2nd Workshop on Formal Methods for Blockchains

FMBC 2020, July 20–21, 2020, Los Angeles, California, USA (Virtual Conference)

Edited by

Bruno Bernardo
Diego Marmsoler

OASICS

Editors

Bruno Bernardo
Nomadic Labs, Paris, France
bruno@nomadic-labs.com

Diego Marmsoler
University of Exeter, UK
d.marmsoler@exeter.ac.uk

ACM Classification 2012
Security and privacy → Logic and verification; Software and its engineering → Formal software verification; Security and privacy → Distributed systems security; Computer systems organization → Peer-to-peer architectures

ISBN 978-3-95977-169-6

Published online and open access by
Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at https://www.dagstuhl.de/dagpub/978-3-95977-169-6.

Publication date
December, 2020

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at https://portal.dnb.de.

Digital Object Identifier: 10.4230/OASIcs.FMBC.2020.0

ISBN 978-3-95977-169-6 **ISSN 1868-8969** **https://www.dagstuhl.de/oasics**

OASIcs – OpenAccess Series in Informatics

OASIcs aims at a suitable publication venue to publish peer-reviewed collections of papers emerging from a scientific event. OASIcs volumes are published according to the principle of Open Access, i.e., they are available online and free of charge.

ISSN 1868-8969

https://www.dagstuhl.de/oasics

Contents

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler
OASICS OpenAccess Series in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Preface

The 2nd Workshop on Formal Methods for Blockchains (FMBC) took place virtually on July 20/21 2020 as part of CAV 2020, the 32nd International Conference on Computer-Aided Verification. Its purpose was to be a forum to identify theoretical and practical approaches applying formal methods to blockchain technology.

This second edition of FMBC attracted 18 submissions (10 long papers, 4 short papers, and 4 extended abstracts) on topics such as verification of smart contracts or analysis of consensus protocols. Each paper was reviewed by at least three program committee members or appointed external reviewers. This led to a selection of 10 papers (7 long and 3 short) that will be presented at the workshop as regular talks, as well as 1 long paper and 4 extended abstracts that will be presented as lightning talks. Additionally, we were very pleased to have an invited keynote by Grigore Rosu (University of Illinois at Urbana-Champaign).

This volume contains the papers selected for regular talks, the extended abstracts and paper selected for lightning talks as well as the abstract of the invited talk. Before inclusion, the papers were reviewed a second time after the workshop by the program committee.

We thank all the authors that submitted a paper, as well as the program committee members and external reviewers for their immense work. We are grateful to Shuvendu Lahiri and Chao Wang, Program Chairs of CAV 2020, and to Zvonimir Rakamaric, Workshop Chair of CAV 2020, for their support and guidance. Finally, we would like to express our gratitude to our sponsor Nomadic Labs for its generous support.

October 2020

Bruno Bernardo
Diego Marmsoler

Program Committee

Wolfgang Ahrendt
Chalmers University of Technology, Sweden

Lacramioara Astefanoei
Nomadic Labs, France

Massimo Bartoletti
University of Cagliari, Italy

Bernhard Beckert
Karlsruhe Institute of Technology, Germany

Bruno Bernardo
Nomadic Labs, France

Achim Brucker
University of Exeter, UK

Silvia Crafa
Universita di Padova, Italy

Zaynah Dargaye
Nomadic Labs, France

Jérémie Decouchant
University of Luxembourg, Luxembourg

Ansgar Fehnker
University of Twente, Netherlands

Georges Gonthier
Inria, France

Maurice Herlihy
Brown University, USA

Florian Kammueller
Middlesex University London, UK

Igor Konnov
Informal, Austria

Andreas Lochbihler
Digital Asset, Switzerland

Diego Marmsoler
University of Exeter, UK

Anastasia Mavridou
NASA Ames, USA

Simão Melo de Sousa
Universidade da Beira Interior, Portugal

Andrew Miller
University of Illinois at Urbana-Champaign, USA

Karl Palmskog
KTH, Sweden

Vincent Rahli
University of Birmingham, UK

Andreas Rossberg
Dfinity Foundation, Germany

Claudio Russo
Dfinity Foundation, USA

César Sanchez
Imdea, Spain

Clara Schneidewind
TU Wien, Austria

Ilya Sergey
Yale-NUS College/NUS, Singapore

Bas Spitters
Aarhus University/Concordium, Denmark

Mark Staples
CSIRO Data61, Australia

Meng Sun
Peking University, China

Simon Thompson
University of Kent, UK

Philip Wadler
University of Edinburgh / IOHK, UK

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler
OASICS OpenAccess Series in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Supporting Reviewers

Luis Arrojado da Horta

Yi Li

João Santos Reis

Ralf Sasse

Søren Eller Thomsen

Formal Design, Implementation and Verification of Blockchain Languages Using K

Grigore Rosu
University of Illinois at Urbana-Champaign, Urbana, IL, USA
`http://fsl.cs.illinois.edu/index.php/Grigore_Rosu`
grosu@illinois.edu

Abstract

The usual post-mortem approach to formal language semantics and verification, where the language is firstly implemented and used in production for many years before a need for formal semantics and verification tools naturally arises, simply does not work anymore. New blockchain languages or virtual machines are proposed at an alarming rate, followed by new versions of them every few weeks, together with programs (or smart contracts) in these languages that are responsible for financial transactions of potentially significant value. Formal analysis and verification tools are therefore needed immediately for such languages and virtual machines. We will present recent academic and commercial results in developing blockchain languages and virtual machines that come directly equipped with formal analysis and verification tools. The main idea is to generate all these automatically, correct-by-construction from a formal language specification.

2012 ACM Subject Classification Software and its engineering → Semantics

Keywords and phrases Blockchain, K Framework

Digital Object Identifier 10.4230/OASIcs.FMBC.2020.1

Category Invited Talk

Bio

Grigore Rosu is a professor in the Department of Computer Science at the University of Illinois at Urbana-Champaign (UIUC), where he leads the Formal Systems Laboratory (FSL), and the founder of Runtime Verification, Inc (RV). His research interests encompass both theoretical foundations and system development in the areas of formal methods, software engineering and programming languages. Before joining UIUC in 2002, he was a research scientist at NASA Ames. He obtained his Ph.D. at the University of California at San Diego in 2000. He was offered the CAREER award by the NSF, the Dean's award for excellence in research by the College of Engineering at UIUC in 2014, and the outstanding junior award by the Computer Science Department at UIUC in 2005. He won the ASE IEEE/ACM most influential paper award in 2016 (for an ASE 2001 paper) and the RV test of time award (for an RV 2001 paper) for papers that helped shape the runtime verification field, the ACM SIGSOFT distinguished paper awards at ASE 2008, ASE 2016, and OOPSLA 2016, and the best software science paper award at ETAPS 2002.

Formal Specification and Verification of Solidity Contracts with Events

Ákos Hajdu
Budapest University of Technology and Economics, Hungary
hajdua@mit.bme.hu

Dejan Jovanović
SRI International, New York City, NY, USA
dejan.jovanovic@sri.com

Gabriela Ciocarlie
SRI International, New York City, NY, USA
gabriela.ciocarlie@sri.com

Abstract

Events in the Solidity language provide a means of communication between the on-chain services of decentralized applications and the users of those services. Events are commonly used as an abstraction of contract execution that is relevant from the users' perspective. Users must, therefore, be able to understand the meaning and trust the validity of the emitted events. This paper presents a source-level approach for the formal specification and verification of Solidity contracts with the primary focus on events. Our approach allows the specification of events in terms of the on-chain data that they track, and the predicates that define the correspondence between the blockchain state and the abstract view provided by the events. The approach is implemented in SOLC-VERIFY, a modular verifier for Solidity, and we demonstrate its applicability with various examples.

2012 ACM Subject Classification Software and its engineering → Formal methods

Keywords and phrases Smart contracts, Solidity, events, specification, verification

Digital Object Identifier 10.4230/OASIcs.FMBC.2020.2

Category Short Paper

Related Version Preprint is available at `https://arxiv.org/abs/2005.10382`.

Supplementary Material `https://github.com/SRI-CSL/solidity`

1 Introduction

Ethereum is a public, blockchain-based computing platform that provides a single-world-computer abstraction for developing decentralized applications [17]. The core of such applications are programs – termed smart contracts [15] – deployed on the blockchain. While Ethereum nodes run a low-level virtual machine (EVM [17]), smart contracts are usually written in a high-level, contract-oriented language, most notably Solidity [14]. The contract code can be executed by issuing transactions to the network, which are then processed by the participating nodes. Results of a completed transaction are provided to the issuing user, and other interested parties observing the contract, through transaction receipts. While the blockchain is publicly available for users to inspect and replay the transactions, the contracts can communicate important state changes, including intermediate changes, by emitting events [1]. Events usually represent a limited abstract view of the transaction execution that is relevant for the users and they can be read off the transaction receipts. The common expectation is that by observing the events, the user can reconstruct the relevant parts of the current state of the contracts. Technically, events can be viewed as special triggers with

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler; Article No. 2; pp. 2:1–2:9
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

arguments that are stored in the blockchain logs. While these logs are programmatically inaccessible from contracts, the users can easily subscribe to and observe the events with the accompanying data. For example, a token-exchange application can monitor the current state of token balances by tracking transfer events in the individual token contracts.

Smart contracts, as any software, are also prone to bugs and errors. In the Ethereum context, any flaws in contracts come with potentially devastating financial consequences, as demonstrated by various infamous examples [2]. While there has been a great interest in applying formal methods to smart contracts [2, 4], events are usually considered merely a logging mechanism that is not relevant for functional correctness. However, since events are a central state-change notification mechanism for users of decentralized applications, it is crucial that the users are able to understand the meaning and trust the validity of the emitted events. In this paper, we propose a source-level approach for the formal specification and verification of Solidity contracts with the primary focus on events. Our approach provides in-code annotations to specify events in terms of the blockchain data they track, and to declare events possibly emitted by functions. We verify that (1) whenever tracked data changes, a corresponding event is emitted, and (2) an event can only be emitted if there was indeed a change. Furthermore, to establish the correspondence between the abstract view provided by events and the actual execution, we allow events to be annotated with predicates (conditions) that must hold before or after the data change. We implemented the proposed approach in the open-source[1] SOLC-VERIFY [9, 8] tool and demonstrated its applicability via various examples. SOLC-VERIFY is based on modular program verification, but we present our idea in a more general setting that can serve as a building block for alternative verification approaches.

Related work. To the best of our knowledge, our approach is the first to enable formal specification and verification of Solidity events in terms of the contract state. MYTHRIL [12] operates over compiled bytecode and focuses on weaknesses defined in the SWC Registry,[2] which currently does not include events. SLITHER [7] supports a few common, built-in patterns related to events (e.g., out-of-order due to reentrancy), but these patterns do not capture functional aspects. VERISOL [16] and VERX [13] target functional verification with invariants, pre- and postconditions, but do not mention events. Such invariants were also studied in the context of instrumentation and runtime validation [10], but not for events; our approach focuses on compile-time verification instead of runtime. In a broader setting, the event mechanism of Solidity is a special case of monitoring used for runtime verification of reactive, event-based systems [6]. In this context, events can be considered as manually written monitors, for which we aim to prove correctness.

2 Background

Solidity. Solidity [14] is a high-level, contract-oriented programming language supporting the rapid development of smart contracts for the Ethereum platform. We briefly introduce Solidity by restricting our presentation to the aspects relevant for events. An example contract (`Registry`) is shown in Figure 1. Contracts are similar to classes in object-oriented programming. A contract can define additional types, such as the `Entry` struct in the example, consisting of a Boolean flag and an integer data. The persistent data stored on

[1] `https://github.com/SRI-CSL/solidity`
[2] `https://swcregistry.io/`

the blockchain can be defined with *state variables*. The example contract declares a single variable `entries`, which is a mapping from addresses to `Entry` structs. Contracts can also define *events*, including possible arguments. The example declares two events, `new_entry` and `updated_entry`, to signal a new or an updated entry, respectively. Both events take the address and the new value for the data as their arguments. Finally, functions are defined that can be called as transactions to act on the contract state. The example defines two functions: `add` and `update`. The `add` function first checks with a `require` that the data corresponding to the caller address (`msg.sender`) is not yet set. If the condition of `require` does not hold, the transaction is reverted. Otherwise, the function sets the data and the flag, and emits the `new_entry` event. The `update` function is similar to `add`, with the exception that the data must already be set, and the new value should be larger than the old one (for illustrative purposes).

Note that Solidity puts no restrictions on the emitted events, and a faulty (or malicious) contract could both emit events that do not correspond to state changes or miss triggering an event on some change [5], potentially misleading users. In the case of the `Registry` contract, the events are emitted correctly, and the user can reproduce the changes in `entries` by relying solely on the emitted events and their arguments.

```
contract Registry {
  struct Entry { bool set; int data; } // User-defined type

  mapping(address=>Entry) entries; // State variable

  /// @notice tracks-changes-in entries
  /// @notice precondition !entries[at].set
  /// @notice postcondition entries[at].set && entries[at].data == value
  event new_entry(address at, int value);

  /// @notice tracks-changes-in entries
  /// @notice precondition entries[at].set && entries[at].data < value
  /// @notice postcondition entries[at].set && entries[at].data == value
  event updated_entry(address at, int value);

  /// @notice emits new_entry
  function add(int value) public {
    require(!entries[msg.sender].set);
    entries[msg.sender].set = true;
    entries[msg.sender].data = value;
    emit new_entry(msg.sender, value);
  }

  /// @notice emits updated_entry
  function update(int value) public {
    require(entries[msg.sender].set && entries[msg.sender].data < value);
    entries[msg.sender].data = value;
    emit updated_entry(msg.sender, value);
  }
}
```

Figure 1 An example contract illustrating Solidity events. Users of the contract can associate an integer value to their address and can later update it with a larger value.

solc-verify. SOLC-VERIFY [9] is a source-level verification tool for checking functional correctness of smart contracts. SOLC-VERIFY takes contracts written in Solidity and provides various in-code annotations to specify functional behavior (e.g., pre- and postconditions, invariants). As an example, consider a typical token contract (illustrated by Figure 2), which gives its creator all the tokens in the constructor and then provides a function to transfer them between users. The functional correctness of the contract logic can be specified by the existing annotation capabilities of SOLC-VERIFY (denoted by `*`). The top-level contract invariant ensures the inductive property that the sum of balances always equals to the total supply. Invariants become postconditions to the constructor and both pre- and postconditions to all public functions. Furthermore, the correctness of the constructor and the transfer function is established with additional postconditions. Besides the illustrated properties, assertions, overflows, preconditions, loop invariants, and modification specifiers are also supported.

SOLC-VERIFY translates the annotated contracts to the Boogie Intermediate Verification Language (IVL). The key idea of the translation is to encode state variables as global heaps and functions as procedures. SOLC-VERIFY relies on the Boogie verifier [3] to perform modular verification by discharging verification conditions to SMT solvers. The verification conditions encode the function body while assuming the preconditions, and then check if postconditions hold. In this process, function calls are replaced by their specification and loops by their invariants (modularity). Finally, the results are back-annotated to the Solidity source.

Goal. Previous versions of SOLC-VERIFY ignored events as they were considered merely a logging mechanism, not directly relevant for functional correctness. However, as argued before, formal specification and verification of events can be relevant. Therefore, this paper presents extensions to the specification and translation capabilities of SOLC-VERIFY that enable reasoning about Solidity events. We propose event-specific annotations (Section 3) and use them to instrument the code during translation with additional conditions to be verified (Section 4).

```
* /// @notice invariant sum(balances) == total
contract Token {
  mapping(address=>uint) balances;
  uint total;

  /// @notice tracks-changes-in balances
  /// @notice tracks-changes-in total
  /// @notice precondition balances[from] == 0
  /// @notice postcondition balances[from] == amount
  /// @notice postcondition total == amount
  event initialized(address from, uint amount);

  /// @notice tracks-changes-in balances
  /// @notice precondition balances[from] >= amount
  /// @notice postcondition balances[from] == before(balances[from]) - amount
  /// @notice postcondition balances[to] == before(balances[to]) + amount
  event transferred(address from, address to, uint amount);

* /// @notice postcondition balances[msg.sender] == _total
  /// @notice emits initialized
  constructor(uint _total) public {
    balances[msg.sender] = total = _total;
    emit initialized(msg.sender, total);
  }

* /// @notice postcondition balances[msg.sender] == old(balances[msg.sender]) - amount
* /// @notice postcondition balances[to] == old(balances[to]) + amount
  /// @notice emits transferred
  function transfer(address to, uint amount) public {
    require(balances[msg.sender] >= amount && msg.sender != to);
    balances[msg.sender] -= amount;
    balances[to] += amount;
    emit transferred(msg.sender, to, amount);
  }
}
```

Figure 2 A token contract illustrating existing specification capabilities of SOLC-VERIFY (marked with *) and the new annotations for events, including postconditions that refer to previous state.

3 Specification of Events

Our approach provides in-code annotations to specify events in terms of the on-chain data that they track for changes. Furthermore, additional predicates can specify the correspondence between the abstract view provided by events and the actual data, before and after the change. With a few exceptions (see later), annotations are expected to be inserted by the developer.

Data changes and checkpoints. Each event can declare a set of contract state variables that it *tracks* for changes. In the `Registry` example (Figure 1), both events track the single state variable `entries`, as specified by the `tracks-changes-in` annotations. In the `Token` example (Figure 2), `transferred` only tracks `balances`, whereas `initialized` tracks `total` as well. Intuitively, we use the tracking of changes to make sure that (1) if a tracked variable changes, a corresponding event must be emitted after; and (2) an event should be emitted only if some of its tracked variables have changed before. As data changes often occur in multiple steps, or conditionally (e.g., updating both members of a struct in the function `add` of Figure 1 or adding and subtracting in `transfer` of Figure 2), events cannot always be emitted directly after a single modifying statement. Therefore, we define the precise semantics of "before" and "after" by introducing *before-* and *after-checkpoints*. Before-checkpoints of an event are determined dynamically by the first change in a variable they track. In contrast, after-checkpoints are defined by static barriers, marking the latest point in code where the emitting should be fulfilled. Currently, we define loop and transaction boundaries (external calls to public functions and function return) as after-checkpoints. The semantics of checkpoints is that an event corresponding to a state variable change must be emitted at some point between before- and after-checkpoints, which also clears the before-checkpoint. Conversely, an event can only be emitted if a tracked variable indeed changed (there was a before-checkpoint).

Event pre- and postconditions. In addition to the set of tracked variables, events can also be annotated with *predicates* that define conditions over the state variables and the arguments of the event. There are two kinds of predicates: *pre-* and *postconditions*. Preconditions capture the values of state variables at the before-checkpoint, while postconditions correspond to the point when the event is emitted. In the `Registry` contract (Figure 1), both events (`new_entry` and `updated_entry`) have the same postcondition, namely that the data at the given address must be set and its value must match the value in the argument. The precondition of `new_entry` is that the data must not yet be set, while for `updated_entry`, it must be set and its value should be smaller than the event argument. Postcondition expressions often need to connect the state at the point of emit and before the change. As an example, consider the `transfer` function of the token contract in Figure 2 that deducts the sender's balance and increases the receiver's. To specify the postcondition of the `Transfer` event, we need to relate the new balances to the previous balances. We provide a special `before` function – to be used in postconditions – that refers to previous values of state variables. Note that each variable appearing in a predicate is implicitly tracked, i.e., no explicit `tracks-changes-in` annotation would be required.

Functions. We require contract functions to be annotated with the events that they possibly emit using the `emits` keyword. For example, the `add` and `update` functions in Figure 1 can emit `new_entry` and `updated_entry`, respectively. If a function calls other functions (including base constructors), the callee's emitted events must also be included in the caller's specifications.

4 Verification

A contract with events and specifications is checked in two steps. First, a syntactical check is performed to ensure that functions only emit events that they specified (via `emits` annotations). Then, we check the data tracking specifications and predicates by translating the contract to the input language of a verifier, and instrumenting the code with the checks and the required bookkeeping. In our implementation, we use the Boogie IVL and verifier [3], but we present our solution in a general way that can be reused by other Solidity verifiers.

Function emits. We first check whether functions only emit those events that are specified via `emits` annotations. This is a syntactic check on the Solidity AST: we find all emit statements in the function and check whether the corresponding events are specified to be emitted. When a function calls other functions *internally* (i.e., from the same contract), we apply a modular check based on the call graph: all events specified to be emitted by the callee must also be specified by the caller. On the other hand, we currently ignore *external* calls (such as `.call()` or `.transfer()`). Such external calls cannot modify state variables or trigger events from the current contract directly (as they are non-public). Indirect modifications and emits are possible by calling back public functions, but those are specified and checked independently (modularity of reasoning [9]). Furthermore, we also treat calls to other contracts' functions as external because addresses are not type checked runtime (only the function signature is checked) [9]. Finally, we also verify at each assignment (to a tracked variable), whether the function specifies a corresponding event to be emitted.

Data tracking and predicates. Verification of data tracking and predicates is performed by instrumenting the contract code with additional variables and statements to save state and to make extra checks at checkpoints. For clarity, we describe the instrumentation on the Solidity level. We illustrate the approach through the example contract in Figure 3, which has two state variables `x` and `y`, and whenever one of them changes, an event is emitted with their current difference. Furthermore, `x <= y` should hold both at the before- and the after-checkpoint. The extra instructions are displayed as labels where they are injected, while the corresponding code can be found in the snippets to the right.

For each state variable that is tracked by any event, we introduce two additional variables in the contract: one with the same type to save the before-state, and a Boolean flag to indicate whether the data has been modified (snippet new-vars in Figure 3). Functions are then instrumented with extra statements to save state, enforce after-checkpoints (barriers) and to perform specification checks when events are emitted. Functions ensure on entry that none of the variables tracked by their specified events have been modified since the checkpoint before the call (snippet assume-clear). In other words, all relevant events must have been emitted before making the call. In modular verification, this assumption becomes a precondition to the function. Before each modification (assignment statement), if the state variable is not modified yet, the current value is stored[3] in the helper variable and the flag for modification is set, introducing a before-checkpoint (snippets y-before and x-before).

At each `emit` statement, several checks are added (snippet emit-spec). First, we check that the data has indeed been modified, otherwise the event should not be emitted. Then we check each pre- and postcondition. By default, preconditions refer to the before-state

---

[3] Saving data (e.g., mappings) with assignments might not yield valid Solidity code. This code is for clarity of presentation and is handled by SOLC-VERIFY internally.

```
contract C {
  uint x;
  uint y;
  [new-vars]
  /// @notice tracks-changes in x
  /// @notice tracks-changes in y
  /// @notice precondition x <= y
  /// @notice postcondition x <= y
  /// @notice postcondition x + diff == y
  event xy_changed(uint diff);

  /// @notice emits xy_changed
  function f(uint n) public {
    [assume-clear]
    require(x <= y);
    [y-before]
    y += n;
    [emit-spec]
    emit xy_changed(y - x);
    [after-chpt]
    for (uint i = 0; i < n; ++i) {
      [x-before]
      x++;
      [emit-spec]
      emit xy_changed(y - x);
      [after-chpt]
    }
    [after-chpt]
  }
}
```

new-vars
```
uint x_old; // Previous state of x
uint y_old; // Previous state of y
bool x_mod; // Modified since last checkpoint
bool y_mod; // Modified since last checkpoint
```

assume-clear
```
require(!x_mod && !y_mod); // Modif. clear
```

after-chpt
```
assert(!x_mod && !y_mod); // Modif. clear
```

y-before
```
// Save y if not saved yet: before-checkpt
if (!y_mod) { y_old = y; y_mod = true; }
```

emit-spec
```
assert(x_mod || y_mod); // Emit without change
assert((x_mod?x_old:x) <= (y_mod?y_old:y)); // Pre
assert(x <= y); // Post
assert(x + (y - x) == y); // Post
x_mod = y_mod = false; // Emitted
```

x-before
```
// Save x if not saved yet: before-checkpt
if (!x_mod) { x_old = x; x_mod = true; }
```

Figure 3 Example contract with instrumentation snippets for checking event specifications.

and postconditions to the current values, except if the variable is explicitly wrapped with `before()`. Note that we refer to the previous value of a variable `v` with `v_mod ? v_old : v` because, in general, there might be variables that were not modified (e.g., `x` at the first emit in Figure 3). After performing the checks, emitting the event clears the flags (before-checkpoints). Finally, before returning, functions enforce after-checkpoints by asserting that no state variable is in a modified state, i.e., the function cannot end in debt with events (snippet after-chpt). In modular verification, this check becomes a postcondition to the function. We also insert an after-checkpoint before the loop and at the end of every iteration (serving as loop invariant).

Discussion. One potential limitation of our approach is that we consider loop boundaries after-checkpoints: some contracts change the data many times in the loop, but only emit a single summarizing event after the loop. This limitation could be alleviated with annotations to "allow delaying" the emit after the loop, but we do not support this as it leads to more complex specification and verification. Note that this limitation comes from modular verification as loops need an invariant. However, if we were to perform bounded model checking or symbolic execution, this might not be a limitation.

Our approach is not tied to Boogie or modular verification. The instrumentation can be performed on the Solidity level, and the correctness of the specification is reduced to checking assertions at particular points in the code. This means that the instrumented code can be fed into any Solidity verifier that can check for assertion failures. The event specifications are deemed correct if and only if there are no related assertion failures.

A possible future use-case of our approach lies in the behavioral analysis of contracts based on logs. Such analyses could reveal relationships individually and across contracts that are not otherwise apparent (e.g., exposing entities that control the blockchain interactions) or attack evidence. Application-level log analysis has been used for a long time for monitoring and security purposes, and most existing techniques assume that application logs can be trusted or, if applications are subverted by attackers, the subversion can be captured [11]. Our approach guarantees the validity of the emitted events, making them even more suitable for such analysis.

5 Conclusion

We presented an approach for the formal specification and verification of Solidity smart contracts that rely on events to communicate with their users, providing an abstract view of their state. We proposed in-code annotations to specify events in terms of the state variables they track for changes. Furthermore, we introduced additional predicates (pre- and postconditions) for specifying conditions on the state before and after the change, establishing the correspondence between the blockchain state and the emitted events. The approach is implemented in SOLC-VERIFY and we demonstrated its applicability with various examples.

