Deep Learning with Relational Logic Representations

Gustav Sourek

Czech Technical University in Prague
souregus@fel.cvut.cz

1 Introduction

As the process of maximizing accuracies on large perceptual datasets slowly reaches its limits, deep learning researchers start to explore increasingly ambitious goals for neural networks to strive for, with thoroughly tuned neural architectures continuously taking over new domains on daily basis.

Despite their significant success, all the existing neural architectures based on static computational graphs processing fixed tensor representations necessarily face fundamental limitations when presented with dynamically sized and structured data. Examples of these are sparse multi-relational structures present everywhere from biological networks and complex knowledge hyper-graphs to logical theories. Likewise, given the cryptic nature of generalization and representation learning in neural networks, potential integration with the sheer amounts of existing symbolic abstractions present in human knowledge remains highly problematic.

Here, we argue that these abilities, naturally present in symbolic approaches based on the expressive power of relational logic, are necessary to be adopted for further progress of neural networks into the more complex domains.

2 Proposed Approach

To marry the advantages of deep and symbolic learning, we propose to follow the strategy of *lifted modelling*. As opposed to standard machine learning approaches, lifted models do not specify a particular model architecture, but rather a *template* from which the particular architectures are being derived as a part of the learning process itself, given the varying context of input samples and knowledge. For example, the famous lifted Markov Logic Network model [Richardson and Domingos, 2006] may express a general template-knowledge that *friends of smokers tend to be smokers*, which then constraints the smoking probabilities of all the particular friends in the given data, as if modeled by a regular Markov network, yet generalizing over diverse social networks.

Importantly, this allows the underlying models to respect the inherent symmetries in the data, not only at the level of single relations, such as the friendship, but possibly complex isomorphic patterns conformed from these. This allows the lifted models to convey a highly *compressed representation* of information as all the isomorphic patterns are parameterized jointly by the single template. While this may sound peculiar in the context of neural networks, we have already seen a great success of one simple incarnation of this concept in *convolutional filters*, effectively following the very principle.

While there is certainly no lack of similar ad-hoc solutions to tweak neural networks towards some facets of the sought after abilities of relational methods, we believe that framing the problem explicitly in relational logic, using the lifted modeling paradigm, yields a more general, highly expressive, and well founded learning formalism providing these abilities in a more direct fashion. We have instantiated these principles in our framework of Lifted Relational Neural Networks [1] (LRNNs) [Šourek *et al.*, 2015].

Similarly to other lifted models, we base the syntax of the template language on weighted relational (horn) clause logic, which we extend with biases and differentiable activations to accommodate the translation into the neural architectures, as briefly depicted in Fig. 1. Semantically, this translation is based on the least Herbrand models of the template and example structures, calculated in either top-down or bottom-up fashion depending on the structure of the domain.

3 Related Work

Realizing the potential of integration of neural and symbolic learning, there is a rich track record of "neural-symbolic" approaches trying to accommodate neural networks into symbolic learning settings [Garcez *et al.*, 2012]. The distinguishing property of our lifted modelling here is the dynamic construction of varying models, which we argue is necessary to exploit the unique relational properties of each example.

The standard territory for lifted models is Statistical Relational Learning (SRL), where lifting of graphical models have been extensively studied [Richardson and Domingos, 2006], however not so much in the context on neural networks.

The need for structured representations naturally arises in language processing (NLP), where recursive neural approaches [Socher *et al.*, 2011] also generate dynamic networks, yet these are merely restricted to exactly follow the tree structure of each example sentence. In a similar fashion, various forms of "graph neural networks" have been proposed for structure-driven embedding and propagation of information in graphs [Niepert *et al.*, 2016] and knowledge bases (KBC), traditionally based on tuple-embedding methods.

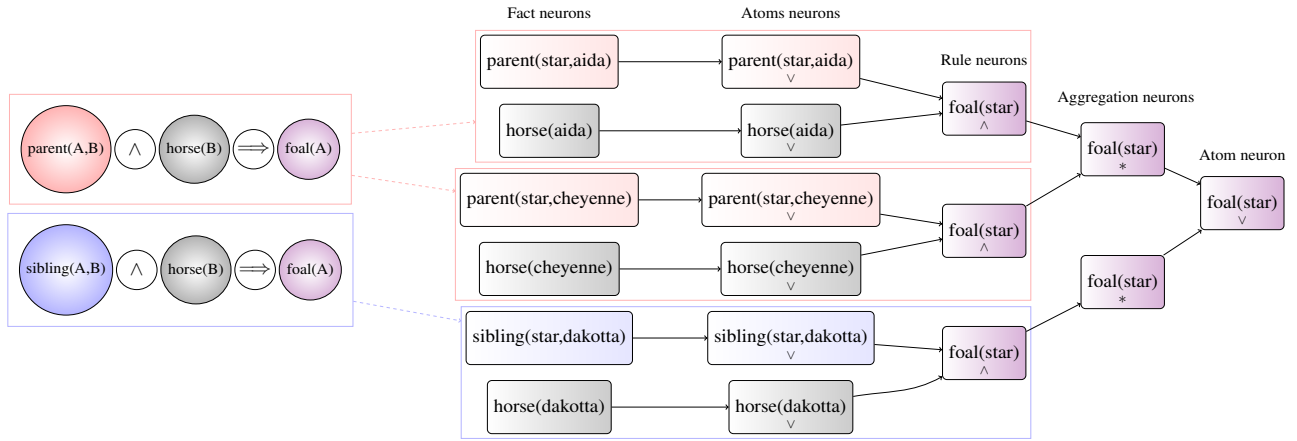[1] https://github.com/gustiks/neuralogic

