

InterEvo-TR: Interactive Evolutionary Test Generation with Readability Assessment*

Pedro Delgado-Pérez¹[0000-0003-1568-9288],
Aurora Ramírez²[0000-0002-1916-6559],
Kevin J. Valle-Gómez¹[0000-0001-6066-9441],
Inmaculada Medina-Bulo¹[0000-0002-7543-2671], and
José Raúl Romero²[0000-0002-4550-6385]

¹ University of Cádiz, Cádiz, Spain
{pedro.delgado, kevin.valle, inmaculada.medina}@uca.es
² University of Córdoba, Córdoba, Spain
{aramirez, jrromero}@uco.es

Keywords: testing tools, evolutionary computing and genetic algorithms, interactive search-based software engineering, readability

Published in: IEEE Transactions on Software Engineering, pp. 1–17, 2022

Impact Factor: JCR-2021 9,322 - Q1 - Position: 3/110 - Area: COMPUTER SCIENCE / SOFTWARE ENGINEERING

DOI: <https://doi.org/10.1109/TSE.2022.3227418>

Abstract. Automated test case generation has proven to be useful to reduce the usually high expenses of software testing. However, several studies have also noted the skepticism of testers regarding the comprehension of generated test suites when compared to manually designed ones. This fact suggests that involving testers in the test generation process could be helpful to increase their acceptance of automatically-produced test suites. In this paper, we propose incorporating interactive readability assessments made by a tester into EvoSuite, a widely-known evolutionary test generation tool. Our approach, INTEREVO-TR, interacts with the tester at different moments during the search and shows different test cases covering the same coverage target for their subjective evaluation. The design of such an interactive approach involves a schedule of interaction, a method to diversify the selected targets, a plan to save and handle the readability values, and some mechanisms to customize the level of engagement in the revision, among other aspects. To analyze the potential and practicability of our proposal, we conduct a controlled experiment in which 39 participants, including academics, professional developers, and student collaborators, interact with INTEREVO-TR. Our results show that the strategy to select and present intermediate results is effective for the purpose of readability assessment. Furthermore, the

* Supported by the European Commission (ERDF), the Spanish Ministry of Science and Innovation (PID2021-122215NB-C33, RTI2018-093608-BC33, RED2018-102472-T, Grant PID2020-115832GB-I00 funded by MICIN/AEI/10.13039/501100011033) and the Andalusian Regional Government (postdoctoral grant DOC_00944).



participants' actions and responses to a questionnaire allowed us to analyze the aspects influencing test code readability and the benefits and limitations of an interactive approach in the context of test case generation, paving the way for future developments based on interactivity.