

FYS 102: Bioinformatics
Homework 6
Due Monday, October 29
(Start now!)

1. Create a keyword tree to solve the following multiple exact pattern matching instance. Show the keyword tree and how it is used to find the patterns.

$p^1 = \text{and}$

$p^2 = \text{great}$

$p^3 = \text{green}$

$p^4 = \text{all}$

$p^5 = \text{the}$

$t = \text{in the great green room there was a telephone and a red balloon}$

2. Create a suffix tree to solve the following exact pattern matching instance. Show the suffix tree and how it is used to find the pattern.

$p = \text{CTC}$

$t = \text{GACTCTCTAGCTTC}$

3. Show how a suffix tree could be used to find the longest common substring in two strings. A substring (vs. a subsequence) is a *contiguous* sequence of characters in the string. Show how your algorithm would be used to find the longest common substring in the strings **stats** and **estate**. Answering this question is going to take some thought!

Start early and have fun!