

# Learning to navigate, read and apply software documentation like a human

Breandan Considine, Xujie Si, Jin Guo

February 24, 2021

## 1 Introduction

Humans are adept information foraging agents. We can quickly find relevant information in a large corpus by recognizing and following textual landmarks. Software projects feature a variety of semi-structured documents containing many such clues indicating where contextually relevant information can be found. In this work, we train an agent to solve a programming task, by learning to navigate and read software artifacts like source code and documentation in order to facilitate a goal-directed programming task.

Our work broadly falls under the umbrella of text-based reinforcement learning. Prior literature falls under two categories: natural or formal language. Reinforcement learning (RL) in the natural domain typically focuses on question answering [4, 6], or interactive text games [9, 3, 11, 8, 2]. RL techniques have begun to show promising results for program synthesis [7, 10, 5]. Our work falls at the intersection of these two domains.

Early work in program learning realized the importance of graph-based representations [1], however required extensive feature-engineering. Adaption across domains, where style, terminology and document structure vary significantly, remains a challenge. More recent work in program synthesis has explored incorporating a terminal [7], graphical [12] or other user interface to explore the space of valid programs. However these approaches generally do not consider the scope or variety of artifacts in a software project. Some attempt to learn a graph [10] but only consider the local structure of source code and still require an explicit parser of some kind.

Unlike prior work, we consider the whole project and related documents. Instead of parsing their contents explicitly, which is computationally expensive and may be too large to fit in memory, we allow the agent to explore

an entire filesystem and construct a graph organically, which can later be decoded to predict a label or sequence at inference time.

Consider a newly-hired developer, who has programming experience, but no prior knowledge about a closed-source project. After onboarding, she gets access to the project and is assigned her first ticket: Fix test broken by 0fb98be. After finding the commit and getting familiar the codebase, she searches on StackOverflow, finds a relevant solution, then copies some code into the editor, makes a few changes, presses run, and the test passes.

In order to accomplish some task such as fixing a test or finding a bug, an information-seeking agent must locate resources in a database to solve the task. It needs to perform simple actions, like `search(query)`, `click(link)`, `read(text)`, `copy(text, from, to)`, and `runCode()`. During each rollout, the agent can explore the environment using the primitives available to build a lazily instantiated knowledge graph.

As the agent is exploring the environment, it collects information about how to solve the task. Similar to navigating a physical environment, it can visit relevant documents and source code artifacts. As the agent explores the project, they gather information in a latent state, which can be decoded at inference time to the proposed solution.

During evaluation, we then measure how well the agent performed. Many loss functions are possible depending on the task, from a simple distance metric on a deleted fragment of code, to a more complex property (e.g. presence or absence of an error, or some property of the output) which must be satisfied by interacting with the runtime environment. Many program analysis and repair tasks are amenable to the described setup, including defect, duplicate or vulnerability detection and correction.

## 2 Experiments

In this work, our primary focus is learning to read software artifacts in a software repository and apply the knowledge contained. Code is a graph database, or knowledge graph whose edges represent semantic relationships between entities. Our research seeks to answer the following research questions. Can the agent can learn to:

1. Navigate the project and locate contextually relevant artifacts?
2. Comprehend the semantic content of contextually relevant artifacts?
3. Apply the knowledge gathered to perform the assigned task?

In our first experiment, we attempt to understand which queries the agent is performing when solving a programming task. Does it search for certain keywords in the context?

In our second experiment, we try to measure the information contained in various filetypes by ablating results in various filetypes. For trajectories which explore various filetypes such as Markdown or Java, what information gain do resources of these filetypes provide?

Our next steps are to build an RL environment which allows an agent to interact with a knowledge base and runtime environment. We will use an in-memory graph database to store and load project artifacts to and from the knowledge base. To model the agent, we will use the actions described above for navigating the environment, searching for text and reading documents, with a transformer-based internal state, pretrained on a corpus of projects in the same language. Finally, to evaluate the agent, we will generate problem instances and use a JIT to rapidly validate generated code fragments.

## References

- [1] Miltiadis Allamanis, Marc Brockschmidt, and Mahmoud Khademi. Learning to represent programs with graphs. *arXiv preprint arXiv:1711.00740*, 2017.
- [2] Prithviraj Ammanabrolu and Matthew Hausknecht. Graph constrained reinforcement learning for natural language action spaces. *arXiv preprint arXiv:2001.08837*, 2020.
- [3] Prithviraj Ammanabrolu and Mark O Riedl. Playing text-adventure games with graph-based deep reinforcement learning. *arXiv preprint arXiv:1812.01628*, 2018.
- [4] Christian Buck, Jannis Bulian, Massimiliano Ciaramita, Wojciech Gajewski, Andrea Gesmundo, Neil Houlsby, and Wei Wang. Ask the right questions: Active question reformulation with reinforcement learning. *arXiv preprint arXiv:1705.07830*, 2017.
- [5] Yanju Chen, Chenglong Wang, Osbert Bastani, Isil Dillig, and Yu Feng. Program synthesis using deduction-guided reinforcement learning. In *International Conference on Computer Aided Verification*, pages 587–610. Springer, 2020.

- [6] Yu Chen, Lingfei Wu, and Mohammed J Zaki. Reinforcement learning based graph-to-sequence model for natural question generation. *arXiv preprint arXiv:1908.04942*, 2019.
- [7] Kevin Ellis, Maxwell Nye, Yewen Pu, Felix Sosa, Josh Tenenbaum, and Armando Solar-Lezama. Write, execute, assess: Program synthesis with a REPL. *arXiv preprint arXiv:1906.04604*, 2019.
- [8] Xiaoxiao Guo, Mo Yu, Yupeng Gao, Chuang Gan, Murray Campbell, and Shiyu Chang. Interactive fiction game playing as multi-paragraph reading comprehension with reinforcement learning. *arXiv preprint arXiv:2010.02386*, 2020.
- [9] Ji He, Jianshu Chen, Xiaodong He, Jianfeng Gao, Lihong Li, Li Deng, and Mari Ostendorf. Deep reinforcement learning with a natural language action space. *arXiv preprint arXiv:1511.04636*, 2015.
- [10] Daniel D Johnson, Hugo Larochelle, and Daniel Tarlow. Learning graph structure with a finite-state automaton layer. *arXiv preprint arXiv:2007.04929*, 2020.
- [11] Karthik Narasimhan, Tejas Kulkarni, and Regina Barzilay. Language understanding for text-based games using deep reinforcement learning. *arXiv preprint arXiv:1506.08941*, 2015.
- [12] Homer Walke, R Kenny Jones, and Daniel Ritchie. Learning to infer shape programs using latent execution self training. *arXiv preprint arXiv:2011.13045*, 2020.