Prefix-Free Subsets of Regular Languages and Descriptional Complexity

Jozef Jirásek
Juraj Šebej

DCFS 2015

Outline

- Maximal prefix-free subsets
- Properties
- Constructing
- Subsets of certain properties
- State complexity
- Non-regular MPFS

Definitions

- DFA $A = (Q, \Sigma, \delta, s, F)$
- We consider incomplete, trim DFAs.

Definitions

- DFA $A = (Q, \Sigma, \delta, s, F)$
- We consider incomplete, trim DFAs.

- $u \leq_p v$ iff $\exists w: uw = v$
- $u <_p v$ iff $u \leq_p v$ and $u \neq v$ $\qquad (w \neq \varepsilon)$
- Set P is *prefix-free* iff $\nexists u, v \in P: u \leq_p v$
- Set P is a *maximal prefix-free subset* of L iff:
  - $P \subseteq L$
  - P is prefix-free
  - $\forall u \in L: \exists v \in P: u \leq_p v$ or $v \leq_p u$

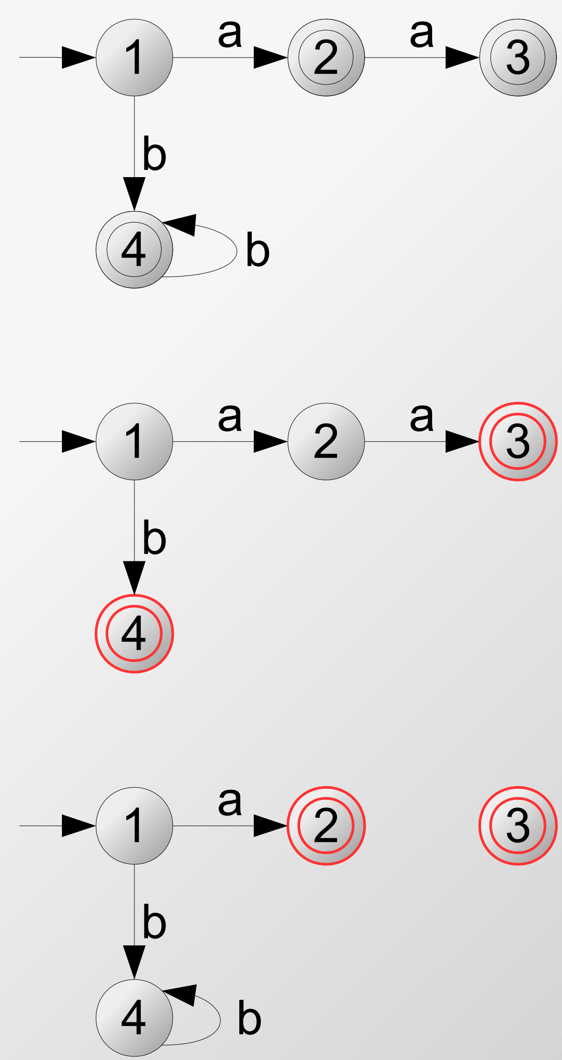
Constructing a MPFS

- Let $A = (Q, \Sigma, \delta, s, F)$; $F' \subseteq F$
- Construct $A_{F'} = (Q, \Sigma, \delta', s, F')$
  - $\delta'(q, a)$ undefined if $q \in F'$
  - as in A otherwise.

- $L(A_{F'})$ is a PFS of $L(A)$
- If $\forall q \in F \setminus F'$ reachable in $A_{F'}$:

