Algorithms for matching patterns with variables

Dependable Systems - Dirk Nowotka

Project Description

A *pattern* is a string (word) that consists of *terminal symbols* (e. g., $\mathtt{a}, \mathtt{b}, \mathtt{c}$), treated as constants, and *variables* (e. g., x_1, x_2, x_3). A pattern is mapped to a word by substituting the variables by strings of terminals. For example, $x_1 x_1 \mathtt{bab} x_2 x_2$ can be mapped to $\mathtt{acacbabcc}$ or $\mathtt{ccbabaa}$ by the substitution $(x_1 \to \mathtt{ac}, x_2 \to \mathtt{c})$ and $(x_1 \to \mathtt{c}, x_2 \to \mathtt{a})$, respectively. If a pattern α can be mapped to a string of terminals w, we say that α matches w.

It is often necessary to solve efficiently the *matching problem*: given a pattern α and a string w, does α match w? Unfortunately, the *matching problem* is $\mathbf{NP}$-complete in general. This is especially bad for some computational tasks on patterns (e.g., in algorithmic learning theory) which implicitly solve the matching problem and are therefore also intractable.

However, there are several classes of patterns, defined by structural properties, for which the matching problem can be solved in polynomial time. For instance, the *non-cross patterns* patterns can be matched in polynomial time; in these patterns no occurrence of a variable $x_2 \neq x_1$ is allowed between any two occurrences of the variable x_1. While showing that the matching problem for a class of patterns can be solved in polynomial time is sometimes simple, designing a very efficient algorithm solving this problem can be more difficult. As such, the goals of this project are:

- to implement the known efficient matching algorithms for several known classes of patterns;
- to identify new classes of patterns with variables which can be matched in polynomial time;
- to design (and implement) for any such new class efficient matching algorithms.

Applicable For

- Bachelorstudents ☑
- Masterstudents ☑

Keywords

Pattern Matching
Algorithms
Data Structures

Contact

Dr. Joel Day, Dr. Florin Manea

@ {jda,flm}@informatik.uni-kiel.de