1. Semi-Thue Processes

2. TMs and Semi-Thue Processes
Definition

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

If $P$ is the semi-Thue production $x \rightarrow y$ we write $u \xrightarrow{P} v$ 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 \( \Pi \) is a semi-Thue process we write \( u \xrightarrow{\Pi} v \) to mean that \( u \xrightarrow{P} v \) for some production \( P \) that belongs to \( \Pi \).
We write

\[ u \xrightarrow{\ast} \Pi v \]

if there is a sequence

\[ u = u_1 \xrightarrow{\Pi} u_2 \xrightarrow{\Pi} \cdots u_{n-1} \xrightarrow{\Pi} u_n = v. \]

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

\[ u \xrightarrow{\ast} \Pi u. \]
When no ambiguity results we may omit the reference to $\Pi$ and write $u \Rightarrow v$ or $u \Rightarrow^*$.

**Example**

Let

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

Then, we have

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

Thus, $aba \Rightarrow^* aaa$, 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

\[ s_1 s_1 s_3 s_2 s_0 s_1 s_2 \]

upward

\[ q_4 \]

will be represented as a single word

\[ hs_1 s_1 s_3 q_4 s_2 s_0 s_1 s_2 h \]

\( h \) is used as a **beginning** and **end** marker, and \( q_4 \) indicates the state of \( M \) and is placed immediately at the left of the scanned square.
Definition

A Post word is a word of the form

\[ huq_i vh \]

where \( 1 \leq i \leq n + 1 \) 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

\[ s_1 s_1 s_3 s_2 s_0 s_1 s_2 \]

\[ \uparrow \]

\[ q_4 \]

may be represented by \( hs_0 s_0 s_1 s_1 s_3 q_4 s_2 s_0 s_1 s_2 s_0 s_0 s_0 h \).
Associating suitable semi-Thue productions with the quadruples of a TM $M$ aims to simulate the effect of quadruples on Post words.

<table>
<thead>
<tr>
<th>Quadruple</th>
<th>semi-Thue Production</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q_i s_j s_k q_\ell$</td>
<td>$q_i s_j \rightarrow q_\ell s_k$</td>
</tr>
</tbody>
</table>
| $q_i s_j R q_\ell$  | $q_i s_j s_k \rightarrow s_j q_\ell s_k$, $0 \leq k \leq K$
|                  | $q_i s_j h \rightarrow s_j q_\ell s_0 h$ |
| $q_i s_j L q_\ell$  | $q_\ell s_k s_j \rightarrow s_0 q_\ell s_k$, $0 \leq k \leq K$
|                  | $hq_i s_j \rightarrow hq_\ell s_0 s_j$ |
TM moves vs. rewriting in $\Pi$:
If we have in TM $q_i s_j s_k q_\ell$ and the production $q_i s_j \rightarrow q_\ell s_k$ in $\Pi$ then
\[
\cdots s_j \cdots \uparrow q_i \vdash \cdots s_k \cdots \uparrow q_\ell
\]
and

\[
huq_i s_j vh \Rightarrow_{\Pi} huq_\ell s_k vh
\]
If $q_i s_j R q_\ell$ is a quadruple in $M$ and $q_i s_j s_k \rightarrow s_j q_\ell s_k$ or $q_i s_j h \rightarrow s_j q_\ell s_0 h$ are semi-Thue productions then in the TM we have:

\[ \cdots s_j s_k \cdots \quad \vdash \quad \cdots s_j s_k \cdots \]

while in $\Pi$:

\[ h u q_i s_j s_k v h \Rightarrow h u s_j q_\ell s_k v h \]

or

\[ h u q_i s_j h \Rightarrow h u s_j q_\ell s_k s_0 h \]
Example

Suppose that $M$ is in the configuration

\[ s_2 s_1 s_0 s_3 \]

represented by the Post word $hs_2 q_4 s_1 s_0 s_3 h$.

