ω-AUTOMATA & ω-REGULAR LANGUAGES

18 Mar 2010

Introduction

The basic idea here is to extend the notion of acceptance of finite state automata to infinite strings inputs. One consequence of this extension is the clear distinction that emerges between deterministic and non-deterministic automata. The usual Scott-Rabin reduction for the case of finite-word languages no longer holds, because non-deterministic automata turn out to be strictly more powerful than deterministic ones. As a consequence certain constructions and proofs such as closure under complementation become more complex and intricate than for the case of regular languages.

Notational conventions

A = a finite alphabet $(a, b, c, \ldots \in A)$

A^* = the set of finite words over A $(u, v, w, \ldots \in A^*)$

A^ω = the set of infinite words over A $(\alpha, \beta, \gamma, \ldots \in A^\omega)$

An infinite word over A is a function $\alpha : \mathbb{N} \to A$ so that $\alpha(i) \in A$ is the letter at the i-th position.

$$\alpha[m..n] = \alpha(m)\alpha(m+1)\ldots\alpha(n) \in A^{n-m+1}$$

For a finite set S and $\sigma : \mathbb{N} \to S$,

$$\inf(\sigma) = \{ s \in S \mid \exists^\omega n \in \mathbb{N} [\sigma(n) = s] \}$$

i.e. $\inf(\sigma)$ is the set of states that occur infinitely often in σ.

Büchi Automata.

A <u>Büchi automaton</u> $\mathcal{B} = \langle S, A, \rightarrow, I, G \rangle$ is a 5-tuple structure consisting of

S = a finite set of states, A = a finite alphabet,

$\rightarrow \subseteq S \times A \times S$ = the transition relation,

$I \subseteq S$ = the set of initial states,

$G \subseteq S$ = the set of accepting states.

Acceptance. $\mathcal{B}$ accepts an infinite word $\alpha \in A^\omega$ if there is an infinite sequence $\sigma = s_0 s_1 \ldots$ such that

There must be at least one state that occurs infinitely often

$$s_0 \xrightarrow{\alpha(0)} s_1 \xrightarrow{\alpha(1)} \cdots \xrightarrow{\alpha(m)} s_{m+1} \rightarrow \cdots$$

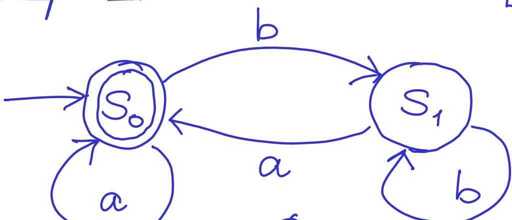
and $\inf(\sigma) \cap G \neq \varphi$.

Called an accepting run

Language accepted by $\mathcal{B}$

$$\mathcal{L}_\omega(\mathcal{B}) = \{ \alpha \in A^\omega \mid \mathcal{B} \text{ accepts } \alpha \}.$$

Example 1 $A = \{a, b\}$ $\quad L_1 = \mathcal{L}_\omega(\mathcal{B}_1) = \{ \alpha \in A^\omega \mid \exists^\omega a \in \alpha \}$

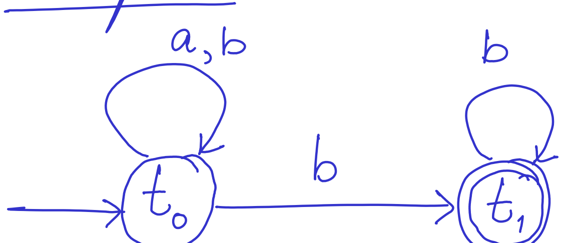


$\mathcal{B}_1$

$\mathcal{B}_1$ is a deterministic automaton

Example 2. Consider $L_2 = \overline{L_1}$ i.e. the complement of L_1

$= A^\omega - L_1$



$\mathcal{B}_2$

$\mathcal{L}_\omega(\mathcal{B}_2) = L_2$ i.e. it accepts every infinite word that has only a finite number of a's.

$\mathcal{B}_2$ is nondeterministic

ω-regular languages and ω-regular expressions.

An ω-regular expression is one formed the usual operators for regular expressions $(\varepsilon | a | r.s | r+s | r^*)$ along with the infinite repetition operation. $^{\omega}$. More compactly we may define an ω-regular expression as being of the form

$$\boxed{t = \sum_{i=1}^{n} r_i s_i^{\omega}}$$

where each r_i and s_i is a regular expression.

The language defined by an ω-regular expression

$$\boxed{\mathcal{L}_\omega(t) = \bigcup_{i=1}^{n} \mathcal{L}(r_i).\mathcal{L}(s_i)^{\omega}}$$

Equivalence of ω-regular expressions

$$t \equiv u \quad \text{iff} \quad \mathcal{L}_\omega(t) = \mathcal{L}_\omega(u).$$

ω-regular languages. $L \subseteq A^{\omega}$ is ω-regular iff there exists an ω-regular expression t such that $L = \mathcal{L}_\omega(t)$.