References

1. A guide to events and logs in Ethereum smart contracts. `https://consensys.net/blog/blockchain-development/guide-to-events-and-logs-in-ethereum-smart-contracts`, 2016.
2. Nicola Atzei, Massimo Bartoletti, and Tiziana Cimoli. A survey of attacks on Ethereum smart contracts. In *Principles of Security and Trust*, volume 10204 of *Lecture Notes in Computer Science*, pages 164–186. Springer, 2017. `doi:10.1007/978-3-662-54455-6_8`.
3. Mike Barnett, Bor-Yuh Evan Chang, Robert DeLine, Bart Jacobs, and K. Rustan M. Leino. Boogie: A modular reusable verifier for object-oriented programs. In *Formal Methods for Components and Objects*, volume 4111 of *LNCS*, pages 364–387. Springer, 2006.
4. Huashan Chen, Marcus Pendleton, Laurent Njilla, and Shouhuai Xu. A survey on Ethereum systems security: Vulnerabilities, attacks, and defenses. *ACM Comput. Surv.*, 53(3), 2020.
5. J. Chen, X. Xia, D. Lo, J. Grundy, X. Luo, and T. Chen. Defining smart contract defects on Ethereum. *IEEE Transactions on Software Engineering*, 2020. Early access.
6. Séverine Colin and Leonardo Mariani. Run-time verification. In *Model-Based Testing of Reactive Systems: Advanced Lectures*, volume 3472 of *Lecture Notes in Computer Science*, chapter 18, pages 525–555. Springer, 2005.
7. Josselin Feist, Gustavo Greico, and Alex Groce. Slither: A static analysis framework for smart contracts. In *Proc. of the 2nd International Workshop on Emerging Trends in Software Engineering for Blockchain*, pages 8–15. IEEE, 2019.
8. Ákos Hajdu and Dejan Jovanović. SMT-friendly formalization of the solidity memory model. In *Programming Languages and System*, volume 12075 of *Lecture Notes in Computer Science*, pages 224–250. Springer, 2020. `doi:10.1007/978-3-030-44914-8_9`.
9. Ákos Hajdu and Dejan Jovanović. solc-verify: A modular verifier for Solidity smart contracts. In *Verified Software. Theories, Tools, and Experiments*, volume 12301 of *Lecture Notes in Computer Science*, pages 161–179. Springer, 2020. `doi:10.1007/978-3-030-41600-3_11`.
10. Ao Li, Jemin Andrew Choi, and Fan Long. Securing smart contract with runtime validation. In *Proc. of the 41st ACM SIGPLAN Conference on Programming Language Design and Implementation*, page 438–453. ACM, 2020.
11. Shiqing Ma, Juan Zhai, Fei Wang, Kyu Hyung Lee, Xiangyu Zhang, and Dongyan Xu. MPI: Multiple perspective attack investigation with semantic aware execution partitioning. In *Proc. of the 26th USENIX Security Symposium*, pages 1111–1128. USENIX Association, 2017.
12. Bernhard Mueller. Smashing Ethereum smart contracts for fun and real profit. In *Proc. of the 9th Annual HITB Security Conference (HITBSecConf)*, 2018.

13 A. Permenev, D. Dimitrov, P. Tsankov, D. Drachsler-Cohen, and M. Vechev. VerX: Safety verification of smart contracts. In *Proc. of the 2020 IEEE Symposium on Security and Privacy*, pages 414–430. IEEE, 2020.

14 Solidity documentation. `https://solidity.readthedocs.io`, 2020.

15 Nick Szabo. Smart contracts. `https://web.archive.org/web/20011102030833/http://szabo.best.vwh.net:80/smart.contracts.html`, 1994.

16 Yuepeng Wang, Shuvendu K Lahiri, Shuo Chen, Rong Pan, Isil Dillig, Cody Born, Immad Naseer, and Kostas Ferles. Formal verification of workflow policies for smart contracts in Azure blockchain. In *Verified Software. Theories, Tools, and Experiments*, volume 12031 of *Lecture Notes in Computer Science*, pages 87–106. Springer, 2020.

17 Gavin Wood. Ethereum: A secure decentralised generalised transaction ledger. `https://ethereum.github.io/yellowpaper/paper.pdf`, 2019.

Populating the Peephole Optimizer of a Smart Contract Compiler

Maria A. Schett
University College London, UK
http://maria-a-schett.net
mail@maria-a-schett.net

Julian Nagele
London, UK
http://jnagele.net
mail@jnagele.net

Abstract

Developing compiler optimizations, especially for new, rapidly evolving smart contract languages, can be onerous and error-prone, but is especially important for smart contracts, where deployment and execution directly translate to monetary cost and which cannot change once deployed. One common optimization technique is the use of *peephole optimizations*, replacement rules that are applied using pattern-matching. These rules are normally constructed using human expertise, which is both time-consuming and far from systematic in exploring opportunities for optimization. In this work we propose a pipeline to automatically populate the peephole optimizer of a smart contract compiler. We apply superoptimization to an existing code base to obtain sequences of instructions, which can be replaced by cheaper, observationally equivalent instructions. We then generate peephole optimization rules by extracting the underlying patterns of these optimizations. We provide a case study of our approach and a prototype implementation for bytecode of the Ethereum Virtual Machine, the tool ppltr, which combines the superoptimizer ebso and the rule generator sorg. Then we evaluate our approach by generating and applying nearly $1k$ peephole optimization rules extracted from $2k$ optimizations obtained from deployed bytecode.

2012 ACM Subject Classification Software and its engineering → Formal methods; Software and its engineering → Compilers

Keywords and phrases Compiler Optimizations, Constraint Solving, Ethereum Bytecode

Digital Object Identifier 10.4230/OASIcs.FMBC.2020.3

Supplementary Material https://github.com/mariaschett/ppltr

1 Introduction

In this work we leverage formal methods to automatically populate the peephole optimizer of a smart contract compiler. A *peephole optimizer* uses pattern matching to optimize a small fragment of code, i.e., a peephole, by applying *optimization rules*. But finding sound optimization rules is a bottleneck as witnessed by the peephole optimizer of the Solidity compiler solc.[1] Currently, solc features fewer than 20 rules compared to LLVM's 1000+ rules. Thus we propose a pipeline to automatically populate the peephole optimizer of a smart contract compiler by combining techniques from constraint solving and rewriting as illustrated in Figure 1.

Smart contract languages typically have a large and accessible code base to use as a basis for finding optimizations, e.g., code deployed to public blockchains or test cases.

[1] github.com/ethereum/solidity/blob/ 019ec63f63bae7bbe89f5b62bb7b202ef5dadce6/libevmasm/PeepholeOptimiser.cpp

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler; Article No. 3; pp. 3:1–3:15
OpenAccess Series in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

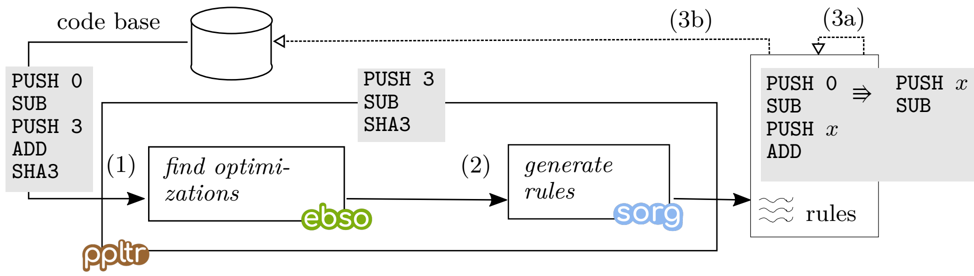

Figure 1 Pipeline to automatically generate peephole optimization rules from a code base.

(1) This allows us to start from an existing code base, to *find optimizations* by using automated tools to synthesize observationally equivalent but cheaper instruction sequences.

This automatic synthesis is possible, because many smart contract languages come with formally defined operational semantics, e.g., the Ethereum yellow paper [22]. Moreover, execution of a smart contract comes with a clear cost model – gas – giving rise to a precise notion of optimality. To give an example, the bytecode for the Ethereum virtual machine `PUSH 0 SUB PUSH 3 ADD SHA3` computes a hash of $3 + (0 - w)$ for some word w already on the stack. As $3 + (0 - w) = 3 - w$ the bytecode corresponding to `PUSH 3 SUB SHA3`, computes the same result *and* cheaper.

(2) From such optimizations, we can *generate rules*. Using concepts from rewriting we generalize "unnecessarily specific" arguments and strip away "unnecessary" context to obtain optimization rules.

For the above example, we generate the rule `PUSH 0 SUB PUSH` x `ADD` $\Rrightarrow$ `PUSH` x `SUB` by generalizing 3 to x.

(3) Finally we can feed back and apply the generated rules to
  (a) the rules themselves, and
  (b) the code base and again start the cycle to find new optimizations.

We demonstrate the applicability of our pipeline in a case study for bytecode of the Ethereum virtual machine (EVM). We implemented a prototype: ppltr, a peephole otimization rule generator. For phase (1), we use the tool ebso [19], a superoptimizer for EVM bytecode. For phase (2), we use sorg, a superoptimization based rule generator. All tools are available open-source under the Apache-2.0 license.[2] We evaluated our approach on bytecode of the 250 most called contracts of the Ethereum blockchain, where we found 2032 distinct optimizations from which we automatically generated 993 optimization rules.

Contributions

1. We propose a pipeline for automatically populating a peephole optimizer, and
2. a sound and complete procedure to generate optimization rules from optimizations.
3. We perform a case study for EVM bytecode with
4. a prototype implementation, together with
5. an evaluation.

[2] Available at `github.com/juliannagele/ebso/tree/v2.1`, `github.com/mariaschett/sorg/tree/v1.1`, and `github.com/mariaschett/ppltr/tree/v1.0`.

2 Approach

We assume a machine model with a *state* over a set of *words* $\mathbb{W}$ with an observational equivalence relation $\equiv$ on states, which may take only parts of the state into account. States are modified based on *instructions* from a set $\mathcal{I}$, where an instruction $\iota \in \mathcal{I}$ deterministically transforms a state σ into some state σ' denoted by $\sigma \xrightarrow{\iota} \sigma'$. Some instructions act only on parts of the state, while others take immediate arguments from $\mathbb{W}$. We write $\iota(w_1, \ldots, w_k)$ for an instruction $\iota \in \mathcal{I}$ which takes k immediate arguments $w_1, \ldots, w_k \in \mathbb{W}$ and say that ι has arity k. For example, in a stack-based machine the instruction `PUSH 3` takes the immediate argument 3, while `SUB` has arity 0, but consumes two arguments from the stack.

A *program* ρ is a sequence of instructions $\iota_0 \cdots \iota_n$. The length of ρ is its number of instructions, denoted by $|\rho|$. We write ε for the *empty* program and $\rho \cdot \tau$ for the concatenation of programs ρ and τ. A program $\rho = \iota_0 \cdots \iota_n$ transforms a family of states $\boldsymbol{\sigma} = (\sigma_j)_{j \leq n+1}$ by stepwise transformation, i.e., $\sigma_0 \xrightarrow{\iota_0} \sigma_1 \xrightarrow{\iota_1} \ldots \xrightarrow{\iota_n} \sigma_{n+1}$, and we write $\sigma_0 \xrightarrow{\rho}\!\!\!\rightarrow \sigma_{n+1}$. Here σ_j is the state after executing j instructions, and σ_0 is the designated start state. We often write states instead of families of states, when the distinction is clear from the context.

We write $\mathsf{cost}(\iota, \sigma)$ for the cost incurred by executing instruction ι on state σ. The cost of executing a program is simply the sum of the cost of its instructions: $\mathsf{cost}(\iota_0 \cdots \iota_n, \boldsymbol{\sigma}) = \sum_{j=0}^{n} \mathsf{cost}(\iota_j, \sigma_j)$. Two programs ρ and τ are *equal*, denoted by $\rho = \tau$, if they are syntactically equal, and *equivalent*, $\rho \equiv \tau$, if they are observationally equivalent, i.e., for states $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma'}$ with $\sigma_0 \equiv \sigma'_0$, $\sigma_0 \xrightarrow{\rho}\!\!\!\rightarrow \sigma_{|\rho|+1}$, and $\sigma'_0 \xrightarrow{\tau}\!\!\!\rightarrow \sigma'_{|\tau|+1}$ we have $\sigma_{|\rho|+1} \equiv \sigma'_{|\tau|+1}$.

▶ **Definition 1.** *Let ρ and τ be programs with $\rho \equiv \tau$ and $\mathsf{cost}(\rho, \boldsymbol{\sigma}) > \mathsf{cost}(\tau, \boldsymbol{\sigma})$ for all states $\boldsymbol{\sigma}$. Then τ is an* optimization *of ρ, and we write $\rho \gtreqqless \tau$.*

In Section 2.1, we will show how we can obtain such optimizations – and in Section 2.2 we will use them to generate optimization rules. To do so, we need to define what constitutes a rule. Therefore we abstract over the immediate arguments of instructions by using a countably infinite set of variables $\mathcal{V}$. We extend $\mathcal{I}$ to $\mathcal{I}^{\mathcal{V}}$ by adding instructions $\iota(x_1, \ldots, x_k)$ for all $x_1, \ldots, x_k \in \mathcal{V}$ and all $\iota \in \mathcal{I}$ of arity $k > 0$.

A program over $\mathcal{I}^{\mathcal{V}}$ is called a *program schema*. To obtain a *maximal schema* of a program schema s every $\iota(w_1, \ldots, w_k)$ in s is replaced by $\iota(x_1, \ldots, x_k)$, where $x_1, \ldots, x_k$ are fresh variables from $\mathcal{V}$. All variables in a program schema s are collected in $\mathcal{V}\mathsf{ar}(s)$.

A *substitution* $\gamma : \mathcal{V} \to \mathbb{W} \cup \mathcal{V}$ maps variables to variables and words. In a *ground* substitution $\overline{\gamma}$ the range is restricted to $\mathbb{W}$, i.e., $\overline{\gamma} : \mathcal{V} \to \mathbb{W}$. We *apply* γ to a schema s by replacing all variables x in s by $\gamma(x)$ and write $s\gamma$ for the result. Note that $s\overline{\gamma}$ is a program. A substitution γ is *at least as general* as a substitution γ', denoted $\gamma \leq \gamma'$, if there is a substitution γ'' such that $\gamma\gamma'' = \gamma'$. If $\gamma \leq \gamma'$ and $\gamma' \nleq \gamma$ then we say γ is *more general* than γ' and write $\gamma < \gamma'$.

We call program schemas s and t *observationally equivalent*, and write $s \equiv t$, if $s\overline{\gamma} \equiv t\overline{\gamma}$ holds for all $\overline{\gamma}$ and write $\mathsf{cost}(s, \boldsymbol{\sigma}) > \mathsf{cost}(t, \boldsymbol{\sigma'})$ if $\mathsf{cost}(s\overline{\gamma}, \boldsymbol{\sigma}) > \mathsf{cost}(t\overline{\gamma}, \boldsymbol{\sigma'})$ for all $\overline{\gamma}$.

▶ **Definition 2.** *Let ℓ and r be program schemas with $\ell \equiv r$ and $\mathsf{cost}(\ell, \boldsymbol{\sigma}) > \mathsf{cost}(r, \boldsymbol{\sigma})$. Then $\ell \Rrightarrow r$ is an (optimization)* rule.

By definition, every optimization $\rho \gtreqqless \tau$ is an optimization rule $\rho \Rrightarrow \tau$. A *context* C is a pair of program schemas (s_1, s_2). We write $C[t]$ for the program schema $s_1 \cdot t \cdot s_2$ and call s_1 a *prefix* and s_2 a *postfix* of $C[t]$. A context (s_1, s_2) is *at least as general* as a context (t_1, t_2), denoted by $(s_1, s_2) \leq (t_1, t_2)$, if there is a context (r_1, r_2) such that $r_1 \cdot s_1 = t_1$ and $s_2 \cdot r_2 = t_2$. If $C \leq C'$ and $C' \nleq C$ then we say C is *more general* than C' and write $C < C'$.

The following definition captures all optimization rules that can produce a given optimization when instantiated.

▶ **Definition 3.** *The* optimization rules *for an optimization $\rho \gtreqless \tau$ are defined as $\mathcal{R}(\rho \gtreqless \tau) = \{\ell \Rrightarrow r \mid \rho = C[\ell\gamma]$ and $\tau = C[r\gamma]$ for some substitution γ and context $C\}$.*

We ensure that applying peephole optimizations is sound by the following lemma.

▶ **Lemma 4.** *If $\rho \equiv \tau$ then $C[\rho] \equiv C[\tau]$ for all contexts C.*

Proof. We show the statement by induction on C. By assumption, the statement holds for the base case $C = (\varepsilon, \varepsilon)$. For the step case $C = (\iota \cdot s_1, s_2)$ observe that every instruction ι is deterministic, i.e., executing ι starting from a state σ leads to a deterministic state σ'. By induction hypothesis, executing $s_1 \rho s_2$ and $s_1 \tau s_2$ from a state σ' leads to an observationally equivalent state σ'', and therefore $\iota \cdot s_1 \cdot \rho \cdot s_2 \equiv \iota \cdot s_1 \cdot \tau \cdot s_2$ holds. We can reason analogously for $C = (s_1, s_2 \cdot \iota)$. ◀

2.1 Find Optimizations

As Definition 1 suggests finding an optimization for a program ρ necessitates finding
1. an observationally equivalent program τ, where
2. the cost of τ is less than the cost of ρ.

We leverage a constraint solver, such as Z3 [8], to automatically find equivalent, but cheaper programs. To this end, we express the above as an SMT problem: given a source program ρ, is there a target program τ such that for all possible inputs, executing ρ and τ results in the same final state, but the cost of τ is less than the cost of ρ? Our encoding is based on the encoding from unbounded superoptimization [11].

Find an Observationally Equivalent Program

To encode observational equivalence we first need a constraint that expresses equality on states: Let $\mathsf{enc\_eq\_state}(\sigma, \sigma')$ be an SMT constraint that evaluates to true, whenever state σ and state σ' are observationally equivalent. The concrete instantiation of this constraint depends on the machine that is modeled. For instance, the state may be modeled as several uninterpreted functions. An encoding for the EVM, modeling the state with a stack, storage, and exceptional halting can be found in Example 16, with the corresponding encoding of enc_eq_state in Example 18.

Based on the operational semantics for every $\iota \in \mathcal{I}$, we need to encode the effect of ι on a state i.e., the relation $\xrightarrow{\iota}$.

▶ **Definition 5.** *Let $\mathsf{enc\_step}(\iota, \sigma, \sigma')$ be an SMT encoding of the effect of an instruction ι as constraints between state σ and state σ'. For a program $\rho = \iota_0 \cdots \iota_n$ and states $\boldsymbol{\sigma}$ we define $\mathsf{enc\_progr}(\rho, \boldsymbol{\sigma})$ as $\bigwedge_{0 \leqslant j \leqslant n} \mathsf{enc\_step}(\iota_j, \sigma_j, \sigma_{j+1})$.*

Again, the concrete encoding of enc_step depends on the machine that is modeled, see Example 17 for our instantiation for the EVM.

Most programs will consume some input words $\vec{x}$. To pass them to the program, we assume an encoding $\mathsf{enc\_init}(\vec{x}, \boldsymbol{\sigma})$ that sets constraints on the start state σ_0 appropriately, e.g., putting the words in $\vec{x}$ in registers or on the stack according to the machine model. Based on the constraint enc_step, we can encode the search space of all possible target programs. To this end we represent the target program as a pair $\tau = \langle \mathsf{instr}, n \rangle$ of an uninterpreted function $\mathsf{instr}(j) : \mathbb{N} \to \mathcal{I}$ and its length $n \in \mathbb{N}$. The function instr acts as a template to be filled by the SMT solver returning the instruction to be used at position j of the target program. After a model has been found, the concrete target program can be reconstructed as $\mathsf{instr}(0) \cdot \mathsf{instr}(1) \cdots \mathsf{instr}(n-1)$.

▶ **Definition 6.** *Given a set of instructions $\mathcal{I}$ we define the SMT encoding for the enumeration of every program of length n as* $\mathsf{enc\_search}(\tau, \boldsymbol{\sigma})$ *as*

$$\forall j.\, 0 \leqslant j < n \rightarrow \bigwedge_{\iota \in \mathcal{I}} \mathsf{instr}(j) = \iota \rightarrow \mathsf{enc\_step}(\iota, \sigma_j, \sigma_{j+1}) \wedge \bigvee_{\iota \in \mathcal{I}} \mathsf{instr}(j) = \iota \tag{1}$$

The first clause states that if we pick ι at position j, then the effect is determined by $\mathsf{enc\_step}(\iota, \sigma_j, \sigma_{j+1})$. The second clause, $\bigvee_{\iota \in \mathcal{I}} \mathsf{instr}(j) = \iota$, ascertains that for every position j some instruction is picked.

▶ **Definition 7.** *The encoding for finding an observationally equivalent program to a given program ρ is*

$$\begin{aligned} &\exists n, \forall \vec{x}.\, \mathsf{enc\_init}(\vec{x}, \boldsymbol{\sigma}) \wedge \mathsf{enc\_init}(\vec{x}, \boldsymbol{\sigma'}) \wedge \\ &\qquad \mathsf{enc\_progr}(\rho, \boldsymbol{\sigma}) \wedge \mathsf{enc\_search}(\tau, \boldsymbol{\sigma'}) \wedge \mathsf{enc\_eq\_state}(\sigma_{|\rho|+1}, \sigma'_n) \end{aligned} \tag{2}$$

The first two constraints initialize states $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma'}$ with the same inputs, the third and fourth constraint encode the effects of the existing program ρ and the sought after target program τ respectively, while the final constraint demands that they are observationally equivalent, i.e., that they result in equivalent states. With this constraint we will find observational equivalent programs. Now we will need to add constraints on the cost.

Find a Cheaper Program

To achieve this we extend Constraint (2) from Definition 7 by a constraint stating that the cost of executing the target program τ is less than the cost of executing the source program ρ: i.e., $\mathsf{cost}(\rho, \boldsymbol{\sigma}) > \mathsf{cost}(\tau, \boldsymbol{\sigma'})$. Here the cost of τ is again defined by summation, i.e., for $\tau = \langle \mathsf{instr}, n \rangle$ we have $\mathsf{cost}(\tau, \boldsymbol{\sigma'}) = \sum_{j=0}^{n-1} \mathsf{cost}(\mathsf{instr}(j), \sigma_j)$.

2.2 Generate Rules

As Definition 3 indicates generating optimization rules from optimizations requires us
1. to find a substitution γ, and
2. to find a context C.

Find a Substitution

In the first step we generalize the immediate arguments of instructions in an optimization $\rho \gtreqless \tau$ by finding a substitution. We capture all possible generalizations of a rule using the following definition.

▶ **Definition 8.** *The* generalized rules *of an optimization rule $\rho \Rrightarrow \tau$ are defined as $\mathcal{G}(\rho \Rrightarrow \tau) = \{\ell \Rrightarrow r \mid \ell\gamma = \rho$ and $r\gamma = \tau$ for some substitution $\gamma\}$.*

▶ **Example 9.** Let $\rho \equiv \tau$ be the optimization from the introduction, i.e., `PUSH 0 SUB PUSH 3 ADD SHA3` $\equiv$ `PUSH 3 SUB SHA3`. Then $\mathcal{G}(\rho \equiv \tau)$ consists of two rules: `PUSH 0 SUB PUSH` x `ADD SHA3` $\Rrightarrow$ `PUSH` x `SUB SHA3` and $\rho \Rrightarrow \tau$ itself. Note that the pair `PUSH` y `SUB PUSH` x `ADD SHA3` and `PUSH` x `SUB SHA3` is not in $\mathcal{G}(\rho \equiv \tau)$. Applying the substitution $\gamma = \{x \mapsto 3, y \mapsto 0\}$ would yield the original optimization, but since `PUSH` y `SUB PUSH` x `ADD SHA3` $\not\equiv$ `PUSH` x `SUB SHA3` they do not constitute an optimization rule.

Example 19 shows further generalized rules for EVM bytecode.

To implement $\mathcal{G}$ we can do an exhaustive search as follows: start from a maximal schema for the given optimization and try all possibilities of mapping the variables back to the original values, checking whether the result yields a rule. The following procedure implements this approach, additionally using an order on the candidate substitutions to prune the search space.

▶ **Definition 10.** *We define the function* generalize *as follows:*

1: ***function*** generalize$(\rho \Rrightarrow \tau)$
2: $\quad\mathcal{R} \leftarrow \varnothing$
3: $\quad\ell_0, r_0 \leftarrow$ *maximal program schemas* ℓ_o *and* r_0 *for* ρ *and* τ *with* $\mathcal{V}\mathsf{ar}(\ell_0) \cap \mathcal{V}\mathsf{ar}(r_0) = \varnothing$
4: $\quad\gamma_0 \leftarrow$ *the substitution* γ_0 *with* $\rho = \ell_0\gamma_0$ *and* $\tau = r_0\gamma_0$
5: $\quad\Gamma \leftarrow \{\gamma \mid \gamma(x) = \gamma_0(x)$ *or* $\gamma(x) = y$ *for* $\gamma_0(x) = \gamma_0(y)$ *and* $x, y \in \mathcal{V}\mathsf{ar}(\ell_0) \cup \mathcal{V}\mathsf{ar}(r_0)\}$
6: $\quad$***for all*** $\gamma \in \Gamma$ ***do***
7: $\quad\quad$***if*** $\ell_0\gamma \equiv r_0\gamma$ ***then***
8: $\quad\quad\quad\mathcal{R} \leftarrow \mathcal{R} \cup \{\ell_0\gamma \Rrightarrow r_0\gamma\}$
9: $\quad\quad\quad\Gamma \leftarrow \Gamma \setminus \{\gamma' \mid \gamma \lessdot \gamma'\}$
10: $\quad\quad$***else***
11: $\quad\quad\quad\Gamma \leftarrow \Gamma \setminus \{\gamma' \mid \gamma' \lessdot \gamma\}$
12: $\quad$***return*** $\mathcal{R}$

Using the order $\lessdot$ on substitutions to prune the search space is key for implementation. Pruning only removes rules covered by others as the following lemma shows.

▶ **Lemma 11.** *For every* $\ell \Rrightarrow r \in \mathcal{G}(\alpha)$ *of a rule* α *there is a* $\ell' \Rrightarrow r' \in$ generalize(α) *and a substitution* γ *such that* $\ell'\gamma = \ell$ *and* $r'\gamma = r$.