Suppose that $M$ contains the quadruple $q_4 s_1 s_3 q_5$. Then, $\Sigma(M)$ contains the production $q_4 s_1 \rightarrow q_5 s_3$ so that

\[ hs_2 q_4 s_1 s_0 s_3 h \Rightarrow_{\Sigma(M)} hs_2 q_5 s_3 s_0 s_3 h. \]
If $M$ contains the quadruple $q_4 s_1 L q_2$, then $\Pi = \Sigma(M)$ contains the production $s_1 q_4 s_1 \to q_2 s_2 s_1$ so that

$$hs_2 q_4 s_1 s_0 s_3 h \Rightarrow_{\Sigma(M)} h q_2 s_2 s_1 s_0 s_3 h.$$
To complete the specification of $\Sigma(\mathcal{M})$:

- Whenever $q_is_j$ are not the first two symbols of a quadruple of $\mathcal{M}$ we place in $\Sigma(\mathcal{M})$ the production $q_is_j \rightarrow q_{n+1}s_j$. Thus, $q_{n+1}$ serves as “halt” state.

- Finally, we place in $\Sigma(\mathcal{M})$ the productions:

$$q_{n+1}s_i \rightarrow q_{n+1}, 0 \leq i \leq K,$$
$$q_{n+1}h \rightarrow q_0h,$$
$$s_iq_0 \rightarrow q_0, 0 \leq i \leq K.$$
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 $w \Rightarrow^* z$, and
- if there is a word $z$ satisfying the above condition, then $z$ is a Post word.
Proof

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

- if $v = 0$ no production of $\Sigma(M)$ 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 $\Sigma(M)$ 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.
If \( i = n + 1 \), then:

- if \( v = 0 \), the only applicable production of \( \Sigma(M) \) is \( q_{n+1}h \rightarrow q_0h \), which yields a Post word;

- if \( v \) begins with a symbol \( s_j \), the only applicable production is \( q_{n+1}s_j \rightarrow q_{n+1} \), which again produces a Post word.
Finally, if $i = 0$, then

- if $u = 0$, no production of $\Sigma(M)$ is applicable;
- if $u$ ends with $s_j$ the only applicable production of $\Sigma(M)$ is $s_jq_0 \rightarrow q_0$, which yields a Post word.
Theorem

Let \( M \) be a nondeterministic TM. For each string \( u \) on the alphabet of \( M \), \( M \) accepts \( u \) if and only if

\[
\begin{align*}
hq_1s_0uh \quad \Rightarrow^*_{\Sigma(M)} \quad hq_0h.
\end{align*}
\]
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 \( s_0 u \) 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

\[
\begin{align*}
hq_1 s_0 u h & \xrightarrow{\ast} h v q_i s_k w h \\
& \xrightarrow{\Sigma(\mathcal{M})} h v q_{n+1} s_k w h \\
& \xrightarrow{\ast} h v q_0 h \\
& \xrightarrow{\Sigma(\mathcal{M})} h q_0 h.
\end{align*}
\]
Proof cont’d

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

$$s_0 \ u$$

$q_1$

$M$ will never halt. Let $w_1 = hq_1s_0uh$ and suppose that

$$w_1 \Rightarrow_{\Sigma(M)} w_2 \Rightarrow_{\Sigma(M)} w_3 \cdots \Rightarrow_{\Sigma(M)} w_m.$$  

Then each $w_j$, $1 \leq j \leq m$ must contain a symbol $q_i$ with $1 \leq i \leq n$. 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 $\Pi$ 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 $\Omega(M)$ be the semi-Thue process that consists of the inverses of all productions of $\Sigma(M)$.
Also, denote the set $\Sigma(M) \cup \Omega(M)$ as $\Theta(M)$.

**Theorem**

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

$$hq_0 h \xrightarrow{\ast_{\Omega(M)}} hq_1 s_0 uh.$$  

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