An Artifact for the Article:

“Dissecting Real-World Cross-Language Bugs”

ABSTRACT

This artifact provides a comprehensive suite to reproduce our study on cross-language bugs in multilingual software. It includes a fully functional version of our novel static cross-language analysis tool, all associated datasets, and the intermediate results generated during the experiments. Users can access the tool via our Docker image or download the complete artifact package. We recommend a system configuration with at least 32 GB of disk space and 16 GB of memory to ensure smooth execution of the experiments.

The artifact is organized into several key components reflecting the different phases of our study. The first component comprises the outputs of our automated tool that identifies and classifies cross-language bug candidates from large-scale code repositories. This includes intermediate data such as multi-language repositories' information and cross-language bug commits.

In addition to the automated analyses, this artifact incorporates a manually validated dataset. Detailed annotations, hand-labeled examples, and comprehensive documentation are provided to support the automated findings, ensuring that the classification and analysis of the cross-language bugs are robust and reliable. This dual approach enhances the overall credibility of the study and offers a rich resource for further research into the challenges of multilingual software development.

For complete installation instructions, usage guidelines, and further technical details, please refer to the README file included in the artifact package. We believe that this comprehensive artifact will serve as a valuable resource for researchers and practitioners seeking to understand and mitigate the complexities inherent in cross-language software systems.