Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Mário Lipovský
Study programme:	Computer Science (Single degree study, bachelor I. deg., full time form)
Field of Study:	Computer Science, Informatics
Type of Thesis:	Bachelor´s thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: Approximate Abundance Histograms and Their Use for Genome Size Estimation

Aim: Technologies for DNA sequencing produce many short sequencing reads coming from a much longer genome. One way of succinctly summarizing this data is to compute a histogram of the number of occurrences of substrings of a fixed length in an input set of sequencing reads. Such histograms can be then used for example to estimate the size of the target genome. Some methods estimate histograms only approximately to reduce the memory required for the computation. The goal of the thesis is to select a method for histogram estimation and to explore the nature of its errors and how these errors influence the accuracy of the genome size estimate.

Supervisor:	doc. Mgr. Bronislava Brejová, PhD.
Department:	FMFI.KI - Department of Computer Science
Head of department:	prof. RNDr. Martin Škoviera, PhD.
Assigned:	26.10.2016
Approved:	26.10.2016

doc. RNDr. Daniel Olejár, PhD.
Guarantor of Study Programme

....
             Student   Supervisor