locally testable

Canonical name	LocallyTestable
Date of creation	2013-03-22 18:59:03
Last modified on	2013-03-22 18:59:03
Owner	CWoo (3771)
Last modified by	CWoo (3771)
Numerical id	7
Author	CWoo (3771)
Entry type	Definition
Classification	msc 68Q45
Classification	msc 68Q42
Synonym	k-testable
Defines	k-testable

A regular language L over an alphabet Σ is *locally testable* if, loosely speaking, testing whether or not an arbitrary word u (over Σ) is in L is completely determined by its subwords of some fixed length. The name *locally testable* comes from the fact that properties of u, and not L, determine the membership of u in L.

To formalize this notion, we first define, for any word u over Σ, the set $\mathrm{sw}_k(u)$ of all subwords of

$$\#u\#$$

of length k, where $\#$ is a symbol not in Σ.

Definition. We say that a regular language L is *k-testable* if for any $u, v \in \Sigma^*$ such that

$$\mathrm{sw}_k(u) = \mathrm{sw}_k(v),$$

we have $u \in L$ iff $v \in L$. The equation above says three things at once:

- the set of subwords of u of length k is equal to the set of subwords of v of length k,
- the prefix of u of length k is equal to the prefix of v of length k, and
- the suffix of u of length k is equal to the suffix of v of length k.

We say that L is *locally testable* if it is k-testable for some $k \geq 0$.

If we denote $\mathscr{T}(k)$ the family of all k-testable languages, and $\mathscr{T}(\infty)$ the family of all locally testable languages, then

$$\mathscr{T}(\infty) = \bigcup_{k=0}^{\infty} \mathscr{T}(k).$$

Note that there are only two 0-testable languages: Σ^* and $\varnothing$.

Proposition 1. *Let $\mathscr{D}$ be the family of definite languages. Then*

1. *$\mathscr{T}(k) \subset \mathscr{T}(k+1)$ for any $k \geq 0$, and the inclusion is strict.*
2. *$\mathscr{D}$ and $\mathscr{T}(k)$ are not comparable for any $k > 0$. In other words, for every k, there is a k-testable language that is not definite, and a definite language that is not k-testable.*
3. *$\mathscr{D} \subset \mathscr{T}(\infty)$, and the inclusion is strict.*

We only sketch the proof here. For the first assertion, note that for every $k \geq 0$,

$$\mathrm{sw}_{k+1}(u) = \mathrm{sw}_{k+1}(v) \quad \text{implies} \quad \mathrm{sw}_k(u) = \mathrm{sw}_k(v).$$

In addition, the language $\{ab^k\}^*$ is $(k+1)$-testable but not k-testable. For the second statement, note that $\{ab^k\}^*$ is not definite for any $k \geq 0$. On the other hand, a finite language containing a single word of length $k+1$ is not k-testable. The last assertion is a direct consequence of the second one.

Thus, the families $\mathscr{T}(k)$ provide us with an example of an infinite chain of subfamilies of the family of regular languages.

With regard to the closure properties in $\mathscr{T}(k)$, it is easily see that $\mathscr{T}(k)$ for all $k \geq 0$ including $k = \infty$, is closed under complementation and intersection, and hence all Boolean operations. The star-closure of $\mathscr{T}(\infty)$ is $\mathscr{R}$, the family of all regular languages.

Remark. Every locally testable language is star-free, but not conversely.

References

[1] A. Salomaa, *Formal Languages*, Academic Press, New York (1973).