Postdoc position: Verifying Graph Neural Networks (M/F)

1 Context

The current project aims to bridge the gap between formal reasoning (logic) and machine learning models.

Figure 1: Setting where a GNN is used to identify whether an account is a human or a bot in a fictitious social network.

Machine learning techniques in general suffer from being verifiable and explanable. Our project aims at providing **tools for verifying and explaining graph neural networks (GNN)**. A graph neural network takes a graph as an input and typically outputs a decision, see Figure 1. GNNs have applications in domains such as social recommendations [SLJ21], drug discovery [XXC+21], material science and chemistry [RNE+22], knowledge graphs [YKS+22], music [KFW24], among others (see [ZCH+20] for an overview). Typically, GNNs are trained on dataset, and are then deployed on specific task, e.g. discovering whether an account in a social network is a bot or not, or discovering whether a given molecule (represented as a graph) is toxic for bees or not. As neural networks in general, GNNs are blackboxes. Sure we understand the output of a given GNN (e.g. "the molecule is toxic") but we do not understand why a GNN provides such an output. This project may help practitioners to improve the understanding of the underlying GNN models they are using.

Related work. There are several works about verification [HKR+20] and explanation [MI22] of neural networks. Usually, they rely on propositional logic or satisfiability modulo theory [KBD+17].

There also exists promising work for graph neural networks as [BKM+20, Gro21]. The current results are mainly about expressivity of GNNs. The seminar paper by Barcelo et al. [BKM+20] and also [Gro21] focus on comparing graded modal logic and GNNs: any graded modal logic formula can be captured by a GNN while the converse - capturing a GNN by a graded modal logic - only holds when the GNN is expressible in first-order logic. Recently, in [NSST24], we proposed a correspondence between a modal logic with counting and GNNs. We also proposed a methodology for verifying GNNs based on this correspondence. Meanwhile, there are also other expressivity results such as the correspondence between a modal logic with counting and universal counting and GNNs with global readout[BLMT24].

2 Research Project

2.1 Expressivity

The objective is to continue to study expressivity power of GNNs by either making new connections with existing logics, or providing new logics or fragments of existing logics corresponding to some GNN classes. Typical research questions we aim at addressing are: what is the GNN counterpart of modal logic K? What is the logical counterpart of some classes of convolutional NNs (that can be seen as GNNs), GNN with spike NN? Graph transformers?

We will also tackle quantization. It means that we will study GNNs as they are really implemented in practice. It means that numbers in GNN are no longer real numbers but are described by a small and fixed number of bits, e.g. 32-bit floating-point numbers. We will study the expressivity of quantized versions of GNNs. A first approach is available in [SST25].

2.2 Algorithms

For different classes of GNNs, we aim at studying the algorithmic verification problem and the synthesis of explanation. Given the link with modal logic, we will also study logical aspects of the involved logics such as axiomatization and proof systems in order to provide algorithms.

Current studied verification problems are decision problems, typically: given a GNN A, a formula ϕ, are the recognized graph by A exactly the models of ϕ? We may relax this yes/no aspect and study quantitative versions. For instance: what is the probability that a given random graph G accepted by A satisfy a property ϕ?

Interestingly, we will also tackle the synthesis problem: given a GNN A, compute a formula ϕ whose models are exactly the graphs recognized by A (or compute a formula ϕ such that the set of models of ϕ is close enough to the set of graphs recognized by A).

We will also tackle the generation of quantized versions of existing GNNs. This is motivated for reducing the memory consumption of GNNs, and then also the energy and time consumption at using GNNs. Given a GNN A, the objective is to construct a quantized GNN A' with a compromise between the size of A' and the loss in expressivity. We may have a situation with a Pareto frontier to study.

2.3 Applications

We will apply our methodology to GNNs for some applications and provide explanations of what a GNN has learnt. To this aim, we will implement some algorithms for problems described in Subsection 2.2. We will try to understand what kind of properties on graphs GNNs have learnt. For instance, in musical applications (see [KFW24]), we may construct a logical formula ϕ explaining why the system has decided to guess a particular voice separation. Our methodology is general and can be applied in other domains, such as bio-informatics or social network analysis.

There are mainly two approaches in learning. The first one is to learn “very learnable but non-explainable” models, i.e. in our case learning GNNs, and then extracting a formula by solving the synthesis problem described in Subsection 2.2. The second one is to directly learn a ”non-very learnable but explainable” models, e.g. learning a formula [Leh09], [tCFJL23]. We aim at comparing the two approaches in terms of accuracy, of learning efficiency and explainability. When is it valuable to learn a GNN and extract a formula, vs. learning a formula?

3 Position

- Position: full-time postdoctoral researcher position
- Duration: 24 months starting ideally from January 1st, 2026
- Renumeration: from 2900€ gross per month, according the ENS salary scale
- Location: Ecole normale supérieure de Lyon (ENS de Lyon) and Laboratoire de l’informatique du parallélisme (LIP)
- PI name: François Schwarzentruber

4 Environment

Ecole normale supérieure de Lyon (ENS de Lyon) is a pleasant location for doing research and meet other people. The lab LIP is located inside ENS de Lyon. ENS de Lyon can be easily reached from the city center of Lyon by metro or bike.

Lyon is known for its gastronomy, and for its well-preserved historic center (a UNESCO World Heritage Site). Lyon offers a high quality of life. Lyon is near Paris (two-hours by train), near the mountains and near the see (two-hours for reaching Montpelier for instance).

