## What's hot in statistical relational AI: The first International Joint Conference on Learning and Reasoning

## Felix Weitkämper LMU München

January 27, 2022

Over the past decade, the annual workshop on statistical relational AI (branded StaRAI) was a popular venue for work in progress as well as substantial technical contributions to the field. While it was previously colocated with one of the broad flagship artificial intelligence conferences AAAI, UAI or IJCAI, this year for the first time it was combined with conferences and workshops on Inductive Logic Programming (ILP), Neural Symbolic Integration (NeSy) and Program induction (AAIP) in a brandnew format, the 1st International Conference on Learning and Reasoning (IJCLR). Originally intended to be held in September 2020 in Athens, it was postponed for a year and eventually held as an online meeting in October 2021. In addition to the traditional formats of the participating events, the IJCLR featured joint sessions of both talks and posters as well as the option to submit papers to a designated journal track.

To give a flavour of the synergies between statistical relational AI and neuralsymbolic reasoning, I report on the first joint session of the IJCLR, devoted to a statistical relational perspective on Graph Neural Networks (GNNs) GNNs, first proposed more than a decade ago, adapt the concepts of deep learning to input data carrying a graph structure. By dealing natively with structured input, GNNs avoid the need for preprocessing the input data into a vectorised form, a step that leads to intransparency and a loss of information. Since their introduction, hundreds of variants of GNNs have been proposed, and GNNs have started to have practical impact in areas such as chemoinformatics and natural language processing. By working with graph-structured input to answer queries both about the graph as a whole and about individual nodes, GNNs are carrying out very similar tasks to statistical relational approaches, suggesting both integration of and comparison between those areas as urgent research issues.

The joint session of the IJCLR featured three papers covering different aspects of this issue:

Sourek, Zelezny and Kuzelka (https://doi.org/10.1007/s10994-021-06017-3) discuss *lifted relational neural networks* as a declarative differential programming environment for the high-level representation of GNN arcitectures. Syn-

tactically a paramterised version of the well-known Datalog language, lifted relational neural networks encode GNN architectures as rules whose literals carry tensor parameters. Learning examples are then encoded as parametrised facts, where the parameter takes unit value for logical facts and tensor values otherwise. The authors then demonstrate how various GNNs and extensions thereof can be expressed in the proposed formalism, and demonstrate experimentally that lifted relational neural networks can actually outperform dedicated contemporary GNN implementations on a wide variety of molecular classification tasks.

Dash, Srinivasan and Vig (https://doi.org/10.1007/s10994-021-05966-z) discuss the problem of how best to incorporate domain-knowledge into GNNs. This is particularly relevant in highly structured domains and where learning examples are scarce, such as the drug evaluation scenario underlying the real-world datasets used for the empirical evaluation of the paper. The authors introduce *vertex-enrichment* to include polyadic relational background knowledge onto a binary graph; in this way, for instance, knowledge of functional groups and rings can be added to the molecular structure encoded as input for a GNN model. In addition to adding the background knowledge taken from the literature, the authors also investigate the usefulness of adding an explicit preprocessing step in which the inductive logic programming system Aleph is used to infer higher predicates from the background knowledge, and then to add them to the vertexenriched GNN. The experimental evidence provided suggests that both steps improve the accuracy of the learned GNN model.

Embar, Srinivasan and Getoor (https://doi.org/10.1007/s10994-021-06007-5) compare GNNs and statistical relational approaches on their performance in answering aggregate graph gueries, more complex queries that aggregate information about multiple nodes to elucidate properties of the graph as a whole. Aggregate graph queries can be used to assess the community structure of a coloured graph, for instance by querying the numbers of intra- and intercommunity edges between graph nodes. The authors discuss various techniques for approximating the answers to such queries, necessitated by the intractability of any direct exact computation. As aggregate graph queries reference the entire graph at once, statistical relational formalisms which explicitly model the joint distribution allow for a straightforward approach through sampling from the joint distribution, while GNN models that do not model a joint distribution have to rely on extrapolating from individual predictions. Empirical results reflect this, and statistical relational methods outperform GNNs on several citation datasets.

Overall, the research presented here has made a good case for the continuing relevance of symbolic relational approaches in the era of deep learning, and in particular underlined that statistical relational representations have a role to play even in settings where accuracy is more important than explainability. On the other hand, the comparisons made here pitted GNNs and symbolic methods in application domains that are considered traditional strongholds of relational techniques. Indeed, predicting molecular properties from their structure was one of the first breakthrough domains of inductive logic programming in the early 90s, and highly complex, aggregate queries seem tailor-made for the expressive and explicit representation of joint probabilities that statistical relational models provide.

The second edition of the IJCLR is already planned for September 2022 as an in-person meeting in Windsor, UK. However, the composition of associated events will change, and the StarAI workshop will no longer be colocated with the IJCLR. Instead, the HLC workshop on human-like computing will be included, alongside ILP, NeSy and AAIP.

The website of the IJCLR 2021 (http://lr2020.iit.demokritos.gr) includes a list of all accepted papers, including those of the associated conferences, and the website of IJCLR 2022 (https://ijclr22.doc.ic.ac.uk/) will be updated with calls for papers and further details in due course.