Figure 1: A template with 2 horn clauses (left) which, given example facts (fact neurons), translates into a simple network consisting of different types of neurons with activations corresponding to differentiable approximations of the corresponding logical connectives.

The most directly related framework here are the recently proposed Neural Theorem Provers [Rocktäschel and Riedel, 2017], an architecture following Prolog's "backward chaining" strategy to generate proof trees, which are then translated into neural networks, much like the earlier proposed framework of LRNNs [Šourek *et al.*, 2015].

4   Contributions and Future Work

We present several unique contributions to the state-of-the-art (SotA). In [Šourek *et al.*, 2015] we demonstrated that with merely very simple templates, it is already possible to achieve SotA results on classic SRL benchmarks. Clearly, the use of templates in the framework allows for natural integration of interpretable background knowledge, but also declarative specification of advanced learning constructs [Šourek *et al.*, 2016]. Namely, these include embeddings of not only constants, but whole latent relational structures, i.e. generic patterns not grounded to any specific domain elements, which was effectively impossible with the previous frameworks.

Moreover, in contrast to standard neural models, these embeddings have a well founded interpretation as soft clusters in conceptual spaces, where each dimension represents the degree with which a certain predicate relates to another.

The expressiveness of relational logic and genericity of the neural translation allows for a more flexible learning paradigm, effectively subsuming a diversity of the previous as well as some of the subsequent approaches, such as the Convolutional Graph Networks [Niepert *et al.*, 2016] and Neural Theorem Provers [Rocktäschel and Riedel, 2017], with a mere declaration of the template and activation functions.

Finally, LRNNs are also the only framework to support proper symbolic structure learning with neural predicate invention [Šourek *et al.*, 2017], allowing to explore non-trivial latent relational patterns. This form of learning was previously only studied in context of inductive logic, as it is effectively impossible with standard greedy optimization routines.

We see future enhancements in targeted optimization of the discussed modelling constructs to allow to scale onto more realistic NLP and KBC applications, where we expect the integration with symbolic knowledge abstractions to play the most significant role for future progress of neural learning.

Acknowledgements

References

[Garcez *et al.*, 2012] Artur S d'Avila Garcez, Krysia B Broda, and Dov M Gabbay. *Neural-symbolic learning systems: foundations and applications*. Springer, 2012.

[Niepert *et al.*, 2016] Mathias Niepert, Mohamed Ahmed, and Konstantin Kutzkov. Learning convolutional neural networks for graphs. In *Proceedings of ICML*, 2016.

[Richardson and Domingos, 2006] Matthew Richardson and Pedro Domingos. Markov logic networks. *Machine learning*, 2006.

[Rocktäschel and Riedel, 2017] Tim Rocktäschel and Sebastian Riedel. End-to-end differentiable proving. In *Advances in Neural Information Processing Systems*, 2017.

[Socher *et al.*, 2011] Richard Socher, Cliff C Lin, Chris Manning, and Andrew Y Ng. Parsing natural scenes and natural language with recursive neural networks. In *Proceedings of ICML*, 2011.

[Šourek *et al.*, 2015] Gustav Šourek, Vojtech Aschenbrenner, Filip Železny, and Ondřej Kuželka. Lifted relational neural networks. In *Proceedings of the NIPS CoCo Workshop: Integrating Neural and Symbolic Approaches*, 2015.

[Šourek *et al.*, 2016] Gustav Šourek, Suresh Manandhar, Filip Železný, Steven Schockaert, and Ondřej Kuželka. Learning predictive categories using lifted relational neural networks. In *Inductive Logic Programming*, 2016.

[Šourek *et al.*, 2017] Gustav Šourek, Martin Svatoš, Filip Železný, Steven Schockaert, and Ondřej Kuželka. Stacked structure learning for lifted relational neural networks. In *Inductive Logic Programming*, 2017.