Bachelor Thesis Review

Faculty of Mathematics and Physics, Charles University

Thesis author	Andrei Lupasco
Thesis title	Deep Neural Networks for Graph Data Processing
Year submitted	2024
Study program	Computer Science
Specialization	Artificial Intelligence

Review author	Vladan Majerech	Reviewer
Department	Katedra teoretické informatiky a matematické logiky	

Overall

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

Thesis is concentrated on problem of neural networks dealing with graph inputs, where typically the network gets different inputs even when the graphs are isomorphic. That could complicate the learning process. Student solved the problem for planar graphs by implanting the inputs to the labels (strings) of the input graphs (two planar graphs have the same label iff they are isomorphic).

At the end of the thesis (nontrivial part) student evaluates benefits of the approach on various datasets comparing with concurent methods.

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				
User Documentation				

The author does not always use mathematically precise formulations. In the introduction to planar graph isomorphism testing he confuses log space with log time. Argumentation using Banach fixed point theorem at 2.4.3 does not work for individual vertices, but it works for the entire graph. He uses $O(|V|!)$ to express lower bound of naive isomorphism testing (the complexity of the algorithm would be $O(|V|! \cdot (|V| + |E|))$, but he probably wanted to say $\Omega(|V|!)$). There is wrong label near vertex D on figure 2.10.

The KHC algorithm which have to compute the planar graph labels has a bug in the parallel case (the version on page 38). It produces the same label independently on which of the P endpoints is an articulation. The theorems in the thesis expect the basics work well and they just try to show the extension does not introduce new problems, but the assumption is not stated (and does not hold).

I hoped the bug in parallel case is corrected when vertex information is implanted into the code. But seems to me, this information is not implanted between the ambivalent positions where the star could be. I am not sure if it is good decision to include vertex information several times making a lot of duplicities.

The proof of Theorem 2 could be corrected if it would talk about bijection among A and a homomorphic image of B (forgetting the implanted information).

As I have mentioned, I am worried the transformation maps the graphs to strings which are not labels as sometimes (rarely) nonisomorphic graph obtain the same label. It would be nice to repeat the experiments after correcting the issue (or showing me I am wrong).

Thesis Code

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

The complexity of input preprocessing is negligable compared to training the neural network, but it could be important for using the trained one. I am a bit worried by the complexity of the `next_edge_tour` implementation (considering all graph edges), but it is used only for S nodes, for R nodes the embedding was computed and vertex neighbor lists are used.

The most difficult part (the SPQR decomposition) was not coded by the student, he rather used the sage library (what is OK).

I have not found in the text part nor in the implementation solution of the case the center of SPQR tree is edge rather then a node. Solution would be to use minimal of codes of the virtual edge representing the bicentre. Maybe the implementation randomly chooses one of the endpoints of the bicentre, but than two isomorphic graphs could obtain different codes (so they could not be labels for another reason).

Overall grade Very Good, worse

Award level thesis No

Date

Signature