

Automatic Grading of Programming Exercises using Property-Based Testing

Clara Benac Earle

Universidad Politécnica de Madrid, Spain

cbenac@fi.upm.es

Lars-Åke Fredlund

Universidad Politécnica de Madrid, Spain

lfredlund@fi.upm.es

John Hughes

Chalmers University of Technology and Quviq AB, Göteborg, Sweden

rjmh@chalmers.se

We present a framework for automatic grading of programming exercises using property-based testing, a form of model-based black-box testing. Models are developed to assess both the functional behaviour of programs and their algorithmic complexity. From the functional correctness model a large number of test cases are derived automatically. Executing them on the body of exercises gives rise to a (partial) ranking of programs, so that a program A is ranked higher than program B if it fails a strict subset of the test cases failed by B. The model for algorithmic complexity is used to compute worst-case complexity bounds. The framework moreover considers code structural metrics, such as McCabe's cyclomatic complexity, giving rise to a composite program grade that includes both functional, non-functional, and code structural aspects. The framework is evaluated in a course teaching algorithms and data structures using Java.

Article published at the 21th Annual Conference on Innovation and Technology in Computer Science Education (ITiCSE 2016), in Arequipa, Peru, July 11 to 13, 2016.