Knowledge graphs for parameter optimisation in large-scale mechanistic models

Knowledge graphs (KGs) are widely used for learning statistical models, and multiple large-scale KGs that utilize statistical relational learning existed already a decade ago (e.g., YAGO, DBpedia, and the Google Knowledge Graph) [1]. Meanwhile, recent studies demonstrate successful combinations of KGs and mechanistic models [2, 3]. This approach is especially relevant when analysing biological data, as high complexity of biological systems calls for the relational approach of KGs, whereas prior knowledge on biophysical and biochemical processes allows assuming mechanistic models. An example of enriching a KG with a mechanistic model is adding attributes to a graph edge such that it would encode “A activates B with a certain rate that depends on the state of the local nodes and on the external parameters” instead of simply “A activates B”.

This project will focus on incorporating a mechanistic model of molecular interactions encoded in the form of high-dimensional ODEs into a KG and optimising the parameters of the model based on distance-scaled data sharing between related nodes. When approaching this task, it could be helpful to consider previous efforts in the area, e.g., [4]. Further, it is of interest to use KGs in order to define parameters for conditions that are not part of the testing data.

The project can be expanded based on the interests of the student. For example, it is possible to introduce stochastic ODEs to the mechanistic model in order to model noise in low-copy molecular numbers.

The student will learn about knowledge graphs, parameter optimisation techniques and mechanistic modelling of biosystems. Skills and knowledge in the following areas will be helpful when working on this project: programming, machine learning and parameter optimization, probability theory.

For further information, please contact Sofia Startceva.
References


