Engineering with Full-scale Formal Architecture: Morello, CHERI, Armv8-A, and RISC-V

Peter Sewell
University of Cambridge
Cambridge, UK
Peter.Sewell@cl.cam.ac.uk

Abstract—Architecture specifications define the fundamental interface between hardware and software. Historically, mainstream architecture specifications have been informal prose-and-pseudocode documents. This talk will describe our work to establish and use mechanised semantics for full-scale instruction-set architectures (ISAs): the mainstream Armv8-A architecture, the emerging RISC-V architecture, the CHERI-MIPS and CHERI-RISC-V research architectures that use hardware capabilities for improved security, and Arm’s prototype Morello architecture – an industrial demonstrator incorporating the CHERI ideas.

We use a variety of tools, especially our Sail ISA definition language and Isla symbolic evaluation engine, to build semantic definitions that are readable, executable as test oracles, support reasoning within the Coq, HOL4, and Isabelle proof assistants, support SMT-based symbolic evaluation, support model-based test generation, and can be integrated with operational and axiomatic concurrency models. These models are all complete enough to boot operating systems and hypervisors, covering the full sequential ISA (though not other SoC components, such as the Arm Generic Interrupt Controller). They range from 5000 to 60000 lines of specification.

For CHERI-MIPS and CHERI-RISC-V, we have used Sail models (and previously L.3 models) as the golden reference during design, working with our systems and computer architecture colleagues in the CHERI team to use lightweight formal specification routinely in documentation, testing, and test generation. We have stated and proved (in Isabelle) some of the fundamental intended security properties of the full CHERI-MIPS ISA.

For Armv8-A, building on Arm’s internal shift to an executable model in their ASL language, we have the complete sequential ISA semantics automatically translated from the Arm ASL to Sail, and for RISC-V, we have hand-written what is now the officially adopted model. For their concurrent semantics, the “user” semantics, partly as a result of our collaborations with Arm and within the RISC-V concurrency task group, have become simplified and well-defined, with multiple models proved equivalent, and we are currently working on the “system” semantics. Our symbolic execution tool for Sail specifications, Isla, supports axiomatic concurrency models over the full ISA.

Morello, supported by the UKRI Digital Security by Design programme, offers a path to hardware enforcement of fine-grained memory safety and/or secure encapsulation in the production Armv8-A architecture, potentially excluding or mitigating a large fraction of today’s security vulnerabilities for existing C/C++ code with little modification. During the ISA design process, we have proved (in Isabelle) fundamental security properties for the complete Morello ISA definition, and generated tests from the definition which were used during hardware development and for QEMU bring-up.

All these tools and models are (or will soon be) available under open-source licences, providing well-validated models for others to use and build on.

This is joint work by many people, including especially, for Sail and RISC-V: Alasdair Armstrong, Brian Campbell, Kathryn E. Gray, Mark Wassell, Jon French, Neel Krishnaswami; for Morello verification and ASL-to-Sail translation: Thomas Bauereiss, Thomas Sewell, Brian Campbell, Alasdair Armstrong, Alastair Reid; for Morello and CHERI-MIPS test generation: Brian Campbell; for:CHERI-MIPS verification: Kynthia Nienhuis; for RISC-V and CHERI-RISC-V specifications: Robert M. Norton, Prashanth Mundkur, Jessica Clark; for MIPS and CHERI-MIPS specifications: Alexandre Joannou, Anthony Fox, Michael Roe, Matthew Naylor; and for Concurrency semantics: Christopher Pulte, Shaked Flur, Will Deacon, Ben Simner, Luc Maranget, Susmit Sarkar, Jean Pichon-Pharabod, Ohad Kammar, Jeehoon Kang, Sung-Hwan Lee, Chung-Kil Hur. All this is in collaboration with the rest of the CHERI team and others in Arm (especially Richard Grisenthwaite, Graeme Barnes, and the Morello team) and in the RISC-V community, with the CHERI team jointly led by Robert N. M. Watson, Simon W. Moore, Peter Sewell, Peter G. Neumann, and Ian Stark.

Fig. 1. Sail models and infrastructure (grayed-out models are partial ISA models in an older version of Sail)