  $\exists w$ accepted from q in $A_{F'}$

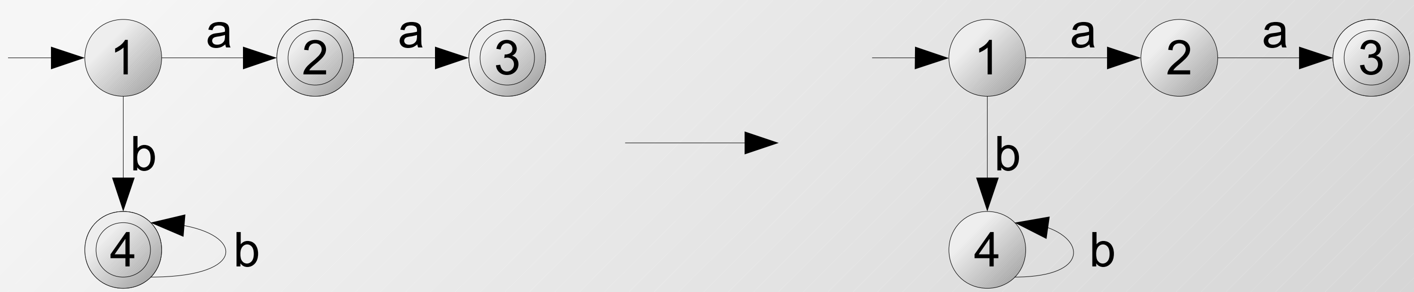
  then $L(A_{F'})$ is a <u>M</u>PFS of $L(A)$.



Constructing some PFS

...remove all strings which are proper prefixes?

- $L_1 = L \setminus \{w \in L \mid \exists v \neq \varepsilon: wv \in L\}$
- $A_1 = (Q, \Sigma, \delta, s, F_1)$
  - $F_1 = \{q \in F \mid \nexists w \neq \varepsilon$ accepted from $q\}$



Results

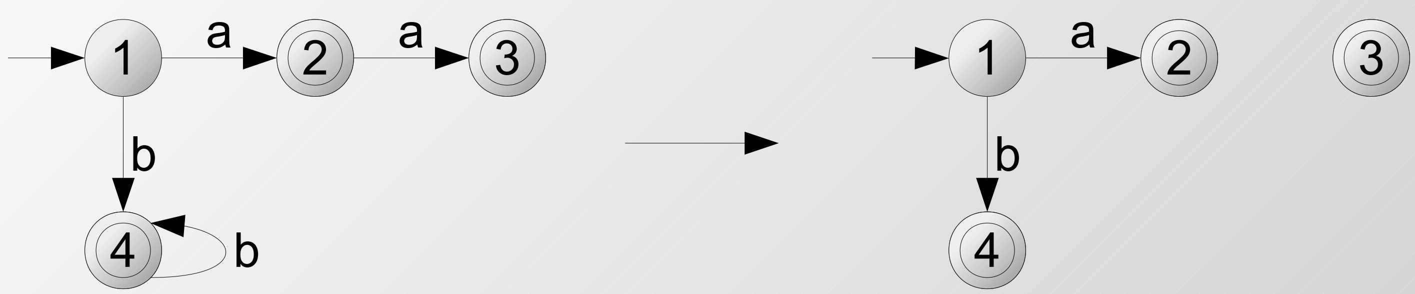
...remove all strings which are proper prefixes?

- $\mathrm{sc}(L_1) \leq \mathrm{sc}(L)$
  - L unary: $\mathrm{sc}(L_1) = 1$ or $\mathrm{sc}(L_1) = \mathrm{n}$
  - $\forall\ 1 \leq k \leq n$ there is a binary L *with* $\mathrm{sc}(L) = n$ and $\mathrm{sc}(L_1) = k$
- L_1 is a PFS of L.
- If L is finite, L_1 is a *largest* MPFS of L.
- Otherwise, ☹

Constructing some PFS

...remove all strings which *have a* proper prefix?

- $L_2 = L \setminus \{w \in L \mid \exists u \in L, v \neq \varepsilon : w = uv\}$
- $A_2 = (Q, \Sigma, \delta_2, s, F)$
  - δ_2 undefined for $q \in F$, as in A otherwise.



Results

...remove all strings which *have a* proper prefix?

