Computer Engineering Lab III (Software)
240-301, Semester 1, 2006-2007

Pattern Matching

Objectives

To introduce students to the string pattern matching algorithms: brute-force, Boyer-Moore, and Knuth-Morris-Pratt.

Background

Background information is in the Powerpoint slides "Pattern Matching". Online copies of the slides and this lab exercise can be found at `http://fivedots.coe.psu.ac.th/Software.coe/LAB/PatMatch`

Learning Aims

Apply the Boyer-Moore and Knuth-Morris-Pratt algorithms.

Required Software: J2SE or a C compiler

Checkpoint 1. Compile and run the Java **or** the C version of the brute-force algorithm.

Implementations of all the algorithms can be found at `http://fivedots.coe.psu.ac.th/Software.coe/LAB/PatMatch/Code` in the subdirectories `/Java` and `/C`. The Java brute-force file is called `BruteSearch.java`, the C file is called `brute.c`.

Details on how to use Java is in the `JavaReadMe.txt` file stored with the pattern matching code in `/Java`. Details on how to use C is in the `cReadMe.txt` file stored with the pattern matching code in `/C`.

Exercise 1

Give the last occurrence function table for the Boyer-Moore algorithm when using the pattern:

"`ststvss`"

The alphabet consists of the lowercase letters {s, t, u, v, w}.

Explain your answer in words (please use English).

Exercise 2

Show how the "`ststvss`" pattern is moved when it is matched using Boyer-Moore against the string:

"`stvstvstvsss`"

Explain your answer in words (please use English).

Exercise 3

Give the failure function for the Knuth-Morris-Pratt algorithm when using the pattern:

"`aiain`"

Explain your answer in words (please use English).

Exercise 4

Show how the "`aiain`" pattern is moved when it is matched using Knuth-Morris-Pratt against the string:

"`aiaiabcaiaiaiain`"

Explain your answer in words (please use English).

IMPORTANT NOTE: Do **not** submit your lab report to the Virtual Classroom. Print out a hard copy and hand it in. The submission deadline is 24th July.