EQUATIONS IN WORDS : AN ALGORITHMIC CONTRIBUTION

Jean Néraud

Abstract

We study the special class of equations in words of type (R, w), where R is a two variable generalized regular expression, without constant, and where w is a constant word. We show that the problem may be solved by applying a $O(|w| \ln^2 |w|)$ time algorithm.

1 Introduction

In Combinatoric on words, the question of deciding whether an arbitrary word (or, equivalently, all the words in a finite family) belongs to a given recursive language L takes a prominent part for the problem it generates. Indeed, in spite of the simplicity of the preceding statement, practical conditions lead to various problems, with a large range of corresponding computational complexity [13]. In a first hand, several general problems are known to be NP-complete, even undecidable, and in another hand, with special instances, famous efficient algorithms have been implemented. Actually, between these two poles, there exists a large gap of open problems. This feature is particularly well illustrated when considering the framework of pattern matching.

The most famous example is certainly the so-called "string matching" problem, which consists in deciding whether a given word u appears as factor in a given "text" w. In other words, the question consists in deciding whether $w \in \Sigma^* u \Sigma^*$, where Σ^* stands for the free monoid generated by Σ, the basic alphabet. With this special case of instance, many famous linear-time algorithms have been implemented (cf e.g. [16], [7], [14]). Actually the implementation of new improvements remains a challenging question.

Another classical question corresponds to construct efficient membership tests to languages of type $L = \Sigma^* L(R) \Sigma^*$, where $L(R)$ stands for the set of all the words which are described by the regular expression R. In [25], an $O(|R||w|)$-time algorithm has been proposed for solving this problem (in a classical way, $|R|$ denotes the length of R). Moreover, with special classes of regular patterns, fast algorithms allow to compute all the occurrences of the pattern R in w (cf e.g. [12], [3], [9]).

Received by the editors March 1993, revised December 1993.
Communicated by M. Boffa.
AMS Mathematics Subject Classification : 68R15.
Keywords : free monoid, word, factor, period, morphism, equation, solution, pattern, algorithm, complexity.