Closure under reversal of languages over infinite alphabets

Daniel Genkin[1] Michael Kaminski[2] Liat Peterfreund[2]

[1]Department of Computer and Information Science, University of Pennsylvania, 3330 Walnut St., Philadelphia PA, 19104 USA

[2]Department of Computer Science, Technion – Israel Institute of Technology, Haifa 32000, Israel

Definition A deterministic one-way k-wPA over an infinite alphabet Σ, is a tuple $\mathfrak{A} = \langle S, s_0, F, T \rangle$ whose components are as follows.

- S is a finite set of states,
- $s_0 \in S$ is the initial state,
- $F \subseteq S$ is a set of accepting states,
- T is a finite set of transitions of the form $\alpha \rightarrow \beta$, where
  - α is of the form (i, σ, P, V, s) or (i, P, V, s), $i \in \{1, \ldots, k\}$, $\sigma \in \Sigma$, $P, V \subseteq \{1, \ldots, i-1\}$, and $s \in S$, and
  - β is of the form $(p, \texttt{action})$, where $p \in S$ and $\texttt{action} \in \{\texttt{move}, \texttt{place}, \texttt{lift}\}$,

such that $\alpha \rightarrow \beta$ and $\alpha \rightarrow \beta'$ imply $\beta = \beta'$.

For a word $\boldsymbol{w} \in \Sigma^*$, a configuration of $\mathfrak{A}$ on $\boldsymbol{w}$ is of the form $\gamma = [i, s, \theta]$, where $i \in \{1, \ldots, k\}$, $s \in S$, and $\theta : \{1, \ldots, i\} \to \{1, \ldots, |\boldsymbol{w}|\}$ indicates the pebble's positions on the input word $\boldsymbol{w}$.

That is, $\theta(j)$ is the position of pebble j.

In what follows, we identify θ with the i-tuple $(\theta(1), \ldots, \theta(i))$. Thus, i can be recovered from θ, but it is convenient to include it into a configuration explicitly.

The *initial configuration* is $\gamma_0 = [1, s_0, (1)]$.

That is, the run starts in the initial state s_0 with pebble 1 placed at the beginning of the input word.

An *accepting configuration* is of the form $[i, s, \theta]$, where $s \in F$.

Let $\boldsymbol{w} = w_1 \cdots w_n \in \Sigma^+$. A transition $(i, \sigma, P, V, s) \to \beta$ *applies* to a configuration $\gamma = [j, s', \theta]$ if

(1) $i = j$ and $s' = s$,

(2) $V = \{h < i : w_{\theta(h)} = w_{\theta(i)}\}$,

(3) $P = \{h < i : \theta(h) = \theta(i)\}$, and

(4) $w_{\theta(i)} = \sigma$.

A transition $(i, P, V, s) \to \beta$ applies to a configuration $\gamma = [j, s', \theta]$, if the above conditions (1)–(3) are satisfied and no transition of the form $(i, \sigma, P, V, s) \to \beta$ applies to γ.

The transition relation $\vdash_{\boldsymbol{w}}$ on the set of all configurations is defined as follows:

$[i, s, \theta] \vdash [i', s', \theta']$ if and only if there is a transition $\alpha \rightarrow (p, \texttt{action})$ that applies to $[i, s, \theta]$ such that $s' = p$ and the following holds.

- For all $j < i$, $\theta'(j) = \theta(j)$,
- if `action` is `move`, then $i' = i$ and $\theta'(i) = \theta(i) + 1$,
- if `action` is `place`, then $i' = i + 1$ and $\theta'(i + 1) = \theta'(i) = \theta(i)$, and
- if `action` is `lift`, then $i' = i - 1$ and θ' is the restriction of θ on $\{1, \ldots, i - 1\}$.

The *language* $L(\mathfrak{A})$ of $\mathfrak{A}$ consists of all words $\boldsymbol{w}$ such that $\gamma_0 \vdash_{\boldsymbol{w}}^{*} \gamma$ for an accepting configuration γ.

To each configuration $\gamma = [i, s, \theta]$ of a deterministic one-way wPA corresponds the vector $\varphi^{\gamma} = (P_1, \ldots, P_i)$, where

$$P_j = \{h < j : \theta(h) = \theta(j)\}\,.$$

That is, P_j is the set of pebbles placed before pebble j which are at the same position as pebble j in configuration γ.

