THEORY OF COMPUTATION
Semi-Thue Processs - 22

Prof. Dan A. Simovici

UMB

Definition

Given a pair of words $x, y \in A^*$ a semi-Thue production or simply a production is an expression of the form $x \to y$.

If P is the semi-Thue production $x \to y$ we write $u \underset{P}{\Rightarrow} v$ to mean that there are words r and s such that

$$u = rxs \text{ and } v = rys.$$

In other words, v is obtained from u by replacement of x by y.

Axel Thue (19 February 1863-7 March 1922) was a Norwegian mathematician known for his work in approximation theory and combinatorics.

Thue was a professor of applied mathematics at the University of Oslo from 1903 to his death. Many of Axel Thue ideas and discoveries have been influential much later in the foundations of Computer Science.

Definition

A semi-Thue process is a finite set of semi-Thue productions.

If Π is a semi-Thue process we write $u \underset{\Pi}{\Rightarrow} v$ to mean that $u \underset{P}{\Rightarrow} v$ for some production P that belongs to Π.

We write

$$u \overset{*}{\underset{\Pi}{\Rightarrow}} v$$

if there is a sequence

$$u = u_1 \underset{\Pi}{\Rightarrow} u_2 \underset{\Pi}{\Rightarrow} \cdots u_{n-1} \underset{\Pi}{\Rightarrow} u_n = v.$$

The sequence $u_1, u_2, \ldots, u_n$ is a derivation of v from u. In particular, taking $n = 1$ we always have

$$u \overset{*}{\underset{\Pi}{\Rightarrow}} u.$$

When no ambiguity results we may omit the reference to Π and write $u \Rightarrow v$ or $u \stackrel{*}{\Rightarrow}$.

Example

Let

$$\Pi = \{ab \to aa, ba \to bb\}.$$

Then, we have

$$aba \Rightarrow abb \Rightarrow aab \Rightarrow aaa.$$

Thus, $aba \stackrel{*}{\Rightarrow} aaa$, and the sequence

$$aba, abb, aab, aaa$$

is a derivation of aaa from aba.

Let $\mathcal{M}$ be a nondeterministic TM with the alphabet $\{s_1, \ldots, s_K\}$ and states $q_1, q_2, \ldots, q_n$. We show how to simulate $\mathcal{M}$ by a semi-Thue process $\Sigma(\mathcal{M})$ on the alphabet

$$s_1, \ldots, s_K, q_1, q_2, \ldots, q_n, q_{n+1}, h.$$

Each state in a computation by $\mathcal{M}$ is specified completely by the current configuration.

Example

The configuration

$$\begin{array}{c} s_1 s_1 s_3 \underset{\substack{\uparrow \\ q_4}}{s_2} s_0 s_1 s_2 \end{array}$$

will be represented as a single word

$$h s_1 s_1 s_3 q_4 s_2 s_0 s_1 s_2 h$$

h is used as a beginning and end marker, and q_4 indicates the state of $\mathcal{M}$ and is placed immediately at the left of the scanned square.

Definition

A Post word is a word of the form

$$huq_ivh$$

where $1 \leqslant i \leqslant n+1$ and $u, v \in \{s_0, s_1, \ldots, s_K\}^*$.

A configuration can be represented by infinitely many Post words because any number of additional blanks may be placed at the left or the right.

Example

The configuration

$$\begin{array}{ccccccc} s_1 & s_1 & s_3 & s_2 & s_0 & s_1 & s_2 \\ & & & \uparrow & & & \\ & & & q_4 & & & \end{array}$$

may be represented by $hs_0s_0s_1s_1s_3q_4s_2s_0s_1s_2s_0s_0s_0h$.

Associating suitable semi-Thue productions with the quadruples of a TM $\mathcal{M}$ aims to simulate the effect of quadruples on Post words.

