semi-Thue system

Canonical name	SemiThueSystem
Date of creation	2013-03-22 17:33:16
Last modified on	2013-03-22 17:33:16
Owner	CWoo (3771)
Last modified by	CWoo (3771)
Numerical id	21
Author	CWoo (3771)
Entry type	Definition
Classification	msc 68Q42
Classification	msc 03D40
Classification	msc 20M35
Classification	msc 03D03
Synonym	rewriting rule
Synonym	rewrite rule
Synonym	rewriting system
Synonym	semi-Thue generable
Related topic	FormalGrammar
Related topic	LabelledStateTransitionSystem
Defines	antecedent
Defines	consequent
Defines	immediately derivable
Defines	derivable
Defines	defining relation
Defines	word problem for semi-Thue systems
Defines	semi-Thue production
Defines	generable by a semi-Thue system

A *semi-Thue system* $\mathfrak{S}$ is a pair (Σ, P) where Σ is an alphabet and P is a non-empty finite binary relation on Σ^*, the Kleene star of Σ.

Elements of P are variously called *defining relations*, *productions*, or *rewrite rules*, and $\mathfrak{S}$ itself is also known as a *rewriting system*. If $(x, y) \in P$, we call x the *antecedent*, and y the *consequent*. Instead of writing $(x, y) \in P$ or xPy, we usually write

$$x \to y.$$

Let $\mathfrak{S} = (\Sigma, P)$ be a semi-Thue system. Given a word u over Σ, we say that a word v over Σ is *immediately derivable* from u if there is a defining relation $x \to y$ such that

$$u = rxs \qquad \text{and} \qquad v = rys,$$

for some words r, s (which may be empty) over Σ. If v is immediately derivable from u, we write

$$u \Rightarrow v.$$

Let P' be the set of all pairs $(u, v) \in \Sigma^* \times \Sigma^*$ such that $u \Rightarrow v$. Then $P \subseteq P'$, and

$$\text{If } u \Rightarrow v, \text{ then } wu \Rightarrow wv \text{ and } uw \Rightarrow vw \text{ for any word } w.$$

Next, take the reflexive transitive closure P'' of P'. Write $a \overset{*}{\Rightarrow} b$ for $(a, b) \in P''$. So $a \overset{*}{\Rightarrow} b$ means that either $a = b$, or there is a finite chain $a = a_1, \ldots, a_n = b$ such that $a_i \Rightarrow a_{i+1}$ for $i = 1, \ldots, n-1$. When $a \overset{*}{\Rightarrow} b$, we say that b is *derivable* from a. Concatenation preserves derivability:

$$a \overset{*}{\Rightarrow} b \text{ and } c \overset{*}{\Rightarrow} d \text{ imply } ac \overset{*}{\Rightarrow} bd.$$

Example. Let $\mathfrak{S}$ be a semi-Thue system over the alphabet $\Sigma = \{a, b, c\}$, with the set of defining relations given by $P = \{ab \to bc, bc \to cb\}$. Then words ac^3b, a^2c^2b and as bc^4 are all derivable from a^2bc^2:

- $a^2bc^2 \Rightarrow a(bc)c^2 \Rightarrow ac(bc)c \Rightarrow ac^2(cb) = ac^3b$,
- $a^2bc^2 \Rightarrow a^2(cb)c \Rightarrow a^2c(cb) = a^2c^2b$, and
- $a^2bc^2 \Rightarrow a(bc)c^2 \Rightarrow (bc)cc^2 = bc^4$.

Under $\mathfrak{S}$, we see that if v is derivable from u, then they have the same length: $|u| = |v|$. Furthermore, if we denote $|a|_u$ the number of occurrences of letter a in a word u, then $|a|_v \leq |a|_u$, $|c|_v \geq |c|_u$, and $|b|_v = |b|_u$. Also, in order for a word u to have a non-trivial word v (non-trivial in the sense that $u \neq v$) derivable from it, u must have either ab or bc as a subword. Therefore, words like a^3 or $c^3b^4a^2$ have no non-trivial derived words from them.

Remarks.

1. Given a semi-Thue system $\mathfrak{S} = (\Sigma, P)$, one can associate a subset A of Σ^* whose elements we call *axioms* of $\mathfrak{S}$. Any word v that is derivable from an axiom $a \in A$ is called a *theorem* (of $\mathfrak{S}$). If v is a theorem, we write $A \vdash_{\mathfrak{S}} v$. The set of all theorems is written $L_{\mathfrak{S}}(A)$, and is called the language (over Σ) generated by A.

2. Let $\mathfrak{S}$ and A be defined as above, and T any alphabet. Call the elements of $T \cap \Sigma$ the *terminals* of $\mathfrak{S}$. The set
$$L_{\mathfrak{S}}(A) \cap T^*$$
is called the *language generated by A over T*, and written $L_{\mathfrak{S}}(A, T)$. It is easy to see that $L_{\mathfrak{S}}(A, T) = L_{\mathfrak{S}}(A, T \cap \Sigma)$.

3. A language L over an alphabet Σ is said to be *generable by a semi-Thue system* if there is a semi-Thue system $\mathfrak{S}$ and a finite set of axioms A of $\mathfrak{S}$ such that $L = L_{\mathfrak{S}}(A, \Sigma)$.

4. Semi-Thue systems are "equivalent" to formal grammars in the following sense:

   > a language is generable by a formal grammar iff it is semi-Thue generable.

   The idea is to turn every defining relation $x \to y$ in P into a production $SxT \to SyT$, where S and T are non-terminals or variables. As such, a production of the form $SxT \to SyT$ is sometimes called a *semi-Thue production*.

5. Given a semi-Thue system $\mathfrak{S} = (\Sigma, P)$, the *word problem* for $\mathfrak{S}$ asks whether or not for any pair of words u, v over Σ, one can determine in a finite number of steps (an algorithm) that $u \overset{*}{\Rightarrow} v$. If such an algorithm exists, we say that the word problem for $\mathfrak{S}$ is solvable. It turns out

there exists a semi-Thue system such that the word problem for it is unsolvable.

6. The word problem for a specific $\mathfrak{S}$ is the same as finding an algorithm to determine whether v is a theorem based on a singleton axiom $\{u\}$ for arbitrary words u, v.

7. *The word problem for semi-Thue systems* asks whether or not, given *any* semi-Thue system $\mathfrak{S}$, the word problem for $\mathfrak{S}$ is solvable. From the previous remark, we see the word problem for semi-Thue systems is unsolvable.

References

[1] M. Davis, *Computability and Unsolvability.* Dover Publications, New York (1982).

[2] H. Hermes, *Enumerability, Decidability, Computability: An Introduction to the Theory of Recursive Functions.* Springer, New York, (1969).