Proof. We fix $\ell \Rrightarrow r \in \mathcal{G}(\alpha)$. Let ℓ_0 and r_0 be the maximal schemas of α. By definition of maximal schema there is a γ' such that $\ell_0\gamma' = \ell$ and $r_0\gamma' = r$. A renaming of γ' is in Γ and thus either generalize(α) will consider it at some point, or it will be removed by either line 9 or line 11.

If it is considered then a renaming of $\ell \Rrightarrow r$ is in generalize(α). If it is removed by line 9, then a substitution γ with $\gamma \lessdot \gamma'$ and and $\ell_0\gamma \equiv r_0\gamma$ was considered. Thus $\ell_0\gamma \Rrightarrow r_0\gamma$ is in generalize(α) and we have $\ell_0\gamma\gamma'' = \ell$ and $r_0\gamma\gamma'' = r$ for some γ'' by $\gamma \lessdot \gamma'$. If γ' was removed by line 11 then a substitution γ with $\gamma' \lessdot \gamma$ and and $\ell_0\gamma \not\equiv r_0\gamma$ was considered. But this contradicts the assumption $\ell \Rrightarrow r \in \mathcal{G}(\alpha)$, because observational equivalence is closed under substitution. ◀

Find a Context

As a second step We strip the generalized rules of any unnecessary pre- and postfix. Again we first capture all possible stripped rules and then give an implementation.

▶ **Definition 12.** *The* stripped rules *of a rule* $\rho \Rrightarrow \tau$ *are defined as* $\mathcal{C}(\rho \Rrightarrow \tau) = \{\ell \Rrightarrow r \mid \rho = C[\ell]$ *and* $\tau = C[r]\}$.

▶ **Example 13.** Continuing Example 9, for the rule `PUSH 0 SUB PUSH` x `ADD SHA3` $\Rrightarrow$ `PUSH` x `SUB SHA3` the stripped rules $\mathcal{C}$ contain the rule `PUSH 0 SUB PUSH` x `ADD` $\Rrightarrow$ `PUSH` x `SUB`, obtained by stripping away the context (ϵ, `SHA3`), and the original rule itself, since applying the empty context (ϵ, ϵ) to a program yields the program itself.

Example 20 shows further rules stripped of their context in EVM bytecode.

To implement $\mathcal{C}$ we follow the same strategy as for $\mathcal{G}$: try all possible contexts in an exhaustive search, checking whether they yield a rule and use an order contexts to prune the search space.

▶ **Definition 14.** *We define the function* strip *as*

1: ***function*** strip$(\rho \Rrightarrow \tau)$
2: $\quad \mathcal{R} \leftarrow \varnothing$
3: $\quad (s_0, t_0) \leftarrow$ *the longest common prefix* s_0 *and the longest common postfix* t_0 *of* ρ *and* τ
4: $\quad \ell_0, r_0 \leftarrow$ *the program schemas* ℓ_0 *and* r_0 *with* $s_0 \cdot \ell_0 \cdot t_0 = \rho$ *and* $s_0 \cdot r_0 \cdot t_0 = \tau$
5: $\quad \Gamma \leftarrow \{C \mid C = (s, t)$ *where* $s' \cdot s = s_0$ *and* $t \cdot t' = t_0$ *for some* $s', t'\}$
6: $\quad$ ***for all*** $C \in \Gamma$ ***do***
7: $\quad\quad$ ***if*** $C[\ell_0] \equiv C[r_0]$ ***then***
8: $\quad\quad\quad \mathcal{R} \leftarrow \mathcal{R} \cup \{C[\ell_0] \Rrightarrow C[r_0]\}$
9: $\quad\quad\quad \Gamma \leftarrow \Gamma \setminus \{C' \mid C < C'\}$
10: $\quad\quad$ ***else***
11: $\quad\quad\quad \Gamma \leftarrow \Gamma \setminus \{C' \mid C' < C\}$
12: $\quad$ ***return*** $\mathcal{R}$

Again, the order on contexts allows us to prune the search space without loss.

▶ **Lemma 15.** *For every* $\ell \Rrightarrow r \in \mathcal{C}(\alpha)$ *of a rule* α *there is a* $\ell' \Rrightarrow r' \in$ strip(α) *and a context* C *such that* $C[\ell'] = \ell$ *and* $C[r'] = r$*.*

Proof. We fix a rule $\ell \Rrightarrow r \in \mathcal{C}(\alpha)$. Let (s_0, t_0) be the longest common prefix and the longest common postfix of α and be ℓ_0, r_0 the program schemas with $s_0 \cdot \ell_0 \cdot t_0 \Rrightarrow s_0 \cdot r_0 \cdot t_0 = \alpha$. A context C' with $C'[\ell_0] = \ell$ and $C'[r_0] = r$ is in Γ and thus either strip(α) will consider it at some point, or it will be removed by either line 9 or line 11.

If it is considered then $\ell \Rrightarrow r$ is in strip(α). If it is removed by line 9, then a context C with $C < C'$ and and $C[\ell_0] \equiv C[r_0]$ was considered. Thus $C[\ell_0] \Rrightarrow C[r_0]$ is in strip(α) and we have $C''[C[\ell_0] = \ell$ and $C''[C[r_0]] = r$ for some C'' by $C < C'$. If C' was removed by line 11 then a context C with $C' < C$ and and $C[\ell_0] \not\equiv C[r_0]$ was considered. Again this contradicts the assumption $\ell \Rrightarrow r \in \mathcal{C}(\alpha)$, because observational equivalence is closed under context. ◀

Soundness and Completeness

Finally, we combine the two functions and for an optimization $\rho \gtreqless \tau$ define sorg$(\rho \gtreqless \tau) = \{$strip$(\ell \Rrightarrow r) \mid \ell \Rrightarrow r \in$ generalize$(\rho \Rrightarrow \tau)\}$.

The rules generated by sorg$(\rho \gtreqless \tau)$ are *sound*: for every $\ell \Rrightarrow r \in$ sorg$(\rho \Rrightarrow \tau)$ there is a substitution γ and a context C such that $C[\ell\gamma] = \rho$ and $C[r\gamma] = \tau$. This directly follows from generalize$(\rho \Rrightarrow \tau) \subseteq \mathcal{G}(\rho \Rrightarrow \tau)$ and strip$(\rho \Rrightarrow \tau) \subseteq \mathcal{C}(\rho \Rrightarrow \tau)$.

The rules generated by sorg$(\rho \gtreqless \tau)$ are also *complete*: for every $\ell \Rrightarrow r \in \mathcal{R}(\rho \gtreqless \tau)$ there is a $\ell' \Rrightarrow r' \in$ sorg$(\rho \gtreqless \tau)$, a substitution γ and a context C such that $C[\ell'\gamma] = \ell$ and $C[r'\gamma] = r$. This directly follows from Lemmas 11 and 15.

3 Case Study: EVM bytecode

To demonstrate the applicability of our pipeline from Figure 1 we implement it in the context of Ethereum for EVM bytecode. We sketch how one could apply the approach to other smart contract languages in Section 5.

The EVM is a virtual machine formally defined in the Ethereum yellow paper [22]. It is based on a stack which holds bit vectors of size 256. The stack may over- or underflow;

both lead the EVM to enter an exceptional halting state. The EVM also features a volatile memory, which is a word-addressed byte array, and a persistent key-value storage, which is a word-addressed word array stored on the Ethereum blockchain.

3.1 Find Optimizations with ebso

We find optimizations using our tool ebso [19], an EVM bytecode superoptimizer. As an input ebso takes an ebso block – a basic block that additionally does not contain instructions whose semantics are not encoded, such as instructions that have an outside effect like `LOG`. Then, encoding the EVM execution state and unbounded superoptimization following Section 2.1, in the best case ebso produces a cheaper, observationally equivalent ebso block.

▶ **Example 16.** We encode the EVM execution state $\boldsymbol{\sigma}$ using four uninterpreted functions $\langle \mathsf{sk}, \mathsf{c}, \mathsf{hlt}, \mathsf{str} \rangle$ to model the stack, stack pointer, exceptional halting and storage:

(i) $\mathsf{sk}(j, \vec{x}, n)$ returns the word from position n, starting from 0, in the stack after executing j instructions on $\vec{x}$,
(ii) $\mathsf{c}(j)$ returns the number of words on the stack after executing j instructions,
(iii) $\mathsf{hlt}(j)$ returns true ($\top$) if exceptional halting has occurred after executing j instructions, and false ($\bot$) otherwise, and
(iv) $\mathsf{str}(j, \vec{x}, k)$ returns the word at key k after executing j instructions on $\vec{x}$.

Note that these functions represent all states throughout an execution, i.e., $\boldsymbol{\sigma}$, while to obtain σ_j for some j, we simply apply them to j thus: $\sigma_j = \langle \mathsf{sk}(j), \mathsf{c}(j), \mathsf{hlt}(j), \mathsf{str}(j) \rangle$. To refer to individual components of states we use subscripts, for instance we write sk_σ to refer to the stack of state σ.

For a program ρ which takes d arguments on the stack we add d fresh variables to represent the input $\vec{x}$ and add the following constraint to $\mathsf{enc\_init}(\vec{x}, \boldsymbol{\sigma})$:

$$\bigwedge_{0 \leqslant i < d} \mathsf{sk}_\sigma(\vec{x}, 0, i) = x_i \wedge \mathsf{c}_\sigma(0) = d \wedge \mathsf{hlt}_\sigma(0) = \bot$$

The storage str is initialized similarly using an Ackermann encoding [1, 14].

To ease readability and save space we do not include the EVM's memory in this encoding of the execution state. It can be represented analogously to the storage.

▶ **Example 17.** Next we instantiate the operational semantics of the instructions. The constraint $\mathsf{enc\_stack}(\iota, \sigma_j, \sigma_{j+1})$ describes the effect that ι has on stack. Here we give as example the instruction `SUB` and refer to [22] or [19] for details. Let $-_{\mathsf{bv}}$ denote subtraction on bit-vectors. Then we have

$$\begin{aligned} \mathsf{enc\_stack}(\texttt{SUB}, \sigma_j, \sigma_{j+1}) := \mathsf{sk}_\sigma(j+1, \vec{x}, \mathsf{c}_\sigma(j+1) - 1) \\ = \mathsf{sk}_\sigma(j, \vec{x}, \mathsf{c}_\sigma(j) - 1) -_{\mathsf{bv}} \mathsf{sk}_\sigma(j, \vec{x}, \mathsf{c}_\sigma(j) - 2) \end{aligned}$$

Using enc_stack we can formulate the constraint enc_step. Here $\delta(\iota)$ and $\alpha(\iota)$ refer to the number of words which ι deletes from, and adds to the stack respectively. For all instructions except `SSTORE` we have:

$$\begin{aligned} \mathsf{enc\_step}(\iota, \sigma_j, \sigma_{j+1}) := \ & \mathsf{enc\_stack}(\iota, \sigma_j, \sigma_{j+1}) \wedge \\ & \mathsf{c}_\sigma(j+1) = \mathsf{c}_\sigma(j) + \alpha(\iota) - \delta(\iota) \wedge \\ & \forall n.\, n < \mathsf{c}_\sigma(j) - \delta(\iota) \rightarrow \mathsf{sk}_\sigma(j+1, \vec{x}, n) = \mathsf{sk}_\sigma(j, \vec{x}, n) \wedge \\ & \mathsf{hlt}_\sigma(j+1) = \mathsf{hlt}_\sigma(j) \vee \mathsf{c}_\sigma(j) - \delta(\iota) < 0 \vee \mathsf{c}_\sigma(j) - \delta(\iota) + \alpha(\iota) > 2^{10} \wedge \\ & \forall w.\, \mathsf{str}_\sigma(j+1, \vec{x}, w) = \mathsf{str}_\sigma(j, \vec{x}, w) \end{aligned}$$

Here the second line updates the counter for the number of words on the stack according to the number of words added and deleted. The third line expresses that all words on the stack below $\mathsf{c}_\sigma(j) - \delta(\iota)$ are preserved. The fourth line captures that exceptions relevant to the stack can occur through either an underflow or an overflow, and that once it has occurred, an exceptional halt state persists. Finally the last line states that all $\iota \neq$ `SSTORE` do not change the storage. The constraint for `SSTORE` is similar updating the storage using the Ackermann encoding.

▶ **Example 18.** The final ingredient we need to instantiate is the equivalence relation on states. For two states at steps j_1 and j_2 where $\sigma_{j_1} = \langle \mathsf{sk}(j_1), \mathsf{c}(j_1), \mathsf{hlt}(j_1), \mathsf{str}(j_1)\rangle$ and $\sigma'_{j_2} = \langle \mathsf{sk}'(j_2), \mathsf{c}'(j_2), \mathsf{hlt}'(j_2), \mathsf{str}'(j_2)\rangle$ and input $\vec{x}$ we define the constraint $\mathsf{enc\_eq\_state}(\sigma_{j_1}, \sigma'_{j_2})$ as

$$
\begin{aligned}
&\mathsf{c}(j_1) = \mathsf{c}'(j_2) \land \mathsf{hlt}(j_1) = \mathsf{hlt}'(j_2)\\
&\land \forall\, w.\, \mathsf{str}(j_1, \vec{x}, w) = \mathsf{str}'(j_2, \vec{x}, w)\\
&\land \forall\, n.\, n < \mathsf{c}(j_1) \to \mathsf{sk}(j_1, \vec{x}, n) = \mathsf{sk}'(j_2, \vec{x}, n)
\end{aligned}
$$

With the presented encoding, ebso, and an SMT solver we can now automatically find optimizations for EVM bytecode. Next, we also want to automatically generate rules.

3.2 Generate Rules with sorg

To generate rules for EVM bytecode we implemented sorg, a superoptimization based rule generator. Like ebso, sorg is implemented in OCaml; sorg depends on ebso for the representation of EVM bytecode and SMT encoding to check observational equivalence.

The main contribution of sorg is to provide notions of program schema, substitutions, and context in order to implement the two main procedures of Section 2.2: generalize and strip. For generalize we implement the procedure from Definition 10, keeping only the most general rules in the result.

▶ **Example 19.** In our evaluation in Section 4, we found the following optimization:

```
SWAP1 POP PUSH 0 PUSH 1 MUL PUSH 0 ⩾ SWAP1 POP PUSH 0 DUP1
```

Generalizing immediate arguments and dropping the prefix `SWAP1 POP` sorg yields two optimization rules: `PUSH` x `PUSH 1 MUL PUSH` x $\Rrightarrow$ `PUSH` x `DUP1` as well as `PUSH 0 PUSH` x `MUL PUSH 0` $\Rrightarrow$ `PUSH 0 DUP1`.

For strip we implement the procedure from Definition 14, keeping only the most stripped rules.

▶ **Example 20.** From the rule `CALLVALUE DUP1 POP` $\Rrightarrow$ `CALLVALUE CALLVALUE POP` sorg can either strip the postfix `POP` or the prefix `CALLVALUE`, obtaining the rules `CALLVALUE DUP1` $\Rrightarrow$ `CALLVALUE CALLVALUE` and `DUP1 POP` $\Rrightarrow$ `CALLVALUE POP`.

One main ingredient of both generalize and strip is a check for observational equivalence. To determine observational equivalence in sorg we use an SMT encoding with components from ebso, similar to Definition 7. For two program schemas ρ and τ, we have $\rho \equiv \tau$ if there are no inputs that distinguish them. That is

$$
\begin{aligned}
\exists \vec{x}.\, &\mathsf{enc\_init}(\vec{x}, \boldsymbol{\sigma}) \land \mathsf{enc\_init}(\vec{x}, \boldsymbol{\sigma'})\\
&\land \mathsf{enc\_progr}(\rho, \boldsymbol{\sigma}) \land \mathsf{enc\_progr}(\tau, \boldsymbol{\sigma'})\\
&\land \neg \mathsf{enc\_eq\_state}(\sigma_{|\rho|+1}, \sigma'_{|\tau|+1})
\end{aligned}
$$

With sorg we can now automatically generate rules, but it remains to glue the tools together and implement a feedback mechanism.

3.3 Coordinate with ppltr

To coordinate our tools ebso and sorg we implemented the tool ppltr, a populator for a peephole optimizer. As ebso and sorg, ppltr is implemented in OCaml. The tool has two main tasks. The first is to manage the interfaces, i.e., to generate ebso blocks from smart contracts, generate ebso blocks for a given size k, prepare optimizations generated by ebso as input for sorg, and analyze and de-duplicate a set of rules produced by sorg. The second main task is to feed back the optimization rules, i.e., to rewrite right-hand sides of the optimization rules themselves, and apply the optimization rules to ebso blocks. To achieve the latter task, ppltr implements a rewrite engine.

4 Evaluation

We evaluate our pipeline by generating optimization rules for EVM bytecode. We collected the 250 most called smart contracts until block 9 786 000 at Apr-01-2020 12:17:26 PM +UTC from the Ethereum blockchain using Google BigQuery.[3]

We split the 250 contracts into 106 798 ebso blocks $\mathcal{E}$. As peephole optimization rules typically span only few instructions, we restrict the size of a block: using a sliding window we split every block larger than 6 instructions into k blocks of at most 6 instructions. To reduce the noise, we remove blocks which are only different in the arguments of `PUSH` keeping only those with words of size smaller than 5 bit. We so obtain 54 301 ebso blocks.

(1) Using ebso find 1580 optimizations from these blocks, run on a cluster with Intel Xeon Gold 6126 CPUs at 2.60 GHz, 2 GB of memory and a time-out of 15 min.

(2) From these optimizations, we generate 1525 rules with sorg, run on the same set-up. For 48 optimizations sorg timed out and could not generate rules and we removed roughly half the rules, as they were duplicates generated from different optimizations.

(3) Thus we arrive at 758 rules $\mathcal{R}_0$, which we use with the rewrite engine of ppltr to

  (a) rewrite the right-hand sides of $\mathcal{R}_0$ reducing 4 rules, and

  (b) rewrite our original ebso blocks in $\mathcal{E}$, which changed 17 255 ebso blocks.

We again use the same window-size and noise reduction to get 25 585 new ebso blocks. Going through the same procedure, we find 452 optimizations with ebso, and generate 435 rules $\mathcal{R}_1$ with sorg with 16 timeouts. Combining the results we get 993 rules $\mathcal{R}_2 = \mathcal{R}_0 \cup \mathcal{R}_1$ which are available at

`github.com/mariaschett/ppltr/blob/v1.0/eval/17-reduced-rules.csv`

We right-reduced 31 rules in $\mathcal{R}_2$ and discarded 967 replicated rules originating from different optimizations. One optimization generated two rules (*cf.* Example 19).

To estimate gas and size saving on a contract level we apply the rules in $\mathcal{R}_2$ to

1. our original 250 most called smart contracts, and
2. extend the data set to the 1000 most called contracts.

---

[3] `cloud.google.com/blog/products/data-analytics/ethereum-bigquery-public-dataset-smart-contract-analytics`.

Table 1 Savings when applying the rules in $\mathcal{R}_2$ on most called contracts.

	accumulated gas savings	accumulated length savings	
250 most called contracts	106 811 g	35 699 instructions	3.94 %
1000 most called contracts	435 002 g	146 376 instructions	4.58 %

Table 1 shows our results. The first column shows the accumulated gas savings over all contracts, and the second column shows the accumulated length savings. Note that results depend on the order in which the rules are applied (*cf.* Section 5). First, we can observe that the rules translate well from 250 to 1000 contracts, achieving roughly 4 times higher savings, which demonstrates that $\mathcal{R}_2$ also extends beyond the original data set, from which it was generated.

Now let us consider the gas savings. In Table 1 we accumulate the cost of all the removed instructions for each contract. How much is actually saved, however, depends on how often the contract is called and which parts are executed. Unfortunately we lack the resources to replay all the transactions to determine the exact savings. Taking into account how often a contract was called, we save 7.41×10^{10} g for the former and 1.02×10^{11} g for the latter. Assuming that about 10 % of a contract is executed per call and that savings are uniformly distributed, this translates to 41 049.33 $ and 56 505.15 $ for a gas price of 27.6 gwei and an ETH-USD course of 200.62 $, which are averages from `etherscan.io/charts`.

While the cost of executing a cheap instruction like `ADD` or `POP` may be negligible, the cost of storing that instruction may not be so. Therefore, we also look at the savings in length: the overall storage space of the bytecode reduces by more than 4.5 %. The contract with the highest length saving was reduced by 19.94 %, removing 345 from originally 1730 instructions.

We also analyze which rules are applied to the contracts. Applying rules may lead to the applicability of other rules, but exploring all rewrite sequences is intractable, and we assume that initial applicability on a contract is a reasonable proxy. Figure 2 groups rules in $\mathcal{R}_2$ by their applicability to the 1000 most called contracts. We can observe a long tail: more than half of the nearly $1k$ rules are applicable only 10 times or less, whereas the top 50 rules are applicable more than 500 times. This suggests that, if a smaller set of rules is desired, this analysis can guide which rules to discard.

Next we inspect the rules within $\mathcal{R}_2$. The five most applied rules for the 1000 most called contracts are listed in Figure 3. Most of these rules are relatively simple and should clearly be applied exhaustively. The fourth rule is perhaps a bit unexpected and may have been missed on manual inspection, but it is cheaper to execute `CALLVALUE` twice than duplicating its result. The last rule hints at a specific compiler produced anti-pattern. Our approach could also be leveraged to detect those.

Figure 4 shows the six rules with the highest gas savings, 17 g and 15 g. We consider two of these rules in more detail. The rule `PUSH 1 MUL PUSH 0 NOT AND` $\Rrightarrow \varepsilon$ combines two observations – that 1 and `PUSH 0 NOT` are neutral elements for multiplication and `AND` respectively.

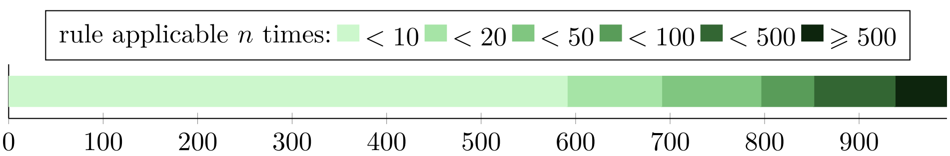

Figure 2 Applicability of rules in $\mathcal{R}_2$ to 1000 most called contracts.

1. `SWAP1 POP POP` ⇛ `POP POP` (×8926)
2. `ISZERO ISZERO ISZERO` ⇛ `ISZERO` (×7893)
3. `PUSH` y `PUSH` x `SWAP1` ⇛ `PUSH` x `PUSH` y (×7742)
4. `CALLVALUE DUP1` ⇛ `CALLVALUE CALLVALUE` (×7740)
5. `SWAP1 SLOAD SWAP1 PUSH` x `EXP SWAP1` ⇛ `PUSH` x `EXP SWAP1 SLOAD` (×5625)

Figure 3 Rules most applied to the 1000 most called contracts.

1. `PUSH 1 MUL DUP3 PUSH 0 NOT AND` ⇛ `DUP3`
   `PUSH 1 MUL PUSH 0 NOT AND` ⇛ ε
2. `PUSH 0 DUP6 DUP5 SUB LT ISZERO` ⇛ `PUSH 1`
   `PUSH 0 NOT AND EQ ISZERO ISZERO` ⇛ `EQ`
   `SWAP1 PUSH 0 NOT AND SWAP1` ⇛ ε
   `PUSH 0 DUP2 PUSH` x `AND LT ISZERO` ⇛ `PUSH 1`

Figure 4 Rules saving most gas.

Depending on the implementation of the peephole optimizer it may be desirable to split this rule which could be achieved by left-reducing the rules. Key to the rule `PUSH 0 DUP6 DUP5 SUB LT ISZERO` ⇛ `PUSH 1` is the less-than comparison `LT` with the smallest element 0 always evaluating to false. The rule does not depend on the result of `DUP6 DUP5 SUB`, and indeed this is replaced by `DUP2 PUSH` x `AND` in the otherwise identical rule in the last line. Generalizing those two rules would require the use of higher-order patterns.

Rules may not only save gas, but also reduce the length of the produced code. These often coincide, and indeed the top 14 length-reducing rules, removing 5 instructions each, subsume the above gas-saving rules. On the other end, there are also rules which save gas but do not reduce the length such as `CALLVALUE DUP1` ⇛ `CALLVALUE CALLVALUE` saving 1 g. In Table 2, we analyze the right-hand sides of $\mathcal{R}_2$. We investigated which instructions were *added*, i.e., do not appear on the left-hand side, and *removed*, i.e., appear on the left- but not the right-hand side of the rule. We group instructions for arithmetic, comparison, bitwise operations, and environment/memory. Unsurprisingly, many more instructions were removed than added, which is expected, because removing instructions always saves gas. The majority of removed instructions is concerned with the stack layout. Surprisingly, also `ISZERO` is often redundant – as also observed in the second rule in Figure 3. Still, instructions are also synthesized on the right-hand side giving rise to optimizations taking the semantic of an instructions into account – potentially also interacting with stack manipulation, for example the rule `SWAP1 LT` ⇛ `GT`.

Finally, we also successfully validated all rules $\mathcal{R}_2$ by running a reference implementation of the EVM, `go-ethereum` version `1.9.14` on pseudo-random input.[4] Therefore, we run the bytecode of every block in $\mathcal{E}$ and the bytecode obtained by applying the rewrite rules to observe that both produce the same final state.

5 Related and Future Work

Chen et al. [7] also developed a tool to rewrite optimization patterns in EVM bytecode. As opposed to our approach, they devised their 24 (anti-)patterns by manual inspection of the code base. Albert et al. [2] synthesize optimized straight-line EVM bytecode for operations on

[4] `github.com/ethereum/go-ethereum`

Table 2 Added and removed instructions by group.

	arith.	comp.	ISZERO	bitwise	DUPi	SWAPi	PUSH	POP	env./mem.
added	10	27	24	12	47	28	134	14	29
removed	80	92	108	83	345	952	182	173	18

the stack with Max-SMT. To gain efficiency, they do not encode the semantics of bit-vector instructions, and instead employ hand-crafted simplification rules. These hand-crafted rules could be inspired by, or even automatically derived from, rules generated by ppltr, which do consider the semantic of bit-vector instructions. Bansal et al. [5] use superoptimization to automatically generate a peephole optimizer for x86 binaries. Aside from the application, the main difference of their approach to ppltr is that it does not process optimizations into rules but instead keeps them in an optimization database in order to reapply them. Moreover it uses an enumeration based superoptimizer, which is more exhaustive, but limits instruction sequences to length 3.

