DOI 10.1007/s00165-007-0064-1

Formal Aspects of Computing (2008) 20: 1–3

Formal Aspects of Computing

Editorial

This issue of *Formal Aspects* is devoted to an experiment conducted as part of the world-wide Grand Challenge in Verified Software. The challenge is to achieve a significant body of verified programs that have precise external specifications, complete internal specifications, and machine-checked proofs of correctness with respect to a sound theory of programming. The first pilot project in the challenge was to mechanise the proof of correctness of the Mondex smart-card for electronic finance. Eight international research groups tackled the problem, and the following papers record the experiences of six of these groups. The pilot project remains open for other groups to contribute.

The certification of the Mondex electronic purse to ITSEC Level E6

Woodcock, Stepney, Cooper, Clark, and Jacob were all involved in the original work on Mondex; the first three were specifiers and the last two evaluators. They recall the work that led to the successful certification of the Mondex electronic purse, and the research that this inspired. The paper contains an introduction to the Mondex specification and refinement, and an overview of the proof.

Mondex, an electronic purse: specification and refinement checks with the Alloy model-finding method

Ramananandro started from the existing specification on Mondex in Z, and constructed a specification in the Alloy specification language, which is based on relational first-order logic with transitive closures. The experiment shows that, if the concerns about finiteness are dropped, then the Mondex specification can be expressed in first-order logic without transitive closures. Ramananandro checked the specification with the Alloy Analyser, a tool for finding models. The Analyser translates the specification into a boolean formula, which it then tries to satisfy. If an assignment of variables is found, then the program translates it back to get a counterexample. This can be accomplished only for bounded numbers of objects: their scope. The specification has been checked for a scope of at most eight objects of each kind. The analysis found several bugs in Mondex: (i) purses can hold unauthentic transaction details; (ii) a wrong case analysis in a proof; (iii) a mistake in a framing schema.

Verification of Mondex electronic purses with KIV: from transactions to a security protocol

Haneberg, Schellhorn, Grandy, and Reif report on three major results from their experiment:

1. They have verified the entire case study using the KIV specification and verification system. They translated Z into algebraic specifications with operational-style ASM rule descriptions as auxiliaries in defining the protocol operations. They tried to follow the specification structure and the data refinement theory (including the use of simulations and invariants) of the original case study as closely as possible. In contrast, the original proof structure was completely abandoned in favour of one that was more appropriate for their theorem prover, KIV.

2. They formalised the underlying theory of data refinement in KIV to check the validity of the proof obligations. This also uncovered a small problem and led to an improvement of the refinement theory: the combination of backwards simulation in the contract approach with invariants. This allowed them to verify the case study as just a single refinement instead of the original two.
3. They extended the case study by adding another refinement to a security protocol based on abstract cryptography.

One person month was needed to formally prove the Mondex case study.

An incremental development of the Mondex system in Event-B

Butler and Yadav carried out Event-B development of the Mondex system using B4free, a proof obligation generator and proof tool for B. Their development is characterised by the use of many levels of refinement that require a larger number of simpler proofs. Their complete development consists of ten levels: an abstract specification and nine levels of refinement. The full development resulted in 703 proof obligations, of which over 97% were proved automatically. The remaining proof obligations were proved interactively using B4free.

The development took approximately 2 weeks of total effort spread over several months. The bulk of the work was spent in constructing models at different levels of abstraction and in finding the right gluing invariants. They did not make use of the original invariants, but used their theorem prover to guide the discovery of new ones. Most of the interactive proof effort was used to discover new invariants rather than to discharge proof obligations.

Modeling and validating Mondex Scenarios described in UML and OCL with USE

Kuhlmann and Gogolla describe the Mondex case study with a simple UML class diagram, including a class representing purses with appropriate attributes and a single operation transfer. The original constraints have been formulated either as OCL class invariants or as OCL pre and postconditions of the transfer operation. The UML class diagram and OCL constraints have been syntactically checked by the USE tool, which validates a specification by testing it with various scenarios. The scenarios are specified as single UML object diagrams, or by constructing a sequence of operation calls leading from an explicitly given start state to a result state. During scenario validation, the validity of invariants and pre and post-conditions are checked by the USE tool. The states obtained are then further explored with OCL queries retrieving interesting objects and properties. The authors have validated the abstract specification of Mondex using both positive and negative test cases.

Specification, proof, and model checking of the Mondex electronic purse using RAISE

George and Haxthausen's approach using RAISE is is similar in spirit to the original Z approach, in that it contains an abstract, an intermediate, and a concrete specification, and is based on refinements between these three. But the detail is quite different.

The abstract level sees the problem as one in accounting, and contains only three values: the total money in circulation, the total money in messages in transit, and the money that has been lost. The intermediate level introduces purses with individual balances, current payment details and statuses, but with an abstract logging mechanism and no message sequence numbers. The concrete level defines the purses' exception logs, the log archiving mechanism and the sequence numbers. The proofs were done using a translator from the RAISE Specification Language (RSL) to PVS and the PVS theorem prover itself.

The proofs are large, but straightforward; about half were proved automatically. The invariants are quite complicated, and their construction required a detailed understanding of the protocol and why it works.

They also used a translator from RSL to SAL and model checked the specification. This required reducing the size of the model in various ways, in particular to two purses and four sequence numbers. With the SAL model checker they were able to check the basic correctness conditions, and also (i) some liveness properties like transfers between purses and loss of money being possible, (ii) that all the invariants hold, and (iii) that confidence conditions (subtype correctness and precondition satisfaction) are true, using a second translation to SAL generated by the RAISE tool.

Mechanising Mondex with Z/Eves

Freitas and Woodcock mechanised the original work using Z/Eves, making only a few very small changes to the original models, so the mechanisation can be regarded as a reference model. They took about eight weeks to complete the mechanisation of the entire specification and refinement and its proof. The Z/Eves theorem prover was available ten years ago, so the experiment demonstrates that the mechanisation of the proof could have been carried out when the original work was done, and that it would have required about 10% of additional effort. Several problems were found: (i) operations involving inauthentic purses; (ii) assumptions regarding finiteness; (iii) four missing properties in the intermediate model; and (iv) an undocumented assumption in the upper refinement.

Cliff Jones and Jim Woodcock