

Understanding A Symbolic Diagram: A Neural Network for Relational Reasoning and Prediction

What are you trying to do?

This project proposes a neural network model that takes in a diagram image which is drawn by human interpretable symbols and performs questions answering and prediction, based on the relations learning by the model.

For example, after training, this model should be able to take in a force diagram which is consisted of objects and arrows indicating the directions of forces, as an input and answer questions such as “is the red circle experiencing any force?” and make prediction such as “in what direction the red circle is moving?”.

How is it done today, and what are the limits of current practice?

This idea is inspired by the paper *A simple neural network module for relational reasoning* in which the author proposed a neural network model that can answer some basic questions regarding the relations of different objects in an image. The limit of this model is that its primary focus is identifying relations, so the model does not learn the actual impact of a relation or make any prediction based on the relations learned.

What is new in your approach and why do you think it will be successful?

First, my approach emphasis on understanding human symbols and the relations they represent. In the example of a force diagram, the traditional approach mentioned above will interpret a rectangle object and an arrow on it as two separate objects, while my model should be able to understand that the rectangle is an object and the arrow is a relation(force) on it.

Second, my model should be able to predict the impact of a relation. For example, given a force diagram where a rectangle has a left arrow on it, the model should be able to predict the moving direction of the rectangle to be left.

Who cares? If you are successful, what difference will it make?

Although I use force diagrams as examples and will probably use them as training tasks, the success of this model should be able to generalize into other domains. For example, with some improvements and potentially combining with a program synthesis function, the model should be able to generate data structure codes in DSL after receiving a diagram representation of the structure.

What are the risks?

One potential risk is the dataset. Since the training tasks require diagrams drawn with specific symbol and rules, existing image datasets for relational reasoning fails to meet the requirement. One solution is to use python to automate generate diagrams. The potential disadvantage of this approach is the validation of data may be a time-consuming problem if the model requires a large set to converge.

The other potential risk is the capability of the model. I plan to base on the open-source implementation of the model proposed in *A simple neural network module for relational reasoning*. There is a notable gap between identifying a relation and predicting based on a relation. My concern is that the existing model may not be enough for such a complicated task. I plan to graduate improving and modifying the model as the training goes.

How much will it cost?

Given the base model is open source and training dataset can be generated, the cost should be minimized.

How long will it take?

Roughly 8 weeks.

What are the mid-term and final “exams” to check for success?

Mid-Term: The model should at least be able to interpret and differentiate objects and their relations in a diagram.

Final: based on the success of mid-term, the model should be able to produce reliable prediction based on the diagram.