Interaction specification mining with adversarial examples
PhD position at CEA LIST
Université Paris-Saclay

1 Description

1.1 Context

Software systems are often developed in a modular way by (re)using components that cooperate with each other to provide a more global service. We will consider component intensive software systems in particular those composed of communicating components via message passing and communication primitives called send/receive actions. An example of such systems are client/server systems, transportation control systems, Internet of Things, connected autonomous vehicles, etc. In practice, the assembly of these components follows poorly documented ad-hoc procedures. Therefore, discovering their formal specifications, i.e., specification mining [1, 2], is of great interest for many software Verification and Validation (V&V) activities.

This doctoral work will be conducted in the frame of the HORIZON Europe project SELFY.

SELFY - SELF ASSESSMENT, PROTECTION & HEALING TOOLS FOR A TRUSTWORTHY AND RESILIENT CCAM

N.B., CCAM stands for Cooperative, Connected and Automated Mobility [1]. SELFY aims to increase CCAM ecosystem's safety, security, robustness, and resilience by researching and developing a toolbox made of collaborative tools, including a V&V tool based on specification mining. The PhD student will be involved in research collaboration with Okayama University (Japan), an international associate partner linked to CEA in the project.

1.2 Work

The objective of the doctoral work is to develop new methods for mining specifications from execution logs of communicating components. These logs are collected via code instrumentation or sniffing or via a the testing architecture. Logged executions are typically sequences of actions occurring at interface of the components which are expected to implement some communication protocol.

In this doctoral work the mining will target specification which are *interactions* such a UML Sequence Diagrams (UML-SD) or Message Sequence Charts (MSC). It will based on a recent work that provides expressive interactions (include rich scheduling and choice operators) with rewriting-based operational unfolding (defining transitions between interaction terms upon occurrence of send/receive action) and implemented into the tool HIBOU [5, 4].

We foresee an interaction mining framework which combines:

- *a test generation process identifying adversarial examples*

  This will be based on key properties such that reachability properties or more generally LTL (Linear Temporal Logic) properties. The tester adversarially guides test generation, searching for counterexamples to these specifications as conducted for instance in [3] for mining finite-state automata to invalidate spurious properties.

- *an efficient interaction mining process*

[1] `https://www.ccam.eu/`

The interaction mining process will apply rewriting strategies based on [4] to infer concise generalizing interactions from the execution logs of the communicating components.We will rely on the algebraic properties of the interactions, either in relation to the scheduling operators themselves (associativity, commutativity, neutral element), or to the projection mechanisms from a global system to its sub-systems

The mining framework will be experimented on uses cases issued from the project SELFY.

1.3 Expected outputs

The position comprises both theoretical work and coding. The expected outputs are various : resolution of research problems, implementation of solutions into evaluated prototypes, publication at top conferences and journals, and participation in the scientific life of the team as well as in the HORIZON Europe project SELFY.

Besides the doctoral work, it will also be possible if interested, to have some teaching activities at Paris-Saclay University or at CentraleSupélec.

2 Requested skills

The applicant must have a master's degree in computer science or a specialization in computer science in an engineering school. The applicant should have:

- excellent programming skills, the knowledge of the programming language Rust (`https://www.rust-lang.org/`) used to implement the HIBOU tool (`https://github.com/erwanM974/hibou_label`) is a plus.
- some knowledge of formal methods, be it rewriting techniques, formal modeling (automata-based, process algebra, etc.), formal verification, model learning, model-based testing.

3 Contact

To candidate please send a detailed CV and at least one reference to the following mails:

- Pascale Le Gall, MICS, CentraleSupélec, Université Paris-Saclay
  email : pascale.legall@centralesupelec.fr
- Boutheina Bannour, CEA-LIST, CEA, Université Paris-Saclay
  email : boutheina.bannour@cea.fr
- Bernard Chenevier, OKAYAMA University
  email : bernard-chenevier@cc.okayama-u.ac.jp

A delay of two or three months is to be expected for the administrative processing of the application at CEA.

References

[1] Glenn Ammons, Rastislav Bodík, and James R. Larus. Mining specifications. In John Launchbury and John C. Mitchell, editors, *Conference Record of POPL 2002: The 29th SIGPLAN-SIGACT Symposium on Principles of Programming Languages, Portland, OR, USA, January 16-18, 2002*, pages 4–16. ACM, 2002.

[2] Sara Belluccini, Rocco De Nicola, Barbara Re, and Francesco Tiezzi. PALM: A technique for process algebraic specification mining. In Brijesh Dongol and Elena Troubitsyna, editors, *Integrated Formal Methods - 16th International Conference, IFM 2020, Lugano, Switzerland, November 16-20, 2020, Proceedings*, volume 12546 of *Lecture Notes in Computer Science*, pages 397–418. Springer, 2020.

[3] Hong Jin Kang and David Lo. Adversarial specification mining. *ACM Trans. Softw. Eng. Methodol.*, 30(2):16:1–16:40, 2021.

[4] Erwan Mahe. *An operational semantics of interactions for verifying partially observed executions of distributed systems. (Sémantique opérationnelle des interactions pour la vérification d'exécutions partiellement observées de systèmes distribués).* PhD thesis, University of Paris-Saclay, France, 2021.

[5] Erwan Mahe, Boutheina Bannour, Christophe Gaston, Arnault Lapitre, and Pascale Le Gall. A small-step approach to multi-trace checking against interactions. In Chih-Cheng Hung, Jiman Hong, Alessio Bechini, and Eunjee Song, editors, *SAC '21: The 36th ACM/SIGAPP Symposium on Applied Computing, Virtual Event, Republic of Korea, March 22-26, 2021*, pages 1815–1822. ACM, 2021.