FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE

Supervisor's statement of a final thesis

Student:	Jan Jirák
Supervisor:	Ing. Jan Trávníček, Ph.D.
Thesis title:	Implementace algoritmů vyhledávání v řetězcích s konstantní pracovní pamětí navíc
Branch of the study:	Computer Science

Date: 15. 6. 2020

Evaluation criterion:	*The evaluation scale: 1 to 4.*
1. Fulfilment of the assignment	***1 = assignment fulfilled,*** *2 = assignment fulfilled with minor objections,* *3 = assignment fulfilled with major objections,* *4 = assignment not fulfilled*

Criteria description:
Assess whether the submitted FT defines the objectives sufficiently and in line with the assignment; whether the objectives are formulated correctly and fulfilled sufficiently. In the comment, specify the points of the assignment that have not been met, assess the severity, impact, and, if appropriate, also the cause of the deficiencies. If the assignment differs substantially from the standards for the FT or if the student has developed the FT beyond the assignment, describe the way it got reflected on the quality of the assignment's fulfilment and the way it affected your final evaluation.

Comments:

The task was to implement algorithms requiring an additional constant amount of memory to store results of the preprocessing and to execute search phase listed in a referenced survey article. The algorithms were implemented except for one which seems to have its pseudocode incorrect in the article. All other algorithms were subjected to extensive testing and later the time measurement of their search phase.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
2. Main written part	*70 (C)*

Criteria description:
Evaluate whether the extent of the FT is adequate to its content and scope: are all the parts of the FT contentful and necessary? Next, consider whether the submitted FT is actually correct – are there factual errors or inaccuracies? Evaluate the logical structure of the FT, the thematic flow between chapters and whether the text is comprehensible to the reader. Assess whether the formal notations in the FT are used correctly. Assess the typographic and language aspects of the FT, follow the Dean's Directive No. 26/2017, Art. 3. Evaluate whether the relevant sources are properly used, quoted and cited. Verify that all quotes are properly distinguished from the results achieved in the FT, thus, that the citation ethics has not been violated and that the citations are complete and in accordance with citation practices and standards. Finally, evaluate whether the software and other copyrighted works have been used in accordance with their license terms.

Comments:
The text is written in English with some sentences hard to understand.

Strings are indexed, unusually, in the text from 0 and, unusually again, factor s[i..j] includes the low index and excludes the high index. However, it is consistent across the text and implementation.
Definition 1.1.11 (Border) should also define what a nontrivial border is.
An example would help with understanding Definition 1.1.12, moreover, the 1-r+1<i<1 does not seem to be correct.
Definition 1.1.13 is missing a definition of what a basic period is.
Definition 1.1.14 is using an undefined notation in its defining formula.

Algorithm 1 should use variable k instead of variable case.
Definition 1.2.3 does not seem to define p properly, is it per?
Algorithm 8, line 3: variable k is unbounded and the semantics of line does not correspond to the matching line in the implementation.
There does not seem to be an increment of variable i in Algorithm 10.

I would appreciate more examples and informal descriptions of researched algorithms in addition to formal ones.

The chapter implementation could be extended with details regarding some differences between the theoretical description of algorithms and their practical implementation.

I'm missing a more detailed description of the measurement implementation, especially details about the measurement script.
Also, the alphabet size of the subject and pattern is not mentioned in the text.

Typography:
The text sometimes overflows to the right. Page 21 underflows vertically.
Sometimes the text is informal (text uses words like "to bother", "to torn").
Articles are missing in the text. Proofreading would very much improve the text quality.
Algorithms sometimes include line numbers sometimes don't.

Typo: page 4, line -1: lenght -> length.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
3. Non-written part, attachments	*89 (B)*

Criteria description:
Depending on the nature of the FT, comment on the non-written part of the thesis. For example: SW work – the overall quality of the program. Is the technology used (from the development to deployment) suitable and adequate? HW – functional sample. Evaluate the technology and tools used. Research and experimental work – repeatability of the experiment.

Comments:
The current implementation of Not-So-Naive algorithm does not match its version in the text.
The code formatting is a little bit inconsistent.
The GalilSeiferas algorithm implementation requires terminating character in both the pattern and the text, however, this is not discussed in the text at all.
The implementation follows the code standards in the Algorithms Library Toolkit.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
4. Evaluation of results, publication outputs and awards	*95 (A)*

Criteria description:
Depending on the nature of the thesis, estimate whether the thesis results could be deployed in practice; alternatively, evaluate whether the results of the FT extend the already published/known results or whether they bring in completely new findings.

Comments:
The implemented algorithms are going to be merged in the Algorithms Library Toolkit after some refactoring. Also, some algorithms may be later modified for linearized trees.

Evaluation criterion:	*The evaluation scale: 1 to 5.*

5. Activity and self-reliance of the student	*5a:* ***1 = excellent activity,*** *2 = very good activity,* *3 = average activity,* *4 = weaker, but still sufficient activity,* *5 = insufficient activity* *5b:* ***1 = excellent self-reliance,*** *2 = very good self-reliance,* *3 = average self-reliance,* *4 = weaker, but still sufficient self-reliance,* *5 = insufficient self-reliance.*

Criteria description:
From your experience with the course of the work on the thesis and its outcome, review the student's activity while working on the thesis, his/her punctuality when meeting the deadlines and whether he/she consulted you as he/she went along and also, whether he/she was well prepared for these consultations (5a). Assess the student's ability to develop independent creative work (5b).

Comments:

The student consulted regularly, he was able to study and understand some non-trivial algorithms alone.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
6. The overall evaluation	*85 (B)*

Criteria description:
Summarize which of the aspects of the FT affected your grading process the most. The overall grade does not need to be an arithmetic mean (or other value) calculated from the evaluation in the previous criteria. Generally, a well-fulfilled assignment is assessed by grade A.

Comments:

The research chapter is very detailed and even though it is sometimes hard to follow it represents a solid survey of the state of the art with references to the original sources. The implementation was extensively tested for correctness. Experimental evaluation was conducted and even though it could have been more thorough it shows that yet restricted to constant memory the researched algorithms are very much competitive with classical ones. I recommend the thesis for defence and recommend to grade it B (very good).

Signature of the supervisor: