

Challenge Problem I

This problem is **not worth any extra credit**, we do not expect you to solve it, as it is harder than any graded problem we will give you in this class. However, please feel free to attempt it! If you present your solution to one of the Head TAs or the Instructors during office hours, we will give you a **sticker and candy!**

You can come to office hours to work on this problem, with friends or by yourself! The teaching staff will be happy to help you (but not too much)!!

You only need to use material taught in class up to the end of the Myhill-Nerode section (beginning of lecture 5). You should not use new tools we teach later, as it betrays the spirit of the problem.

1 The Sphinx's Riddle

Alone in desert, the Sphinx decides that she has grown tired of asking travelers the standard riddles that they always seem to have already heard. One day, a Traveler with a penchant for theoretical computer science comes by, and asks for a special riddle. Thinking fondly of her time in CS Theory, she asks: "Traveler, do you recall the Myhill-Nerode theorem?"

"Of course," the Traveler beams, equally as fond of their time in CST.

"Then, you will like this. Fix language $L \subseteq \Sigma^*$. Recall that for a string x , $[x]_L$ is the Myhill-Nerode equivalence class containing x . Now, define the set $S_{L,n} = \{[x]_L : x \in \Sigma^* \text{ and } |x| \leq n\}$."

"I see! This set captures the set of equivalence classes that contain at least one string of length at most n ," the Traveler exclaims.

"Exactly. Now, define the set $D_{L,n} = S_{L,n} \setminus S_{L,n-1}$."

"Ah, and this set represents a "frontier" of equivalence classes that contain at least one string of length n , but no strings of length at most $n - 1$."

"Now I want you to consider the functions $s_L(n) = |S_{L,n}|$ and $d_L(n) = |D_{L,n}|$."

"Okay, so the function s_L charts the 'growth' of the set of equivalence classes as the maximum string length increases and d_L gives us the 'speed' of this growth."

"Wonderful, you understand these functions well! Now, to win my coveted treasure, you must answer three questions..."

"Very well," the Traveler responds, eager.

The eyes of the Sphinx glow as she enumerates her queries:

- (1) Let there be some N such that for every $n > N$, $d_L(n) = 0$, can we conclude L is regular?
- (2) Let us say for some constant k where for every n , $d_L(n) \leq k$, is L regular?
- (3) Let there be some N such that for every $n > N$, $s_L(n) \leq \sqrt{n}$, is L regular?

As the Sphinx closes her eyes, satisfied, the traveler gets to work.

2 The Sphinx's Riddle (formally)

Fix language $L \subseteq \Sigma^*$. Recall that for a string x , $[x]_L$ is the Myhill-Nerode equivalence class containing x .

Definition 1. We say the growth set of L for length n is $S_{L,n} = \{[x]_L : x \in \Sigma^* \text{ and } |x| \leq n\}$ and the difference set of L for length n is $D_{L,n} = S_{L,n} \setminus S_{L,n-1}$.

Definition 2. We say the growth function of L is $s_L(n) = |S_{L,n}|$ and the difference function of L is $d_L(n) = |D_{L,n}|$.

Q1. Prove or Disprove: If $\exists N$ such that for all $n > N$, $d_L(n) = 0$, then L is regular.

Q2. Prove or Disprove: If there is some k where for every n , $d_L(n) \leq k$, then L is regular.

Q3. Prove or Disprove: If $\exists N > 0$ such that $\forall n > N$, $s_L(n) \leq \sqrt{n}$, then L is regular.