- $\mathrm{sc}(L_2) \leq \mathrm{sc}(L) + 1$
  - $\forall\ 1 \leq k \leq n + 1$
    there is a unary L *with* $\mathrm{sc}(L) = n$ and $\mathrm{sc}(L_2) = k$
- L_2 is a MPFS of L.
- If L_2 is infinite, L does not have any finite MPFS.
- If L_2 is finite, it is a *smallest* MPFS of L.

Finding a largest finite MPFS

L has both finite and infinite MPFS. We can find the smallest (L_2), can we find the largest finite MPFS?

Finding a largest finite MPFS

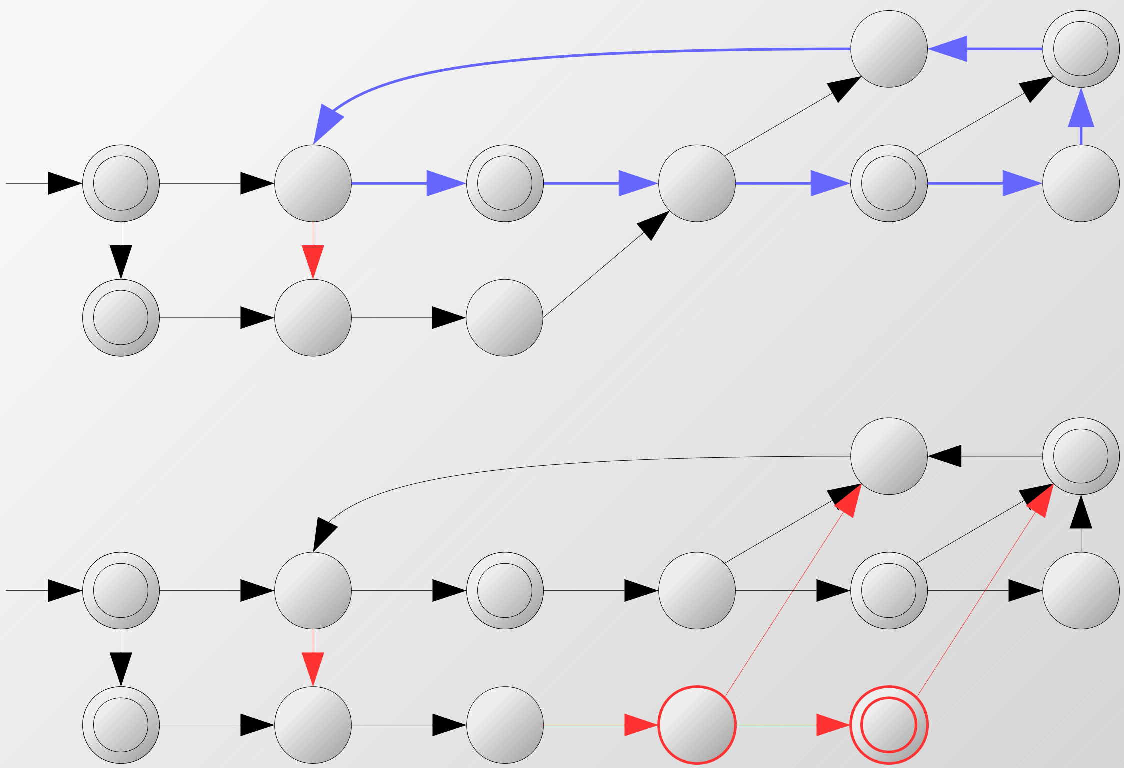
L has both finite and infinite MPFS. We can find the smallest (L_2), can we find the largest finite MPFS?

- Yes we can!
- Polynomial-time algorithm which either:
  - finds $F' \subseteq F$, such that $L(A_{F'})$ is a largest finite MPFS of L; or
  - determines that no largest finite MPFS of L exists.
    - Only infinite MPFS.
    - Finite MPFS of unlimited size.
- $sc(L(A_{F'})) \leq sc(L) + 1$
  - Reached for every $1 \leq k \leq n + 1$ on unary languages.
- Construction in paper.

Infinite MPFS

Task: Find an infinite regular MPFS of $L = L(A)$.