If $\gamma \vdash \gamma'$, then $\varphi^{\gamma'}$ can be computed from φ^{γ}, according to the automaton transitions.

Thus, we may assume that the left hand side of a transition is of the form (i, σ, V, s) or (i, V, s).

By adding some extra states and modifying the transitions appropriately, we can normalize the k-wPA behavior such that for each $i \in \{2, \ldots, k\}$ it acts as follows.

- A pebble is never lifted, but falls down when moving from the right end of the input. Thus, action `lift` is redundant.
- Only pebble 1 can enter a final state and only after it falls down from the right end of the input. In such a case, the accepting configuration consists of the corresponding accepting state only.
- Immediately after pebble i moves without falling down, pebble $i+1$ is placed.
- Immediately after pebble i falls down, pebble $i-1$ moves.

We denote the set of letters occurring in a word $\boldsymbol{u}$ by $[\boldsymbol{u}]$. That is, if $\boldsymbol{u} = u_1 \cdots u_n$, then $[u] = \{u_1, \ldots, u_n\}$.

Example The language

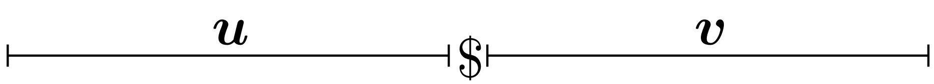
$$L_{\mathtt{diff}} = \{\sigma_1 \cdots \sigma_n : n \geq 1,\ \sigma_i \neq \$,\ \text{for each } i = 1, \ldots, n, \text{ and } \sigma_i \neq \sigma_j, \text{ whenever } i \neq j\}.$$

is accepted by 2-wPA.

Example The language

$$L_{\mathtt{diff\$diff}} = \{\boldsymbol{u}\$\boldsymbol{v} : \boldsymbol{u}, \boldsymbol{v} \in L_{\mathtt{diff}}\}$$

is accepted by 2-wPA.



Example The language

$$L_{\subseteq} = \{\boldsymbol{u}\$\boldsymbol{v} : \boldsymbol{u}, \boldsymbol{v} \in L_{\mathsf{diff}} \text{ and } [\boldsymbol{u}] \subseteq [\boldsymbol{v}]\}$$

is accepted by 2-wPA.

Theorem The language

$$L_{\supseteq} = \{\boldsymbol{u}\$\boldsymbol{v} : \boldsymbol{u}, \boldsymbol{v} \in L_{\mathsf{diff}} \text{ and } [\boldsymbol{v}] \subseteq [\boldsymbol{u}]\}$$

is not accepted by wPA.

Proposition There exists a positive integer ℓ_2 such that for all $\boldsymbol{w} \in \Sigma^+$, $w = w_1 \cdots w_n$, the following holds. If

$$[2, s_{j_1}, (p_1, j_1)] \vdash [2, s_{j_1+1}, (p_1, j_1 + 1)] \vdash \cdots \vdash [2, s_{j_2}, (p_1, j_2)],$$

where $w_j \neq w_{p_1}$ for all $j_1 \leq j \leq j_2$, then the sequence of states $s_{j_1+\ell_2}, \ldots, s_{j_2}$, is periodic with period ℓ_2.

Corollary Let $\boldsymbol{z}' = \boldsymbol{x}\boldsymbol{y}'$ and $\boldsymbol{z}'' = \boldsymbol{x}\boldsymbol{y}''$, $\boldsymbol{x} = x_1 \cdots x_n$, where

$$[\boldsymbol{x}] \cap ([\boldsymbol{y}'] \cup [\boldsymbol{y}'']) = \emptyset \,,$$

$$|\boldsymbol{y}'|, |\boldsymbol{y}''| \geq \ell_2 \,,$$

and

$$|\boldsymbol{y}''| \equiv_{\ell_2} |\boldsymbol{y}'| \,.$$

If

$$[2, s, (p, |\boldsymbol{x}|)] \vdash_{\boldsymbol{z}'} [2, t, (p, |\boldsymbol{x}\boldsymbol{y}'|)] \,,$$

then

$$[2, s, (p, |\boldsymbol{x}|)] \vdash_{\boldsymbol{z}''} [2, t, (p, |\boldsymbol{x}\boldsymbol{y}''|)] \,.$$

