{X}$ is the same problem with the yes and no answers reverse.

Ex. $\overline{\text{PRIMES}}$ = { 0, 1, 4, 6, 8, 9, 10, 12, 14, 15, … }

PRIMES = { 2, 3, 5, 7, 11, 13, 17, 23, 29, … }

co-NP. Complements of decision problems in NP.

Ex. TAUTOLOGY, NO-HAM-CYCLE, PRIMES.

NP. YES instances have succinct witness/certificate

co-NP. NO instances have succinct disqualifier

(e.g., an assignment s.t. Φ evaluates to false → Φ is not a TAUTOLOGY)

NP = co-NP ?

Fundamental question. Does NP = co-NP?

- Do yes instances have succinct certificates iff no instances do?
- Consensus opinion: no.

Theorem. If NP $\neq$ co-NP, then P $\neq$ NP.

Pf idea.

- P is closed under complementation.
- If P = NP, then NP is closed under complementation.
- In other words, NP = co-NP.
- This is the contrapositive of the theorem.

Good Characterizations

Good characterization. [Edmonds 1965] NP $\cap$ co-NP.

- If problem X is in both NP and co-NP, then:
  - for yes instance, there is a succinct certificate
  - for no instance, there is a succinct disqualifier
- Provides conceptual leverage for reasoning about a problem.

Ex. Given a bipartite graph, is there a perfect matching.

- If yes, can exhibit a perfect matching.
- If no, can exhibit a set of nodes S such that $|N(S)| < |S|$.

Good Characterizations

Observation. $P \subseteq NP \cap$ co-NP.

- Proof of max-flow min-cut theorem led to stronger result that max-flow and min-cut are in P.
- Sometimes finding a good characterization seems easier than finding an efficient algorithm.

Fundamental open question. Does $P = NP \cap$ co-NP?

- Mixed opinions.
- Many examples where problem found to have a non-trivial good characterization, but only years later discovered to be in P.
  - linear programming [Khachiyan, 1979]
  - primality testing [Agrawal-Kayal-Saxena, 2002]

Fact. Factoring is in $NP \cap$ co-NP, but not known to be in P.

↑ if poly-time algorithm for factoring, can break RSA cryptosystem

FACTOR is in NP ∩ co-NP

FACTORIZE. Given an integer x, find its prime factorization.

FACTOR. Given two integers x and y, does x have a nontrivial factor less than y?

Theorem. FACTOR $\equiv_P$ FACTORIZE.

Theorem. FACTOR is in NP ∩ co-NP.

Pf.

- **Certificate:** a factor p of x that is less than y.
- **Disqualifier:** the prime factorization of x (where each prime factor is greater than y)
  - Can validate prime factorization of x since PRIMES is in P.

iClicker

Assume P $\neq$ NP. Which of the following problems are not in NP $\cap$ co-NP?

A. FACTOR

B. PRIMES

C. 3COLOR

D. BIPARTITE MATCHING (Given a bipartite graph G and integer k is there a matching that contains at least k edges)?

E. All of the problems are in NP $\cap$ co-NP

iClicker

Assume P ≠ NP. Which of the following problems are not in NP ∩ co-NP?

Should have assumed coNP ≠ NP

A. FACTOR
- FACTOR is in NP ∩ co-NP (but not known to be in P)

  (prime factorization is either a witness/disqualifier)

B. PRIMES
- Primes is in P (contained in NP ∩ co-NP)

C. 3COLOR
- If 3COLOR is in co-NP then co-NP = NP since 3COLOR is NP-Complete
- More precise problems statement: ``Assume coNP ≠ NP"

D. BIPARTITE MATCHING (Given a bipartite graph G and integer k is there a matching that contains at least k edges)?
- Primes is in P (contained in NP ∩ co-NP)

E. All of the problems are in NP ∩ co-NP
- This is actually true if NP = co-NP!
- (Full Credit if you picked choice E since we don't know)