Well-quasi-orders and Regular $\omega$-languages

Mizuhito Ogawa

$^a$Japan Advanced Institute of Science and Technology
1-1 Asahidai Tatsunokuchi Nomi Ishikawa 923-1292 Japan
mizuhito@jaist.ac.jp

Abstract

In "On regularity of context-free languages, Theoretical Computer Science Vol.27, pp.311-332, 1983", Ehrenfeucht et al. showed that a set $L$ of finite words is regular if and only if $L$ is $\leq$-closed under some monotone well-quasi-order (WQO) $\leq$ over finite words. We extend this result to regular $\omega$-languages. That is,

(1) an $\omega$-language $L$ is regular if and only if $L$ is $\preceq$-closed under a periodic extension $\preceq$ of some monotone WQO over finite words, and

(2) an $\omega$-language $L$ is regular if and only if $L$ is $\preceq$-closed under a WQO $\preceq$ over $\omega$-words that is a continuous extension of some monotone WQO over finite words.

Key words: $\omega$-language, well-quasi-order, regularity.

1 Preliminaries

Throughout the paper, we will use $A$ for a finite alphabet, $A^*$ for a set of all (possibly empty) finite words on $A$, and $A^\omega$ for a set of all $\omega$-words on $A$. A concatenation of two words $u, v$ is denoted by $uv$, an element-wise concatenation of two sets $U, V$ of words by $UV, VV, \ldots, V^i$ by $V^i$, and $VV, \ldots$ by $V^\omega$.

The length of a finite word $u$ is denoted by $|u|$. As a convention, we will use $\epsilon$ for the empty word, $u, v, w, \cdots$ for finite words, $\alpha, \beta, \cdots$ for $\omega$-words, $a_1, a_2, \cdots$ for elements in $A$, $i, j, k, l, \cdots$ for indices, and $U, V, \cdots$ (capital letters) for sets. We sometimes use $x, y, \cdots$ for elements of a set.

A regular $\omega$-language is a set of $\omega$-words that are accepted by a (nondeterministic) $Büchi$ automaton $A = \{Q, q_0, \Delta, F\}$, where $Q$ is a finite set of states,
Lemma 1.2. For a finite congruence that saturates \( L \), \( L \) saturates \( L \).

Lemma 1.4. A congruence \( \sim \) is finite if there are only finitely many \( \sim \)-classes. Details are given elsewhere [3].

Definition 1.1. Let \( L \subseteq A^\omega \) and let \( \sim \) be a congruence over \( A^* \). We say that \( \sim \) saturates \( L \) if for each \( \sim \)-class \( U,V \), \( U.V^\omega \cap L \neq \emptyset \) implies \( U.V^\omega \subseteq L \).

Lemma 1.2. For a \( \text{B"uchi} \) automaton \( A \) and \( u,v \in A^* \), we define \( u \sim_A v \) if 
\[
(q \xrightarrow{u} q' \iff q \xrightarrow{v} q') \land (q \xrightarrow{F}\_u q' \iff q \xrightarrow{F}\_v q')
\]
for each \( q,q' \in Q \). Then \( \sim_A \) is a finite congruence that saturates \( L(A) \).

Theorem 1.3. \( L \subseteq A^\omega \) is regular if and only if some finite congruence saturates \( L \).

Lemma 1.4. Let \( \sim \) be a finite congruence over \( A^* \).

1. Let \( \alpha = u_1 u_2 \cdots \in A^\omega \) and let \( u(i,j) = u_1 u_{i+1} \cdots u_{j-1} \) where \( u_i \in A^* \). There exist a \( \sim \)-class \( V \) and \( i_1 < i_2 < \cdots \) such that \( u(i_j,i_k) \in V \) for each \( j,k \) with \( j < k \).

2. Let \( U,V \) be \( \sim \)-classes. There exist \( \sim \)-classes \( U',V' \) such that \( U.V^\omega \subseteq U'.V'^\omega \), \( U'.V' \subseteq U' \), and \( V'.V' \subseteq V' \).

Proof

1. Since \( \sim \) has only finitely many \( \sim \)-classes, this is a direct consequence of (infinite) Ramsey Theorem.

