DELFT UNIVERSITY OF TECHNOLOGY
Faculty of Electrical Engineering, Mathematics and Computer Science

CSE2315 Lab Course
Solutions 8

deadline: April 7, 2020, 15:45

1. Consider the problem HAMILTONIAN CIRCUIT (HAMC):

$$\{(V, E) \mid \text{The graph } G = (V, E) \text{ contains a simple cycle } \pi \text{ such that } \forall v \in V : v \in \pi\}$$

In a different notation we would write:

Name: HAMILTONIAN CIRCUIT (HAMC)
Input: A graph $G = (V, E)$.
Question: Is there a simple cycle π in G such that all vertices $v \in V$ are in that cycle?

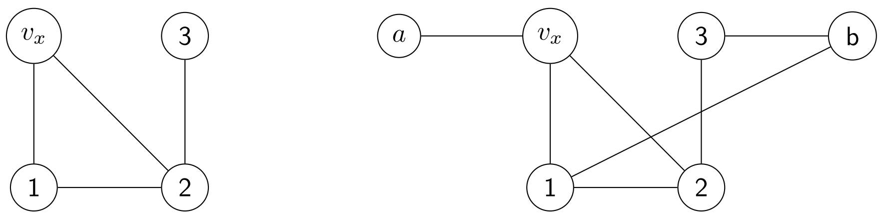
Now take the following reduction from HAMC to HAMP, given an instance $G = (V, E)$ of HAMC:

I. Copy G and add 2 new nodes a, b to V
II. Connect a to some vertex $v_x \in V$ and connect b to all neighbours of v_x.

Show that this reduction is **not** reliable.

Note: this can be quite challenging! Try and apply the reduction on some graphs and see what happens first, get a feeling for how the reduction influences the graphs.

Solution: Take for example the following:



The left is a no-instance of HAMC, whereas the right is a yes-instance of HAMP (take $(a, v_x, 1, b, 3, 3)$).

2. After many hours of hard work, Prof. P. E. Kwalsenpee has managed to create an algorithm `TSP-Decider` that can solve the decision variant of TSP in polynomial time. That is, given a set $S = \{s_i\}_{i=1}^{n}$ of cities, a $n \times n$ intercity distance matrix $D = [d_{i,j}]$ where $d_{i,j} \in \mathbb{Z}^+$ and a positive integer k, the algorithm will accept if and only if there is a tour through all vertices in S s.t. the total cost of the tour is $\leq k$.

(a) Explain how we can also use `TSP-Decider` to decide the HAMILTONIAN CIRCUIT problem in polynomial time.

Solution: Given an instance $G = (V, E)$ of HAMC. Create a city s_i for every node $v_i \in V$. Construct D such that the distance between two cities is 1 if the corresponding nodes in G are directly connected, otherwise, let the distance 2. Now take $k = |V|$ and run the algorithm. If the algorithm accepts, accept, otherwise reject. Clearly such an instance of TSP can be constructed in polynomial time and since `TSP-Decider` runs in polynomial time on this instance, HAMILTONIAN CIRCUIT can be decided in polynomial time.

- 1pt for correctly modeling S
- 2pt for correctly modeling D.
- 1pt for correctly modeling k.

(b) Given this decision algorithm by the professor and an instance (S, D, k) of TSP, give a polynomial time algorithm that returns an actual TSP-tour through the cities with total cost k, if such a tour exists. If not, the algorithm should return `nil`.

Solution:

if $\text{TSP-DECIDER}(S, D, k)$ rejects **then**
  return nil
else ▷ A solution exists, now we need to find it
  for $d_{i,j} \in D$ **do**
    $t \leftarrow d_{i,j}$; $d_{i,j} \leftarrow k + 1$
    if $\text{TSP-DECIDER}(S, D, k)$ rejects **then** ▷ That entry is crucial
      $d_{i,j} \leftarrow t$
    end if
  end for
end if
return D ▷ The path is now indicated by all entries with value $\leq k$.

3. An important problem in analysing social networks is the question whether there are persons in the network who occur in more than one large clique of the network. Therefore, we consider the following problem:

Name IMPORTANT NODE

Instance an undirected graph $G = (V, E)$, and two positive integers $K > 1$ and $L > 1$.

Question do there exist at least K nodes $v \in V$ such that v occurs in at least two different cliques of G, each containing at least L nodes?

Show that IMPORTANT NODE is an NP problem.

Solution: Without loss of generality we can assume $K, L \leq |V| = n$ (else, such an instance is a no-instance).
Guess a set V' of nodes and a set $\{C_1, C_2, \dots C_m\}$ of subsets of nodes with $m \leq 2n$. (You may also read *guess* as *given*. The first is a formulation to solve the problem on a non-deterministic machine, the second is a formulation to verity a solution on a deterministic machine. Both work for the purposes of showing the problem is in NP.)
Then verify

(a) $V' \subseteq V$: Time cost $O(n^2)$.

(b) $|V'| \geq K$: Time cost $O(n + \log n) = O(n)$.

(c) each $C_i \subseteq V$: Time cost $O(n^3)$.

(d) each C_i is a clique in G: Time cost $O(n.n^2.|E|) = O(n^3.|E|)$.

(e) each C_i has at least L nodes:Time cost $O(n.(n + \log n)) = O(n^2)$.

(f) for each $v \in V'$ there exist two different cliques C_i and C_j such that $v \in C_i$ and $v \in C_j$: Time cost $O(2n.n) = O(n^2)$ per node v, so total time cost $O(n^3)$.

Hence, the total verification costs of a yes instance $(G = (V, E), K, L)$ are $O(|input|^4)$-time. Therefore, IMPORTANT NODE is an **NP** problem.