deletion operation on languages

Canonical name	DeletionOperationOnLanguages
Date of creation	2013-03-22 18:56:23
Last modified on	2013-03-22 18:56:23
Owner	CWoo (3771)
Last modified by	CWoo (3771)
Numerical id	8
Author	CWoo (3771)
Entry type	Definition
Classification	msc 68Q70
Classification	msc 68Q45
Synonym	deletion-closed
Synonym	deletion-closure
Related topic	ShuffleOfLanguages
Related topic	InsertionOperationOnLanguages
Defines	deletion
Defines	deletion closed
Defines	deletion closure

Let Σ be an alphabet and u, v be words over Σ. A *deletion* of v from u is a word of the form u_1u_2, where $u = u_1vu_2$. If w is a deletion of v from u, then u is an insertion of v into w.

The deletion of v from u is the set of all deletions of v from u, and is denoted by $u \longrightarrow v$.

For example, if $u = (ab)^6$ and $v = aba$, then

$$u \longrightarrow v = \{(ba)^4b, ab(ba)^3b, (ab)^2(ba)^2b, (ab)^3bab, (ab)^4b\}.$$

More generally, given two languages L_1, L_2 over Σ, the *deletion* of L_2 from L_1 is the set

$$L_1 \longrightarrow L_2 := \bigcup\{u \longrightarrow v \mid u \in L_1, v \in L_2\}.$$

For example, if $L_1 = \{(ab)^n \mid n \geq 0\}$ and $L_2 = \{a^nb^n \mid n \geq 0\}$, then $L_1 \longrightarrow L_2 = L_1$, and $L_2 \longrightarrow L_1 = L_2$. If $L_3 = \{a^n \mid n \geq 0\}$, then $L_2 \longrightarrow L_3 = a^mb^n \mid n \geq m \geq 0\}$, and $L_3 \longrightarrow L_2 = L_3$.

A language L is said to be *deletion-closed* if $L \longrightarrow L \subseteq L$. L_1, L_2, and L_3 from above are all deletion closed, as well as the most obvious example: Σ^*. $L_2 \longrightarrow L_3$ is not deletion closed, for $a^3b^4 \longrightarrow ab^3 = \{a^2b\} \nsubseteq L_2 \longrightarrow L_3$.

It is easy to see that arbitrary intersections of deletion-closed languages is deletion-closed.

Given a language L, the intersection of all deletion-closed languages containing L, denoted by $\mathrm{Del}(L)$, is called the *deletion-closure* of L. In other words, $\mathrm{Del}(L)$ is the smallest deletion-closed language containing L.

The deletion-closure of a language L can be constructed from L, as follows:

$$\begin{aligned}
L_0 &:= L \\
L_{n+1} &:= L_n \longrightarrow (L_n \cup \{\lambda\}) \\
L' &:= \bigcup_{i=0}^{\infty} L_i
\end{aligned}$$

Then $\mathrm{Del}(L) = L'$.

For example, if $L = \{a^p \mid p \text{ is a prime number }\}$, then $\mathrm{Del}(L) = \{a^n \mid n \geq 0\}$, and if $L = \{a^mb^n \mid m \geq n \geq 0\}$, then $\mathrm{Del}(L) = \{a^mb^n \mid m, n \geq 0\}$.

Remarks.

- If L_1, L_2 are regular languages, so is $L_1 \longrightarrow L_2$. In other words, the family $\mathscr{R}$ of regular languages is closed under the deletion binary operation.
- In addition, $\mathscr{R}$ is closed under deletion closure: if L is regular, so is $\mathrm{Del}(L)$.
- However, $\mathscr{F}$, the family of context-free languages, is neither closed under deletion, nor under deletion closure.

References

[1] M. Ito, *Algebraic Theory of Automata and Languages*, World Scientific, Singapore (2004).