The postdoctoral researcher will participate in the research team MC2 (Models of computation, Complexity, Combinatorics). So it will be easy for the postdoctoral reasearcher to discuss with researchers that are interested in graph theory and logic.

The postdoctoral researcher will collaborate with François Schwarzentruber (leader of the chair) and will have many opportunities to interact with the external collaborators. The chair provides significant resources to finance trips to conferences and to host visiting researchers.

Furthermore, the postdoctoral researcher will also have the opportunity to discuss with researchers in machine learning (LIP, OCKHAM team).

5 How to Apply

Interested candidates should contact François Schwarzentruber and provide:

- a detailed CV, containing the applicant's full publication list
- university transcripts for Master's / engineering degree,
- names of 2-3 references,
- a brief description of how the project topics relate to their prior experience and interests.

Applications should be submitted by email to the following email address:

`francois.schwarzentruber@ens-lyon.fr`

References

[BKM+20] Pablo Barceló, Egor V Kostylev, Mikaël Monet, Jorge Pérez, Juan Reutter, and Juan-Pablo Silva. The logical expressiveness of graph neural networks. In *8th International Conference on Learning Representations (ICLR 2020)*, Virtual conference, Ethiopia, April 2020.

[BLMT24] Michael Benedikt, Chia-Hsuan Lu, Boris Motik, and Tony Tan. Decidability of graph neural networks via logical characterizations. In Karl Bringmann, Martin Grohe, Gabriele Puppis, and Ola Svensson, editors, *51st International Colloquium on Automata, Languages, and Programming, ICALP 2024, July 8-12, 2024, Tallinn, Estonia*, volume 297 of *LIPIcs*, pages 127:1–127:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024.

[Gro21] Martin Grohe. The logic of graph neural networks. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021, Rome, Italy, June 29 - July 2, 2021*, pages 1–17. IEEE, 2021.

[HKR+20] Xiaowei Huang, Daniel Kroening, Wenjie Ruan, James Sharp, Youcheng Sun, Emese Thamo, Min Wu, and Xinping Yi. A survey of safety and trustworthiness of deep neural networks: Verification, testing, adversarial attack and defence, and interpretability. *Comput. Sci. Rev.*, 37:100270, 2020.

[KBD+17] Guy Katz, Clark W. Barrett, David L. Dill, Kyle Julian, and Mykel J. Kochenderfer. Reluplex: An efficient SMT solver for verifying deep neural networks. In Rupak Majumdar and Viktor Kuncak, editors, *Computer Aided Verification - 29th International Conference, CAV 2017, Heidelberg, Germany, July 24-28, 2017, Proceedings, Part I*, volume 10426 of *Lecture Notes in Computer Science*, pages 97–117. Springer, 2017.

[KFW24] Emmanouil Karystinaios, Francesco Foscarin, and Gerhard Widmer. Perception-inspired graph convolution for music understanding tasks. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI 2024, Jeju, South Korea, August 3-9, 2024*, pages 7681–7689. ijcai.org, 2024.

[Leh09] Jens Lehmann. Dl-learner: Learning concepts in description logics. *J. Mach. Learn. Res.*, 10:2639–2642, 2009.

[MI22] João Marques-Silva and Alexey Ignatiev. Delivering trustworthy AI through formal XAI. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, pages 12342–12350. AAAI Press, 2022.

[NSST24] Pierre Nunn, Marco Sälzer, François Schwarzentruber, and Nicolas Troquard. A logic for reasoning about aggregate-combine graph neural networks. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI 2024, Jeju, South Korea, August 3-9, 2024*, pages 3532–3540. ijcai.org, 2024.

[RNE+22] Patrick Reiser, Marlen Neubert, André Eberhard, Luca Torresi, Chen Zhou, Chen Shao, Houssam Metni, Clint van Hoesel, Henrik Schopmans, Timo Sommer, and Pascal Friederich. Graph neural networks for materials science and chemistry. *Communications Materials*, 3(93), 2022.

[SLJ21] Amirreza Salamat, Xiao Luo, and Ali Jafari. Heterographrec: A heterogeneous graph-based neural networks for social recommendations. *Knowl. Based Syst.*, 217:106817, 2021.

[SST25] Marco Sälzer, François Schwarzentruber, and Nicolas Troquard. Verifying quantized graph neural networks is pspace-complete, 2025.

[tCFJL23] Balder ten Cate, Maurice Funk, Jean Christoph Jung, and Carsten Lutz. Sat-based PAC learning of description logic concepts. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI 2023, 19th-25th August 2023, Macao, SAR, China*, pages 3347–3355. ijcai.org, 2023.

[XXC+21] Jiacheng Xiong, Zhaoping Xiong, Kaixian Chen, Hualiang Jiang, and Mingyue Zheng. Graph neural networks for automated de novo drug design. *Drug Discovery Today*, 26(6):1382–1393, 2021.

[YKS+22] Zi Ye, Yogan Jaya Kumar, Goh Ong Sing, Fengyan Song, and Junsong Wang. A comprehensive survey of graph neural networks for knowledge graphs. *IEEE Access*, 10:75729–75741, 2022.

[ZCH+20] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. Graph neural networks: A review of methods and applications. *AI Open*, 1:57–81, 2020.