Corollary Let $\boldsymbol{w}, \boldsymbol{w}' \in L_{\mathtt{diff\$diff}}$, $\boldsymbol{w} = \boldsymbol{u}'\boldsymbol{v}\$\boldsymbol{x}$ and $\boldsymbol{w}' = \boldsymbol{u}'\boldsymbol{u}''\boldsymbol{v}\$\boldsymbol{x}$ be such that $|\boldsymbol{u}''| \equiv_{\ell_2} 0$ and $|\boldsymbol{v}| \geq \ell_2$. If

$$[1, s_0, (1)] \vdash^*_{\boldsymbol{w}} [1, t, (|\boldsymbol{u}'|)] \,,$$

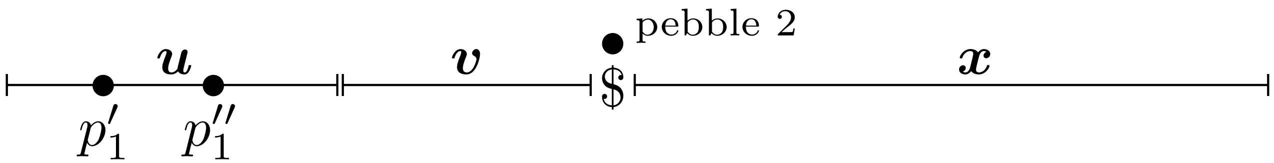
then

$$[1, s_0, (1)] \vdash^*_{\boldsymbol{w}'} [1, t, (|\boldsymbol{u}'|)] \,.$$

Definition Let ℓ be a positive integer and let $\boldsymbol{u}, \boldsymbol{v} \in L_{\mathsf{diff}}$, $\boldsymbol{u} = u_1 \cdots u_m$ and $\boldsymbol{v} = v_1 \cdots v_n$, be such that $[\boldsymbol{u}] \subseteq [\boldsymbol{v}]$: $u_i = v_{j_i}$, $i = 1, \ldots, m$. We say that $\boldsymbol{u}$ is *ℓ-spread* in $\boldsymbol{v}$, if for all $i = 1, \ldots, m$, $j_i > j_{i-1}$ and $j_i \equiv_\ell j_{i-1}$, where $j_0 = 0$.

Proposition Let $\boldsymbol{w} = \boldsymbol{u}\boldsymbol{v}\$\boldsymbol{x} \in L_{\mathsf{diff\$diff}}$, where $\boldsymbol{u}$ is ℓ_2-spread in $\boldsymbol{x}$, and let $1 < p_1' < p_1'' \leq |\boldsymbol{u}|$. If

then
$$[2, s, (p_1', |\boldsymbol{u}\boldsymbol{v}\$|)] \vdash^* [1, t, (p_1')]\,,$$
$$[2, s, (p_1'', |\boldsymbol{u}\boldsymbol{v}\$|)] \vdash^* [1, t, (p_1'')]\,.$$



Corollary Let $\boldsymbol{w} = \boldsymbol{u}\boldsymbol{v}\$\boldsymbol{x} \in L_{\mathsf{diff\$diff}}$ be such that $\boldsymbol{u}$ is ℓ_2-spread in $\boldsymbol{x}$ and $|\boldsymbol{v}| \geq \ell_2$, and let $p_1' < p_1'' \leq |\boldsymbol{u}|$ be equivalent modulo ℓ_2. If

then
$$[2, s, (p_1', p_1')] \vdash^* [2, t, (p_1' + 1, p_1' + 1)]\,,$$
$$[2, s, (p_1'', p_1'')] \vdash^* [2, t, (p_1'' + 1, p_1'' + 1)]\,.$$