We believe our approach is also applicable for different smart contract languages. Facebook's Move [6] is a gas-metered and verification friendly designed language with an existing code base, such as for example from `github.com/libra/libra/tree/master/language/move-lang/functional-tests/tests`. The machine model of Move is stack-based with typed locals. To adapt the presented approach the SMT encoding would need to be extended to incorporate types and locals. Michelson [15], the smart contract language for the Tezos blockchain, also comes with a detailed formal semantics. Like the EVM it is a stack-based language, but features high-level data types, like lists, sets, and maps. To use the presented approach these data types need to be handled in the SMT encoding and SMT solvers do support complex theories such as sets and lists. Moreover, type information could be used to prune search space, resulting in a positive performance impact.

To automatically integrate the rules generated by ppltr into a compiler a DSL like the one used by GCC[5] or Alive [16] might prove useful. Such an automatic integration would be especially welcome when one wants to re-populate the optimizer of a compiler, e.g. because new instructions are available, such as the addition of shift-operators to the EVM.

Hirai [10] used the meta-tool Lem [17] to formalize the semantics of the EVM. This formalization was extended by Amani et al. [3] by a program logic using the interactive proof assistant Isabelle/HOL to provide an approach to the verification of Ethereum smart contracts. Another formalization of the EVM semantics by Hildenbrandt et al. [9] use the K-framework [20], a rewriting-based framework for defining programming language design and semantics. One of these formalizations could be used to verify the correctness of our encoding, or possibly even generate it automatically.

Our definitions in Section 2.2 are based on concepts from term rewriting [4] and thus we also look at the machinery of term rewriting. Termination of the rules ensures we can apply them exhaustively without looping. Intuitively all rules in $\mathcal{R}_2$ are terminating, since left-hand sides have a higher cost than right-hand sides, and indeed the termination prover WANDA [12] shows termination of all 993 rules in $\mathcal{R}_2$.[6] Confluence guarantees a unique result regardless of how the rules are applied. To check confluence one analyses critical pairs, situations where

[5] `gcc.gnu.org/onlinedocs/gccint/The-Language.html`
[6] We chose WANDA as its support for types allowed us to leverage that arguments of PUSH are words, which greatly aided the automated proof.

application of one rule potentially destroys the possibility for applying another one. The confluence checker CSI [18] reports 82 765 critical pairs, 14 973 of which are joinable and thus harmless. The remaining 67 792 are not, so the rules in $\mathcal{R}_2$ are not confluent. This is not surprising, since there are different ways to achieve the same with the same cost, e.g. `PUSH` x `PUSH` x and `PUSH` x `DUP1`. This may be resolved by defining an additional precedence on the rules, e.g., based on the size of their bytecode. To make a terminating set of rules confluent, one can use *completion* – automatically if we employing tools such as Ctrl [21]. Finally, one could imagine more expressive rules such as `PUSH` x `PUSH` y `ADD` $\Rrightarrow$ `PUSH` z *where* $z = x + y$. Such rules allow to capture constant folding. To do so, rules in *constrained rewriting* [13] come with constraints over a theory as used in SMT solvers.

6 Conclusion

We propose a pipeline to populate the peephole optimizer of a smart contract compiler with three phases to

(1) find optimizations, from which we

(2) generate rules, and

(3) a feedback mechanism to apply the rules.

We demonstrate our approach for EVM bytecode using the tools ebso, sorg, and ppltr, generating 993 peephole optimization rules from the 250 most called contracts of the Ethereum blockchain. We successfully applied our rules to the 1000 most called contracts and discarded 146 376 instructions, saving 435 002 g and 4.5 % storage space. An advantage of our approach lies in its modularity. On the one hand in the modularity of the phases. One could, for example, obtain additional optimizations in a different manner and incorporate them easily. On the other hand, there is the modularity inherent to peephole optimization rules being applied to short programs: it enables an iterative approach to encoding and optimizing instructions based on feasibility and profitability.

Our approach is tailored towards new, rapidly evolving languages and their compilers with clear cost models such as gas metering, and we believe readily applies to languages other than EVM bytecode such as Move and Michelson.

References

1. Wilhelm Ackermann. *Solvable cases of the decision problem.* Studies in logic and the foundations of mathematics. North-Holland Publishing Co., Amsterdam, 1954.
2. Elvira Albert, Pablo Gordillo, Albert Rubio, and Maria A Schett. Synthesis of super-optimized smart contracts using Max-SMT. In *Proc. 32nd CAV*, volume 12224 of *LNCS*. Springer, 2020. doi:10.1007/978-3-030-53288-8_10.
3. Sidney Amani, Myriam Bégel, Maksym Bortin, and Mark Staples. Towards verifying ethereum smart contract bytecode in Isabelle/HOL. In *Proc. 7th CPP*, pages 66–77. ACM, 2018. doi:10.1145/3167084.
4. Franz Baader and Tobias Nipkow. *Term Rewriting and All That.* Cambridge University Press, New York, NY, USA, 1998.
5. Sorav Bansal and Alex Aiken. Automatic generation of peephole superoptimizers. In *Proc. 12th ASPLOS*, pages 394–403. ACM, 2006. doi:10.1145/1168857.1168906.
6. Sam Blackshear, Evan Cheng, David L Dill, Victor Gao, Ben Maurer, Todd Nowacki, Alistair Pott, Shaz Qadeer, Dario Russi, Stephane Sezer, Tim Zakian, and Runtian Zhou. Move: A language with programmable resources. developers.libra.org/docs/assets/papers/libra-move-a-language-with-programmable-resources/2020-04-09.pdf.

7 Ting Chen, Zihao Li, Hao Zhou, Jiachi Chen, Xiapu Luo, Xiaoqi Li, and Xiaosong Zhang. Towards saving money in using smart contracts. In *Proc. 40th ICSE-NIER*, pages 81–84. ACM, 2018. doi:10.1145/3183399.3183420.

8 Leonardo De Moura and Nikolaj Bjørner. Z3: An efficient SMT solver. In *Proc. 14th TACAS*, volume 4963 of *LNCS*, pages 337–340. Springer, 2008. doi:10.1007/978-3-540-78800-3_24.

9 Everett Hildenbrandt, Manasvi Saxena, Xiaoran Zhu, Nishant Rodrigues, Philip Daian, Dwight Guth, Brandon Moore, Yi Zhang, Daejun Park, Andrei Ştefănescu, and Grigore Roşu. KEVM: A complete semantics of the ethereum virtual machine. In *Proc. 31st CSF*, pages 204–217. IEEE, 2018. doi:10.1109/CSF.2018.00022.

10 Yoichi Hirai. Defining the Ethereum virtual machine for interactive theorem provers. In *Financial Cryptography and Data Security*, volume 10323 of *LNCS*, pages 520–535. Springer, 2017. doi:10.1007/978-3-319-70278-0_33.

11 Abhinav Jangda and Greta Yorsh. Unbounded superoptimization. In *Proc. Onward! 2017*, pages 78–88. ACM, 2017. doi:10.1145/3133850.3133856.

12 Cynthia Kop. *Higher Order Termination*. PhD thesis, Vrije Universiteit, Amsterdam, 2012.

13 Cynthia Kop and Naoki Nishida. Constrained term rewriting tool. In *Proc. 20th LPAR*, volume 9450 of *LNCS*, pages 549–557, 2015. doi:10.1007/978-3-662-48899-7.

14 Daniel Kroening and Ofer Strichman. *Decision Procedures: An Algorithmic Point of View*. Springer, Berlin, 2008. doi:10.1007/978-3-540-74105-3.

15 Nomadic Labs. Michelson: the language of smart contracts in Tezos. tezos.gitlab.io/whitedoc/michelson.html.

16 Nuno P. Lopes, David Menendez, Santosh Nagarakatte, and John Regehr. Practical verification of peephole optimizations with Alive. *Commun. ACM*, 61(2):84–91, 2018. doi:10.1145/3166064.

17 Dominic P. Mulligan, Scott Owens, Kathryn E. Gray, Tom Ridge, and Peter Sewell. Lem: Reusable engineering of real-world semantics. In *Proc. 19th ICFP*, pages 175–188. ACM, 2014. doi:10.1145/2628136.2628143.

18 Julian Nagele, Bertram Felgenhauer, and Aart Middeldorp. CSI: New evidence – A progress report. In *Proc. CADE-26*, volume 10395 of *LNAI*, pages 385–397, 2017. doi:10.1007/978-3-319-63046-5.

19 Julian Nagele and Maria A. Schett. Blockchain superoptimizer. In *Preproc. 29th LOPSTR*, pages 166–180, 2019. arXiv:2005.05912.

20 Grigore Roşu and Traian Florin Şerbănuţă. An overview of the K semantic framework. *Journal of Logic and Algebraic Programming*, 79(6):397–434, 2010. doi:10.1016/j.jlap.2010.03.012.

21 Sarah Winkler and Aart Middeldorp. Completion for logically constrained rewriting. In *Proc. 3rd FSCD*, volume 108 of *LIPIcs*, pages 30:1–30:18, 2018. doi:10.4230/LIPIcs.FSCD.2018.30.

22 Gavin Wood. Ethereum: A secure decentralised generalised transaction ledger, 2018. Byzantium Version e94ebda, ethereum.github.io/yellowpaper/paper.pdf.

Tezla, an Intermediate Representation for Static Analysis of Michelson Smart Contracts

João Santos Reis
Release Lab, Nova-Lincs, University of Beira Interior, Portugal
joao.reis@ubi.pt

Paul Crocker
Release Lab, C4, University of Beira Interior, Portugal
crocker@di.ubi.pt

Simão Melo de Sousa
Release Lab, C4, Nova-Lincs, University of Beira Interior, Portugal
desousa@di.ubi.pt

Abstract

This paper introduces Tezla, an intermediate representation of Michelson smart contracts that eases the design of static smart contract analysers. This intermediate representation uses a store and aims to preserve the semantics, flow and resource usage of the original smart contract. This enables properties like gas consumption to be statically verified. We provide an automated decompiler of Michelson smart contracts to Tezla. In order to support our claim about the adequacy of Tezla, we develop a static analyser that takes advantage of the Tezla representation of Michelson smart contracts to prove simple but non-trivial properties.

2012 ACM Subject Classification Theory of computation → Program analysis

Keywords and phrases Intermediate representation, Static analysis, Tezos blockchain, Michelson

Digital Object Identifier 10.4230/OASIcs.FMBC.2020.4

Related Version `https://arxiv.org/abs/2005.11839`

Supplementary Material Source code of the implementation available at: `https://gitlab.com/releaselab/fresco/tezla`.

Funding This work was supported and funded by the Tezos Foundation by the project FRESCO (Formal Verification of Tezos Smart Contracts).

1 Introduction

The term "smart contract" was proposed by Nick Szabo as a way to formalize and secure relationships over public networks [26]. In a blockchain, a smart contract is an application written in some specific language that is embedded in a transaction (hence the program code is immutable once it is on the blockchain). Some examples of smart contracts applications are the management of agreements between parties without resorting to a third party (escrow) and to function as a multisignature account spending requirement. Smart contracts have the ability to transfer/receive funds to/from users or from other smart contracts and can interact with other smart contracts.

There are reports of bugs and consequently attacks in smart contracts that have led to losses of millions of dollars worth of assets. One of the most famous and most costly of these attacks was on the Distributed Autonomous Organization (DAO), on the Ethereum blockchain [8]. The attacker managed to withdraw approximately 3.6 million ethers from the contract.

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler; Article No. 4; pp. 4:1–4:12
OpenAccess Series in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Given the fact that a smart contract in a blockchain cannot be updated or patched, there is an increasing interest in providing tools and mechanisms that guarantee or potentiate the correctness of smart contracts and to verify certain properties. However, current tools and algorithms for program verification that are based, for example, on deductive verification and static analysis, are usually designed for classical store-based languages in contrast with Michelson [15], the smart contract language for the Tezos Blockchain [11], which is stack-based.

To facilitate the usage of such tools to verify Michelson smart contracts, we present Tezla, a store-based intermediate representation language for Michelson, and its respective tooling. We provide an automated translator of Michelson smart contracts to Tezla. The translator was designed and implemented in a way that aims to preserve the semantics, flow, and resource usage of the original smart contract, so that properties like gas consumption can be faithfully verified at the Tezla representation level. To support our work, we present a case study of a demo platform for the static analysis of Tezos smart contracts using the Tezla intermediate representation alongside with an example analysis.

The paper is structured as follows. In section 2, we introduce the Tezla intermediate representation and the translation mechanism of Michelson code to Tezla. Section 3 addresses the static analysis platform case study that targets Tezla-represented smart contracts. In section 4, we talk about the related work. Finally, section 5 concludes with a general overview of this contribution and future lines of work.

2 Tezla

Tezla aims to facilitate the adoption of existing static analysis tools and algorithms. As such, Tezla is an intermediate representation of Michelson code that uses a store instead of a stack, enforces the Static Single-Assignment Form (SSA) [20] and aims to preserve information about gas consumption. We will see in the next section how such characteristics ease the translation of a Tezla program into its Control Flow Graph (CFG) forms and the construction of data-flow equations.

Compiled languages (like Albert [5], LIGO [1], SmartPy [16], Lorentz [25], etc.) also provide a higher-level abstraction over Michelson. However, as it happens with most compiled languages, the produced code may not be as concise or compact as expected which, in the case of smart contracts, may result in a higher gas consumption and, consequentially, undesired costs. Tezla was designed to have a tight integration with the Michelson code to be executed, not as a language that compiles to Michelson neither as a higher level language to ease the writing of Michelson smart contracts.

Tezla adapts the Michelson syntax and semantics in order to transform the stack usage to a traditional store usage. As such, we encourage the reader to head to the Michelson documentation [14] for more information about the Michelson language and its syntax and semantics.

Due to its large extent, the full syntax and semantics of the Tezla representation are not presented here but can be found at `https://gitlab.com/releaselab/fresco/tezla-semantics`.

In a general way, Michelson push-like instructions are translated into variable assignments, whereas instructions that consume stack values are translated to expressions that use as arguments the variables that match the values from the stack. Furthermore, lists, sets and maps deconstruct and lifting of `option` and `or` types that happen implicitly are represented through explicit expressions added to Tezla.

Since the operational effect of stack manipulation is transposed into variable assignments, we also expose in a Tezla represented contract the stack manipulation as instructions that act as no-op instructions in the case of a semantics that do not take resource consumption into account[1]. In the case of a resource aware semantics, these instructions will semantically encode this consumption.

The following section describes in detail the process of transforming a Michelson smart contract to a Tezla representation.

2.1 Push-like instructions and stack values consumption

Instructions that push N values to the stack are translated to N variable assignments of those values. The translation process maintains a Michelson program stack that associates each stack position to the variable to which that position value was assigned to. When a stack element is consumed, the corresponding variable is used to represent the value. A very simple example is provided in listings 1 and 2.

Listing 1 Stack manipulation example – Michelson code.

```
PUSH nat 5;
PUSH nat 6;
ADD;
```

Listing 2 Stack manipulation example – Tezla code.

```
v1 := PUSH nat 5;
v2 := PUSH nat 6;
v3 := ADD v1 v2;
```

The block on listing 1 is translated to the Tezla representation shown in listing 2.

From the previous example, we can also observe that Michelson instructions that consume N stack variables are translated to an expression that consumes those N values. Concretely, the instruction `ADD` that consumes two values (say, `a` and `b`), from the stack is translated to `ADD a b`.

2.2 Branching and deconstructions

Michelson provides developers with branching structures that act on different conditions. As Tezla aims at being used as an intermediate representation for static analysis, there are some properties we would like to maintain. One such property is static single-assignment form (SSA-form) [20], so that we obtain data flow information in a way that simplifies analyses and code optimization. This is guaranteed as Tezla-represented smart contracts are, by construction, in SSA-form, since each assignment uses new variables.

In order to deal with branching, the Tezla representation makes use of ϕ-functions (see [20]) that select between two values depending on the branch. As an illustration consider the Michelson example in listing 3.

The `IF_CONS` instruction tests if the top of the stack is a non-empty list, and deconstructs the list in the true branch by putting the head and the tail of the list on top of the stack.

In this example, the `IF_CONS` instructions checks the top of the stack and if it is a non-empty list it inserts the sum of a `int` value already on the stack with the head of the list at the lists's head. If the list is empty, it inserts the value into the empty list. Here, each branch of the `IF_CONS` instruction will result in a stack with a list of integers, whose values depends on which branch was executed. This translates to the Tezla representation presented on listing 4.

[1] This is the case of the semantics presented in this paper.

Listing 3 Branching example – Michelson code.

```
IF_CONS
  { DUP ;
    DIP { CONS ; SWAP } ;
    ADD ; CONS }
  { NIL int ; SWAP ; CONS } ;
  DUP ;
  PAIR;
```

Listing 4 Branching example - Tezla code.

```
IF_CONS v1
{
  v2 := hd v1;
  v3 := tl v1;
  v4 := DUP v2;
  v5 := CONS v2 v3;
  SWAP;
  v6 := ADD v4 v0;
  v7 := CONS v6 v5
}
{
  v8 := NIL int;
  SWAP;
  v9 := CONS v0 v8
};
v10 := ϕ(v7, v9);
v11 := DUP v10;
v12 := PAIR v11 v10;
```

The variable `v10` will receive its value through a ϕ-function that returns the value of `v7` if the true branch is executed, or the value of `v9` otherwise.

From this example, it is possible to observe that the deconstruction of a list is explicit through two variable assignments. This is also the behaviour of `IF_NONE` and `IF_LEFT` instructions, where the unlifting of **`option`** and **`or`** types happens explicitly through an assignment.

2.3 Loops, maps and iterations

Michelson also provides language constructs for looping and iteration over the elements of lists, sets and maps. Sets in Michelson are defined as ordered lists, whereas maps are defined as lists of key-value pairs ordered by key. These are treated using the same ϕ-functions mechanism in order to preserve SSA-form. We can observe this on the example listing 5.

This example uses a `LOOP_LEFT` (loop with an accumulator) to sum 1 to a **`nat`** (starting with the value 0) until that value becomes greater than 100 and casts the result to an **`int`**. This example translates to the code presented in listing 6.

Note that the `LOOP_LEFT` variable is assigned to the value of `v1` if it is the first time that the loop condition is checked, or `v12` if the program flow comes from the loop body. Moreover, notice that the same explicit deconstruction of an **`or`** (union type) variable is applied here, where `v3` gets assigned the value of the unlifting of the loop variable in the beginning of the loop body and `v13` at the end of the loop. Similar behaviour applies to the other looping and iteration instructions.

2.4 Parameter and Storage

We now present an example of a complete Michelson smart contract (listing 7).

The contract takes an **`int`** as parameter and adds 1 to that value, which is later put in the storage. This contract translates to the Tezla code of figure 8.

Listing 5 Loop example – MICHELSON code.

```
PUSH nat 0 ;
LEFT nat ;
LOOP_LEFT
 { DUP ;
   PUSH nat 100 ;
   COMPARE ;
   GE ;
   IF
    { PUSH nat 1 ;
      ADD ; LEFT nat }
    { RIGHT nat } } ;
INT ;
```

Listing 6 Loop example – TEZLA code.

```
v0 := PUSH nat 0;
v1 := LEFT nat v0;
LOOP_LEFT v2 := ϕ(v1, v12)
{
  v3 := unlift_or v2;
  v4 := DUP v3;
  v5 := PUSH nat 100;
  v6 := COMPARE v5 v4;
  v7 := GE v6;
  IF v7
  {
    v8 := PUSH nat 1;
    v9 := ADD v8 v3;
    v10 := LEFT nat v9;
  }
  {
    v11 := RIGHT nat v3;
  }
  v12 := ϕ(v10, v11);
}
v13 := unlift_or v2;
v14 := INT v13;
```

Listing 7 Example contract – MICHELSON code.

```
parameter (int) ;
storage (int) ;
code { CAR ;
       PUSH int 1;
       ADD;
       NIL operation ;
       PAIR; }
```

Listing 8 Example contract – TEZLA code.

```
v0 := CAR parameter_storage;
v1 := PUSH int 1;
v2 := ADD v1 v0;
v3 := NIL operation;
v4 := PAIR v3 v2;
return v4;
```

In this example, we can observe that a MICHELSON contract has a parameter and storage. The initial stack of any MICHELSON smart contract is a stack that contains a single pair whose first element is the input parameter and second element is the contract storage. As such, we introduce a variable called `parameter_storage` that contains the value of that pair.

The final stack of any MICHELSON smart contract is also a stack that contains a single pair whose first element is a list of internal operations that it wants to emit and whose second element is the resulting storage of the smart contract. We identify the variable containing this pair through the addition of a `return` instruction.

3 Building static analyses for Tezla smart contracts

In this section, we present the experiments conducted in order to test and demonstrate the applicability of the TEZLA intermediate representation to perform static analysis.

3.1 SoftCheck

We build and organise these static analyses upon a generic data-flow analysis platform called SOFTCHECK [18]. SOFTCHECK provides an internal and intermediate program representation, called SCIL, rich enough to express high-level as well as low-level imperative programming constructs and simple enough to be adequately translated into CFGs.

SoftCheck is organised upon a generic monotone framework [12] that is able to extract a set of data-flow equations from (1) a suitable representation of programs and; (2) a set of monotone functions; and then to solve them. SoftCheck is written in OCaml and makes use of functor interfaces to leverage its genericity (see figure 1).

By generic we mean that, given a translation from a programming language to SCIL. SoftCheck gives the ability to instantiate its underlying monotone framework by means of a functor interface. Then all defined static analyses are automatically available for the given programming language.

On the other hand, once written as a set of properties that define the domain of the analysis and the monotone functions on that domain, a particular static analysis can be incorporated (again, through instantiating a functor) as an available static analysis for all interfaced programming languages.

SoftCheck offers several standard data-flow analysis such as very busy expressions, available expressions, tainted analysis etc.

We propose in the next sections to detail how we have interfaced Tezla with SCIL, how we have designed a simple but useful data-flow analysis within SoftCheck and how we have tested this analysis on the Michelson smart contracts running in the Tezos blockchain.

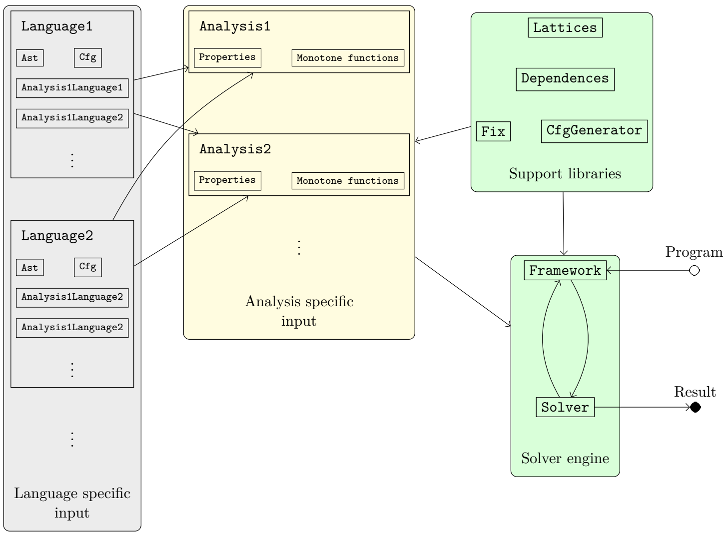

Figure 1 SoftCheck in a picture (adapted from [21]).

3.2 Constructing a Tezla Representation of a Contract

To obtain the Tezla representation of a smart contract, we first developed a parser to obtain an abstract syntax representation of a Michelson smart contract. This parser was implemented in OCaml and Menhir and respects the syntax described in the Tezos documentation [15]. It allows us to obtain a data type that fully abstracts the syntax (with the exception of annotations). The reason behind the implementation of our own parser was to obtain a data type that would better suit and ease the adoption of the integration with SoftCheck. Therefore, to improve the integration between these two forms, Tezla data types were built upon the data types of Michelson.

Control-flow graphs are a common representation among static analysis tools. We provide a library for automatic extraction of such representation from any Tezla-represented smart contract. This library is based upon the control-flow generation template present in

SoftCheck. As such, control-flow graphs generated with this library can be used with SoftCheck without further work. To instantiate the control-flow graph generation template, we simply provided the library with a module with functions that describe how control flows between each node.

3.3 Sign Detection: An Example Analysis

At this point, the SoftCheck platform is ready to be used to develop data flow analyses targeting Tezla represented smart contracts.

Here we devise an example of a static analysis for sign detection. The abstract domain consists of the following abstract sign values: 0 (zero), 1 (one), 0+ (zero or positive), 0- (zero or negative), + (positive), - (negative), $\top$ (don't know) and $\perp$ (not a number). These values are organised according to the lattice on figure 2.

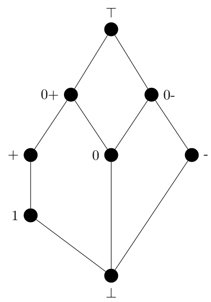

Figure 2 Sign lattice.

Using SoftCheck, we implemented a simple sign detection analysis of numerical values. By definition, `nat`s have a lowest precision value of 0+, while `int`s can have any value. Every other data type has a sign value of $\perp$.

This implementation does not propagate information to non-simple types (`pair`, `or`, etc.), but it does perform some precision refinements on branching.

To implement such an analysis, we provided SoftCheck, in addition to the previously defined Tezla control-flow graph library, a module that defines how each instruction impacts the sign value of a variable. Then, using the integrated solver mechanism based on the monotone framework, we are able to run this analysis on any Tezla represented smart contract.

We now present an example. Listings 9 and 10 show the code of a smart contract and its Tezla representation. This contract multiplies its parameter by -5 if the parameter is equal to 0, or by -2 otherwise, and stores the result in the storage. Figure 3 shows the control-flow graph of representation of that contract.

Running this analysis on the previously mentioned contract produced the results shown in Figure 4. In these results we can observe the known sign value of each variable at the exit of each block of the control-flow graph in Figure 3. For brevity, we omitted non-numerical variables from the result.

It it possible to observe from the results that the analysis takes into account several details. For instance, the sign of values of type `nat` are, by definition, always zero or positive. The analysis also refines the sign values on conditional branches according to the test. In this case, we can observe that in blocks 6 and 7 (true branch) the sign value of `v1` must be `0`, as the test corresponds to `0 == v1`. Complementary to this, in blocks 8 and 9 the value of `v1` assumes the sign value of `+`, since being a `nat` value its value must be `0+` and we know that its values is not zero because the test `0 == v1` failed.

