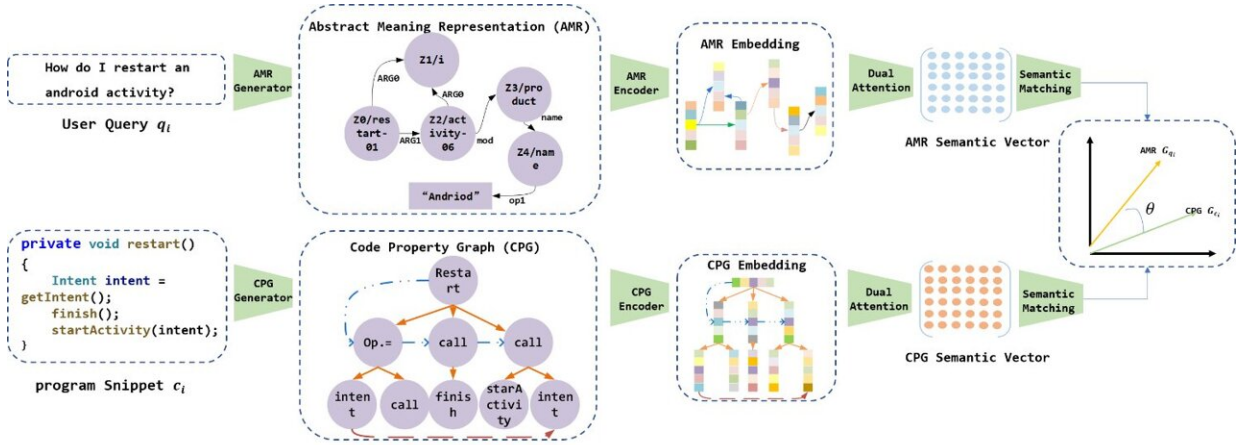
A multi-relational graph perspective on semantic similarity in program retrieval

May 29 2024



An overview of the proposed method, including: 1) Multi-relational graph construction; 2) Multi-relational graph embedding; 3) Semantic similarity calculation. Credit: *Frontiers of Computer Science* (2023). DOI: 10.1007/s11704-023-2678-8

Program retrieval remains a cornerstone of software development, crucial for boosting productivity throughout the development lifecycle. Amidst diverse program retrieval models, many have ignored the disparities between natural language queries and code, resulting in a prominent semantic gap.

Moreover, programs and queries carry rich structural and semantic

information. Yet, prevailing approaches often overlook the cohesion among different aspects of source code and treat queries as sequences, neglecting their inherent structural characteristics.

To solve these problems, a research team led by Yunwei Dong published their research in *Frontiers of Computer Science*.

The team proposed a framework that formulates program retrieval as a multi-relational graph similarity problem. Furthermore, a dual-level attention is applied to assign weights to nodes in multi-relational graphs by intra-relation and inter-relation level attention.

To begin, the multi-relational graph construction module focuses on representing programs and queries using code property graphs (CPG) and abstract meaning representations (AMR). This strategic approach facilitates a more comprehensive and nuanced portrayal of program and query semantics.

Then the dual-level attention graph neural network is leveraged to learn semantic information for AMR and CPG. Finally, a semantic similarity calculation module is designed to calculate the similarity of query-program pairs. Compared with the existing research results, the proposed method performs relatively well among all baselines.

Future research endeavors could concentrate on optimizing multi-relational graphs by minimizing extraneous information, thereby diminishing graph complexity. Additionally, a promising avenue lies in the deliberate integration of external knowledge, such as knowledge graphs, aiming to enhance the representation of program semantics.

More information: Qianwen Gou et al, Semantic similarity-based program retrieval: a multi-relational graph perspective, *Frontiers of Computer Science* (2023). DOI: 10.1007/s11704-023-2678-8

Provided by Frontiers Journals

Citation: A multi-relational graph perspective on semantic similarity in program retrieval (2024, May 29) retrieved 29 June 2024 from https://techxplore.com/news/2024-05-multi-graph-perspective-semantic-similarity.html

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.