Proposition For each $\boldsymbol{w} = \boldsymbol{u}\boldsymbol{v}\$\boldsymbol{x} \in L_{\mathit{diff}\$\mathit{diff}}$ such that $\boldsymbol{u}$ is ℓ_2-spread in $\boldsymbol{x}$ and $|\boldsymbol{v}| \geq \ell_2$, there exist positive integers $m_{\boldsymbol{w}}$ and $\ell_{\boldsymbol{w}}$ for which the following holds. If

$$[2, s_{j_1}, (p_1, p_1)] \vdash^* [2, s_{j_2}, (p_1 + 1, p_1 + 1)] \vdash^* \cdots \vdash^* [2, s_{j_{|\boldsymbol{u}|-p_1}}, (|\boldsymbol{u}|, |\boldsymbol{u}|)] ,$$

then the sequence of states $s_{p_1+m_{\boldsymbol{w}}}, \ldots, s_{|\boldsymbol{u}|-p_1}$ is periodic with period $\ell_{\boldsymbol{w}}$.

Corollary There exist a positive integer ℓ_1 such that the following holds. Let $\boldsymbol{w} = \boldsymbol{u}\boldsymbol{v}\$\boldsymbol{x} \in L_{\mathit{diff}\$\mathit{diff}}$, where $\boldsymbol{u}$ is ℓ_2-spread in $\boldsymbol{x}$ and $|\boldsymbol{v}| \geq \ell_2$. If

$$[2, s_{j_1}, (p_1, p_1)] \vdash^* [2, s_{j_2}, (p_1 + 1, p_1 + 1)] \vdash^* \cdots \vdash^* [2, s_{j_{|\boldsymbol{u}|-p_1}}, (|\boldsymbol{u}|, |\boldsymbol{u}|)] ,$$

then the sequence of states $s_{p_1+\ell_1}, \ldots, s_{|\boldsymbol{u}|-p_1}$ is periodic with period ℓ_1.

Proof of the theorem Assume to the contrary that $L(\mathfrak{A}) = L_{\supseteq}$. Let

$$\boldsymbol{w}' = \boldsymbol{u}'\boldsymbol{u}''\boldsymbol{v}\$\boldsymbol{x} \in L_{\supseteq} \cap L_{\subseteq}\,,$$

where $\boldsymbol{u}'$ is ℓ_2-spread in $\boldsymbol{x}$, $|\boldsymbol{v}| \geq \ell_2$, and $|\boldsymbol{u}'| = |\boldsymbol{u}''| = \ell_1$; and let

$$\boldsymbol{w} = \boldsymbol{u}'\boldsymbol{v}\$\boldsymbol{x}\,.$$

Since $\ell_1 \equiv_{\ell_2} 0$,

$$[1, s_0, (1)] \vdash^*_{\boldsymbol{w}} [1, t, (|\boldsymbol{u}'|)]$$

implies

$$[1, s_0, (1)] \vdash^*_{\boldsymbol{w}'} [1, t, (|\boldsymbol{u}'|)]$$

and, since $|\boldsymbol{u}'| = |\boldsymbol{u}''| = \ell_1$,

$$[1, s_0, (1)] \vdash^*_{\boldsymbol{w}'} [1, t, (|\boldsymbol{u}'\boldsymbol{u}''|)]\,.$$

In addition, the runs of $\mathfrak{A}$ from state t on the (same) suffix $\boldsymbol{v}\$\boldsymbol{x}$ of $\boldsymbol{w}$ and $\boldsymbol{w}'$ are the same. In particular, they terminate in the same state. However, $\boldsymbol{w}'$ is accepted by $\mathfrak{A}$, whereas $\boldsymbol{w}$ is not. $\square$

Removing the distinguished separator symbol $

Let

$$L'_{\subseteq} = \{\sigma\boldsymbol{u}\sigma\boldsymbol{v} : \sigma\boldsymbol{u}, \sigma\boldsymbol{v} \in L_{\mathtt{diff}} \text{ and } [\boldsymbol{u}] \subseteq [\boldsymbol{v}]\}$$

and

$$L'_{\supseteq} = \{\sigma\boldsymbol{u}\sigma\boldsymbol{v} : \sigma\boldsymbol{u}, \sigma\boldsymbol{v} \in L_{\mathtt{diff}} \text{ and } [\boldsymbol{v}] \subseteq [\boldsymbol{u}]\}\,.$$

Then $L'_{\supseteq}$ is the reversal of $L'_{\subseteq}$.