Listing 9 Example contract for sign analysis – MICHELSON code.

```
parameter nat ;
storage int ;
code { CAR ;
       DUP ;
       PUSH nat 0 ;
       COMPARE ;
       EQ ;
       IF { PUSH int -5 ; MUL }
          { PUSH int -2 ; MUL } ;
       NIL operation ;
       PAIR }
```

Listing 10 Example contract for sign analysis – TEZLA code.

```
v0 := CAR parameter_storage;
v1 := DUP v0;
v2 := PUSH nat 0;
v3 := COMPARE v2 v1;
v4 := EQ v3;
IF v4
{
    v5 := PUSH int -5;
    v6 := MUL v5 v0;
}
{
    v7 := PUSH int -2;
    v8 := MUL v7 v0;
};
v9 := phi(v6, v8);
v10 := NIL operation;
v11 := PAIR v10 v9;
return v11;
```

Due to the TEZLA nature, we were able to take advantage of existing tooling, such as the SOFTCHECK platform, and effortlessly design the run a data-flow analysis. This enables and eases the development of static analysis that can be used to verify smart contracts but also to perform code optimisations, such as dead code elimination. Albeit simple, the sign analysis can be used to instrument such dead code elimination procedure.

3.4 Experimental Results and Benchmarking

TEZLA and all the tooling are implemented in OCaml and are available at [13]. TEZLA accepts Michelson contracts that are valid according to the Tezos protocol 006 Carthage. We conducted experimental evaluations that consisted in transforming to TEZLA and running the developed analyses on a batch of smart contracts.

To do so, we implemented a tool that allows the extraction of smart contracts available in the Tezos blockchain. With that tool, we extracted 142 unique smart contracts. We tested these unique contracts alongside 21 smart contracts we have implemented ourselves.

We successfully converted all smart contracts with a coverage result of all Michelson instructions except for 9 instructions that were not used in any of these 163 contracts. On those, we ran the available analyses and obtained the benchmarks presented on table 1. These experiments were performed on a machine with an Intel i7–8750H (2.2 GHz) processor with 6 cores and 32 GB of RAM.

In the absence of an optimisation tool that takes advantage of the information computed by the analysis, we do not produce any optimisations from the analyses results. To do so, currently one must manually inspect the reports produced by the analysis. These reports, the source code of contracts under evaluation, as well as the respective analysis results and other performed static analyses are available at [22, 19].

4 Related Work

Albert [5] is an intermediate language for the development of Michelson smart contracts. This language provides an high-level abstraction of the stack and some of the language datatypes. This language can be compiled to Michelson through a compiler written in Coq that targets Mi-Cho-Coq [4], a Coq specification of the Michelson language.

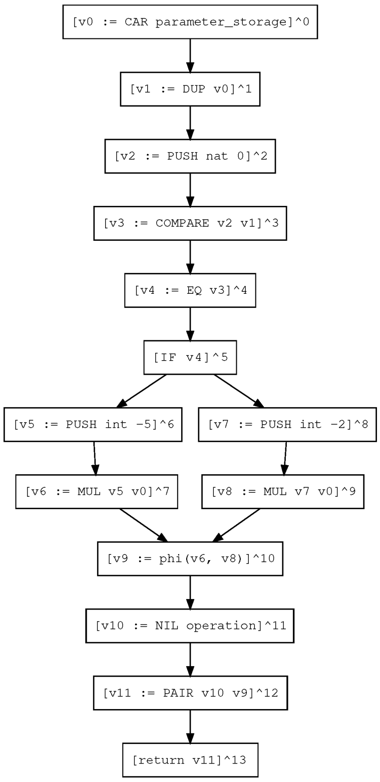

Figure 3 Generated CFG, by the SoftCheck tool.

Table 1 Benchmark results.

Average time	0.48 s	Worst-case (number of instructions)	2231 (6.08 s)
Worst-case (time)	9.87 s (926 instructions)	Average time per instrucion	0.0009

Several high-level languages [1, 2, 16, 7, 25] that target Michelson have been developed. Each one presents a different mechanism that abstracts the low-level stack usage. However, a program analysis tool that would target one of these languages should not be easily reusable to programs written in the other languages.

Scilla [23, 24] is an intermediate language that aims to be a translation target of high-level languages for smart contract development. It introduces a communicating automata-based computational model that separates the communication and programming aspects of a contract. The purpose of this language is to serve as a basis representation for program analysis and verification of smart contracts. We believe that Tezla is at a different level than Scilla, as we could use a Tezla representation to be mid step between having a Scilla representation and the Michelson code.

Slither [10], presented in 2019, is a static analysis framework for Ethereum smart contract. It uses the Solidity smart contract compiler generated Abstract Syntax Tree to transform the contract into an intermediate representation called SlithIR. This representation also uses a

```
0: {
    v0: 0+
}
1: {
    v0: 0+,
    v1: 0+
}
2: {
    v0: 0+,
    v1: 0+,
    v2: 0
}
3: {
    v0: 0+,
    v1: 0+,
    v2: 0,
    v3: 0-
}
4: {
    v0: 0+,
    v1: 0+,
    v2: 0,
    v3: 0-
}
5: {
    v0: 0+,
    v1: 0+,
    v2: 0,
    v3: 0-
}
6: {
    v0: 0,
    v1: 0,
    v2: 0,
    v3: 0-,
    v5: -
}
7: {
    v0: 0,
    v1: 0,
    v2: 0,
    v3: 0-,
    v5: -,
    v6: 0
}
8: {
    v0: +,
    v1: +,
    v2: 0,
    v3: 0-,
    v7: -
}
9: {
    v0: +,
    v1: +,
    v2: 0,
    v3: 0-,
    v7: -,
    v8: -
}
10: {
    v0: 0+,
    v1: 0+,
    v2: 0,
    v3: 0-,
    v5: -,
    v6: 0,
    v7: -,
    v8: -,
    v9: 0-
}
11: {
    v0: 0+,
    v1: 0+,
    v2: 0,
    v3: 0-,
    v5: -,
    v6: 0,
    v7: -,
    v8: -,
    v9: 0-
}
12: {
    v0: 0+,
    v1: 0+,
    v2: 0,
    v3: 0-,
    v5: -,
    v6: 0,
    v7: -,
    v8: -,
    v9: 0-
}
13: {
    v0: 0+,
    v1: 0+,
    v2: 0,
    v3: 0-,
    v5: -,
    v6: 0,
    v7: -,
    v8: -,
    v9: 0-
}
```

Figure 4 Generated report for the sign analysis.

SSA form and a reduced instruction set to facilitate the implementation of program analyses of smart contracts. However, the Slither intermediate representation is not able to accurately model some low-level information like gas computations, which we took into account when designing Tezla. Also, this work does not contemplate a formal semantics of SlithIR.

Solidifier [3] is a bounded model checker for Ethereum smart contracts that converts the original source code to Solid, a formalisation of Solidity that runs on its own execution environment. Solid is translated to Boogie, an intermediate verification language that is used by the bounded model checker Corral, which is then used to look for semantic property violations.

Durieux et. al [9] presented a review on static analysis tools for Ethereum smart contracts. This work presents an extensive list of 35 tools, of which 9 respected their inclusion criteria: the tool is publicly available and supports a command-line interface; takes as input a Solidity contract; requires nothing but the source code of the contract; the tool claims to be able to identify vulnerabilities and bad practices in the contract. The authors then used those tools to test several vulnerabilities on a sample set of 47,587 smart contracts. This work presents some interesting results, as it was able to detect 97% of the smart contracts as vulnerable, as well as identify two categories of DASP10 as not able to be detect by the tools.

5 Conclusion

To the best of our knowledge, this is the first work towards a static analysis framework for Tezos smart contracts. Tezla positions itself as an intermediate representation obtained from a Michelson smart contract, the low-level language of Tezos smart contracts. This representation abstracts the stack usage through the usage of a store, easing the adoption of mechanisms and frameworks for program analysis that assume this characteristic, while maintaining the original semantics of the smart contract.

We have presented a case study on how this intermediate representation can be used to implement a static analysis by using Tezla along side the SoftCheck platform. This has shown how effortlessly one can perform static analysis on Michelson code without forcing developers to use a different language or implement `ad hoc` static analysis tooling for a stack-based language.

Michelson smart contracts have a mechanism of contract level polymorphism called entrypoints, where a contract can be called with an entrypoint name and an argument. This mechanism takes the form of a parameter composed as nesting of `or` types with entrypoint name annotations. This parameter is then checked at the top of the contract in a nesting of `IF_LEFT` instructions, running the desired entry point this way. This mechanism is optional and transparent to smart contracts without entry points. As such, they are also transparent to Tezla. We therefore plan to extend Tezla to deal with entrypoints and generate isolated components for each entrypoint of a smart contract, which allow us to obtain clearer control flow graphs and analysis results. This allows us to analyse each entry point separately and possibly obtain more fine-grained results.

5.1 Future Work

At the moment of this paper writing, there is an initial work on an static analysis of Tezla represented smart contracts to detect potentially costly loops.

Future plans include a proof of correctness of the Michelson to Tezla transformation through a proof of equivalence of the Tezla semantics in respect to Michelson semantics. We aim to do so by developing a Tezla semantics using the Why3 deductive program verification platform and using the work done in WhylSon [6] to prove the semantic equivalence of Michelson and Tezla. Furthermore, this semantics should be accountable of gas consumption, so that we can provide a sound Tezla resource analysis in respect to the original Michelson code. This will also make way to the development of a platform for principled static analysis of Michelson smart contracts.

We plan to study which problems and properties are of interest so that we can integrate existing tools and algorithms for code optimization, resource usage and security analysis and correctness verification.

Another direction to tackle is the interfacing of Tezla with other static analysis platforms such as those provided by the MOPSA project [17] which, among other capabilities, provides a means to integrate static analyses. The integration with different static analysis platforms makes way to a more diverse universe of possible static analysis. Furthermore, it reinforces the statement that Tezla is an intermediate representation suitable not only for SoftCheck but for other platforms.

References

1 Gabriel Alfour. LIGO. URL: `https://ligolang.org/`.

2 Stephen Andrews and Richard Ayotte. Fi - Smart coding for smart contracts. URL: `https://fi-code.com/`.

3 Pedro Antonino and A. W. Roscoe. Formalising and verifying smart contracts with Solidifier: A bounded model checker for Solidity. *arXiv:2002.02710 [cs]*, February 2020. `arXiv:2002.02710`.

4 Bruno Bernardo, Raphaël Cauderlier, Zhenlei Hu, Basile Pesin, and Julien Tesson. Mi-Cho-Coq, a framework for certifying Tezos Smart Contracts. In *1st Workshop on Formal Methods for Blockchains*, September 2019. `arXiv:1909.08671`.

5 Bruno Bernardo, Raphaël Cauderlier, Basile Pesin, and Julien Tesson. Albert, an intermediate smart-contract language for the Tezos blockchain. In *4th Workshop on Trusted Smart Contracts*, January 2020. `arXiv:2001.02630`.

6 Luís Pedro Arrojado da Horta, João Santos Reis, Mário Pereira, and Simão Melo de Sousa. WhylSon: Proving your Michelson Smart Contracts in Why3. *arXiv:2005.14650 [cs]*, May 2020. `arXiv:2005.14650`.

7 DaiLambda. SCaml. URL: `https://gitlab.com/dailambda/scaml`.

8 Michael del Castillo. The DAO Attacked: Code Issue Leads to $60 Million Ether Theft, June 2016. URL: `https://www.coindesk.com/dao-attacked-code-issue-leads-60-million-ether-theft`.

9 Thomas Durieux, João F. Ferreira, Rui Abreu, and Pedro Cruz. Empirical Review of Automated Analysis Tools on 47,587 Ethereum Smart Contracts. In *42nd International Conference on Software Engineering (ICSE '20)*, February 2020. `doi:10.1145/3377811.3380364`.

10 Josselin Feist, Gustavo Greico, and Alex Groce. Slither: A static analysis framework for smart contracts. In *Proceedings of the 2nd International Workshop on Emerging Trends in Software Engineering for Blockchain*, WETSEB '19, pages 8–15, Montreal, Quebec, Canada, May 2019. IEEE Press. `doi:10.1109/wetseb.2019.00008`.

11 L M Goodman. Tezos-a self-amending crypto-ledger White paper. Technical report, Tezos, 2014.

12 John B. Kam and Jeffrey D. Ullman. Monotone data flow analysis frameworks. *Acta Informatica*, 7(3):305–317, 1977. `doi:10.1007/BF00290339`.

13 RELEASE Lab. FRESCO - formal verification and static analysis of tezos compliant smart contracts, gitlab. URL: `https://gitlab.com/releaselab/fresco`.

14 Nomadic Labs. Michelson Reference. URL: `https://michelson.nomadic-labs.com`.

15 Nomadic Labs. Michelson: The language of Smart Contracts in Tezos. URL: `http://tezos.gitlab.io/whitedoc/michelson.html`.

16 Francois Maurel and Smart Chain Arena. SmartPy. URL: `https://smartpy.io/`.

17 A. Miné, A. Ouadjaout, and M. Journault. Design of a modular platform for static analysis. In *Proc. of 9h Workshop on Tools for Automatic Program Analysis (TAPAS'18)*, Lecture Notes in Computer Science (LNCS), page 4, August 2018. URL: `http://www-apr.lip6.fr/~mine/publi/mine-al-tapas18.pdf`.

18 João Reis. Softcheck, a generic and modular data-flow analyser. URL: `https://gitlab.com/joaosreis/softcheck`.

19 João Reis. TezCheck Analysis Results - GitLab. URL: `https://gitlab.com/releaselab/fresco/tezcheck-results`.

20 B. K. Rosen, M. N. Wegman, and F. K. Zadeck. Global value numbers and redundant computations. In *Proceedings of the 15th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages - POPL '88*, pages 12–27, San Diego, California, United States, 1988. ACM Press. `doi:10.1145/73560.73562`.

21 João Santos Reis. *SoftCheck, uma plataforma de construção de análises estáticas para a segurança de programas, genérica e composicional.* PhD thesis, University of Beira Interior, Covilhã, Portugal, July 2018.

22 João Santos Reis. TezCheck - softcheck interface for tezos, gitlab. URL: `https://gitlab.com/releaselab/fresco/tezcheck`.

23 Ilya Sergey, Amrit Kumar, and Aquinas Hobor. Scilla: A Smart Contract Intermediate-Level LAnguage. *arXiv:1801.00687 [cs]*, January 2018. `arXiv:1801.00687`.

24 Ilya Sergey, Vaivaswatha Nagaraj, Jacob Johannsen, Amrit Kumar, Anton Trunov, and Ken Chan Guan Hao. Safer smart contract programming with Scilla. *Proceedings of the ACM on Programming Languages*, 3(OOPSLA):1–30, October 2019 . `doi:10.1145/3360611`.

25 Serokell and TQ Group. Lorentz. URL: `https://gitlab.com/morley-framework/morley/-/tree/master/code/lorentz`.

26 Nick Szabo. Formalizing and Securing Relationships on Public Networks. *First Monday*, 2(9), September 1997. `doi:10.5210/fm.v2i9.548`.

A Blockchain Model in Tamarin and Formal Analysis of Hash Time Lock Contract

Colin Boyd
NTNU - Norwegian University of Science and Technology, Trondheim, Norway
colin.boyd@ntnu.no

Kristian Gjøsteen
NTNU - Norwegian University of Science and Technology, Trondheim, Norway
kristian.gjosteen@ntnu.no

Shuang Wu
NTNU - Norwegian University of Science and Technology, Trondheim, Norway
shuang.wu@ntnu.no

Abstract

Formal analysis and verification methods can aid the design and validation of security properties in blockchain based protocols. However, to generate a reasonable and correct verification, a proper model for the blockchain is needed. In this paper, we give a blockchain model in Tamarin. Based on our model we analyze and give a formal verification for the hash time lock contract, an atomic cross chain trading protocol. The result shows that our model is able to identify an underlying assumption for the hash time lock contract and that the model is useful for analyzing blockchain based protocols.

2012 ACM Subject Classification Security and privacy → Formal security models

Keywords and phrases Blockchain model, Tamarin, Hash time lock contract, Formal verification

Digital Object Identifier 10.4230/OASIcs.FMBC.2020.5

Supplementary Material The source code can be found at
`https://github.com/ShuangWu121/Tamarin-code-for-HTLC-verification`

1 Introduction

In a blockchain based protocol, the blockchain serves as a reliable public ledger to deliver ordered outcomes to all its agents. Protocols can be executed by using smart contracts and the execution states are recorded on the blockchain. The blockchain essentially performs as a distributed trusted party to reduce the direct trust between the entities in the system.

In order to formally verify the security properties of protocols built on top of blockchains, a proper model for blockchains is needed. The model must capture the interesting properties of blockchains, without becoming too complicated. A blockchain is more than a public ledger. The dynamics of the growing chain provide a time reference: the relatively stable growth of the blockchain height offers a "global time". With respect to this global time, a blockchain enables a time lock function used as a restriction specifying that a transaction cannot be added to blockchain before a set time (actually a given chain height). Thus in order to capture properties of time lock contracts a blockchain model should include the following features.

Model time. The blockchain model should contain a global time reference in the system. Different blockchains contain different global time references. The model should be able to capture time-relevant risks, such as race conditions.

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler; Article No. 5; pp. 5:1–5:13
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Model the time lock restriction. The time out event of a time lock should be triggered by the time reference. It should be possible to model the risk introduced by a time lock that times out earlier or later than is expected.

Clarify the underlying assumptions. If a certain property of the blockchain fails, a protocol built on top of it will not be safe either.

Related work

In 2014, Andrychowicz et al. [2] modeled a multiparty computation contract in Bitcoin by using timed automata. Back then the time lock functionality in Bitcoin was limited and consequently the structure of the contract is different today. Bursuc and Kremer [4] used Tamarin to model the blockchain as a public ledger, and analysed the ZKCP [7] protocol built on top of it. But in their model, the executions are not time-relevant. Turuani et al. [10] give a formal model in ASLan++ of the two-factor authentication protocol used by the Electrum Bitcoin wallet. Bentov et al. [3] propose a real-time cryptocurrency exchange service, and they give an informal cryptographic proof for the security of a hash time lock protocol, with a probabilistic modeling of forking. Sun and Yu [9] give a formal verification model for five kinds of security issues in the Ethereum blockchain using Coq.

Our contributions

To address the above challenges, we build a blockchain model in Tamarin [8]. The model defines a public ledger and a global time reference for the system, with time lock functionality built on top. We also define the security properties of an atomic cross chain trading protocol and give a formal proof for the security of the hash time lock contract (HTLC). To our knowledge, this is the first HTLC analysis by formal verification tools. The proof clarifies a "hidden" security assumption: the growth speed of the two blockchains need to be stable, otherwise security will fail. We further use our model to analyze an older version of the hash time lock contract, and Tamarin is able to find a flaw. Even if the assumption looks trivial and the flaw is somewhat obscure, this demonstrates that our model is able to address the above challenges and can be used for formal verification of blockchain based protocols.

2 Background

2.1 Hash time lock contract

The goal of the hash time lock contract (HTLC) is to exchange different cryptocurrencies between two players in a decentralized way. Consider Alice who wants to exchange Bitcoin for Altcoin, and Bob who wants to exchange Altcoin for Bitcoin. They could do the following:

1. Alice creates a transaction that is locked by a hash value $\mathsf{h} := \mathsf{H}(\mathsf{sk})$ to send Bob 1 Bitcoin. Bob can take the funds only if he can provide the hash pre-image. This is Alice's *commitment transaction*.
2. After Alice's commitment transaction has been confirmed on the Bitcoin blockchain, Bob creates a transaction (contract) to send 1 Altcoin to Alice, locked using the same hash value $\mathsf{h} := \mathsf{H}(\mathsf{sk})$. This is Bob's commitment transaction.
3. Alice takes Bob's Altcoin by providing her signature and the pre-image of the hash lock. Bob learns the hash pre-image and unlocks the Bitcoin that Alice sent to him.

In order to avoid an interrupted protocol leaving players' funds locked forever, the commitment transactions are also locked by time locks. After the time lock times out, the transaction can be redeemed by the sender. The time lock of Alice's commitment transaction

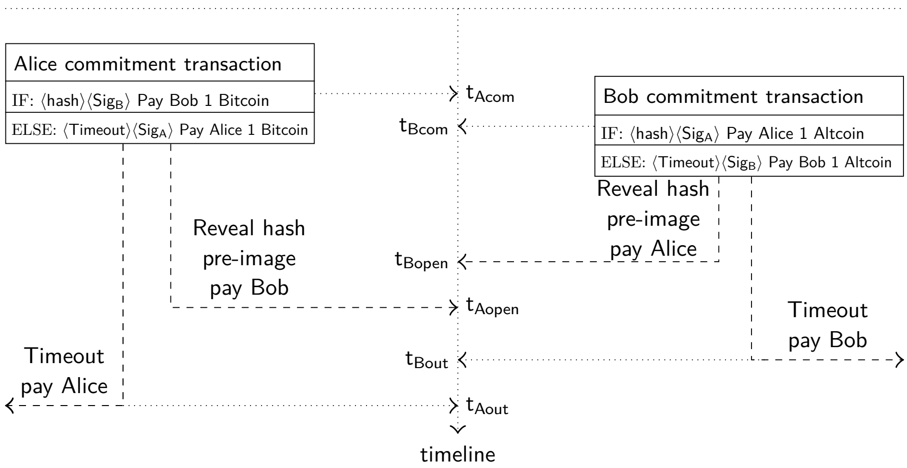

Figure 1 Hash time lock contract execution.

should be longer than Bob's commitment transaction, since in the case that Alice takes Bob's Altcoin at the last moment before Bob's commitment transaction timed out, her commitment is still locked by the time lock and Bob still has time to take Alice's Bitcoin. A successful execution of a hash time lock contract can be seen in figure 1. Notice that in the figure we use the same structure (script) to describe Altcoin and Bitcoin, but in fact we just consider two Bitcoin-like blockchains. As long as the blockchain supports both timelock and hash lock functionalities, the hash time lock contract protocol can be used.

The above description is the latest version of the hash time lock contract [6], where a time lock restricts when a transaction can be spent by its following transaction. Thus the two potential outputs of a commitment transaction are specified inside the commitment transaction. The previous version [5, 1] utilizes a time lock that only restricts when a transaction can be added to blockchain. The time lock is then not specified in the commitment transaction, but in the redeem transaction. In this case the redeem transaction must be signed by multiple signatures, thus the procedure involves two players exchanging signatures on transactions.

2.2 About Tamarin

Tamarin [8] is an automatic symbolic protocol verification tool. Given a protocol, the user specifies the roles running the protocol and their behaviors, the adversary model and the security properties by using the Tamarin programming language. Tamarin applies malicious adversarial behavior to the roles and uses a backward search method to generate counter-examples to the security claims. Tamarin ends up with either a proof that demonstrates that the given protocol satisfies the security properties, or Tamarin would give an attack for a failed security claim.

In Tamarin, the communication messages, fresh randomness and the states of the protocols are represented by symbolic terms called facts. There are two special facts to model the interaction with the untrusted environment: $\mathsf{In}(*)$, $\mathsf{Out}(*)$,representing the protocol's input and output from and to the environment. All the messages forwarded by $\mathsf{In}(*)$ and $\mathsf{Out}(*)$ can be learned by the adversary. The fact $\mathsf{K}(\mathsf{x})$ denotes the adversary learning x. Some facts are linear, which means that they can be used only once. The protocols and the specifications of the adversaries are modeled by using multiset rewriting rules. These rules and facts define a labeled transition system. Security properties are either defined in terms of traces of the transition system or the observational equivalence of two transition systems.

A role in the protocol is specified by Tamarin multiset rewriting rules. A rule consists of three elements: $(\mathsf{L}, \mathsf{A}, \mathsf{R})$:$[\mathsf{L}] - [\mathsf{A}] \rightarrow [\mathsf{R}]$, the left side facts L (states, messages of the protocol) are the premises of the rule, the right side facts R are rule conclusions, and the actions in the middle square brackets A are to label the traces. A rule can be executed as long as its premises exist in the current system states. Then the facts in the premises will be removed from the current system states, while the facts in conclusion will be added. Users can also add restrictions to enforce that only traces satisfying the restrictions are considered by Tamarin's backward search.

We illustrate Tamarin syntax by introducing a toy Diffie-Hellman key exchange protocol:

```
rule Server_1:
[ Fr(~a) ]-->[ S_1( ~a, 'g'^~a), Out( 'g'^~a) ]

rule Client:
[ Fr(~b), In( X ) ]--[ Key(X ^~b) ]->[ Out('g' ^~b ) ]

rule Server_2:
[ S_1( a, 'g'^a ), In( Y ) ]--[ Key( Y^a ) ]->[ ]
```

In the first step, the server generates fresh randomness ~a (the symbol ~ denotes a fresh nonce, the function Fr($*$) means generating a fresh nonce), sends g^a to client by the fact Out('g' ^~a), and it records the inner state by the fact S_1(~ a, 'g' ^~a) . This state will be used in next step of the server with the name Server_2.

The client receives the message from server by fact In(X), it then generates the session key according to the Diffie-Hellman key exchange protocol. This trace and its parameters are recorded by the action Key(X^~b), this will later be used to claim the security property of the protocol. The server's next step generates a similar action.

The security properties to be evaluated are defined by lemmas. In the above example we want to claim there is no adversary that can learn the secret key.

```
lemma Key_secrecy:
" All key #i . Key( key )@i ==> not Ex #j . K( key )@j "
```

The lemma Key_secrecy specifies that in all the traces that have an action Key(key), no adversary could learn the input of the action, namely, the value key, expressed by statement that there is no fact K(key) in the trace.

3 Tamarin Blockchain model

3.1 Simplification

A complete blockchain model would be too complex for Tamarin to work with, if it could even be expressed. We have simplified the structures of the transactions and blocks to make our blockchain model simpler, while still expressive enough to capture the essential elements for describing attacks on the protocols, and thus making verification possible. We let the blocks only include zero or one transactions, and forks are not allowed, thus we only consider the blocks that are already stable. The consensus protocol and cost are not modeled in our work. A transaction contains six elements: the id of the transaction that is being spent, the sender's address (we simply denote the addresses as public keys), the input signature (or script), output address (or script), the block sequence and the id of this transaction.