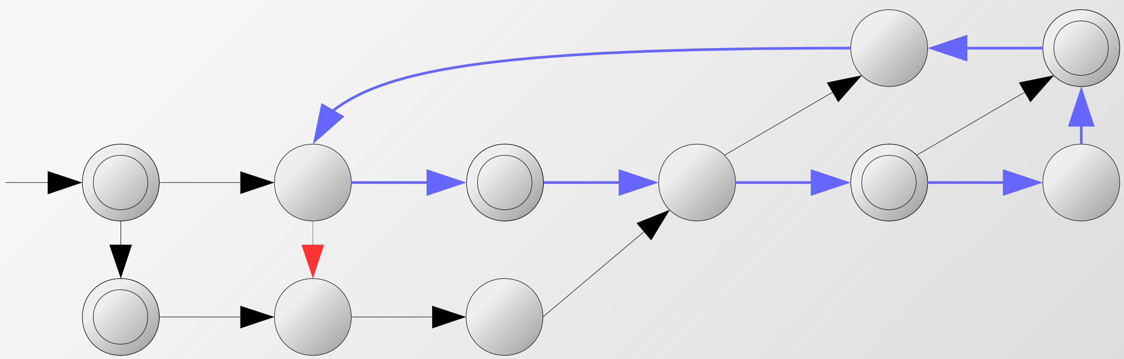
- L infinite $\Rightarrow$ A contains a cycle
- (1) transition out of the cycle:

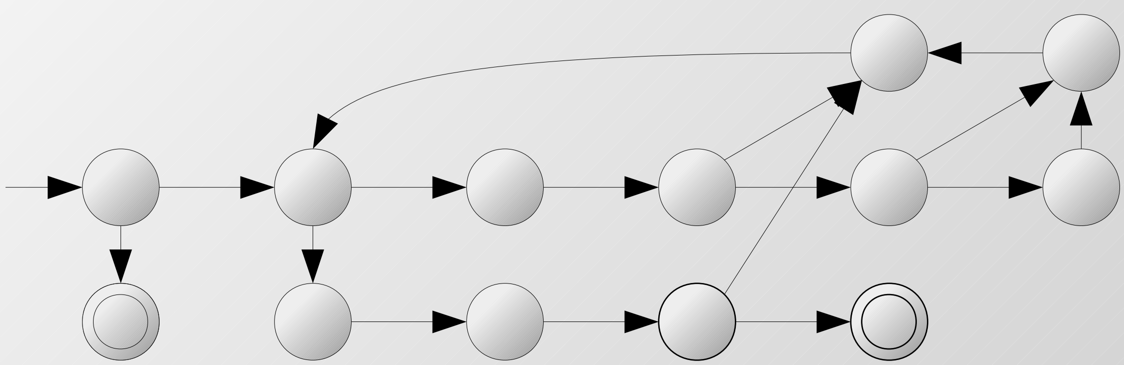


Infinite MPFS

Task: Find an infinite regular MPFS of $L = L(A)$.

- L infinite $\Rightarrow$ A contains a cycle
- (1) transition out of the cycle:





Infinite MPFS

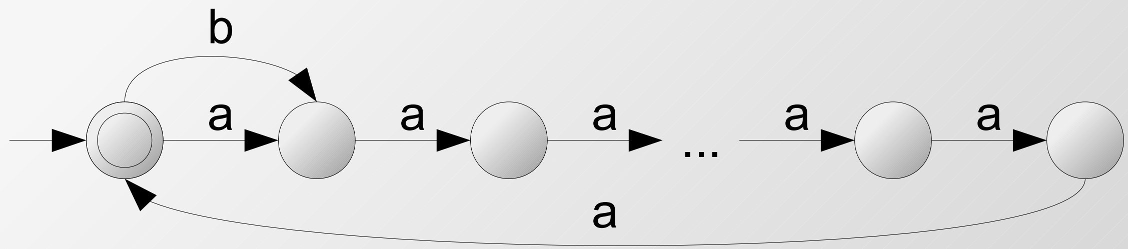
Task: Find an infinite regular MPFS of $L = L(A)$.

