On the Gap Between Separating Words and Separating Their Reversals

Farzam Ebrahimnejad

Combinatorics, Automata and Number Theory – CANT, December 2016

Department of Computer Engineering,
Sharif University of Technology

Introduction

Definition

- For a DFA D, We denote the set of states of D by Q_D.
- For a state $q \in Q_D$, we define $\delta_D(q, w)$ to be the state in Q_D at which we end if we start reading w from q.
- Assuming the start state of D is $q_0 \in Q_D$, we define $\delta_D(w) = \delta_D(q_0, w)$.

Definition

We say a DFA D *separates* two distinct words $w, x \in \Sigma^*$, if it accepts w but rejects x. Furthermore, we let $\mathrm{sep}(w, x)$ be the minimum number of states required for a DFA to separate w and x.

Introduction

Definition

We say a DFA D *separates* two distinct words $w, x \in \Sigma^*$, if it accepts w but rejects x. Furthermore, we let $\mathrm{sep}(w, x)$ be the minimum number of states required for a DFA to separate w and x.

Remark

- *If D separates w and x, then $\delta_D(w) \neq \delta_D(x)$.*
- $\mathrm{sep}(w, x) = \mathrm{sep}(x, w)$.

The Separating Words Problem

- Good upper and lower bounds on

$$S(n) := \max_{w \neq x \wedge |w|,|x| \leq n} \operatorname{sep}(w, x)?$$

- $S(n) = O(n^{2/5} (\log n)^{3/5})$
- $S(n) = \Omega(\log n)$

In 2011, Demaine, Eisenstat, Shallit, and Wilson published a paper titled “Remarks on Separating Words” that surveys the latest results about this problem, and while proving several new theorems, they also introduced three new open problems.

The Problem

Must $\mathrm{sep}(w^R, x^R) = \mathrm{sep}(w, x)$? No, for $w = 1000, x = 0010$, we have

$$\mathrm{sep}(w, x) = 3$$

but

$$\mathrm{sep}(w^R, x^R) = 2.$$

The Problem

Problem

Is $\left|\mathrm{sep}(w,x) - \mathrm{sep}(w^R, x^R)\right|$ *unbounded?*

The Proof

Preliminaries

Lemma

$\forall u, v, w, x \in \Sigma^* : \text{sep}(uwv, uxv) \geq \text{sep}(w, x)$

Proof.

1. $\text{sep}(wv, xv) \geq \text{sep}(w, x)$
2. $\text{sep}(uwv, uxv) \geq \text{sep}(wv, xv)$

□

Preliminaries

Lemma

$\forall u, v, w, x \in \Sigma^* : \mathrm{sep}(uwv, uxv) \geq \mathrm{sep}(w, x)$

Proof.

1. $\mathrm{sep}(wv, xv) \geq \mathrm{sep}(w, x)$
2. $\mathrm{sep}(uwv, uxv) \geq \mathrm{sep}(wv, xv)$

□

- Must $\mathrm{sep}(uwv, uxv) = \mathrm{sep}(w, x)$?

Preliminaries

Lemma

$\forall u, v, w, x \in \Sigma^* : \mathrm{sep}(uwv, uxv) \geq \mathrm{sep}(w, x)$

Proof.

1. $\mathrm{sep}(wv, xv) \geq \mathrm{sep}(w, x)$
2. $\mathrm{sep}(uwv, uxv) \geq \mathrm{sep}(wv, xv)$

□

- Must $\mathrm{sep}(uwv, uxv) = \mathrm{sep}(w, x)$? No, we have $\mathrm{sep}(100, 001) = 2$ but $\mathrm{sep}(1000, 0010) = 3$.

Preliminaries

- Let $f_n = 0^n, g_n = 0^{n+(2n+1)!}$.
- $\mathrm{sep}(f_n, g_n) = n + 2$.

Preliminaries

- Let $f_n = 0^n, g_n = 0^{n+(2n+1)!}$.
- $\mathrm{sep}(f_n, g_n) = n + 2$.
- By the previous lemma, we have

$$\mathrm{sep}(uf_nv, ug_nv) \geq n + 2.$$

The Basic Idea

For all $k \in \mathbb{N}$, we will construct two words

$$w_k = u_k f_n v_k,$$

$$x_k = u_k g_n v_k,$$

for some $n \in \mathbb{N}$ and $u_k, v_k \in \{0, 1, 2\}^+$, s.t.

$$\mathrm{sep}(w, x) - \mathrm{sep}(w^R, x^R)$$

approaches infinity as k approaches infinity.