We set the relative growing speeds of the two blockchains to be the same. This simplification will not change the primary mechanism of the protocol because if the speed of Alice's blockchain is two time faster than Bob's blockchain, the time lock of Alice will be twice as long to ensure that it is longer than Bob's time lock in real time.

3.2 Tamarin blockchain model rules

We describe the rules of our blockchain model in two parts: the ledger rules and the global time rules. The ledger rules add a transaction to a block. The global time rules generate the time state called 'Tick' to specify the time point of a block being added to the blockchain.

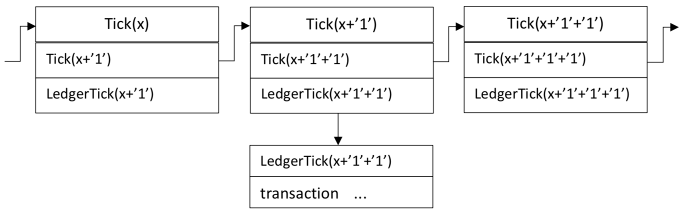

Figure 2 Tick chain.

In the global time rules, each time Tick has a unique parameter time. (It also has another parameter to tie a Tick to a specific blockchain, so that block chains can grow at different speeds. For simplicity we leave it out of Figure 2.) When generating a new Tick, an older Tick that has the largest time will be consumed and the time will be increased by one. Thus the Ticks form a time state transition chain that we call a Tickchain. Given the uniqueness of each Tick and since "time" is always increasing, each Tick can be considered as an empty block and the Tickchain can serve as base for a blockchain. We refer the blockchain in our Tamarin model as Tickchain and its blocks are called Tickblocks. In order to model adding a transaction to a certain Tickblock, a LedgerTick with a parameter Height equal to time is generated along with a new Tick. The ledger rules consume a LedgerTick to create a new transaction. In this way we bind a transaction to a Tickblock. The parameter Height also implies a sequence of transactions. After the executions of a protocol, there may be some LedgerTicks left without being consumed, which means that no transaction was added to the corresponding Tickblock.

Global time rules. There are two rules: Tick_start and Tick to create a blockchain. (We also use Tick to name the rule that generates a Tick state.) The rule Tick_start initiates the clock and the rule Tick updates the clock, i.e. increase the clock by adding '1'. There are three facts involved in the global time rules: $\mathsf{Chain}(\mathsf{BC})$, $\mathsf{Tick}(\mathsf{BC}, \mathsf{time})$ and $\mathsf{LedgerTick}(\mathsf{BC}, \mathsf{height})$. $\mathsf{Chain}(\mathsf{BC})$ specifies which blockchain. $\mathsf{Tick}(\mathsf{BC}, \mathsf{x})$ and $\mathsf{LedgerTick}(\mathsf{BC}, \mathsf{x})$ denote a certain block with the block height x. $\mathsf{Tick}(\mathsf{BC}, \mathsf{x})$ will be consumed by the Tick rules to updates the clock by iteration. $\mathsf{LedgerTick}(\mathsf{BC}, \mathsf{x})$ will be consumed by the ledger rules to link a transaction to a block. All these facts are linear facts that can be only consumed one time.

Global time rules

Tick_Start	Tick
Input: $\mathsf{Chain}(\mathsf{BC})$	Input: $\mathsf{Tick}(\mathsf{BC}, \mathsf{time})$
Output: $\mathsf{Tick}('1'), \mathsf{LedgerTick}('1')$	Output: $\mathsf{Tick}(\mathsf{BC}, \mathsf{time} +' 1')$, $\mathsf{LedgerTick}(\mathsf{BC}, \mathsf{time} +' 1')$

Ledger rules. The ledger rules model the nodes in blockchain network: the nodes get transaction information from the network, check its validity and then record the transaction to the blockchain.

There are two types of transaction in our model: SimpleTx(BC, InTx, InSig, OutPk, tx, height) to model the transactions without the hash and time lock, and CommitTx(BC, InTx, InSig, OutScript, tx, height) to model the transactions locked by a hash and a time lock. In these two transactions, BC denotes which blockchain the transaction belongs to; InTx is a nonce that identifies a previous unspent transaction owned by the sender; InSig is the sender's signature. tx is a nonce that identifies this transaction; height specifies in which block this transaction has been recorded. In the simple transaction, the OutPk is the receiver's address, while the OutScript in a commitment transaction is a hash time lock contract script, specifying the hash value, time lock value and receiver's address.

There are five rules to model the blockchain behaviors, Mine_Coin, Simple_Tx, Commit_Tx, Commit_open and Commit_timeout. The purpose of these rules are to generate blocks that contain different types of transactions and append the block to the blockchain. Mine_Coin creates the original coins of the blockchain. Simple_Tx spends a simple transaction and creates a new unspent simple transaction. Commit_Tx creates a transaction that is locked by a hash and a time lock. Commit_open models the transaction that is spent by revealing the hash pre-image. The Commit_timeout model the commit transaction that is spent by sender redeeming the transaction in the case of timeout.

Ledger rules

Mine_Coin

Input: Fr(∼ n), PK(A, pkA), ledgerTick(BC, t)
Output: SimpleTx(BC,′ 0′,′ 0′, pkA, n, t)

Simple_Tx

Input: SimpleTx(BC, InTx, InSig, Pk, n, height), In(⟨tx, Sig, pkB⟩), ledgerTik(BC, t)
Output: SimpleTx(BC, n, Sig, pkB, tx, t)

Commit_Tx

Input: SimpleTx(BC, InTx, InSig, Pk, n, height), In(⟨Sig, ⟨pkA, timelock, hash, pkB⟩⟩), LedgerTick(BC, t)
Output: Commit_Tx(BC, n, Sig, ⟨pkA, timelock, hash, pkB⟩, tx, t)

Commit_open

Input: Commit_Tx(BC, InTx, InSig, ⟨pkA, timelock, hash, pkB⟩, n, height), In(⟨⟨Script1, Script2⟩, PKaddress⟩), LedgerTick(BC, t),
Output: SimpleTx(BC, n, Sig, pkB, tx, t)

Commit_timeout

Input: Commit_Tx(BC, InTx, InSig, ⟨pkA, timelock, hash, pkB⟩, n, height) In(⟨Script1, PKaddress⟩) LedgerTik(BC, t)
Output: SimpleTx(BC, n, Script1, Pk, tx, t)

Restrictions. The no double spending property of blockchain is guaranteed by restriction rule. When a transaction has been spent, an action Spend(BC, tx, M, t) will be recorded in the Tamarin system to identify the event. On a single blockchain, for a transaction x, there can only exist one Spend(BC, x, M, t).

```
restriction DoubleSpending:
"All BC x n m t1 t2 #i #j .Spend(BC,x,n,t1)@i &Spend(BC,x,m,t2)@j==>#i=#j"
```

To help Tamarin reason more efficiently, we add one more restriction HappenBefore. This restriction simply tells Tamarin that a transaction that has a larger block number should happen later than a transaction that has a smaller block number.

```
restriction HappenBefore:
"All BC t1 t2 #i .HappenBefore(BC,t1,t2)@i==>Ex x .t2=t1+x"
```

4 Model HTLC in Tamarin

We model the two roles Alice and Bob in the hash time lock contract. Alice is the contract initiator, Bob is the responder. Alice is not allowed to set up a hash time lock contract with herself. The roles send data to blockchain network by using fact Out($*$), i.e. they send data to the environment directly. It models the blockchain network as public, where the adversary learns anything sent to and received from the network.

4.1 HTLC rules

Alices' rules. Alice is defined by two rules: Alice_send and Alice_receive. The rule Alice_send broadcasts Alice's commitment transaction and redeem transaction to the blockchain network. The rule Alice_receive broadcasts the transaction to open Bob's commitment transaction. Note that even though Alice broadcasts her commitment transaction and its redeem transaction at the same time, the redeem transaction cannot be added to the blockchain until the time lock of Alice's commitment transaction expires.

Alice_send

The rule takes a simple transaction tx, Alice's secret keys, Bob's address and a fresh nonce as input. It outputs Alice's commitment transaction, Alice's redeem transaction, and a state Alice_1_record to record the hash pre-image. It spends the simple transaction tx with signature SigA and outputs a commitment transaction that has ⟨pk(ltkA1), timelock_A, hash, pkB3⟩ as output. The two potential ways to spend this commitment transaction are: 1) Redeem by Alice: when the time lock timelock_A timed out, Alice could redeem the commitment transaction by providing the signature of the public key pk(ltkA1). 2) Opened by Bob: Bob can take the funding by providing the pre-image of the hash lock and the signature of pkB3. The rule outputs the redeem transaction at the same time, since Alice desires to broadcast the redeem transaction earlier so that it can be added on the blockchain as soon as the time lock expires. The Tamarin code is listed below.

```
rule Alice_send:
  let
    timelock_A='1'+'1'
    hash=HTLChash(~hsk)
    SigA=sign(<'BC1',tx,pk(ltkA),<pk(ltkA1),timelock_A,hash,pkB3>>,ltkA)
    CommitTxAlice=TXhash(<tx,SigA,<pk(ltkA1),timelock_A,hash,pkB3>>)
    SigA1=sign(<'BC1',CommitTxAlice,<pk(ltkA1),timelock_A,hash,pkB3>,pkA2>,ltkA1)
  in
    [ !SimpleTx('BC1','0','0',pk(ltkA),tx,t) ,!PK(A,pk(ltkA1)),!PK(A,pkA2),!PK(B,pkB3),
    Fr(~hsk)]
  --[ InEq(A,B) ]->
    [ Out(<tx,SigA,<pk(ltkA1),timelock_A,hash,pkB3>>),Out(<CommitTxAlice,SigA1,pkA2>)
      ,Alice_1_record(hash,~hsk)]
```

Alice_receive

The rule takes a commitment transaction that has Alice as receiver, a state record Alice_1_record and Alice's address as inputs. It outputs a transaction spending the commitment transaction and a fact Reveal(hsk). The rule opens the pre-image of hash and provide the signature SigA3 for pkA3. This spending transaction will be added to the blockchain by the ledger rule Commit_open, it transfers the funding to the target address.

```
rule Alice_receive:
  let
    SigA3=sign(<'BC2',CommitTxBob,<pkB1,timelock_B,hash,pk(ltkA3)>,pkA4>,ltkA3)
  in
    [!CommitTx('BC2',tx0,SigB0,<pkB1,timelock_B,hash,pk(ltkA3)>,CommitTxBob,t)
    ,Alice_1_record(hash,hsk),!PK(A,pkA4)]
  --[ Alice_receive(CommitTxBob) ]->
    [ Out(<CommitTxBob,<hsk,SigA3>,pkA4,hsk>)]
```

Bob's rules. Bob is specified by the rules Bob_send, Bob_receive and a restriction Not_Spend. The rule Bob_send generates the commitment transaction and its redeem transaction, and broadcasts them to the blockchain network. The rule Bob_receive is to open Alice's commitment transaction and the restriction Not_Spend checks that Alice's commitment transaction has not been spent.

Bob_send

The rule takes Alice's commitment transaction, a simple transaction, Alice's receiving address, Bob's redeem address and Bob's secret key as input. The restriction Not_Spend checks if Alice's commitment transaction is just added to the blockchain. If it is, the rule will output Bob's commitment transaction, its redeem transaction and Bob_1_record to record the hash lock and the transaction id of Alice's commitment transaction. The output script in Bob's commitment transaction is $\langle \mathsf{pk}(\mathsf{ltkB}), \mathsf{timelock\_B}, \mathsf{hash}, \mathsf{pkA3}\rangle$. The hash is the same with the hash lock in Alice's commitment transaction.

```
rule Bob_send:
  let
    timelock_B='1'
    SigB=sign(<'BC2',tx,pk(ltkB),<pk(ltkB1),timelock_B,hash,pkA3>>,ltkB)
    CommitTxBob=TXhash(<tx,SigB,<pk(ltkB1),timelock_B,hash,pkA3>>)
    SigB1=sign(<'BC2',CommitTxBob,<pk(ltkB1),timelock_B,hash,pkA3>,pkB2>,ltkB1)
  in
    [!SimpleTx('BC2','0','0',pk(ltkB),tx,t1),!PK(B,pk(ltkB1)) ,!PK(B,pkB2),!PK(A,pkA3)
    ,!CommitTx('BC1',tx_0,SigA_0,<pkA,timelock_A,hash,pkB>,CommitTxAlice,t)]
  --[Not_Spend(CommitTxAlice)]->
    [ Bob_1_record(hash,CommitTxAlice)
     ,Out(<tx,SigB,<pk(ltkB1),timelock_B,hash,pkA3>>)
     ,Out(<CommitTxBob,SigB1,pkB2>)]
```

Bob_receive

The rule takes a state record Bob_1_record, the hash pre-image In(hsk), Bob's address and Alice's commitment transaction as inputs. It outputs the transaction to open Alice's commitment transaction. By providing the signature of the public key pk(ltkB3) and the pre-image of the hash lock, Bob transfers the funding to his address pkB4.

```
rule Bob_receive:
  let
    SigB3=sign(<'BC1',CommitTxAlice,<pkA1,timelock_A,hash,pk(ltkB3)>,pkB4>,ltkB3)
  in
    [ Bob_1_record(hash,CommitTxAlice),In(hsk),!PK(B,pkB4)
    ,!CommitTx('BC1',tx0,SigA0,<pkA1,timelock_A,hash,pk(ltkB3)>,CommitTxAlice,t)]
  --[ Bob_receive(CommitTxAlice) ]->
    [ Out(<CommitTxAlice,<hsk,SigB3>,pkB4,hsk>) ]
```

5 Tamarin Security analysis

5.1 Preliminaries

We describe a transaction as a tuple of six elements: $\mathsf{TX}\{\mathsf{BC}, \mathsf{InTx}, \mathsf{InSig}, \mathsf{Output}, \mathsf{n}, \mathsf{height}\}$, where BC is the blockchain which this transaction belongs to, InTx is the ID of the input transaction, InSig is the input signature, n is the id of this transaction, and height specifies which block contains this transaction. For a simple transaction, the parameter Output is simply a public key, while for a commitment transaction, the Output will be a tuple $\langle \mathsf{pk}_1, \mathsf{timelock}, \mathsf{hash}, \mathsf{pk}_2\rangle$ that contains two public keys pk_1 and pk_2, a time lock and a hash lock. The commitment transaction can be spent by revealing the hash pre-image and the signature of pk_2 or providing the signature of pk_1 if the time lock timed out. The parameter height is ignored if a transaction is not recorded on the blockchain yet.

For a specific time lock, we denotes its value as Δ. It restricts a commitment transaction can be spent only if there is at least Δ blocks appended after the block that contains this commitment transaction. We specify the corresponding real time duration of generating these Δ blocks as δ. The relationship is typically simple, for instance, in the Bitcoin blockchain the approximate time to generate 20 blocks is 200 minutes, so with timelock $\Delta = 20$ we get real time $\delta = 200$. The reason why the real time also involved in the formula is that we are dealing with two blockchains. Each of the two blockchains can be seen as a time reference, but these two time references might get out of sync, thus we need a single global clock.

When a commitment transaction is added on the blockchain, we denote the event as $\{\Gamma^{\mathsf{h_{sk}},\Delta_A}_{\mathsf{Acom}}, \mathsf{t}, \mathsf{Tick}\}$, which means Alice's commitment transaction is recorded on blockchain at time point t in block sequence Tick, locked by Δ_A and a hash with pre-image $\mathsf{h_{sk}}$. The open and timeout of the commitment transaction are specified as $\{\Gamma^{\mathsf{h_{sk}},\Delta}_{\mathsf{Aopen}}, \mathsf{t}, \mathsf{Tick}\}$ and $\{\Gamma^{\mathsf{h_{sk}},\Delta}_{\mathsf{Ared}}, \mathsf{t}, \mathsf{Tick}\}$, respectively.

5.2 Security claim

For Alice, the hash time lock contract should satisfy the first two properties. For Bob, the protocol should guarantee the last two security properties:

Property 1. Bob cannot open Alice's commitment transaction and take her funding unless Bob has created a commitment transaction to Alice.

$$\forall\{\Gamma^{\mathsf{h_{sk}},\Delta}_{\mathsf{Aopen}}, \mathsf{t_{Aopen}}, \mathsf{Tick_{Aopen}}\} ==> \exists\{\Gamma^{\mathsf{h_{sk}},\Delta}_{\mathsf{Bcom}}, \mathsf{t_{Bcom}}, \mathsf{Tick_{Bcom}}\}$$

The equation claims that for all the events that Alice's commitment transactions have been opened, there must exist an event that Bob made a commitment transaction before. The commitment transaction made by Bob should use the same hash lock generated from $\mathsf{h_{sk}}$ and it sends funding to Alice's address.

```
lemma Security_1_Alice:
" All A tx1 SigA pkA1 timelock_A hash pkB3 CommitTxAlice
      TickAcom TickAopen #tAcom #tAopen #Apk1 .


    !PK(A,pkA1)@Apk1
    &!CommitTx('BC1',tx1,SigA,<pkA1,timelock_A,hash,pkB3>,CommitTxAlice,TickAcom)@tAcom
    &Spend('BC1',CommitTxAlice,'CommitOpen',TickAopen)@tAopen



  ==>Ex tx2 SigB pkB1 timelock_B pkA3 CommitTxBob
        TickBcom #tBcom #Apk2 .
```

```
        !PK(A,pkA3)@Apk2
        &!CommitTx('BC2',tx2,SigB,<pkB1,timelock_B,hash,pkA3>,CommitTxBob,TickBcom)@tBcom
        &#tBcom<#tAopen
"
```

Tamarin verifies the first security claim is true.

Property 2. Bob can redeem his funding only if the time lock of his commitment transaction timed out.

$$\forall(\{\Gamma^{h,\Delta_B}_{Bcom}, t_{Bcom}, Tick_{Bcom}\} \wedge \{\Gamma^{h,\Delta_A}_{Bred}, t_{Bred}, Tick_{Bred}\}) ==> t_{Bred} > t_{Bcom} + \delta_B$$

Since in a single blockchain, a transaction recorded early has a smaller height than those recorded later. We reduce this security property to:

$$\forall(\{\Gamma^{h,\Delta_B}_{Bcom}, t_{Bcom}, Tick_{Bcom}\} \wedge \{\Gamma^{h,\Delta_A}_{Bred}, t_{Bred}, Tick_{Bred}\}) ==> Tick_{Bred} > Tick_{Bcom} + \Delta_B$$

The equation claims that the duration between the time point Bob's commitment transaction is added to the blockchain and the time point Bob's redeem transaction is added to the blockchain is always larger than the duration of its time lock. Bob cannot redeem his commitment transaction before it timed out. This property guarantees that there is no race condition between Bob's redeem transaction and the transaction of Alice to open Bob's commitment transaction.

```
lemma Security_2_Alice:
" All tx2 SigB pkB1 timelock_B hash pkA3 CommitTxBob TickBcom #tBcom
        TickBTout #tBTout .

       !CommitTx('BC2',tx2,SigB,<pkB1,timelock_B,hash,pkA3>,CommitTxBob,TickBcom)@tBcom
       &Spend('BC2',CommitTxBob,'CommitTout',TickBTout)@tBTout
 ==>Ex x. TickBTout=TickBcom+timelock_B+x
"
```

Tamarin verifies the above security claim is true.

Property 3. After Alice takes Bob's funding, Bob has time to take Alice's funding before Alice's commitment transaction time out.

$$\forall(\{\Gamma^{h,\Delta_A}_{Acom}, t_{Acom}, Tick_{Acom}\} \wedge \{\Gamma^{h,\Delta_B}_{Bopen}, t_{Bopen}, Tick_{Bopen}\} ==> t_{Acom} + \delta_A > t_{Bopen}$$

The equation claims that if Alice takes Bob's funding at the last moment before it timed out, Bob should always have some time left before Alice's commitment transaction is timed out. This property avoids the risk of the race condition between Bob opening Alice's commitment transaction and Alice redeems her commitment transaction.

```
lemma Security_3_Bob:
"
All CommitTxAlice hash timelock_A pkA1 tx1 SigA pkB3 TickAcom #tAcom
      CommitTxBob timelock_B pkB1 tx2 SigB pkA3 TickBcom #tBcom
        #tBopen1 #tATout1 #tBopen #tATout .

    !CommitTx('BC1',tx1,SigA,<pkA1,timelock_A,hash,pkB3>,CommitTxAlice,TickAcom)@tAcom
    &!CommitTx('BC2',tx2,SigB,<pkB1,timelock_B,hash,pkA3>,CommitTxBob,TickBcom)@tBcom


    &Spend('BC2',CommitTxBob,'CommitOpen',TickBcom+timelock_B)@tBopen1
    &LedgerTick('BC2',TickBcom+timelock_B)@tBopen

    &LedgerTick('BC1',TickAcom+timelock_A+'1')@tATout
    &Spend('BC1',CommitTxAlice,'CommitTout',TickAcom+timelock_A+'1')@tATout1
==> #tBopen<#tATout
"
```

Tamarin gives a counterexample to this security claim, because the growth speed of the blockchains may differ. We explain in detail in the next subsection.

Property 4. Alice could redeem her funding only if her commitment transaction timed out.

$$\forall(\{\Gamma^{h,\Delta_A}_{Acom}, t_{Acom}, Tick_{Acom}\} \wedge \{\Gamma^{h,\Delta_A}_{Ared}, t_{Ared}, Tick_{Ared}\}) ==> t_{Ared} > t_{Acom} + \delta_A$$

The equation removes the same race condition risk that Alice has as described in security property 2.

```
lemma Security_4_Bob:
"
All tx1 SigA pkA1 timelock_A hash pkB3 CommitTxAlice TickAcom #tAcom #tATout
      TickATout .

   !CommitTx('BC1',tx1,SigA,<pkA1,timelock_A,hash,pkB3>,CommitTxAlice,TickAcom)@tAcom
   &Spend('BC1',CommitTxAlice,'CommitTout',TickATout)@tATout
==>Ex x. TickATout=TickAcom+timelock_A+x
"
```

Tamarin verifies this security claim is true.

5.3 Discussion on property 3

The failure of property 3 claims that after Alice taking Bob's funding, there exists a case that Bob has no time to open Alice commitment transaction before it expires. Tamarin shows that this attack happens in the case that the blockchain on which Alice made a commit transaction grows faster than is expected. Thus Alice's commitment transaction expires earlier even before Bob's commitment transaction expires. Therefore Alice has the chance to redeem her funding and also take Bob's funding.

Therefore, we need to have a blockchain that not only has liveness and consistency but also keeps a stable growth speed for the block height. Based on the Tamarin result, we add an extra restriction to restrict the growing speed of the blockchain "BC2" to be at least as fast as "BC1", and then evaluate the security property again.

```
restriction stable_growing_blockchain:
"All height #i .Tick('BC1',height)@i==>Ex #j.Tick('BC2',height)@j"
```

Tamarin now proves that property 3 holds. Notice that in the real scenario we expected both two blockchains should have stable growing speed, but this condition is not necessary for HTLC. The result shows that as long as "BC" grows relatively no slower than "BC1", HTLC is secure. The reason is Alice holds the pre-image of the hash, she only needs to observe the height of "BC2" to take Bob's funding before its timelock expires, she doesn't need to worry Bob will take her funding since he doesn't know the hash pre-image. While for Bob, if he publishes his commitment transaction, he needs to make sure Alice cannot withdraw her funding earlier than Alice taking his funding.

6 Analysis of the old version of HTLC

The old version of the hash time lock contract was used when the time lock functionality could only constrain the time point that a certain transaction is allowed to be added to blockchain. In this case, the time lock is specified in the redeem transaction rather than the commitment transaction. To make an agreement for the time lock duration of the redeem transaction, the two players need to exchange their signatures on the redeem transaction. The multi-signatures are checked by the nodes before they add the transaction into a block. The signature exchanging procedure is done before the players publish their commitment transaction, otherwise, they might be unable to redeem their commitment transactions.

We claim the same four security properties from section 5 for the old version hash time lock contract. Tamarin verifies that the protocol satisfies the security claims given that Alice is allowed to only use a fixed duration of timelock in the contract. However, in reality, Alice might use the timelock with different durations. In this case, there is an attack that allows Alice to redeem her funding earlier than the time period that Bob has signed.

The attack is as follows: Alice will initiate two hash time lock contracts with Bob, these two hash time lock contracts are the same except the second one has longer time lock than the first one. She aborts the first one when she gets Bob's signature on her redeem transaction. Bob will also abort the contract since Alice doesn't publish her commitment transaction. Alice initiates the second contract with Bob, using the same hash lock, but longer time lock. (Bob could in principle notice that he has signed the same hash before, but this requires Bob to keep track of earlier contracts, which is impractical.) In this scenario, after both players publish their commitment transactions to blockchains, Alice can use the redeem transaction of the fist hash time lock contract to unlock her commitment transaction in the second hash time lock contract. Because the second redeem transaction has a shorter time lock, she can redeem the commitment transaction earlier than Bob's expectation.

When we enable different timelocks in our Tamarin model, Tamarin finds the attack and shows that security property 3 fails even with the synchronization between the two blockchain growth speeds.

7 Conclusion

In this paper, we give a formal model for blockchain in Tamarin. Using this model we give a formal verification for security of the hash time lock contract. The verification result from Tamarin shows that the security of HTLC is based on the security assumptions of the underlying blockchain, but also requires that the responder blockchain (the blockchain that Bob operates on) needs to grow at least as fast as the initiator's blockchain. This result demonstrates that our Tamarin blockchain model can be used to find security issues in blockchain-based protocols. We note that the verification process of Tamarin needs human guidance to some extent, which could be improved in future work. Also, the model can be improved to allow forks to be more comprehensive.

References

1. Marcin Andrychowicz, Stefan Dziembowski, Daniel Malinowski, and Łukasz Mazurek. Fair two-party computations via bitcoin deposits. In Rainer Böhme, Michael Brenner, Tyler Moore, and Matthew Smith, editors, *Financial Cryptography and Data Security*, pages 105–121, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.
2. Marcin Andrychowicz, Stefan Dziembowski, Daniel Malinowski, and Łukasz Mazurek. Modeling bitcoin contracts by timed automata. In Axel Legay and Marius Bozga, editors, *Formal Modeling and Analysis of Timed Systems*, pages 7–22, Cham, 2014. Springer International Publishing.
3. Iddo Bentov, Yan Ji, Fan Zhang, Lorenz Breidenbach, Philip Daian, and Ari Juels. Tesseract: Real-time cryptocurrency exchange using trusted hardware. In Lorenzo Cavallaro et al., editors, *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, CCS 2019*, pages 1521–1538. ACM, 2019. doi:10.1145/3319535.3363221.
4. Sergiu Bursuc and Steve Kremer. Contingent payments on a public ledger: models and reductions for automated verification. In *ESORICS 2019 - The 24th European Symposium on Research in Computer Security*, Luxembourg, Luxembourg, September 2019. URL: https://hal.archives-ouvertes.fr/hal-02269063.

