

Structural testing criteria for concurrent programs considering loop executions

Silvia Diaz Diaz ¹, Paulo Sérgio Lopes de Souza ¹

¹University of Sao Paulo

Abstract. Parallel programs are imperative for improving performance and problem solving, having an increasing demand on implementing efficient parallel programming techniques. This entails new challenges on software testing to ensure their quality and reliability. Structural testing is a technique that allows the identification of concurrency defects by analyzing the internal structure of the program. However, the non-determinism of concurrent programs has implications in the testing activity, requiring the use of structured methods to reveal defects. Testing criteria support the selection of test cases in a systematic form by statically analysing elements of concurrent programs. We found that there are currently gaps in the definition of testing criteria contemplating scenarios with elements that are dynamically evaluated, such as the execution of communication primitives inside loops. The objective of this project is to define structural testing criteria to guide the selection of test cases, improving the reliability of concurrent programs by revealing non-determinism related errors present in repetition structures. We developed a Concurrent Defects Taxonomy, identifying and classifying concurrency types of defects found in related literature. The analysis of such defects, paths inside loops, number of loop iterations, and nested loops allow us to model the proposed structural testing criteria. We define new sets and associations related to communication and synchronization flows for message-passing programs, establishing a model for testing criteria. We implemented the proposed test model in ValiMPI, a testing tool prototype, considering the new concepts defined in our test model, generating required elements and evaluating coverage after constructing loop paths. For the application evaluation of criteria we perform an empirical study with statistical validation, indicating the results for cost, effectiveness and strength. Our experimental evaluation demonstrated that the proposed testing criteria generates required elements that support the identification of concurrency defects occurring in different loop iterations, when having communicational events with non-deterministic behavior.

Resumo. A programação paralela é imperativa para melhorar o desempenho e a resolução eficiente de problemas, tendo uma demanda crescente na implementação de técnicas de programação paralela. Isso implica novos desafios no teste de software para garantir a qualidade e confiabilidade. O teste estrutural é uma técnica que permite a identificação de defeitos de concorrência, analisando a estrutura interna do programa. No entanto, os programas concorrentes são não determinísticos, com desafios na atividade de teste, exigindo o uso de métodos estruturados para revelar defeitos. Os

critérios de teste suportam a seleção de casos de teste de forma sistemática, analisando estaticamente elementos de programas concorrentes. Foi encontrado que atualmente existem lacunas na definição de critérios de teste contemplando cenários com elementos dinâmicos, como a execução de primitivas de comunicação dentro de loops. O objetivo deste projeto é definir critérios estruturais para orientar a seleção de casos de teste, revelando erros relacionados ao não-determinismo e melhorando a confiabilidade de programas concorrentes. Foi desenvolvida uma Taxonomia de Defeitos Concorrentes, identificando e classificando os tipos de defeitos de concorrência encontrados na literatura relacionada. A análise de tais defeitos, a seleção de caminhos de loop, o número de iterações de loop e loops aninhados permitem modelar os critérios de testes estruturais propostos. Foram definidos novos conjuntos e associações relacionadas aos fluxos de comunicação e sincronização de programas de passagem de mensagens, estabelecendo um modelo para os critérios de teste. O modelo de teste proposto foi implementado no protótipo de ferramenta de teste chamada ValiMPI, considerando as associações definidas para os critérios propostos, gerando elementos necessários e cobertura de avaliação após a identificação dos nós de loop. Para a avaliação da aplicação dos critérios, foi realizado um estudo empírico com testes estatísticos, indicando os resultados para custo, efetividade e strength. A avaliação experimental demonstrou que os critérios de teste propostos geram elementos necessários que suportam a identificação de defeitos presentes em diferentes iterações dos loops, quando existem eventos de comunicação com comportamento não determinístico.