2. Note that for each \( \sim \)-class \( U_1, \cdots, U_m, W, U_1, \cdots, U_n \cap W \neq \emptyset \) implies \( U_1, \cdots, U_n \subseteq W \). Since \( \sim \) has only finitely many \( \sim \)-classes, from (infinite) Ramsey Theorem there exist a \( \sim \)-class \( V' \) and \( i_1 < i_2 < \cdots \) such that \( V^{i_1}_{i'} \subseteq V' \) for each \( j,k \) with \( j < k \) and \( V'.V' \subseteq V' \). Let \( U' \) be a \( \sim \)-class that includes \( U.V^{i_1} \). Then \( U.V^\omega \subseteq U'.V'^\omega \), \( U'.V' \subseteq U' \), and \( V'.V' \subseteq V' \).

We denote a quasi-order (QO, i.e., reflexive transitive binary relation) over a set \( S \) by \((S, \leq)\). If \( S \) is clear from the context, we simply denote by \( \leq \). As a convention, a QO over finite words is denoted it by \( \leq \), and a QO over \( \omega \)-words is denoted by \( \preceq \).
For each $u, v \in A^*$, $u_i \leq v_i$ for any $i$ implies $u_1u_2u_3 \cdots \leq v_1v_2v_3 \cdots$.

For each $\alpha \in A^\omega$, there exist $u, v \in A^*$ such that $\alpha \leq u.v^\omega$ and $\alpha \geq u.v^\omega$.

For instance, the embedding over $\omega$-words is the periodic extension of the embedding over finite words. Note that a periodic extension of a monotone WQO over $A^*$ is a WQO over $A^\omega$. We will prove Theorem 2.2 below.

Let $L \subseteq A^\omega$. $L$ is regular if and only if $L$ is $\leq$-closed under a periodic extension $(A^\omega, \leq)$ of a monotone WQO $(A^*, \leq)$.

For $u, v \in A^*$, if $uv^\omega \in U.V^\omega$, $U.V \subseteq U$, and $V.V \subseteq V$, there exist $w_1 \in U$ and $w_2 \in V$ such that $w_1w_2^\omega = uv^\omega$.

Proof Let $uv^\omega = u'v_1^i v_2^j \cdots$ satisfying $u' \in U$ and $v_i^j \in V$, and let $w(i, j) = v_i^j \cdots v_{i-1}^j$ for $i < j$. Let $k_j \equiv |w(1, j)| \ (mod \ |v|)$. Then there exist $k_{j1}$ and $k_{j2}$ such that $k_{j1} < k_{j2}$ and $k_{j1} \equiv k_{j2} \ (mod \ |v|)$. Since there are infinitely many such pairs, we can assume that $|u| \leq |u'.w(1, j_1 - 1)|$. Let $w_1 = u'.w(1, j_1 - 1)$ and $w_2 = w(j_1, j_2 - 1)$. Since $U.V \subseteq U$ and $V.V \subseteq V$, $w_1 \in U$, $w_2 \in V$ and $uv^\omega = w_1w_2^\omega$.

Lemma 2.4 For a B"uchi automaton $A$ and $\alpha \in A^\omega$, let $[\alpha] = \{U.V^\omega \mid \alpha \in U.V^\omega \}$ where $U, V$ are $\sim_A$-classes. We define $\alpha \preceq' \beta$ if $[\alpha] \cap [\beta] \neq \emptyset$. Then,

1. $L(A)$ is $\leq'$-closed.
2. $u_i \sim_A v_i$ for each $i$ imply $u_1u_2 \cdots \preceq' v_1v_2 \cdots$.

Proof From Lemma 1.2, $\sim_A$ saturates $L$ and $U.V^\omega \subseteq L$ for each $U.V^\omega \in [\alpha]$. Thus $L$ is $\leq'$-closed.
From Lemma 1.4 (i), there exist a \( \sim_A \)-class \( V \) and \( i_1 < i_2 < \cdots \) such that \( u(i_j, i_k) \in V \) for each \( j < k \). Let \( U \) be a \( \sim_A \)-class such that \( u(1, i_1) \in U \). (We bow the notation from Lemma 1.4 (i).) Since \( \sim_A \) is a congruence, \( v(1, i_j) \in U \) and \( v(i_j, i_k) \in V \) for each \( j < k \). Thus \( u_1 w_2 \cdots \in U.V^\omega \) implies \( v_1 v_2 \cdots \in U.V^\omega \), and \( v \leq U \).