Example 1. $L_1 = \mathcal{L}_\omega\left((b^*a)^{\omega} + (b^*a)^* a^{\omega}\right)$

Example 2. $L_2 = \mathcal{L}_\omega\left((a+b)^* b^{\omega}\right)$

<u>Limit languages</u>. Let $U \subseteq A^*$ be a language of finite strings, then $\lim(U) = \{\alpha \in A^\omega \mid \exists^\omega n \in \mathbb{N}\, [\alpha[0..n] \in U]\}$ is called the limit language of U.

<u>Fact</u>. $\alpha \in \lim(U)$ iff infinitely many prefixes of α belong to U.

<u>Theorem 1.</u> A language is accepted by a <u>deterministic</u> Büchi automaton iff it is the limit language of a regular language on the same alphabet.

$\vdash (\Rightarrow)$ Let $L \subseteq A^\omega$ be accepted by a DBA

$$\mathcal{D} = \langle S, A, \longrightarrow, s_0, G \rangle$$

with $G \subseteq S$. Treat G as the set of final states of a DFA and consider the language of finite words accepted by $\mathcal{D}$. Clearly $L = \lim(\mathcal{L}(\mathcal{D}))$.

$(\Leftarrow)$ Conversely for $L = \lim(U)$ where $U \subseteq A^*$ is regular. there must exist a DFA $\mathcal{D} = \langle S, A, \rightarrow, s_0, F \rangle$ such that $\mathcal{L}(\mathcal{D}) = U$. Treating the set F as the set G we can use the same automaton as a Büchi automaton which accepts L. $\dashv$

Theorem (NBAs are more powerful than DBAs). There exists no DBA which accepts the language $L_2 = \mathcal{L}_\omega((a+b)^* b^\omega)$.

$\vdash$ Suppose there exists a DBA accepting L_2. Then there exists a a regular language U such that $L_2 = \lim(U)$.

Since $b^\omega \in L_2$, there exists a finite prefix $b^{n_1} \in U$. Further since $b^{n_1} a b^\omega \in L_2$, there exists a finite prefix $b^{n_1} a b^{n_2} \in U$.

Continuing in this fashion there exist an infinite sequence of $\mathbb{N}$, $M = \{n_1, n_2, \ldots\}$ such that for all $m > 0$,

$$b^{n_1} a b^{n_2} a \ldots a b^{n_m} a b^\omega \in L_2$$

and

$$\{ b^{n_1} a b^{n_2} a \ldots a b^{n_m} \mid m \in M \} \subseteq U$$

But from this it follows that the infinite word $\alpha \in L_2$ where $\alpha = b^{n_1} a b^{n_2} a \ldots a b^{n_m} a b^{n_{m+1}} a \ldots$. But α has an infinite number of occurrences of the letter 'a', which is impossible since $L_2 = \overline{L_1}$ where L_1 is the ω-regular lang with an infinite number of occurrences of 'a'. $\dashv$

<u>Theorem</u>. $L \subseteq A^\omega$ is ω-regular iff it is accepted by a NBA.

$\vdash (\Leftarrow)$ Assume L is accepted by a NBA $\mathcal{B} = \langle S, A, \to, I, G \rangle$.
For any $\alpha \in L$, the accepting run has an infinite number of occurrences of some $g \in G$. For any states s, t let
$$U_{st} = \{ u \in A^* \mid s \xrightarrow{u} t \}.$$

<u>Claim</u>. U_{st} is regular

$\vdash$ Use the reachable sub-automaton with $S' \subseteq S$
$$\langle S', A, \to, s, \{t\} \rangle$$
$\dashv$

Now there exists a regular expression $r_{U_{st}}$ defining this language. Similarly for any good state g that is visited infinitely often there exists a language
$$V_{gg} = \{ v \in A^* \mid g \xrightarrow{v} g \}$$
and a corresponding regular expression $S_{V_{gg}}$.

Clearly then we have
$$L = \mathcal{L}_\omega(\mathcal{B}) = \sum_{\substack{s \in I \\ g \in G}} r_{sg} \cdot S_{gg}$$
which is ω-regular

($\Rightarrow$) Assume L is ω-regular. Then clearly there exist regular expressions r_i, s_i such that

$$L = \mathcal{L}_\omega\left(\sum_{i=1}^{n} r_i . s_i^\omega \right)$$

<u>Claim 1.</u> If r is a regular expression then $\mathcal{L}_\omega(r^\omega)$ is Büchi recognizable

<u>Claim 2</u>. If r is regular and L is Büchi recognizable then $\mathcal{L}(r).L$ is Büchi recognizable

<u>Claim 3</u>. The union of Büchi-recognizable languages is Büchi recognizable.

From these claims it follows that L is Büchi recognizable.

$\dashv$