ROOP: Translating Sequential Code Fragments to Distributed Code Fragments Using Deep Reinforcement Learning

Authors : Arun Sanjel, Greg Speegle

Abstract : Every second, massive amounts of data are generated, and Data Intensive Scalable Computing (DISC) frameworks have evolved into effective tools for analyzing such massive amounts of data. Since the underlying architecture of these distributed computing platforms is often new to users, building a DISC application can often be time-consuming and prone to errors. The automated conversion of a sequential program to a DISC program will consequently significantly improve productivity. However, synthesizing a user's intended program from an input specification is complex, with several important applications, such as distributed program synthesizing and code refactoring. Existing works such as Tyro and Casper rely entirely on deductive synthesis techniques or similar program synthesis approaches. Our approach is to develop a data-driven synthesis technique to identify sequential components and translate them to equivalent distributed operations. We emphasize using reinforcement learning and unit testing as feedback mechanisms to achieve our objectives.

Keywords : program synthesis, distributed computing, reinforcement learning, unit testing, DISC
Conference Title : ICMLDA 2023 : International Conference on Machine Learning and Data Analysis
Conference Location : New York, United States
Conference Dates : June 05-06, 2023