Computability and Complexity

Department of Computer Science, University of Copenhagen,
Jakob Grue Simonsen & Christian Wulff-Nilsen

Written exam for 4 hours, April 12th, 2018

All usual aids such as books, notes, exercises, and the like may be used during the exam, but no calculators, computers, cell phones, or similar equipment. The exam must be answered in Danish or English and may be answered in pencil.

This exam has 5 pages and consists of 4 questions, each of approximately the same weight.

Question 1

For each of the following four languages L_1, L_2, L_3, and L_4, determine whether it is regular, context-free, or neither. Provide appropriate proofs.

All languages have alphabet $\Sigma = \{0, 1\}$. $\mathbb{N} = \{1, 2, 3, \ldots\}$.

Part 1.1 $L_1 = \{01^n0 : n \in \mathbb{N}\}$.

Part 1.2 $L_2 = \{01^n01^{n+1}0 : n \in \mathbb{N}\}$.

Part 1.3 $L_3 = \{01^n01^{n+1}01^{n+2}0 : n \in \mathbb{N}\}$.

Part 1.4 Let the *parity* of a position p in a string w over Σ be the number of zeroes occurring in w at or before p minus the number of ones occurring in w before p).

For example if $w = 0010$, the parities at each position $p \in \{0, 1, 2, 3\}$ are

- $p = 0 : 1$
- $p = 1 : 2.$
- $p = 2 : 1.$
- $p = 3 : 2.$

The language L_4 is defined as

$L_4 = \{w : \text{all positions in } w \text{ have non-negative parity, and the last position in } w \text{ has parity } 0\}$

Thus, $0010 \notin L_4$ (last position has parity 2), $01100 \notin L_4$ (position 2 has parity -1). But $001101 \in L_4$, and $0101001011 \in L_4$.

[Hint: Find a simpler description of L_4.]

Question 2

Let FASTER_{TM} be the problem of determining, for a given pair of deterministic TMs M_1 and M_2, if on every input string, M_1 executes for at most as many steps as M_2. We do not require the TMs to halt on all inputs: if M_2 does not halt on a string w, we say that M_1 executes for at most as many steps as M_2 on w regardless of whether M_1 halts on w or not. Also, we do not require M_1 and M_2 to recognize the same language.

Part 2.1 Formulate FASTER_{TM} as a language and show that it is co-Turing-recognizable. [Hint: for a pair $\langle M_1, M_2\rangle$, consider simulating M_1 and M_2 for up to i steps on the jth string over all pairs of positive integers (i, j).]

Part 2.2 Is FASTER_{TM} Turing-recognizable? Explain your answer.

Part 2.3 Consider the language FASTER_{DFA} which is defined as FASTER_{TM} but with pairs of DFAs instead of pairs of deterministic TMs.

Is FASTER_{DFA} decidable? Explain your answer.

Question 3

In the following, we consider only finite non-empty sets and finite collections of sets.

Given a collection $\mathcal{C}$ of sets, a *hitting set* for $\mathcal{C}$ is a set S that intersects each set in $\mathcal{C}$, i.e., $S \cap C \neq \emptyset$ for each $C \in \mathcal{C}$.

We will consider a stronger property than this. A 2-*hitting set* for $\mathcal{C}$ is a set S that intersects at least two elements in each set C of $\mathcal{C}$, unless C has fewer than two elements, in which case S intersects all elements of C. In short, $|S \cap C| \geq \min\{|C|, 2\}$ for each $C \in \mathcal{C}$.

Let $2{-}\mathit{HITTINGSET}$ be the language

$$2{-}\mathit{HITTINGSET} = \{\langle \mathcal{C}, K\rangle | \mathcal{C} \text{ is a collection of sets and there is a 2-hitting set } S \text{ for } \mathcal{C} \text{ of size } K\}.$$

In the definition of this language, we make the standard assumption that numbers are represented in binary and we assume an encoding of $\mathcal{C}$ requiring only polynomial size in the total size of all sets of $\mathcal{C}$. Feel free to pick any such encoding. For instance, if there are k elements in the union of all sets, each element can be represented with a unique number requiring $\Theta(\log k)$ bits; each set is represented as the concatenation of these numbers with a special $\Theta(\log k)$ bit separator symbol after the last element of the set.

Part 3.1 Show that $2{-}\mathit{HITTINGSET}$ belongs to NP.

Part 3.2 Show that $2{-}\mathit{HITTINGSET}$ is NP-complete. [Hint: consider $\mathit{VERTEX} - \mathit{COVER}$, and in the reduction introduce sets $\{u, v, e\}$ and $\{e\}$ for each edge e.]

Part 3.3 Consider the related language $(2, 10){-}\mathit{HITTINGSET}$, defined as

$$(2,10){-}\mathit{HITTINGSET} = \{\langle \mathcal{C}, K\rangle | \mathcal{C} \text{ is a collection of sets and there is a 2-hitting set } S \text{ for } \mathcal{C}' \text{ for some } \mathcal{C}' \subseteq C \text{ with } |S| = K \text{ and } |\mathcal{C}'| = 10\}.$$

Show that $(2, 10){-}\mathit{HITTINGSET}$ belongs to P. [Hint: Argue that each $\mathcal{C}'$ has a 2-hitting set of size at most 20. Consider the two cases $K \geq 20$ and $K < 20$ separately.]

Question 4

Recall that a directed graph G is *acyclic* (and called a *DAG*) if there are no directed paths in G that start and end at the same vertex.

Consider the following problem:

$$\text{PATH}_{\text{acyclic}} = \{(G, s, t) : \text{G is a } DAG \text{ and there is a directed path from } s \text{ to } t \text{ in } G\}$$

Part 4.1 Prove that $\text{PATH}_{\text{acyclic}}$ is NL-complete.

[Hint: Given a directed graph G and nodes s and t, find some way of creating a (possibly larger) DAG G' and nodes s' and t' in G' with nice properties.]

Part 4.2 A directed graph G is said to be *out-deterministic* if, for each vertex u, there is at most one vertex v such that (u, v) is an edge of G (i.e., there is at most one outgoing edge from u). Define

$$\text{OD-PATH} = \{(G, s, t) : G \textit{ is an out-deterministic directed graph, and there is a path from } s \text{ to } t \text{ in } G\}$$

Prove that OD-PATH is in L.

END OF EXAM SET