Course on *Learning with Knowledge Graphs* and *Causal Reasoning for Artificial Intelligence*

Volker Tresp

Ludwig Maximilian University of Munich and Siemens AG
volker.tresp@siemens.com

Abstract. The the first part I will cover recent work on learning with knowledge graphs. In the second part I will cover the two main causal theories of relevance for AI: Pearl's theory of causal and counterfactual inference and Rubin's causal model.

1 Deep X: Deep Learning with Deep Knowledge

Deep X In many applications, the full potential of *deep learning* can only unfold in combination with *deep knowledge*. Deep knowledge can mean that instances or entities are described by many dimensions, e.g., patients are described by their general health profiles in conjunction with extensive molecular profiles. Here we focus on deep knowledge in the form of deeply structured *knowledge graphs*.

Knowledge Graphs The most prominent example is the Google Knowledge Graph [7], approaching 100 billion statements and describing world facts as triples, such as *(Obama, exPresidentOf, US)*. Knowledge graphs are closely related to relational databases and graph databases, supplemented with type constraints and concept hierarchies. Relational databases are ubiquitous in industry in general, and graph databases are extensively used in communication and social networks. Knowledge graphs are considered easier to extend and to maintain than relational databases and are becoming increasingly popular in many industries. Knowledge graphs can be used for linking information sources, for querying, and in question answering. Different analytic functions can be realized, such as trend analysis, the visualization of views, and the calculation of statistics.

Machine Learning with Knowledge Graphs Knowledge graphs can learn. Relational machine learning can be used to derive triples that are not part of the knowledge graph, such as *(Obama, gender, Male), (Obama, race, Caucasian)* [6, 5]. (Well, 50% correct!) Furthermore, machine learning can derive priors for text and image understanding and thus support the automatic filling of knowledge graphs [2]. Finally, latent entity representations derived from machine learning can support other applications.

Modelling Events Events in time can be modelled by adding a time index to a triple. This concept is very useful in the development of medical decision support systems where a semantic knowledge graph represents a patient's background (existing conditions, age, genetic profile, ...) and an episodic knowledge graph represents patient-specific events like treatments, outcomes, lab measurements, and administered medications [3, 11].

Perception: "You only see what you know" Deep Learning is currently the leading computational approach to image analysis. But perception is more: perception requires a decoding of sensory inputs in the context of an agent's understanding of the world. So the Goethe quote "Man sieht nur, was man weiß" might be quite appropriate! In [1], it was shown how regional convolutional neural networks (R-CNNs) can be combined with knowledge graphs, which describe prior knowledge about concepts and their dependencies, to map an image to a set of triples.

Cognitive Deep X It has been argued that our conscious mind emerges from thousands of lower-level processes operating in parallel: "The human brain has a modular organization consisting of identifiable component processes that participate in the generation of a cognitive state."[4] We argue that some modules might adequately be modeled by deep neural networks, but for others, like memory functions, knowledge graphs and their tensor models might be more suitable [8, 9, 10].

2 Causal Reasoning for Artificial Intelligence

Causality is a surprisingly contentious topic. For example, in 1918 Bertrand Russell (Nobel Prize in Literature) stated: "The law of causality, I believe, is a relic of bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm". In the mean time many causal theories have evolved in science, engineering, psychology, philosophy, and other fields. The two main theories relevant for AI are Judea Pearl's (Turing Award Winner) theory of causal and counterfactual inference (more popular in AI and machine learning) and Donald Rubin's Causal Models (RCMs) (more popular in statistics). Both theories provide insights into how one should address causality in practice.

I got interested in the field because I had to deal with causal issues in clinical applications. I want to convey some of the deeper issues in causality but also try to address the issue of how to apply the theoretical insights in practice!

References

[1] Stephan Baier, Yunpu Ma, and Volker Tresp. Improving visual relationship detection using semantic modeling of scene descriptions. In *ISWC*, 2017.

[2] Xin Dong, Evgeniy Gabrilovich, Geremy Heitz, Wilko Horn, Ni Lao, Kevin Murphy, Thomas Strohmann, Shaohua Sun, and Wei Zhang. Knowledge Vault: A Web-scale Approach to Probabilistic Knowledge Fusion. In *ACM SIGKDD*, 2014.

[3] Cristóbal Esteban, Danilo Schmidt, Denis Krompaß, and Volker Tresp. Predicting sequences of clinical events by using a personalized temporal latent embedding model. In *IEEE International Conference on Healthcare Informatics*, 2015.

[4] Michael S Gazzaniga. Organization of the human brain. *Science*, 245(4921), 1989.

[5] Maximilian Nickel, Kevin Murphy, Volker Tresp, and Evgeniy Gabrilovich. A review of relational machine learning for knowledge graphs. *Proceedings of the IEEE*, 2015.

[6] Maximilian Nickel, Volker Tresp, and Hans-Peter Kriegel. A Three-Way Model for Collective Learning. In *ICML*, 2011.

[7] Amit Singhal. Introducing the Knowledge Graph: things, not strings. *Official Google Blog*, 2012.

[8] Volker Tresp, Cristóbal Esteban, Yinchong Yang, Stephan Baier, Yunpu Ma, and Denis Krompaß. Learning with memory embeddings. *arXiv preprint arXiv:1511.07972*, 2015.

[9] Volker Tresp and Yunpu Ma. The tensor memory hypothesis. *NIPS 2016 Workshop on Representation Learning in Artificial and Biological Neural Networks (MLINI)*, 2016.

[10] Volker Tresp, Yunpu Ma, Stephan Baier, and Yinchong Yang. Embedding learning for declarative memories. In *ESWC*, 2017.

[11] Yinchong Yang, Peter A Fasching, and Volker Tresp. Predictive modeling of therapy decisions in metastatic breast cancer with recurrent neural network encoder and multinomial hierarchical regression decoder. In *IEEE International Conference on Healthcare Informatics*, 2017.