- L infinite $\Rightarrow$ A contains a cycle
- (2) no transition out of any cycle:

Only finite prefix-free sets!

State complexity

- For every $n \geq 2$, there exists a language L with $sc(L) = n$, such that every infinite MPFS of L has $sc \geq 2n$.



- For every $n \geq 4$, there exists a language L with $sc(L) = n$, which has infinite MPFS of unlimited sc.

$$L = \mathtt{b}^+\mathtt{a} \cup \{\mathtt{a}^i \mid i \geq n - 3\}$$

$$P_k = \mathtt{b}^+\mathtt{a} \cup \{\mathtt{a}^k\}$$

$$sc(P_k) = k$$

Non-regular MPFS of regular languages

Regular language L with a non-regular MPFS:

- $L = a^*b^*$
- $P = \{a^n b^n \mid n > 0\}$

Non-regular MPFS of regular languages

L has a non-regular MPFS ⇒ *L* has uncountably many MPFS.

- Let P be a non-regular MPFS of a regular language L.
  - $P^{+} = \{w \in P \mid \exists u_w \in L\colon w <_p u_w\}$
  - $P^{-} = \{w \in P \mid \nexists u \in L\colon w <_p u\}$

Non-regular MPFS of regular languages

L has a non-regular MPFS ⇒ *L* has uncountably many MPFS.

- Let P be a non-regular MPFS of a regular language L.
  - $P^+ = \{w \in P \mid \exists u_w \in L: w <_p u_w\}$
  - $P^- = \{w \in P \mid \nexists u \in L: w <_p u\}$
- Assume that P^+ is finite:
  - Let $R = L \setminus \{[w] \mid w \in P^+\}$
  - Let $R^- = \{w \in R \mid \nexists u \in R: w <_p u\}$
  - Claim: $R^- = P^-$
  - Thus $P = P^+ \cup P^-$ is regular! ↯

Non-regular MPFS of regular languages

L has a non-regular MPFS ⇒ *L* has uncountably many MPFS.

- Let P be a non-regular MPFS of a regular language L.
  - $P^+ = \{w \in P \mid \exists u_w \in L: w <_p u_w\}$ (infinite)
  - $P^- = \{w \in P \mid \nexists u \in L: w <_p u\}$
- Pick $S \subseteq P^+$
  - Let $S' = S \cup \{wu_w \mid w \in P^+ \setminus S\}$
  - Extend S' to a MPFS S''.

Non-regular MPFS of regular languages

L has a non-regular MPFS ⇒ *L* has uncountably many MPFS.

- Let P be a non-regular MPFS of a regular language L.
  - $P^+ = \{w \in P \mid \exists u_w \in L: w <_p u_w\}$ (infinite)
  - $P^- = \{w \in P \mid \nexists u \in L: w <_p u\}$
- Pick $S \subseteq P^+$
  - Let $S' = S \cup \{wu_w \mid w \in P^+ \setminus S\}$
  - Extend S' to a MPFS S''.
- Picking two different S_1, S_2 results in two different S_1'', S_2''.
- Uncountably many choices ⇒ uncountably many MPFS. □

Non-regular MPFS of regular languages

L has a non-regular MPFS ⇒ L has uncountably many MPFS.

- Let P be a non-regular MPFS of a regular language L.
  - $P^+ = \{w \in P \mid \exists u_w \in L : w <_p u_w\}$ (infinite)
  - $P^- = \{w \in P \mid \nexists u \in L : w <_p u\}$
- Pick $S \subseteq P^+$
  - Let $S' = S \cup \{wu_w \mid w \in P^+ \setminus S\}$
  - Extend S' to a MPFS S".
- Picking two different S_1, S_2 results in two different S_1'', S_2''.
- Uncountably many choices ⇒ uncountably many MPFS. □
- Further results on non-regular MPFS in an upcoming paper!

Thank you for your attention!

Questions?