

Burghardt, Jochen

A tree pattern matching algorithm with reasonable space requirements. (English)

Zbl 0645.68077

Trees in algebra and programming, Proc. 13th Colloq. CAAP, Nancy/France 1988, Lect. Notes Comput. Sci. 299, 1-15 (1988).

Summary: [For the entire collection see [Zbl 0635.00017](#).]

This paper presents an algorithm which solves the pattern matching problem for trees in linear matching time. The algorithm consists of two phases. In the first phase, a matching automaton is built from a given finite set of trees (pattern set). In the second phase, one or more subject trees are put into the automaton, yielding all positions where one of the patterns matches.

The size of the automaton as well as its construction time are theoretically bounded by $O(2^{\text{pattern size}})$, however, in most practical cases they grow only linear with the pattern size, which is better than O'Donnell's algorithm [(*) *Ch. M. Hoffmann and M. J. O'Donnell*, J. Assoc. Comput. Mach. 29, 68-95 (1982; [Zbl 0477.68067](#))].

Since the algorithm makes no use of specific tree properties (such as bottom up traversal, like (*)) it works in fact not only for trees but for a rather large class of objects called "symbol fields", which include for example n-dimensional arrays of almost arbitrary shapes, n-ary trees and strings with wildcards.

After giving some basic definitions in section 2, we introduce the algorithm for patterns without variables. In section 4 we explain the extensions to be made in order to deal with variables. Finally we give an outline of the correctness proof and the complexity estimation in sections 5 and 6.

MSC:

[68R10](#) Graph theory (including graph drawing) in computer science

[68Q25](#) Analysis of algorithms and problem complexity

Cited in 1 Document

Keywords:

pattern matching; linear matching time; matching automaton; n-dimensional arrays; n-ary trees