5 Christian Decker and Roger Wattenhofer. A fast and scalable payment network with bitcoin duplex micropayment channels. In Andrzej Pelc and Alexander A. Schwarzmann, editors, *Stabilization, Safety, and Security of Distributed Systems*, pages 3–18, Cham, 2015. Springer International Publishing.

6 Thaddeus Dryja Joseph Poon. The Bitcoin Lightning Network: Scalable off-chain instant payments, 2016. URL: `https://lightning.network/lightning-network-paper.pdf`.

7 Gregory Maxwell. Zero knowledge contingent payments, 2011. URL: `https://en.bitcoin.it/wiki/ZeroKnowledgeContingentPayment`.

8 Simon Meier, Benedikt Schmidt, Cas Cremers, and David Basin. The tamarin prover for the symbolic analysis of security protocols. In *Proceedings of the 25th International Conference on Computer Aided Verification - Volume 8044*, CAV 2013, page 696–701, Berlin, Heidelberg, 2013. Springer-Verlag.

9 Tianyu Sun and Wensheng Yu. A formal verification framework for security issues of blockchain smart contracts. *Electronics*, 9:255, February 2020. `doi:10.3390/electronics9020255`.

10 Mathieu Turuani, Thomas Voegtlin, and Michael Rusinowitch. Automated verification of electrum wallet. In Jeremy Clark, Sarah Meiklejohn, Peter Y.A. Ryan, Dan Wallach, Michael Brenner, and Kurt Rohloff, editors, *Financial Cryptography and Data Security*, pages 27–42, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg.

Authenticated Data Structures as Functors in Isabelle/HOL

Andreas Lochbihler
Digital Asset, Zurich, Switzerland
`http://www.andreas-lochbihler.de`
andreas.lochbihler@digitalasset.com

Ognjen Marić
Digital Asset, Zurich, Switzerland
ognjen.maric@digitalasset.com

Abstract

Merkle trees are ubiquitous in blockchains and other distributed ledger technologies (DLTs). They guarantee that the involved systems are referring to the same binary tree, even if each of them knows only the cryptographic hash of the root. Inclusion proofs allow knowledgeable systems to share subtrees with other systems and the latter can verify the subtrees' authenticity. Often, blockchains and DLTs use data structures more complicated than binary trees; *authenticated data structures* generalize Merkle trees to such structures.

We show how to formally define and reason about authenticated data structures, their inclusion proofs, and operations thereon as datatypes in Isabelle/HOL. The construction lives in the symbolic model, i.e., we assume that no hash collisions occur. Our approach is modular and allows us to construct complicated trees from reusable building blocks, which we call Merkle functors. Merkle functors include sums, products, and function spaces and are closed under composition and least fixpoints. As a practical application, we model the hierarchical transactions of Canton, a practical interoperability protocol for distributed ledgers, as authenticated data structures. This is a first step towards formalizing the Canton protocol and verifying its integrity and security guarantees.

2012 ACM Subject Classification Theory of computation → Logic and verification; Theory of computation → Higher order logic; Theory of computation → Cryptographic primitives

Keywords and phrases Merkle tree, functor, distributed ledger, datatypes, higher-order logic

Digital Object Identifier 10.4230/OASIcs.FMBC.2020.6

Related Version An extended version is available at `https://www.canton.io/publications/iw2020.pdf`.

Supplementary Material The formalization is available in the Archive of Formal Proofs [18]: `http://isa-afp.org/entries/ADS_Functor.html`.

1 Introduction

Authenticated data structures (ADSs) allow systems to use succinct digests to ensure that they are referring to the same data structure, even if each system knows only a part of the data structure. The benefits are twofold. First, this saves storage and bandwidth: the systems can store only the structure's parts that are relevant for them, and transmit just digests, not the whole structure. Blockchains use ADSs for this reason, both in the core design and in various optimizations (e.g., Bitcoin's lightweight clients). Second, ADSs can keep parts of the structure confidential to the subset of the systems involved in processing the structure. For example, distributed ledger technology (DLT) promises to keep multiple organizations synchronized on their shared business data. Synchronization requires transactions, i.e., atomic changes to the shared state. Yet organizations often do not want to share their full state with all involved parties. Some DLT protocols such as the Canton interoperability

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler; Article No. 6; pp. 6:1–6:15
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

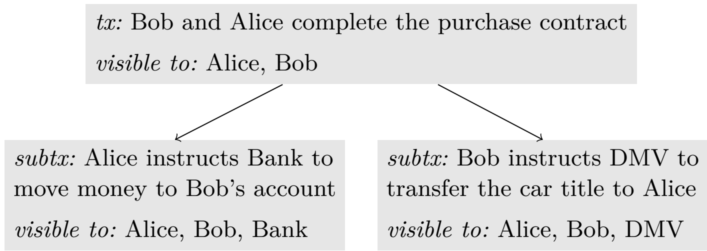

Figure 1 A hierarchical Canton transaction. DMV is the department of motor vehicles.

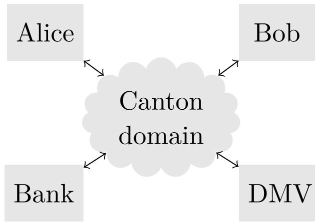

Figure 2 Example topology of a Canton-based distributed ledger.

protocol [7] and Corda [8] leverage ADSs to provide both transactions and varying levels of confidentiality. Formal reasoning about blockchains and DLTs thus often requires mechanised theories of ADSs. In fact, the formalization of Canton was the starting point for this work.

Merkle trees [20] are the prime example of an ADS. They are binary trees of digests, i.e., cryptographic hashes. Leaves contain data hashes, and inner nodes combine their children's hashes using a hash function h. An *inclusion proof*, also known as a Merkle proof, shows that a tree t includes a subtree st. It consists of the roots of t and st and the siblings of nodes on the path between these roots. The proof is valid if the hash of every node on the path is h of the children's hashes. It is sound, i.e., does prove inclusion, if h is collision-resistant. It keeps the rest of the tree confidential if h is preimage-resistant and the hashed data contains sufficient entropy.

ADSs [21] generalize these ideas to arbitrary finite tree data structures, whose hierarchies can conveniently encode more complex relationships between data. Our main example are the hierarchical transactions [4] in the Canton protocol. Suppose that Alice wants to sell a car title to Bob. Figure 1 shows the corresponding Canton transaction for exchanging the money and the title. (We take significant liberties in the presentation of Canton in this paper and focus on parts relevant for the construction of ADSs and for reasoning about them.) The transaction is generated from a smart contract (written in the DAML [10] programming language) implementing the purchase agreement.

The transactions' hierarchical nature benefits Canton in three crucial ways. First, complex transactions can be composed from simpler building blocks, which are transactions themselves. The purchase transaction above composes two such sub-transactions: the money transfer and the title transfer. Second, participants learn only the contents of subtransactions they are involved in. Above, the Bank only sees the money transfer, but not what Alice bought; similarly, the DMV does not learn the car's price. This also improves scalability, as everyone processes only the subtransactions they are involved in. Third, the hierarchy enables correct delegation in Canton's built-in authorization logic even in a Byzantine setting. Canton encodes this hierarchy, enriched with some additional data, in ADSs, and exchanges inclusion proofs for subtransactions. We give more details throughout the paper, but summarize the resulting requirements on the formalization here:

1. It must support ADS digests, to check that two inclusion proofs refer to the same ADS. This allows the example transaction to commit atomically, even if the Bank and the DMV see only a part of it.
2. Proofs must enable proving inclusion for multiple subtrees simultaneously, not just single subtree as standard. Canton uses such inclusion multi-proofs to save bandwidth.
3. Inclusion proofs refering to the same ADS must be mergeable into one multi-proof. In the example of Figure 1, Alice receives inclusion proofs for the entire transaction as well as both sub-transactions, and merges them to a single data structure, the entire transaction.

In this work, we show how to modularly define ADSs as datatypes in Isabelle/HOL. The modular approach is our main theoretical contribution. It allows us to construct complicated trees from small reusable building blocks, for which properties are easy to prove. To that end, we consider authenticated data structures as so-called *Merkle functors* and equip them with appropriate operations and their specifications. The class of Merkle functors includes sums, products, and function spaces, and is closed under composition and least fixpoints. Hence, the construction works for any inductive datatype (sums of products and exponentials). Concrete functors are defined as algebraic datatypes using Isabelle/HOL's datatype package [3]. This shallow embedding is a significant practical benefit, as it enables the use of Isabelle's rich reasoning infrastructure for datatypes. The construction lives in the symbolic model, i.e., we assume that no hash collisions occur. Finally, we show that the theory is suitable for constructing concrete real-world instances such as Canton's transaction trees. Our formalization is available in the Isabelle AFP [18].

The rest of the paper is structured as follows. In Section 2, we provide the background on Canton and use it to motivate our abstract interface for ADSs. Section 3 shows how to construct such interfaces for tree-like structures in a modular fashion. Section 4 demonstrates how to create inclusion proofs for general rose trees and Canton transactions in particular. We discuss the related work in Section 5 and conclude in Section 6.

2 Operations on Authenticated Data Structures

We now present the interfaces for ADSs, motivated by their application to Canton. Figure 2 shows a suitable Canton-based deployment for our example transaction. The participants transact using Canton, a distributed commit protocol similar to a two-phase commit protocol. The protocol is run over a Canton *domain* operated by a third party that acts as the commit coordinator. While the participants may be Byzantine, the domain is assumed to be honest-but-curious. That is, it is trusted to correctly execute the protocol, but it should not learn the contents of a transaction (e.g., how much Alice pays to Bob). Unlike in most other DLT solutions, participants share business data only on a need-to-know basis [6]. In particular, the domain receives business data only in encrypted form or as a digest. The domain may only learn the metadata that allows the protocol to tolerate Byzantine participants.

These privacy requirements motivate the hierarchical transactions that Canton uses, which are encoded in *transaction trees*. The tree for the example transaction from Figure 1 is shown in Figure 3. Each (sub)-transaction of Figure 1 is turned into a *view* in Figure 3, which consists of the view *data* and view *metadata*. For example, the node labeled by 1 in Figure 3 is the view corresponding to the top-level transaction in Figure 1. Its first two children contain the view's data and metadata. The metadata lists who is affected by the view and should therefore participate in the commit protocol (here, Alice and Bob), and is shared with Alice, Bob and the domain. The view data contains the confidential data with the actual state updates, and is shared only with Alice and Bob. This view also has two *subviews*, which correspond to the sub-transactions in Figure 1 as expected. Views can have an arbitrary number of subviews; e.g., the views labeled by 1.1 and 1.2 have no subviews.

Additionally, the two leaf children of the tree root store metadata describing transaction-wide parameters that apply to all views. The first is visible to the domain and the participants involved in the transaction; the second only to the latter. Formally, the transaction tree can be modelled by the following datatypes, for some types *common-metadata*, *participant-metadata*, *view-metadata*, and *view-data* whose contents are irrelevant for this paper.

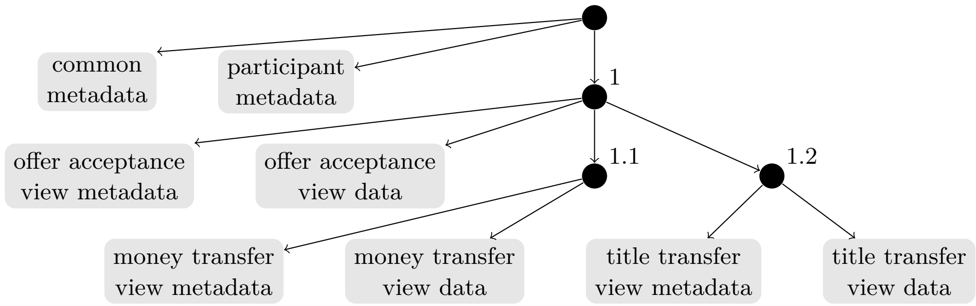

Figure 3 Simplified Canton transaction tree for car title sale of Figure 1.

datatype *view* = *View* ‹*view-metadata*› ‹*view-data*› (*subviews*: ‹*view list*›)
datatype *transaction* =
  Transaction ‹*common-metadata*› ‹*participant-metadata*› (*views*: ‹*view list*›)

In Figure 3, the *Transaction* and *View* constructors become the inner nodes (black circles) and the data sits at the leaves (grey rectangles).

The participants and the domain can use a root hash of an ADS over a *Transaction* to ensure that they are all referring to the same transaction tree. When constructing ADS hashes, we need to consider ADSs with multiple roots (i.e., forests) rather than just a single root like in a Merkle tree. For example, computing the hash of an inner node in a Merkle tree requires taking a hash over both of its children, i.e., over the forest constructed from its two children. The concrete hash operation depends on the shape of the forest (a pair in this case). The new root is again a degenerate forest of a single tree with a single root hash. This view underlies our modular construction principle in Section 3.

In this paper, we use the following Isabelle notations: Type variables $'a$, $'b$ are prefixed by $'$ like in Standard ML. Type constructors like *list* are usually written postfix as in *string list*. Exceptions are the function space $\Rightarrow$, sums $+$, and products $\times$, all written infix. The notation $t :: \tau$ denotes that the term t has the type τ. In our construction, we will use the following decorations. Raw data to be arranged in an ADS is written as usual, e.g., $'a$, $'a$ *list*. Hashes and forests of hashes carry a subscript $_h$ as in $'a_h$. We leave hashes for now abstract as type variables and define them only in Section 3. Since the root hash identifies an ADS, we represent ADSs by their hashes.

A root hash makes communication more efficient, but we require more. For example, the Bank does not know the contents or participants of view 1.2; the domain hides the latter. Still, the Bank must ensure that the view 1.1 is really included in the transaction tree. In general, the views visible to a participant are called the participant's *projection* of the transaction. Canton aims to achieve the following integrity guarantee [4]: There exists a shared ledger that adheres to the underlying DAML smart contracts such that its projection to each honest participant consists exactly of the updates that have passed the participant's local checks. This requires the ability to prove that a substructure is included in a root hash.

Inclusion proofs are therefore the main workhorse in our formalization and the focus of this paper. We denote the type of inclusion proofs over the source type with the subscript $_m$, e.g., $'a_m$, $('a_m, 'a_h)$ $tree_m$. We need two operations on inclusion proofs:

1. Computing the (forest of) root hashes of an inclusion proof, in order to identify the ADS to which the inclusion proof corresponds.
2. Merging two inclusion proofs with the same root hash.

Accordingly, we introduce two type synonyms for these operations:

type_synonym $('a_m, 'a_h)$ *hash* $= \langle 'a_m \Rightarrow 'a_h \rangle$
type_synonym $'a_m$ *merge* $= \langle 'a_m \Rightarrow 'a_m \Rightarrow 'a_m$ *option*$\rangle$

We model the merge operation as a partial function using the *option* that returns *None* iff the inclusion proofs have different (forests of) root hashes. We require that merging is idempotent, commutative, and associative. The merge operation makes inclusion proofs with the same hash into a semi-lattice, where the induced order treats an inclusion proof as smaller than another if it reveals less. In that case, we say that the smaller is a *blinding* of the larger inclusion proof.

type_synonym $'a_m$ *blinding-of* $= \langle 'a_m \Rightarrow 'a_m \Rightarrow bool \rangle$

▶ **Definition 1.** *A* Merkle interface *consists of three operations* $h :: ('a_m, 'a_h)$ *hash and* $m :: 'a_m$ *merge and* $bo :: 'a_m$ *blinding-of with the following properties:*

1. *Merge respects hashes, i.e.,* $(h\ a = h\ b) = (\exists\, ab.\ m\ a\ b = Some\ ab)$.
2. *Merge is idempotent, i.e.,* $m\ a\ a = Some\ a$.
3. *Merge is commutative, i.e.,* $m\ a\ b = m\ b\ a$.
4. *Merge is associative, i.e.,* $m\ a\ b \gg= m\ c = m\ b\ c \gg= m\ a$, *where* $(\gg=)$ *is the monadic bind on the option type.*
5. *Blinding is induced by merge, i.e.,* $bo\ a\ b = (m\ a\ b = Some\ b)$.

So merge is the least upper bound in the blinding relation:

$$(m\ a\ b = Some\ ab) = (bo\ a\ ab \wedge bo\ b\ ab \wedge (\forall\, u.\ bo\ a\ u \longrightarrow bo\ b\ u \longrightarrow bo\ ab\ u))$$

Also, the equivalence closure of the blinding relation gives the equivalence classes of the inclusion proofs under the hash function: *equivclp* bo = *vimage2p* $h\ h\ (=)$ where *equivclp* R denotes the equivalence closure of R and *vimage2p* $f\ g\ R = (\lambda x\ y.\ R\ (f\ x)\ (g\ y))$ the preimage of a relation under a pair of functions.

Isabelle/HOL's term language is not expressive enough to automatically create the ADS and inclusion proof types of arbitrary tree-shaped data, define the interface's operation, or build inclusion proofs for subtrees of tree-shaped data. Instead, in the next two sections, we show how to systematically construct these types and operations.

3 Modularly Constructing Forests of Authenticated Data Structures

In this section, we develop the theory to modularly construct ADSs, their inclusion proofs as HOL datatypes, and Merkle interfaces over them. We start with the concept of a blindable position (Section 3.1), which models an ADS node, and show how we obtain ADSs for Canton's transaction trees by introducing blindable positions in the right spots of the datatype definitions (Section 3.2).

We have shown how the Merkle interface specification is preserved by type composition (Section 3.3). It is, however, not inductive and therefore not preserved by datatype constructions. We thus generalize it and show that functor composition and least fixpoint preserve the generalization (Section 3.4). Finally, we show that sums, products and function spaces preserve the generalization (Section 3.5) and compose these preservation results to obtain the Merkle interface properties for Canton transactions (Section 3.6).

3.1 Blindable position

A *blindable position* represents a node (inner node or leaf) in an ADS. Recall that "blinding" allows an inclusion proof to hide the node contents by using just the root hash of the node. In this work, we model such hashes symbolically, that is, as injective functions, and assume that no hash collisions occur. We do not assume surjectivity though: some hashes do not correspond to any value. We model such values as garbage coming from a countable set (the naturals). This suffices as digests contain only a finite amount of information.

datatype $'a_h$ $\mathit{blindable}_h$ = *Content* $\langle 'a_h \rangle$ | *Garbage* $\langle \mathit{nat} \rangle$

Since the hash function is injective, we can identify the values $'a$ with a subset of the hashes, namely those of form *Content*. Accordingly, we could also have written $'a$ $\mathit{blindable}_h$ instead of $'a_h$ $\mathit{blindable}_h$. However, as an ADS contains hashes of hashes, $'a_h$ is more accurate here. For example, a degenerate Merkle tree with a single leaf, which stores some data x, has the root hash *Content* x.

What does an inclusion proof for this tree look like? It can take two forms. Either it reveals x, i.e., the leaf is not blinded, or it does not reveal x, i.e., the leaf is blinded. The following datatype formalizes these cases.

datatype $('a_m,\ 'a_h)$ $\mathit{blindable}_m$ = *Unblinded* $\langle 'a_m \rangle$ | *Blinded* $\langle 'a_h$ $\mathit{blindable}_h \rangle$

Similar to $\mathit{blindable}_h$, inclusion proofs may be nested, e.g., if a Merkle tree leaf contains another Merkle tree as data. We therefore use the inclusion proof type variable $'a_m$ instead of $'a$. In the second case, the hash could be garbage, so we use $'a_h$.

Note that our $\mathit{blindable}_h$ hashes are typed: hashes of those ADSs that store *int*s and those that store *string*s in their leaves always differ. In the real world, they can be equal as hashes are just bitstrings. However, for systems which follow security best practices, type flaw attacks lead to different hashes unless a hash collision occurs. Garbage hashes adequately model such confusion possibilities: a hash of the *int* Leaf would be treated as garbage in the type of hashes for the ADS of *string*s. This is adequate for inclusion proofs because we care about the contents of a hash only if the position is unblinded and thus of shape *Content*.

Having introduced the types for blindable positions, we now define the corresponding operations and show that they satisfy the specification *merkle-interface*. The hash operation *hash-blindable* :: $('a_m,\ 'a_h)$ *hash* $\Rightarrow$ $(('a_m,\ 'a_h)$ $\mathit{blindable}_m$, $'a_h$ $\mathit{blindable}_h)$ *hash* converts an inclusion proof into the root hash of the tree. It is parameterized by a hash function h_a that converts nested inclusion proofs $'a_m$ into their root hashes $'a_h$. Its definition is straightforward: for unblinded nodes, apply h_a, and for blinded nodes, just take the contained hash. Similarly, the blinding order *blinding-of-blindable* :: $('a_m,\ 'a_h)$ *hash* $\Rightarrow$ $'a_m$ *blinding-of* $\Rightarrow$ $('a_m,\ 'a_h)$ $\mathit{blindable}_m$ *blinding-of* is parametrized by the hash h_a and the blinding order bo_a for the nested inclusion proofs, as well as the blindable inclusion proofs to be compared. If both of the compared inclusion proofs unblind the contents, then we compare the contents using bo_a. Otherwise, the first argument is a blinding of the second one only if it is blinded, and if its hash matches the hash of the second argument. Merging of blindable positions is also similar. If both positions are unblinded, *merge-blindable* tries to merge the contents. If both are blinded, it succeeds iff the hashes are the same. Otherwise, it checks that the hashes are the same and, if so, returns the unblinded version. It is straightforward to show the following lemma.

▶ **Lemma 2.** *If h_a, bo_a, and m_a jointly form a Merkle interface, then so do hash-blindable h_a, blinding-of-blindable h_a bo_a, and merge-blindable h_a m_a.*

3.2 Example: Canton transaction trees

We now illustrate how to use $blindable_h$ and $blindable_m$ to define the ADSs and inclusion proofs for the Canton transaction trees from Section 2. As shown in Figure 3, the transaction tree contains a node for the transaction tree as a whole, every view, and every leaf (*common-metadata*, *participant-metadata* *view-metadata*, and *view-data*). Yet, the datatype declarations do not contain the information what should become a separate node in the ADS. To make the construction systematic, we start from an isomorphic representation of *view* and *transaction*, where we mark the blindable positions with the type constructor *blindable*, which is just the identity functor:

datatype *view* = *View*
‹((*view-metadata blindable* × *view-data blindable*) × *view list*) *blindable*›
datatype *transaction* = *Transaction*
‹((*common-metadata blindable* × *participant-metadata blindable*) × *view list*) *blindable*›

To define the hashes and inclusion proofs, we simply replace each type constructor τ with its counterparts τ_h and τ_m. For views, this looks as follows. Here $\times_h$, $\times_m$, $list_h$, and $list_m$ are type synonyms for $\times$ and *list*; Section 3.5 introduces them formally. We abuse notation by writing $view\text{-}metadata_h$ and $view\text{-}metadata_m$ for the blindable position of *view-metadata*.

type_synonym $view\text{-}metadata_h$ = ‹*view-metadata* $blindable_h$›
type_synonym $view\text{-}data_h$ = ‹*view-data* $blindable_h$›
datatype $view_h$ = $View_h$ ‹(($view\text{-}metadata_h \times_h view\text{-}data_h) \times_h view_h\ list_h$) $blindable_h$›
type_synonym $view\text{-}metadata_m$ = ‹(*view-metadata*, *view-metadata*) $blindable_m$›
type_synonym $view\text{-}data_m$ = ‹(*view-data*, *view-data*) $blindable_m$›
datatype $view_m$ = $View_m$
‹(($view\text{-}metadata_m \times_m view\text{-}data_m) \times_m view_m\ list_m$,
($view\text{-}metadata_h \times_h view\text{-}data_h) \times_h view_h\ list_h$) $blindable_m$›

These types nest hashes and inclusion proofs: A view node, e.g., nests hashes and inclusion proofs for the metadata, the data, and all the subviews. In particular, the $view_h$ and $view_m$ datatypes recurse through the $blindable_h$ and $blindable_m$ type constructors. This works because $blindable_h$ and $blindable_m$ are bounded natural functors (BNFs) [3]. In fact, this transformation works for any datatype declaration thanks to the compositionality of BNFs. The construction for transaction trees is similar.

3.3 Composition

Having defined the types of ADSs, we next must define the operations on ADSs and prove that they form a Merkle interface. Doing so directly is possible, but prohibitively complex. Instead, we modularize the proofs following the structure of the types. We can derive preservation lemmas for all involved type constructors analogous to *merkle-blindable*.

The preservation lemmas are compositional by construction: if $'a_h\ \tau_h/('a_m,\ 'a_h)\ \tau_m$ and $'b_h\ \sigma_h/('b_m,\ 'b_h)\ \sigma_m$ satisfy *merkle-interface*, then so does their composition $'a_h\ \tau_h\ \sigma_h/(('a_m,\ 'a_h)\ \tau_m,\ 'a_h\ \tau_h)\ \sigma_m$. For example, we can define the instance for blindable nodes of type *view-data* compositionally. First, we exploit the fact that every nullary functor satisfies *merkle-interface* with the discrete ordering (=), hash *id* and *merge* defined only for equal operands. Second, we compose *view-data*, viewed as a nullary functor with *blindable*. For example, we define:

abbreviation *hash-view-data* :: ‹($view\text{-}data_m$, $view\text{-}data_h$) *hash*› **where**
‹*hash-view-data* ≡ *hash-blindable id*›

We perform the same constructions on *view-metadata*, and then use composition for the pair *view-metadata* $\times$ *view-data*, to get the following (the operations for products will be introduced in Section 3.5).

▶ **Lemma 3.** *The following three operations form a Merkle interface:*
- *hash-prod hash-view-metadata hash-view-data*
- *blinding-of-prod blinding-of-view-metadata blinding-of-view-data*
- *merge-prod merge-view-metadata merge-view-data*

3.4 Inductive generalization for least fixpoints

The *view* datatype is the least fixpoint of the functor

$$'a\ F = ((\textit{view-metadata blindable} \times \textit{view-data blindable}) \times\ 'a\ \textit{list})\ \textit{blindable}$$

and so are $\textit{view}_h$ and $\textit{view}_m$ of analogous functors F_h and F_m. Composition gives us a preservation theorem for F, but we need another one for least fixpoints.

