CSE340: Theory of Computation (Final Exam – Part 2)

21st November, 2021

Subjective Questions **Total Points: 55**

Question 1. (10 points) Design the context-free grammar (CFG) for the following language,

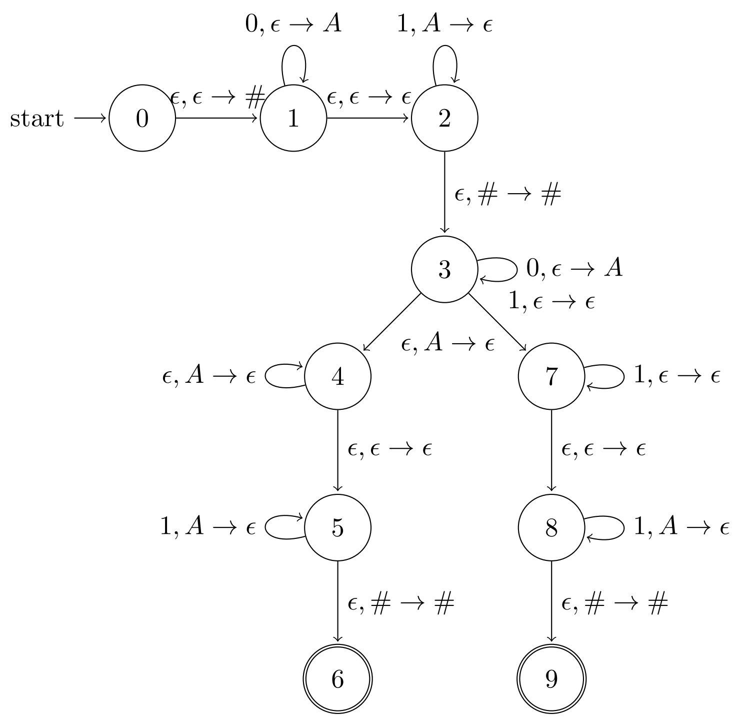
$$L = \{w \in \{a, b\}^* \mid w \text{ has more } a\text{'s than } b\text{'s}\}$$

Solution:

$$\begin{aligned} S &\longrightarrow STA \mid TS \mid aA \\ T &\longrightarrow aTb \mid bTa \mid TT \mid \epsilon \\ A &\longrightarrow aA \mid \epsilon \end{aligned}$$

Question 2. (10 points) Design a PDA for the language $L = \{0^p 1^q 0^r 1^s \mid p = q \text{ and } r \neq s\}$

Solution:



Question 3. (10 points) An *weight function* in a graph $G = (V, E)$ is a function $w : E \to \mathbb{N}$ that assigns a non-negative integer to every edge in G. A graph with a weight function is called a weighted graph. The *weight* of a walk in G is the sum of the weights of the edges in the walk.

$$\mathsf{TSP} = \{\langle G, w, B\rangle \mid G \text{ is a weighted directed graph with weight function } w \text{ and } G \text{ has a walk of weight at most } B\text{, that visits every vertex in } G \text{ and returns to the starting vertex}\}$$

Prove that TSP is NP-complete.

(Hint: You can give a reduction from the HamPath problem.)

Solution:

Showing that TSP is in NP

Certificate: A sequence of vertices $\sigma = \langle u_1, u_2, \ldots, u_k\rangle$

Verifier's Algorithm:

Input: $\langle G = (V, E), w, B, \sigma\rangle$

1. Check if the sequence σ forms a walk in G. If not then Reject.
2. Check if every vertex in G is part of σ. If not then Reject.
3. Check if $u_1 = u_k$. If not then Reject.
4. Check if $\sum_{i=1}^{k-1} w(u_i, u_{i+1}) \leq B$. If not then Reject, else Accept.

The above algorithm clearly runs in polynomial time.

Choosing a suitable NP-complete problem

Consider the problem HamPath discussed in lectures.

$$\mathsf{HamPath} = \{\langle G, s, t\rangle \mid G \text{ is a directed graph having a Hamiltonian path from } s \text{ to } t\}.$$

We will show that

$$\mathsf{HamPath} \leq_p \mathsf{TSP}.$$

The reduction function f

Input: An instance of HamPath, $\langle G = (V, E), s, t\rangle$

1. Construct graph $H = (V_H, E_H)$ as follows: $V_H = V \cup \{x\}$, $E_H = E \cup \{(x, s), (t, x)\}$. That is, we add a vertex x and the directed edges (x, s) and (t, x) to the graph G to obtain H.
2. For every edge $e \in E_H$, set $w(e) = 1$.
3. Set $B = |V| + 1$.

Output: $\langle H, w, B\rangle$

Time complexity of the reduction

The construction of $\langle H, w, B\rangle$ can be achieved in linear time in the size of G.

