Bachelor Thesis Review

Faculty of Mathematics and Physics, Charles University

Thesis author	Nikita Saydametov
Thesis title	Segmentation of PDF Documents
Year submitted	2024
Study program	Computer Science
Specialization	IOI

Review author	Mgr. Lucia Tódová	Reviewer
Department	Department of Software and Computer Science Education	

Overall

	good	OK	poor	insufficient
Assignment difficulty	X			
Assignment fulfilled		X		
Total size *. . . text and code, overall workload*	X			

The task of the student was to create a program with a user-friendly GUI that would be capable of segmenting input PDF documents and outputting individual components, such as images or text. The task is clearly and properly described, and the individual steps are straightforward and coherent.

From the implementation point of view, the task is satisfied. The functions of the program work as described and the interface allows for convenient manipulation with the files. My main objection to the implementation is the user documentation, which does not have sufficient information about code execution. A list of all required dependencies, as well as information about the exact operating system the program was tested on, would be valuable.

From the point of view of the thesis text, I find certain areas lacking, and the overall form to be weaker. My main objection is to the quality of references and to a multitude of grammar and stylistic mistakes in the text (see more thorough explanation in "Thesis Text"), which are not acceptable in academic writing.

Overall, I find both the assignment, and the length of the text and the code appropriate for a Bachelor thesis.

Thesis Text

	good	OK	poor	insufficient
Form *. . . language, typography, references*			X	
Structure *. . . context, goals, analysis, design, evaluation, level of detail*			X	
Problem analysis		X		
Developer documentation		X		
User Documentation	X			

The thesis text is lacking in multiple areas.

My main concern, and probably the most problematic aspect of the whole thesis, is the lack of references and their quality. While it is acceptable to reference online links and pages of e.g. products that are not backed up by a peer-reviewed article, online blogs and Wikipedia are not considered trustworthy sources and should not be used as such. Additionally, a lot of information provided in the thesis is not backed up by references altogether.

The English of the student is also lacking. The thesis would greatly benefit from a lot more proof-reading, as it contains multiple nonsensical sentences, incorrect usage of prepositions and articles, and even spelling mistakes. Formatting issues in the thesis could also have been addressed.

At the beginning, the structure of the text is quite clear, and the problem analysis is sufficient. After moving on to the chapter "Available solutions", I had a couple of reservations. Specifically, neither of the technologies used in the implementation is mentioned in the chapter, and both are only discussed after introducing the implementation. This is quite confusing - a more clear approach would be to mention all available solutions, and then provide discussion on why the specific technologies were chosen.

A minor mistake I would additionally like to point out is the referencing of figures, which are placed in the text but almost never referenced. This is also incompatible with academic writing.

Apart from the build problems mentioned earlier, the user documentation is clear and sufficient. Although the developer documentation does not have the typical structure of a documentation and could be found unclear and a bit confusing at times, it contains all the necessary information to comprehend the code. Additionally, the comments in the code are sufficient.

Thesis Code

	good	OK	poor	insufficient
Design ... *architecture, algorithms, data structures, used technologies*	X			
Implementation ... *naming conventions, formatting, comments, testing*	X			
Stability	X			

The implementation and the code quality is satisfactory. I agree with the used technologies and also with the decisions made during the development of the program. The architecture is simple and efficient, and the comments and documentation that belong to individual functions are clear.

While the code itself is more of a combination of existing techniques rather than an implementation of something new, I find it sufficient for a Bachelor thesis and I find the execution satisfactory.

Minor objections regarding the code include inconsistent naming conventions (using camel case and snake case in the same source file) and insufficient error messages - although there are error messages always present, they sometimes do not specify the nature of the error, which makes it harder to make the program work.

Overall, I find the code sufficient and do not have any major objections.

Overall grade   Good (better)
Award level thesis   No

Date   **3.6.2024**   Signature