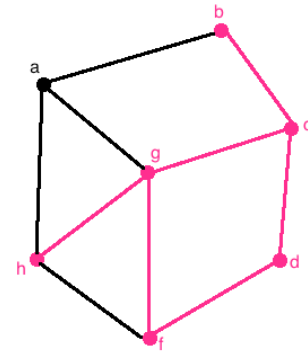




Topics in Graph-Walking Program Induction

Choose your favourite graph data sample where all nodes are labelled according to some scheme. Given a starting point and a set of endpoints from that graph, how can we identify the traversal query (algorithm) that was used to reach the endpoints?

Graph (network) data is ubiquitous and often very complex. Unlike relational database data (that can too be viewed as graph data), it usually does not benefit from a fixed, explicit structure and is therefore somewhat more difficult to work with.



At DISCO we've already developed a basic theory of graph-walking program induction. We are now looking to extend this theory, implement our algorithms for real-world use, and test them on large datasets, both synthetic and natural. Here are a few topics to give you an idea, each worthy its own thesis:

1. **Developing Simple Colour Program Library in Python.** The aim is to implement our algorithms as a separate Python library built on top of the NetworkX library.
2. **Extending the Graph-Walking Program Induction Theory.** A theory project.
3. **Graph-Walking Programs Meet Information Theory.** Some of our algorithms have a pretty bleak worst-case scenario runtime. However, given the properties of the classes of graphs we are likely to encounter in practice, the average runtime is certainly much better. This begs a proof. Basic background in information theory would be helpful for this project.

Who is this for? Bachelor's and master's students interested in diving into the above-mentioned problems. Some familiarity with graphs, graph terminology, and graph algorithms will make for a speedier start. Interest in research would be put to a good use.

Interested? Please contact us for more details!

Contact

- Peter Belcak: belcak@ethz.ch, ETZ G63