**Definition 2.5** [1] For \( u, v \in A^* \), we define \( u \equiv_L v \) if \( w(w_1 u v w_v) \in L \iff w(w_1 v w_2 w_v) \in L \) and \( w_1 u w_2 w_v \in L \iff w_1 v w_2 w_v \in L \) for each \( w, w_1, w_2 \in A^* \).

**Proof of Theorem 2.2**

*Only-if part:* Assume \( L \) is regular. Let \( A \) be a Büchi automaton such that \( L = L(A) \). Since \( \sim_A \) is a finite congruence, \((A^* , \sim_A)\) is a monotone WQO. Define \( \preceq \) as the transitive closure of \( \equiv \) (defined in Lemma 2.4), then \((A^*, \preceq)\) is a periodic extension of \((A^*, \sim_A)\) and \( L(A) \) is \( \preceq \)-closed.

*If part:* Assume that \( L \) is \( \preceq \)-closed where \( \preceq \) is a periodic extension of a monotone WQO \( \preceq \). First, we show that \( \equiv_L \) is a finite congruence. Assume that \( \{ u_i \} \) is an infinite set in \( A^* \) such that \( u_i \not\equiv_L u_j \) for \( i \neq j \). Since \((A^*, \preceq)\) is a WQO, there exists an infinite ascending subsequence \( \{ u_k \} \).

Let \( F(u) = \{ (v, v_1, v_2, w_1, w_2, w) \in A^* \times A^* \times A^* \times A^* \times A^* \mid v(v_1 u v_2) \in L \land w_1 u w_2 w_v \in L \} \). Since \( \preceq \) is a periodic extension of \( \preceq \) and \( L \) is \( \preceq \)-closed, each \( F(u) \) is \( \preceq \)-closed. Since \( u_k \not\equiv_L u_j \) for \( i \neq j \), \( F(u_k) \subsetneq F(u_j) \), thus \( F(u_k) \subsetneq F(u_j) \).

Then there exists an infinite sequence in which each pair of different elements is incomparable. Since \( \preceq \)-closed is a WQO over \( A^* \times A^* \times A^* \times A^* \times A^* \), this is a contradiction.

Second, we show that \( \equiv_L \) saturates \( L \). Assume that some \( \equiv_L \)-classes \( U, V \) satisfy \( U.V^\omega \cap L \neq \phi \) and \( U.V^\omega \not\subseteq L \). From Lemma 1.4 (ii), we can assume that \( U.V \subseteq U \land V.V \subseteq V \).

Let \( \alpha \in U.V^\omega \land L \) and \( \beta \in U.V^\omega \land L \). Since \((A^w, \preceq)\) is a periodic extension, from Lemma 2.3 there exist \( u, u' \in U \) and \( v, v' \in V \) such that \( \alpha = u v^\omega \) and \( \beta = u' v'^\omega \). By definition of \( \equiv_L \), \( u v^\omega \in L \land u' v'^\omega \in L \) are contradictory.

## 3 Second theorem

**Definition 3.1** For a monotone QO \((A^*, \preceq)\), a QO \((A^w, \preceq)\) is a continuous extension if the following conditions are satisfied.

1. For each \( u, v \in A^* \) and \( \alpha, \beta \in A^w \), \( u \leq v \) and \( \alpha \leq \beta \) imply \( u \alpha \leq v \beta \).
(2) Let \( u_j, v_j \in A^* \) for each \( j \) and let \( \alpha_i = v_1 \ldots v_{i-1} u_i \ldots \) for each \( i \) and \( \alpha_\infty = v_1 v_2 \ldots \). For \( \beta \in A^\omega \), if \( u_i \leq v_i \) and \( \alpha_i \leq \beta \) for each \( i \), then \( \alpha_\infty \leq \beta \), and if \( u_i \geq v_i \) and \( \alpha_i \geq \beta \) for each \( i \), then \( \alpha_\infty \geq \beta \).