Proof of correctness

If there is a Hamiltonian path P in G from s to t, then adding the edges (x, s) and (t, x) to P, creates a closed walk of length $|V| + 1$ in H that visits every vertex in H.

On the other hand, suppose there is a walk C of weight at most $|V| + 1$ in H, that visits every vertex in H and returns to the starting vertex. Firstly C must visit every vertex exactly once since in H there are exactly $|V| + 1$ vertices. This implies that C is a cycle. Secondly C must include the edges (x, s) and (t, x), since these are the only two edges incident on the vertex x. Now if we remove (x, s) and (t, x) from C, we get a path P' from s to t, of length $|V| - 1$, that visits every vertex in V. Hence this is a Hamiltonian path in G.

Question 4. Consider the following two languages

$$\begin{array}{rcl} L_1 & = & \{\langle M_1, M_2 \rangle \mid M_1, M_2 \text{ are two TMs and } L(M_1) \subseteq L(M_2)\} \\ L_2 & = & \{\langle D_1, D_2 \rangle \mid D_1, D_2 \text{ are two DFAs and } L(D_1) \subseteq L(D_2)\} \end{array}$$

(a) (7 points) Show that L_1 is undecidable.

Solution: Consider the undecidable language

$$\mathsf{E_{TM}} = \{\langle M \rangle \mid M \text{ is a TM and } L(M) = \emptyset\}.$$

Claim 1. $\mathsf{E_{TM}} \leq_m L_1$.

We will construct a computable function f that takes as input $\langle M \rangle$ and produces an output $\langle M_1, M_2 \rangle$ such that $L(M) = \emptyset \iff L(M_1) \subseteq L(M_2)$.

The reduction function f

Input: $\langle M \rangle$

1. Set $M_1 := M$.
2. Set M_2 to be a TM that rejects all strings ($L(M_2) = \emptyset$).

Output: $\langle M_1, M_2 \rangle$

Proof of correctness

Now,

$$L(M) = \emptyset \iff L(M_1) = \emptyset \iff L(M_1) \subseteq L(M_2)$$

Therefore, $\mathsf{E_{TM}} \leq_m L_1$. This proves that L_1 is undecidable.

(b) (6 points) Show that L_2 is decidable.

Solution:

Algorithm for L_2

Input: $\langle D_1, D_2 \rangle$

1. Construct a DFA D_3, such that $L(D_3) = L(D_1) \cap L(D_2)$.
   (This is possible since we know that regular languages are closed under intersection.)
2. Apply the algorithm for EQ_{DFA} (discussed in lectures), to the input $\langle D_1, D_3 \rangle$ and Output the answer.

Proof of correctness

Now,

$$L(D_1) \subseteq L(D_2) \iff L(D_1) \cap L(D_2) = L(D_1)$$

This is precisely what we check in the above algorithm.

(Note: One can also use union or set difference instead of intersection to design an algorithm for L_2.)

Question 5. In this question, for two languages L_1 and L_2, if $L_1 \leq_m L_2$ then we denote it by $L_1 \longrightarrow L_2$. Consider the relation between the languages given by the following diagram.

$$\begin{tikzcd}
 & \overline{B} \arrow[dl] \arrow[r] & D & \\
A_{TM} \arrow[r] & A & \overline{A} \arrow[r] & B \arrow[ul] \\
 & C \arrow[ul] \arrow[r] & \overline{A_{TM}} \arrow[ur] &
\end{tikzcd}$$

Consider the following 4 class of languages:

- Dec: decidable languages
- TR': undecidable but Turing recognizable languages
- coTR': undecidable but co-Turing recognizable languages
- NTR: neither Turing recognizable nor co-Turing recognizable languages

Now for each of the following languages, mention which of the above class does it belong to by giving proper justification.

(a) (3 points) A

> **Solution:** TR'
> Since $A_{TM} \leq_m A$, therefore A is undecidable. Moreover since $\overline{B} \leq_m A_{TM}$, B is co-Turing recognizable. Also we have $\overline{A} \leq B$. Hence A is Turing recognizable. Therefore A is in TR'.

(b) (3 points) B

> **Solution:** coTR'
> Since $\overline{A_{TM}} \leq_m B$, therefore B is undecidable. Moreover since $\overline{B} \leq_m A_{TM}$, B is co-Turing recognizable. Therefore B is in coTR'.

(c) (3 points) C

> **Solution:** Dec
> Since $C \leq_m A_{TM}$ and $C \leq_m \overline{A_{TM}}$, it is both Turing recognizable and co-Turing recognizable. Hence C is in Dec.

(d) (3 points) D

> **Solution:** NTR
> Since B is undecidable and both $B \leq_m D$ and $\overline{B} \leq_m D$, therefore D is in NTR.