Foundations of Graph Neural Networks
(by E. V. Kostylev)

MIMUW PhD Open, 21–23 May 2025

Exam Paper

Instructions

1. Submit to `egork@uio.no` by 23:59 CET, 11 July 2025. If you do not receive a confirmation from me within 24 hours, send me another email without an attachment. If you need your grade earlier than the end of July, let me know in advance and submit at least 1 week before your desired date.
2. You can use whatever resources they want, but you should not collaborate with other students and anyone else (the most relevant resource is the course slides, which are available at the Web page of the course: `https://phdopen.mimuw.edu.pl/index.php?page=l25w2`).
3. If you believe that there is a mistake or ambiguity in the exam, you can either send me an email for clarification or just write in the solution that you believe it is such and such, and you solve a slightly different problem (but if I find that your belief is unreasonable, you will be downgraded).
4. There are 6 questions in the exam. Preliminary marking scheme is such that a solution to one question, possibly with some minor mistakes, costs one point, and the number of points translate directly to the grade (5!, 5, 4.5, 4, 3.5, 3, Fail).

Question 1

Consider the *basic sum-plus GNNs with sigmoid activation*, which update node labellings as follows (here and throughout the exam paper, as usual, ℓ ranges over layers and v over graph nodes):

$$\mathbf{x}_v^{(\ell)} := f\left(\mathbf{A}^{(\ell)}\mathbf{x}_v^{(\ell-1)} + \mathbf{C}^{(\ell)}\left(\sum_{u\in\mathcal{N}_G(v)} \mathbf{x}_u^{(\ell-1)}\right) + \mathbf{b}^{(\ell)}\right),$$

where the sigmoid is the logistic function $f(x) = \frac{1}{1+e^{-x}}$. Write such a GNN with input dimension 2 that realises the node-level binary (i.e., Boolean) classification function 'there is a node with the first component greater or equal to 1 at distance exactly 3 from the given node.' You may assume that all nodes in input graphs are labelled with 0-1 vectors (i.e., correspond to logical structures with two unary predicates).

Question 2

Does the same distinguishing power imply the same expressive power? Does the latter imply the former? Justify your answers formally.

Question 3

Recall that in the lectures we considered a simplified version of GCNs, that update node labellings as follows:

$$\mathbf{x}_v^{(\ell)} := \mathit{ReLU}\left(\mathbf{A}^{(\ell)}\left(\operatorname{avg}_{u\in\mathcal{N}_G(v)\cup\{v\}}\mathbf{x}_u^{(\ell-1)}\right)\right).$$

Consider another version of GCNs (closer to the real ones) that use slightly different updates:

$$\mathbf{x}_v^{(\ell)} := \mathit{ReLU}\left(\mathbf{A}^{(\ell)}\left(\sum_{u\in\mathcal{N}_G(v)\cup\{v\}} \frac{\mathbf{x}_u^{(\ell-1)}}{\sqrt{\deg(v)\cdot\deg(u)}}\right)\right),$$

where $\deg(w)$ denotes the degree (i.e., number of neighbours) of a node w. Compare these two classes of GNNs in terms of distinguishing and expressive power, focusing on node-level binary classifiers. Justify your answers.

Question 4

One of the main theorems of our course states that 'A node classifier is realisable by both an AC-GNN and a first-order logic (FO) formula if and only if it is realisable by a Graded Modal Logic (GML) formula.' We proved the backward direction of this theorem by simulating every GML formula with a GNN with update of the following form:

$$\mathbf{x}_v := trReLU\left(\mathbf{A}\mathbf{x}_v^{(\ell-1)} + \mathbf{C}\left(\sum_{u \in \mathcal{N}_G(v)} \mathbf{x}_u^{(\ell-1)}\right) + \mathbf{b}\right).$$

Note that the GNNs in this class are *homogeneous*—that is, each of them has the same matrices and bias vector on each layer. Demonstrate that we can prove this backward direction using the same class of GNNs except that they use usual $ReLU$ in place of $trReLU$. If you encounter difficulties doing so, try again with the homogeneity requirement removed.

Question 5

As we briefly mentioned in lectures, the main theorem referred to in Question 4 has an analogue for node-level GNNs; however, it is still a hypothesis, as we only know how to prove the backward direction. Formulate this analogue, including the relevant fragment of FO. Recall now, that, at the node level, the proof of the forward direction of the theorem relies on a Van Benthem-style theorem by Martin Otto: 'The counting bisimulation invariant fragment of FO is precisely GML.' Formulate the analogue of this theorem for the node level, including the relevant variant of bisimulation.

Question 6

Consider the class of all node-level AC-GNNs in which the aggregation function is restricted to (element-wise) max. Describe a reasonably large non-trivial class of (simple, undirected, node-labelled) graphs for which this class of GNNs is universal in terms of expressive power.