**Theorem 3.2** Let \( L \subseteq A^\omega \). \( L \) is regular if and only if \( L \) is \( \preceq \)-closed under a WQO \((A^\omega, \preceq)\) that is a continuous extension of a monotone WQO \((A^*, \leq)\).

For the embedding \( \leq \) over finite words, let \((A^*, \leq^o)\) be defined as \( u \leq^o v \) if and only if \( u \leq v \) and \( \text{elt}(u) = \text{elt}(v) \), where \( \text{elt}(u) = \{a_i \mid u = a_1 a_2 \ldots a_j\} \). Since the embedding \( \leq \) over finite words is a WQO from Higman’s lemma, \( \leq^o \) is also a WQO. Then the embedding over \( A^\omega \) is a continuous extension of \( \leq^o \). Note that the embedding over \( A^\omega \) is a continuous extension of the embedding \( \leq \) over finite words. Actually, any continuous extension of the embedding \( \leq \) over finite words is a trivial WQO (i.e., \( A^\omega \times A^\omega \)). For instance, given \( \alpha, \beta \in A^\omega \). Let \( \alpha(1, i) \) be the prefix of \( \alpha \) of the length \( i \) and \( \alpha_i = \alpha(1, i) \), \( \beta \) for each \( i \). Since \( \alpha(1, i) \geq \epsilon, \alpha_i \geq \beta \) for each \( i \). Thus, by definition of continuity, \( \alpha_\infty = \alpha \geq \beta \). Hence, for any \( \alpha, \beta \in A^\omega \), we conclude \( \alpha \preceq \beta \).

**Definition 3.3** Let \( u, v \in A^* \) and let \( L \subseteq A^\omega \). We write

- \( u \sim_1^L v \) if and only if \( \forall w \in A^*, \forall \alpha \in A^\omega: \text{wu} \alpha \in L \iff \text{w} v \alpha \in L \),
- \( u \sim_2^L v \) if and only if \( \forall w \in A^*: \text{wu} \alpha \in L \iff \text{w} v \alpha \in L \), and
- \( u \sim_L v \) if and only if \( u \sim_1^L v \) and \( u \sim_2^L v \).

**Proof of Theorem 3.2**

*Only-if part:* Assume \( L \) is regular. Let \( \mathcal{A} \) be a B"uchi automaton such that \( L = L(\mathcal{A}) \). Since \( \sim_\mathcal{A} \) is a finite congruence, \((A^*, \sim_\mathcal{A})\) is a monotone WQO. Define \( \preceq \) as the transitive closure of \( \preceq' \) (defined in Lemma 2.4), then \( L(\mathcal{A}) \) is \( \preceq \)-closed. Since \( \preceq \) is symmetric, \((A^\omega, \preceq)\) is a continuous extension of \((A^*, \preceq_\mathcal{A})\) from Lemma 2.4 (ii). For the index \( n \) of \( \sim_\mathcal{A} \), the number of \( \preceq \)-classes is bound by \( 2n^2 \). Thus \( \preceq \) is a WQO.

*If part:* First, we show that \( \simeq_\mathcal{A} \) is a finite congruence. Assume that \( \{u_i\} \) is an infinite set in \( A^* \) such that \( u_i \not\simeq_\mathcal{A} u_j \) for \( i \neq j \). Since \((A^*, \preceq)\) is a WQO, there exists an infinite ascending subsequence \( \{u_{k_i}\} \).

Let \( F(u) \subseteq A^* \times A^\omega \times A^* \) be a set such that \( (w, \alpha, v) \in F(u) \iff wu \alpha \in L \land wv \alpha \in L \). Then, each \( F(u) \) is \( \leq \times \leq \preceq \)-closed and hence \( F(u_{k_i}) \subseteq F(u_{k_j}) \) for \( i < j \). Since \( u_{k_i} \not\simeq_\mathcal{A} u_{k_j} \) for \( i \neq j \), \( F(u_{k_i}) \neq F(u_{k_j}) \), thus \( F(u_{k_i}) \subseteq F(u_{k_j}) \). Then there exists an infinite sequence in which each pair of different elements is incomparable. Since \( \leq \times \preceq \times \leq \) is a WQO over \( A^* \times A^\omega \times A^* \), this is a contradiction.