The Function C_n

Lemma

$\forall n \in \mathbb{N}, w_0 \in \Sigma^+ : \exists w \in w_0(0^+w_0)^*$ *s.t.*

$$\mathrm{sep}(wf_nw, wg_nw) \geq 2n + 2.$$

We denote the w corresponding to w_0 by $C_n(w_0)$.

The Regular Language G_k

Definition

For $k \in \mathbb{N}$, we define

$$
\begin{aligned}
L_k := \{\, 1^{2i}2 \mid i \in \mathbb{N} \wedge i \leq k \,\} \\
\cup \{\, 1^{i_1}21^{i_2}2 \cdots 21^{i_{s-1}}21^{i_s}2 \mid \; & s, i_1, \ldots, i_s \in \mathbb{N} \\
& \wedge i_1 + i_2 + \cdots + i_s = 2k + 1 \\
& \wedge i_1, i_2, \ldots, i_{s-1} \equiv 0 \pmod 2 \,\}.
\end{aligned}
$$

Also, we define $G_k := L_k^*$.

The Regular Language G_k

Definition

For a regular language $L \subseteq \Sigma^*$, we define $\mathrm{sc}(L)$ to be the minimum number of states needed for a DFA to accept L.

The Regular Language G_k

Definition

For a regular language $L \subseteq \Sigma^*$, we define $\mathrm{sc}(L)$ to be the minimum number of states needed for a DFA to accept L.

Lemma

$\forall k \in \mathbb{N}$:

$$\mathrm{sc}(G_k) \geq 2^k \wedge \mathrm{sc}(G_k^R) \leq 5k + 3$$

The Regular Language G_k

Definition

For a regular language $L \subseteq \Sigma^*$, we define $\mathrm{sc}(L)$ to be the minimum number of states needed for a DFA to accept L.

Lemma

$\forall k \in \mathbb{N}$:

$$\mathrm{sc}(G_k) \geq 2^k \wedge \mathrm{sc}(G_k^R) \leq 5k + 3$$

Lemma

$\exists z_k \in (G_k - \{ \epsilon \})$ *s.t.*

$$\mathrm{sep}(z_k, \{ 1, 2 \}^* - G_k) \geq 2^k$$

Mapping $\{0,1,2\}^*$ to $\{0,1\}^*$

Definition

$\mathrm{tl} : \{0,1,2\}^* \longrightarrow \{0,1\}^*$

- $0 \to 0$
- $1 \to 11$
- $2 \to 01$

Mapping $\{0,1,2\}^*$ to $\{0,1\}^*$

Lemma

$\mathrm{sep}(\mathrm{tl}(w), \mathrm{tl}(x)) \geq \mathrm{sep}(w, x)$

Proof.

Let D be a DFA that separates $\mathrm{tl}(w)$ and $\mathrm{tl}(x)$. We construct a new DFA E with $Q_E = Q_D$ that separates w and x. For all $q \in Q_E$, we set

$$\delta_E(q, 0) = \delta_D(q, 0), \delta_E(q, 1) = \delta_D(q, 11), \delta_E(q, 2) = \delta_D(q, 01).$$

□

The Main Result

Theorem

For all $k, n \in \mathbb{N}$, there exist two distinct words $w, x \in \{0, 1\}^*$ such that

$$\mathrm{sep}(w, x) \geq \min(2n + 2, 2^{k/2}),$$

but

$$\mathrm{sep}(w^R, x^R) \leq n + 10k + 10.$$

The Main Result

Theorem

The difference

$$\left|\operatorname{sep}(w, x) - \operatorname{sep}(w^R, x^R)\right|$$

is unbounded.

Proof.

Let $k \in \mathbb{N}$. We set $n = 2^{k/2-1} - 1$. By the previous theorem, there exist $w, x \in \Sigma^*$ such that

$$\mathrm{sep}(w, x) \geq \min(2n + 2, 2^{k/2}) = 2^{k/2},$$

and

$$\mathrm{sep}(w^R, x^R) \leq n + 10k + 10 = (2^{k/2-1} - 1) + 10k + 10.$$

Hence

$$\begin{aligned} \mathrm{sep}(w, x) - \mathrm{sep}(w^R, x^R) &\geq 2^{k/2} - \left(2^{k/2-1} + 10k + 9\right) \\ &= 2^{k/2-1} - 10k - 9, \end{aligned}$$

which tends to infinity as k tends to infinity. $\square$

Conclusion

Open Problem

Is

$$\operatorname{sep}(w, x) / \operatorname{sep}(w^R, x^R)$$

unbounded?

Questions?

Thank you!