Quadruple	semi-Thue Production
$q_i \ s_j \ s_k \ q_\ell$	$q_i s_j \to q_\ell s_k$
$q_i \ s_j \ R \ q_\ell$	$q_i s_j s_k \to s_j q_\ell s_k, 0 \leqslant k \leqslant K$
	$q_i s_j h \to s_j q_\ell s_0 h$
$q_i \ s_j \ L \ q_\ell$	$s_k q_i s_j \to q_\ell s_k s_j, 0 \leqslant k \leqslant K$
	$h q_i s_j \to h q_\ell s_0 s_j$

TM moves vs. rewriting in Π:

If we have in TM $q_i \; s_j \; s_k \; q_\ell$ and the production $q_i s_j \to q_\ell s_k$ in Π then

$$\begin{array}{c} \cdots s_j \cdots \\ \uparrow \\ q_i \end{array} \quad \vdash \quad \begin{array}{c} \cdots s_k \cdots \\ \uparrow \\ q_\ell \end{array}$$

and

$$h u q_i s_j v h \underset{\Pi}{\Rightarrow} h u q_\ell s_k v h$$

If $q_i\ s_j\ R\ q_\ell$ is a quadruple in $\mathcal{M}$ and $q_i s_j s_k \to s_j q_\ell s_k$ or $q_i s_j h \to s_j q_\ell s_0 h$ are semi-Thue productions then in the TM we have:

$$\begin{array}{ccc} \cdots \underset{\underset{q_i}{\uparrow}}{s_j} s_k \cdots & \vdash & \cdots s_j \underset{\underset{q_\ell}{\uparrow}}{s_k} \cdots \end{array}$$

while in Π:

$$huq_i s_j s_k vh \underset{\Pi}{\Rightarrow} hus_j q_\ell s_k vh$$

or

$$huq_i s_j h \underset{\Pi}{\Rightarrow} hus_j q_\ell s_k s_0 h$$

Example

Suppose that $\mathcal{M}$ is in the configuration

$$\begin{array}{l} s_2\,s_1\,s_0\,s_3 \\ \quad\;\uparrow \\ \quad\; q_4 \end{array}$$

represented by the Post word $hs_2q_4s_1s_0s_3h$.
Suppose that $\mathcal{M}$ contains the quadruple $q_4\ s_1\ s_3\ q_5$. Then, $\Sigma(\mathcal{M})$ contains the production $q_4s_1 \to q_5s_3$ so that

$$hs_2q_4s_1s_0s_3h \underset{\Sigma(\mathcal{M})}{\Rightarrow} hs_2q_5s_3s_0s_3h.$$

If $\mathcal{M}$ contains the quadruple $q_4\ s_1\ L\ q_2$, then $\Pi = \Sigma(\mathcal{M})$ contains the production $s_2 q_4 s_1 \to q_2 s_2 s_1$ so that

$$hs_2 q_4 s_1 s_0 s_3 h \underset{\Sigma(\mathcal{M})}{\Rightarrow} hq_2 s_2 s_1 s_0 s_3 h.$$

To complete the specification of $\Sigma(\mathcal{M})$:

- whenever $q_i s_j$ are not the first two symbols of a quadruple of $\mathcal{M}$ we place in $\Sigma(\mathcal{M})$ the production $q_i s_j \rightarrow q_{n+1} s_j$. Thus, q_{n+1} serves as "halt" state.
- Finally, we place in $\Sigma(\mathcal{M})$ the productions:

$$q_{n+1} s_i \rightarrow q_{n+1}, 0 \leqslant i \leqslant K,$$
$$q_{n+1} h \rightarrow q_0 h,$$
$$s_i q_0 \rightarrow q_0, 0 \leqslant i \leqslant K.$$

Theorem

Let $\mathcal{M}$ be a deterministic TM and let w be a Post word on the alphabet of $\Sigma(\mathcal{M})$. Then,

- *there is at most one word z such that $w \underset{\Sigma(\mathcal{M})}{\Rightarrow} z$, and*
- *if there is a word z satisfying the above condition, then z is a Post word.*

Proof

Let $w = huq_i vh$. For $1 \leqslant i \leqslant n$ we have:

- if $v = 0$ no production of $\Sigma(\mathcal{M})$ applies to w;
- if v begins with s_j and there is a (necessarily unique) quadruple that begins with $q_i s_j$, then there exists a unique applicable production of $\Sigma(\mathcal{M})$ and the result will be a Post word;
- if v begins with s_j and there is no quadruple that begins with $q_i s_j$, then the one applicable production is $q_i s_j \rightarrow q_{n+1} s_j$, that yields another Post word.