Second, we show that \( \simeq_\mathcal{A} \) saturates \( L \). Assume that some \( \simeq_\mathcal{A} \)-classes \( U, V \)
satisfy $U.V^\omega \cap L \neq \emptyset$ and $U.V^\omega \not\subseteq L$. From Lemma 1.4 (2), we can assume that $V.V \subseteq V$.

Let $\alpha = u\nu_1\nu_2 \cdots$ be a minimal element (wrt $\preceq$) in $U.V^\omega \cap L$, and let $\beta = u'u_1\nu'_2 \cdots \in U.V^\omega \setminus L$ such that $u, u' \in U$ and $\nu_i, \nu'_i \in V$. Let $\{\vec{v}_i\}$ be sets of minimal elements of $V$ wrt $\preceq$. Since $(V, \preceq)$ is a WQO, $\{\vec{v}_i\}$ are finite.

Let $\alpha'(j, j+k) = v_j \cdots v_{j+k}$. Since $\vec{v}_i$ are finitely many, from (infinite) Ramsey Theorem there exist $l$ and an ascending sequence $0 < j_1 < j_2 < \cdots$ such that $\alpha'(j_m, j_{m+1} - 1) \geq \vec{v}_l$ for any $m > 0$.

Let $\alpha_m = u \alpha'(1, j_1 - 1) \vec{v}_l^{m-1} \alpha'(j_m, j_{m+1} - 1) \cdots$. Obviously, $\alpha_m \preceq \alpha$ and $\alpha_m \in U.V^\omega \cap L$. Since $\alpha$ is minimal in $U.V^\omega \cap L$, $\alpha_m \preceq \alpha$. By definition of the continuous extension, $\alpha_\infty = u \alpha'(1, j_1 - 1) \vec{v}_l^\omega \succeq \alpha$. Thus since $L$ is $\preceq$-closed, $\alpha_\infty \in U.V^\omega \cap L$.

Let $\beta'(j, j+k) = v'_j \cdots v'_{j+k}$. Since $\vec{v}_l$ are finitely many, from (infinite) Ramsey Theorem there exist $\nu'$ and an ascending sequence $0 < j'_1 < j'_2 < \cdots$ such that $\beta'(j'_m, j'_{m+1} - 1) \geq \vec{v}_l$ for any $m > 0$. Let $\beta_\infty = u' \beta'(1, j_1 - 1) \vec{v}_l^\omega$. By definition of the continuous extension, $\beta_\infty \not\preceq \beta$. Since $L$ is $\preceq$-closed, $\beta \not\in L$ implies $\beta_\infty \not\in L$. Thus $\beta \in U.V^\omega \setminus L$.

Since $u \succeq_L u'$ and $\vec{v}_j \succeq_L \vec{v}_l$ for each $j$, repeated applications of $\succeq_L^1$ and an application of $\succeq_L^2$ imply that $\alpha_\infty \in L \iff \beta_\infty \in L$. This contradicts $\alpha_\infty \in L$ and $\beta_\infty \not\in L$. \hfill \blacksquare

**Example 3.4** Either the periodic or continuous assumption cannot be dropped. Let $\beta = abaabaaaabaa \cdots$ and let $L(\beta)$ be the set of $\omega$-words that have a common suffix with $\beta$. For $\alpha \in A^\omega$, let $p^\beta_\alpha(\alpha) = 1$ if $\alpha \in L(\beta)$ and let $p^\beta_\alpha(\alpha) = 0$ if $\alpha \not\in L(\beta)$. Define $\alpha \preceq \alpha' \iff p^\beta_\alpha(\alpha) \leq p^\beta_\alpha(\alpha')$. Then $\preceq$ is a WQO over $\omega$-words and $L(\beta)$ is $\preceq$-closed, but $L(\beta)$ is not regular.

**Acknowledgements**

The author thanks Jean-Eric PIN for valuable comments at the previous presentation. This work is partially supported by PRESTO, Japan Science and Technology Corporation.

**References**