Proof cont'd

If $i = n + 1$, then:

- if $v = 0$, the only applicable production of $\Sigma(\mathcal{M})$ is $q_{n+1}h \to q_0 h$, which yields a Post word;
- if v begins with a symbol s_j, the only applicable production is $q_{n+1}s_j \to q_{n+1}$. which again produces a Post word.

Proof cont'd

Finally, if $i = 0$, then

- if $u = 0$, no production of $\Sigma(\mathcal{M})$ is applicable;
- if u ends with s_j the only applicable production of $\Sigma(\mathcal{M})$ is $s_j q_0 \to q_0$, which yields a Post word.

Theorem

Let $\mathcal{M}$ be a nondeterministic TM. For each string u on the alphabet of $\mathcal{M}$, $\mathcal{M}$ accepts u if and only if

$$hq_1 s_0 uh \underset{\Sigma(\mathcal{M})}{\overset{*}{\Rightarrow}} hq_0 h.$$

Proof

Let $\{s_1, \ldots, s_K\}$ be the alphabet of $\mathcal{M}$. First, suppose that $\mathcal{M}$ accepts u. If $\mathcal{M}$ begins in the configuration

$$\begin{array}{l} s_0\, u \\ \uparrow \\ q_1 \end{array}$$

it will eventually reach a state q_i scanning a symbol s_k where no quadruple of $\mathcal{M}$ begins with $q_i s_k$. Then, we will have

$$\begin{aligned} hq_1 s_0 uh &\overset{*}{\underset{\Sigma(\mathcal{M})}{\Rightarrow}} hvq_i s_k wh \\ &\underset{\Sigma(\mathcal{M})}{\Rightarrow} hvq_{n+1} s_k wh \\ &\overset{*}{\underset{\Sigma(\mathcal{M})}{\Rightarrow}} hvq_0 h \\ &\overset{*}{\underset{\Sigma(\mathcal{M})}{\Rightarrow}} hq_0 h. \end{aligned}$$

Proof cont'd

Suppose now that $\mathcal{M}$ does not accept u. Then, beginning with the configuration

$$\underset{q_1}{\overset{\uparrow}{s_0}}\, u$$

$\mathcal{M}$ will never halt. Let $w_1 = hq_1s_0uh$ and suppose that

$$w_1 \underset{\Sigma(\mathcal{M})}{\Rightarrow} w_2 \underset{\Sigma(\mathcal{M})}{\Rightarrow} w_3 \cdots \underset{\Sigma(\mathcal{M})}{\Rightarrow} w_m.$$

Then each w_j, $1 \leqslant j \leqslant m$ must contain a symbol q_i with $1 \leqslant i \leqslant n$. Hence, there can be no derivation of a Post word containing q_0 from w_1, so, in particular there is no derivation of hq_0h from w_1.

Definition

Let Π be a semi-Thue process. The inverse of a production $x \rightarrow y$ is the production $y \rightarrow x$.

Example

The inverse of the production $aab \rightarrow ba$ is the production $ba \rightarrow aab$.

Let $\Omega(\mathcal{M})$ be the semi-Thue process that consists of the inverses of all productions of $\Sigma(\mathcal{M})$.
Also, denote the set $\Sigma(\mathcal{M}) \cup \Omega(\mathcal{M})$ as $\Theta(\mathcal{M})$.

Theorem

Let $\mathcal{M}$ be a nondeterministic TM. For each u in the alphabet of $\mathcal{M}$, $\mathcal{M}$ accepts u if and only if

$$hq_0h \underset{\Omega(\mathcal{M})}{\overset{*}{\Rightarrow}} hq_1s_0uh.$$

Proof: This is an immediate consequence of the previous theorem.