Yet, the Merkle interface specification is not inductive and thus not preserved by fixpoints. We now generalize it. Simultaneously, we make the generalization more amenable to Isabelle's proof automation by focusing on the blinding order and characterizing merge as its join. Our generalization splits the Merkle interface into three:

1. The interface *blinding-respects-hashes* assumes that $bo \leq \textit{vimage2p}\ h\ h\ (=)$ where $(\leq)$ denotes inclusion on binary predicates.
2. The interface *blinding-of-on* formalizes the order properties of the blinding relation bo: Reflexivity $bo\ x\ x$, transitivity $bo\ x\ y \Longrightarrow bo\ y\ z \Longrightarrow bo\ x\ z$, and antisymmetry $bo\ x\ y \Longrightarrow bo\ y\ x \Longrightarrow x = y$ hold for all $x \in A$ and all y, z: The restriction of x to the set A makes the statement inductive, as A can be instantiated to the set of smaller values in structural induction proofs.
3. The interface *merge-on* extends *blinding-of-on* applied to the type's universal set $\textit{UNIV}$ with the characterization of merge as the join, but now again restricted by a set A. In the unrestricted case $A = \textit{UNIV}$, *merge-on* is equivalent to the Merkle interface.

We are now ready to define the class of Merkle functors. For readability, we only spell out the case of unary functors. The generalization to n-ary functors is as expected.

▶ **Definition 4** (Merkle functor)**.** *A unary BNF F_h and binary BNF F_m constitute a unary Merkle functor if there exist operations:*
- $\textit{hash}'_F :: (('a_h,\ 'a_h)\ F_m,\ 'a_h\ F_h)\ \textit{hash}$ *and*
- $\textit{blinding-of}_F :: ('a_m,\ 'a_h)\ \textit{hash} \Rightarrow\ 'a_m\ \textit{blinding-of} \Rightarrow ('a_m,\ 'a_h)\ F_m\ \textit{blinding-of}$ *and*
- $\textit{merge}_F :: ('a_m,\ 'a_h)\ \textit{hash} \Rightarrow\ 'a_m\ \textit{merge} \Rightarrow ('a_m,\ 'a_h)\ F_m\ \textit{merge}$

with the following properties

$$\textit{Monotonicity}\qquad \frac{bo \leq bo'}{\textit{blinding-of}_F\ h\ bo \leq \textit{blinding-of}_F\ h\ bo'}$$

$$\textit{Congruence}\qquad \frac{\forall\, a{\in}A.\ \forall\, b.\ m\ a\ b = m'\ a\ b}{\forall\, x \in \{y.\ \textit{set}_1\text{-}F_m\ y \subseteq A\}.\ \forall\, b.\ \textit{merge}_F\ h\ m\ x\ y = \textit{merge}_F\ h\ m'\ x\ y}$$

Hashes $$\frac{\textit{blinding-respects-hashes } h \ bo}{\textit{blinding-respects-hashes } (\textit{hash}_F \ h) \ (\textit{blinding-of}_F \ h \ bo)}$$

Blinding order $$\frac{\textit{blinding-of-on } A \ h \ bo}{\textit{blinding-of-on } \{x. \ \textit{set}_1\textit{-}F_m \ x \subseteq A\} \ (\textit{hash}_F \ h) \ (\textit{blinding-of}_F \ h \ bo)}$$

Merge $$\frac{\textit{merge-on } A \ h \ bo \ m}{\textit{merge-on } \{x. \ \textit{set}_1\textit{-}F_m \ x \subseteq A\} \ (\textit{hash}_F \ h) \ (\textit{blinding-of}_F \ h \ bo) \ (\textit{merge}_F \ h \ m)}$$

where $\textit{hash}_F \ h = \textit{hash}'_F \circ \textit{map-}F_m \ h \ \textit{id}$ *for the BNF mapper* $\textit{map-}F_m$*, and where the BNF setter* $\textit{set}_1\textit{-}F_m \ x$ *returns all atoms of type* $'a_m$ *in* $x :: ('a_m, 'a_h) \ F_m$*.*

Every Merkle functor preserves the Merkle interface specification: set $A = \textit{UNIV}$ in the merge property and use the equivalence between the Merkle interface and *merge-on*. With this, we now state the main theoretical contribution of this paper.

▶ **Theorem 5.** *Merkle functors of arbitrary arity are closed under composition and least fixpoints.*

Proof. (Sketch) Closure under composition is obvious from the shape of the properties and the fact that BNFs are closed under composition. For closure under least fixpoints, we consider a functor F and its least fixpoint T through one of F's arguments. say **datatype** $T = T \ \langle T \ F \rangle$, and similarly for T_h and T_m. The operations are defined as follows, where we omit all Merkle operation parameters for type parameters that are not affected.

- The hash operation $\textit{hash-}T'$ is defined by primitive recursion:

  $$\textit{hash-}T' \ (T_m \ x) = T_h \ (\textit{hash-}F' \ (\textit{map-}F_m \ \textit{hash-}T' \ x)).$$

- The blinding order $\textit{blinding-of-}T$ is defined inductively by the following rule:

  $$\frac{\textit{blinding-of-}F \ \textit{hash-}T \ \textit{blinding-of-}T \ x \ y}{\textit{blinding-of-}T \ (T_m \ x) \ (T_m \ y)}$$

  Monotonicity ensures that $\textit{blinding-of-}T$ is well-defined.
- Merge $\textit{merge-}T$ is defined by well-founded recursion over the subterm relation on T_m:

  $$\textit{merge-}T \ (T_m \ x) \ (T_m \ y) = \textit{map-option} \ T_m \ (\textit{merge-}F \ \textit{hash-}T \ \textit{merge-}T \ x \ y)$$

  Congruence ensures that $\textit{merge-}F$ calls $\textit{merge-}T$ recursively only on smaller arguments.

Monotonicity and preservation of *blinding-respects-hashes* are proven by rule induction on $\textit{blinding-of-}T$. Congruence, *blinding-of-on*, and *merge-on* are shown by structural induction on the argument that is constrained by A. ◀

Isabelle/HOL lacks the abstraction over type constructors necessary to formalize this theorem. As our approach also translates to theorem provers with more expressive type systems (e.g., Lean, Coq), the theorem could be formalized there. For Isabelle/HOL, we adopt an approach similar to Blanchette et al. [3]. We axiomatize a binary Merkle functor and carry out the construction and proofs for least fixpoints and composition, illustrating how the definition and proofs generalize to functors with several type arguments. The example ADS constructions in Section 3.6 then merely adapt these proofs to the concrete functors at hand.

3.5 Concrete Merkle functors

We now present concrete Merkle functors. They show that the class of Merkle functors is sufficiently large to be of interest. In particular, it contains all inductive datatypes (least fixpoints of sums of products). We have formalized all of the following.

- The discrete functor from Section 3.3 with hash operation *id* and the discrete blinding order (=) is a nullary Merkle functor.
- Blindable positions $\mathit{blindable}_h$ and $\mathit{blindable}_m$ are a unary Merkle functor.
- Sums and products are binary Merkle functors. We set $\times_h = \times_m = \times$ and $+_h = +_m = +$. The hash operations *hash-prod* and *hash-sum* are the mappers *map-prod* and *map-sum*, respectively. The blinding orders *blinding-of-prod* and *blinding-of-sum* are the relators *rel-prod* and *rel-sum*. The merge operation *merge-of-prod* attempts to merge each component separately, while *merge-of-sum* can only merge left and left, or right and right values. (Formally, $\times_m$ and $+_m$ should take four type arguments. However, as sums and products do not themselves contain blindable positions, the type arguments $'a_h$ and $'b_h$ are ignored in inclusion proofs and we therefore omit them.)
- The function space $'a \Rightarrow 'b$ is a unary Merkle functor in the codomain. Like for sums and products, $\Rightarrow_h = \Rightarrow_m = \Rightarrow$ and no additional type arguments are added. Hashing is function composition and the blinding order is pointwise.

3.6 Case study: Merkle rose trees and Canton's transactions

Theorem 5 shows that all datatypes built from the Merkle functors in the previous section are Merkle functors. We apply the construction sketched in the proof to concrete datatypes that build on top of each other. For example, lists, rose trees [24], and Canton transactions are all Merkle functors. We prove that $'a$ *list* is a Merkle functor with the help of an isomorphic data type that is the least fixpoint $\mu X.\ 1 + 'a \times X$ and following the fixpoint construction of Theorem 5. We transfer the definitions and theorems to *list* using the transfer package [16]. Rose trees are then given by the datatype

datatype $'a$ *rose-tree* = *Tree* ‹($'a \times 'a$ *rose-tree list*) *blindable*›

Applying the construction gives us Merkle rose trees with the corresponding operations and their properties.

datatype $'a_h$ $\mathit{rose\text{-}tree}_h = \mathit{Tree}_h$ ‹($'a_h \times_h 'a_h$ $\mathit{rose\text{-}tree}_h$ list_h) $\mathit{blindable}_h$›
datatype $('a_m, 'a_h)$ $\mathit{rose\text{-}tree}_m = \mathit{Tree}_m$
‹($'a_m \times_m ('a_m, 'a_h)$ $\mathit{rose\text{-}tree}_m$ list_m, $'a_h \times_h 'a_h$ $\mathit{rose\text{-}tree}_h$ list_h) $\mathit{blindable}_m$›

From here, it is only a small step to transactions in Canton. Views are isomorphic to Merkle rose trees where the data at the nodes is instantiated, i.e., composed, with the Merkle functor corresponding to *view-metadata blindable* × *view-data blindable*. Then, transactions compose the Merkle functor for *common-metadata blindable* × *participant-metadata blindable* × - *list* with views. We have lifted our machinery from these raw Merkle functors to the datatypes view_m and $\mathit{transaction}_m$ using the lifting and transfer packages [16].

4 Creating Inclusion Proofs

So far, given a tree-like data type $'t$, we showed how to systematically construct the corresponding type of ADSs $'t_h$ and their inclusion proofs $'t_m$. To make use of this construction in practice, we must also be able to create values of type $'t_m$ from values of type $'t$. As in the case of our composition and fixpoint theorem, HOL's lack of abstraction over type constructors makes it impossible to express this process in HOL in its full generality. Instead, we sketch how it works on rose trees, as these are the most general type of tree in terms of branching. The construction can be easily adapted for other kinds of trees.

There are three basic operations:

- Digesting, *hash-source-tree*, returns the root hash for a rose tree.
- Embedding, *embed-source-tree* returns the inclusion proof that proves inclusion of the whole tree.
- Fully blinding, *blind-source-tree* returns the inclusion proof that proves no inclusion at all (the root is blinded).

Digesting and fully blinding conceptually do the same thing, but their return types ($'a_h$ $\mathit{rose\text{-}tree}_h$ and $('a_m,\ 'a_h)\ \mathit{rose\text{-}tree}_m$) differ. As rose trees are parameterized by their node label type, digesting, embedding, and fully blinding take parameters which digest, embed, or fully blind the node labels. The expected properties hold: the embedded and fully blinded versions of the same rose tree have the same hash, namely the digest of the rose tree, and the former is a blinding of the latter.

The more interesting operations concern creating an inclusion proof for a subtree of a tree. For example, with Canton's hierarchical transactions, we would like to prove that a subtransaction is really part of the entire transaction. Such a proof consists of the subtree itself, together with a path connecting the tree's root to the subtree's root. As noticed by Seefried [23], this corresponds to a zipper [15] focused on the subtree. This connection enables simple manipulation of such proofs in a functional programming style, well-suited to HOL. The zippers for rose trees are captured by the following types.

type_synonym $'a\ \mathit{path\text{-}elem} = \langle 'a \times 'a\ \mathit{rose\text{-}tree\ list} \times 'a\ \mathit{rose\text{-}tree\ list}\rangle$
type_synonym $'a\ \mathit{path} = \langle 'a\ \mathit{path\text{-}elem\ list}\rangle$
type_synonym $'a\ \mathit{zipper} = \langle 'a\ \mathit{path} \times 'a\ \mathit{rose\text{-}tree}\rangle$

Given a zipper that focuses on a node, we define the operations that turn rose trees into zippers and vice versa.

$$\mathit{tree\text{-}of\text{-}zipper}\ ([],\ t) = t$$
$$\mathit{tree\text{-}of\text{-}zipper}\ ((a,\ l,\ r) \cdot z,\ t) = \mathit{tree\text{-}of\text{-}zipper}\ (z,\ \mathit{Tree}\ (a,\ l\ @\ t \cdot r))$$

$$\mathit{zipper\text{-}of\text{-}tree}\ t \equiv ([],\ t)$$

The zippers for Merkle rose trees, i.e., inclusion proofs for rose trees, have the exact same shape, except that all the type constructors are subscripted by $_m$ and have another type parameter capturing the type of hashes (e.g., $('a_m,\ 'a_h)\ \mathit{zipper}_m$). Like for rose trees, we define operations that blind and embed a path respectively. This way, zippers on rose trees can be turned into zippers on Merkle rose trees. As expected, starting with a rose tree zipper, blinding and embedding its path yields a Merkle rose tree with the same hash. Furthermore, reconstructing a Merkle rose tree from an embedded rose tree zipper gives the same result as first reconstructing the rose tree and then embedding it into a Merkle rose tree. Finally, we show that reconstruction of trees from zippers respects the blinding relation if the Merkle operations on the labels satisfy *merkle-interface*:

$$\mathit{blinding\text{-}of\text{-}tree}\ h\ bo\ (\mathit{tree\text{-}of\text{-}zipper}_m\ (p,\ t))\ (\mathit{tree\text{-}of\text{-}zipper}_m\ (p,\ t')) =$$
$$\mathit{blinding\text{-}of\text{-}tree}\ h\ bo\ t\ t'$$

Inclusion proofs derived from zippers prove inclusion of a single subtree of the rose tree. The general case of several subtrees can be reduced to the single-subtree case using merging. When we want to create an inclusion proof for several subtrees, we create an inclusion proof for each individual subtree and then merge them into one.

To that end, we have defined operations to turn a rose tree into a zipper focused on the root and into zippers into its subtrees. Then, the function *zippers-rose-tree* enumerates the inclusion proof zippers for all nodes of a rose tree using those two operations. This allows us to easily model the messages that the initiator of a transaction sends in the first phase of Canton's commit protocol. The initiator constructs all zippers for the views in the transaction tree, and then turns each such zipper into an inclusion proof. Finally, the initiator merges each view proof with the proof from the zipper for the transaction metadata and ships it to the recipients.

At the end of the two-phase commit protocol, the domain's commit message contains an inclusion proof of the view metadata for all the views that the participant should have received. The participant can decide whether it has received all views it was supposed to receive, it compares this inclusion proof against the merged inclusion proofs that it had received from the initiator, using the inclusion proof order *blinding-of-transaction* on transactions.

5 Related Work

Miller et al. developed a lambda calculus with authentication primitives for generic tree structures [21]. The calculus was formalized in Isabelle/HOL by Brun and Traytel [5]. In the calculus, the programmer annotates the structures with authentication tags. Given a value of such a structure, and a function operating on it, their presented method automatically creates a correctness proof accompanying a result. The proof allows a verifier that holds only a digest of values with authentication tags (but not the values themselves) to check the function's result for correctness. The proof is a stream of inclusion proofs, one for each tagged value that the function operates on. Merging of inclusion proofs is not considered, although the streams can be optimized by sharing. Unlike Brun and Traytel [5] who use a deep embedding with the Nominal library, our embedding is shallow. Furthermore, our ADSs can provide inclusion proofs for multiple sub-structures simultaneously. However, we do not aim to derive generic correctness proofs for functions on the data structures.

Several other works formalize (binary) Merkle trees. White [25] formalized sparse Merkle trees [9] as part of a Coq model of a cryptographic ledger. An asset belongs to an address if the address encodes a path in the sparse Merkle tree from the root node to a leaf with the asset. A merge operation allows a single Merkle tree to provide several inclusion proofs. Our generic development can be instantiated to cover this structure. Yu et al. [26] use Merkle constructions on different binary trees to implement logs with inclusion and exclusion proofs. The constructions are proved correct using a pen-and-paper approach. The proved properties are then used in the Tamarin verification tool to analyze a security protocol. Ogawa et al [22] formalize binary Merkle trees as used in a timestamping protocol. They automatically verify parts of the protocol using the Mona theorem prover.

As part of the Everest project, HACL* contains a formal verification of balanced binary Merkle trees [13]. The balanced trees represent a sequence of hashes, which is padded with dummy values to a power of 2. A reduction proof shows that hash collisions between root hashes can be traced back to hash collisions of the underlying hash function. The main focus is on a refinement to an efficient executable implementation. It would be interesting to investigate whether and how their reduction-proof approach to dealing with hash collisions can be generalized compositionally to our general ADS setting.

Seefried [23] observed that inclusion proofs in a Merkle tree correspond to Huet-style zippers [15], where the subtrees in zipper context have been replaced by the Merkle root hashes. McBride showed that zippers represent one-hole contexts [19]. In this analogy, our inclusion multi-proofs correspond to contexts with arbitrarily many holes. These many-hole zippers must not be confused with Kiselyov's zippers [17] and Hinze and Jeuring's webs [14], which are derived from the traversal operation rather than the data structure.

6 Conclusion and Future Work

We have presented a modular construction principle for authenticated data structures over tree-shaped HOL datatypes (i.e., functors), and basic operations over these structures. The class of supported functors includes sums, products, and functions, and is closed under composition and least fixpoints. The supported operations are root hash computations and merging of inclusion proofs. We showed how to instantiate the construction to rose trees, as well as to real-world structures used in Canton, a Byzantine fault tolerant commit protocol.

The ongoing formalization of the Canton protocol will continue to test our abstractions and trigger further improvements. As noted earlier, ADSs not only improve storage efficiency, but also provide confidentiality. For example, Canton uses them to keep parts of a transaction confidential to a subset of the transaction's participants. However, reasoning about confidentiality is not straightforward. As hashing is injective, we can simply write *inv h* in HOL to invert hash functions. In fact, our current model does not even distinguish between the authenticated data structure and its digest because of this. A sound confidentiality analysis must therefore restrict the adversary using an appropriate calculus, e.g., a Dolev-Yao style deduction relation [11]. The analysis must take into account situations such as a Merkle tree node with two children with identical hashes; unblinding one child automatically unblinds the other. However, our representation distinguishes between the two, which might represent a problem. Another situation where this might be a problem is when merging inclusion proofs for commutative structures. One option is to consider Merkle functors as quotients with respect to a normalization function that collects all unblinding information and propagates the unblinding across the whole inclusion proof. The normalized inclusion proofs then serve as the canonical representatives. We have not yet worked out whether such a construction can still be modular and whether the quotients are still BNFs [12].

Moreover, our representation of hashes as terms makes hashing injective. While this is "morally equivalent" to standard cryptographic assumptions, an alternative (followed by [5]) would be to prove results about authentication as a disjunction: either the result holds, or a hash collision was found. The advantage of such a statement would be that hash collisions become explicit, which simplifies the soundness argument for the formalization. As is, nothing prevents us from conceptually "evaluating" the hash function on arbitrarily many inputs, which would not be cryptographically sound. To make hash collisions explicit, we must make hashes explicit, i.e., use a type like *bitstring*s instead of terms. We do not expect problems with extending our constructions to such a model, but it is unclear how severely the indirection through *bitstring*s impacts our proofs, in particular the Canton formalization.

We have based our construction on bounded natural functors (BNFs) as they are the semantic domain for datatypes in Isabelle/HOL and closed under least fixpoints. Fortunately, our Merkle constructions and proof need very little of the BNF structure and therefore generalize straightforwardly to other systems. For example, Lean's quotients of polynomial functors (QFPs) [1] are more general than BNFs and also closed under fixpoints. The concept of a Merkle functor can be directly expressed on QPFs as the BNF setter in Def. 4 can be replaced by the predicate lifting for QPFs. The closure proofs for composition and least fixpoint also work with predicate lifting. Moreover, the meta-theory can be formalized in Lean's more expressive type system, even for functors of arbitrary arity, and then instantiated for the concrete functor at hand. So in Lean, we would not have to redo the proof for every ADS. This also applies to other systems like Agda and Coq. Furthermore, the construction of concrete functors can be mimicked in any system that supports mutually recursive algebraic datatypes and higher-order functions, as all our ADS are built from sums, products, function spaces, and nested recursion through other datatypes, e.g., $blindable_h$ and $blindable_m$. (Nested datatype recursion can be reduced to mutual recursion [2], so mutually recursive algebraic datatypes suffice.)

References

1 Jeremy Avigad, Mario Carneiro, and Simon Hudon. Data types as quotients of polynomial functors. In John Harrison, John O'Leary, and Andrew Tolmach, editors, *ITP 2019*, volume 141 of *LIPIcs*, pages 6:1–6:19. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2019. `doi:10.4230/LIPIcs.ITP.2019.6`.

2 Stefan Berghofer and Markus Wenzel. Inductive datatypes in HOL – lessons learned in formal-logic engineering. In Y. Bertot, G. Dowek, A. Hirschowitz, C. Paulin, and L. Théry, editors, *TPHOLs 1999*, volume 1690 of *LNCS*, pages 19–36. Springer, 1999. `doi:10.1007/3-540-48256-3_3`.

3 Jasmin Christian Blanchette, Johannes Hölzl, Andreas Lochbihler, Lorenz Panny, Andrei Popescu, and Dmitriy Traytel. Truly modular (co)datatypes for Isabelle/HOL. In *Interactive Theorem Proving (ITP 2014)*, pages 93–110, 2014.

4 Sören Bleikertz, Andreas Lochbihler, Ognjen Marić, Simon Meier, Matthias Schmalz, and Ratko G. Veprek. A structured semantic domain for smart contracts. Computer Security Foundations poster session (CSF 2019), `https://www.canton.io/publications/csf2019-abstract.pdf`, 2019.

5 Matthias Brun and Dmitriy Traytel. Generic authenticated data structures, formally. In *Interactive Theorem Proving (ITP 2019)*, pages 10:1—10:18, 2019.

6 Canton: A private, scalable, and composable smart contract platform. `https://www.canton.io/publications/canton-whitepaper.pdf`, 2019.

7 Canton: Global synchronization beyond blockchain. `https://www.canton.io/`, 2020.

8 Corda: Open source blockchain platform for business. `https://www.corda.net/`, 2020.

9 Rasmus Dahlberg, Tobias Pulls, and Roel Peeters. Efficient sparse merkle trees: Caching strategies and secure (non-)membership proofs. Cryptology ePrint Archive, Report 2016/683, 2016. URL: `https://eprint.iacr.org/2016/683`.

10 Digital Asset. Daml programming language. `https://daml.com`, 2020.

11 D. Dolev and A. Yao. On the security of public key protocols. *IEEE Transactions on Information Theory*, 29(2):198–208, 1983.

12 Basil Fürer, Andreas Lochbihler, Joshua Schneider, and Dmitriy Traytel. Quotients of bounded natural functors. In N. Peltier and V. Sofronie-Stokkermans, editors, *Automated Reasoning (IJCAR 2020)*, volume 12167 of *LNAI*, pages 58–78. Springer, 2020. `doi:10.1007/978-3-030-51054-1_4`.

13 Hacl* verified merkle tree library. `https://github.com/project-everest/hacl-star/tree/master/secure_api/merkle_tree`, 2020.

14 Ralf Hinze and Johan Jeuring. Weaving a web. *J. Funct. Program.*, 11(6):681—689, 2001. `doi:10.1017/S0956796801004129`.

15 Gérard Huet. The zipper. *Journal of Functional Programming*, 7(5):549—554, 1997.

16 Brian Huffman and Ondřej Kunčar. Lifting and transfer: A modular design for quotients in Isabelle/HOL. In *Certified Programs and Proofs (CPP 2013)*, pages 131—-146, 2013.

17 Oleg Kiselyov. Zippers with several holes. `http://okmij.org/ftp/Haskell/Zipper2.lhs`, 2011.

18 Andreas Lochbihler and Ognjen Marić. Authenticated data structures as functors. *Archive of Formal Proofs*, April 2020. URL: `http://isa-afp.org/entries/ADS_Functor.html`.

19 Conor McBride. The derivative of a regular type is its type of one-hole contexts, 2001.

20 Ralph C. Merkle. A digital signature based on a conventional encryption function. In *Advances in Cryptology (CRYPTO 1987)*, pages 369–378, 1987.

21 Andrew Miller, Michael Hicks, Jonathan Katz, and Elaine Shi. Authenticated data structures, generically. In *Principles of Programming Languages (POPL 2014)*, pages 411—423, 2014.

22 Mizuhito Ogawa, Eiichi Horita, and Satoshi Ono. Proving properties of incremental Merkle trees. In *Automated Deduction (CADE 2005)*, pages 424–440, 2005.

23 Sean Seefried. Merkle proofs for free! Functional Programming Sydney, `http://code.ouroborus.net/fp-syd/past/2017/2017-04-Seefried-Merkle.pdf`, 2017.

24 D. B. Skillicorn. Parallel implementation of tree skeletons. *Journal of Parallel and Distributed Computing*, 39(2):115–125, 1996. `doi:10.1006/jpdc.1996.0160`.

25 Bill White. A theory for lightweight cryptocurrency ledgers. `https://github.com/input-output-hk/qeditas-ledgertheory`, 2015.

26 Jiangshan Yu, Vincent Cheval, and Mark Ryan. DTKI: a new formalized PKI with no trusted parties. *The Computer Journal*, 59(11):1695–1713, November 2016. arXiv: 1408.1023.

Mechanized Formal Model of Bitcoin's Blockchain Validation Procedures

Kristijan Rupić
Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
kristijan.rupic@fer.hr

Lovro Rožić
Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
lorozic33@gmail.com

Ante Derek
Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
ante.derek@fer.hr

Abstract

We present the first mechanized formal model of Bitcoin's transaction and blockchain data structures including the formalization of the blockchain validation procedures. Our formal model, though still a simplified representation of an actual Bitcoin blockchain, includes regular and coinbase transactions, segregated witnesses, relative and absolute locktime, the Bitcoin Script language expressions together with a denotational semantics, transaction fees and block rewards. We formally specify the details of validity checks performed when adding new blocks to the blockchain. We assume perfect cryptography and use the symbolic approach for modeling hash functions and digital signatures.

To demonstrate the utility of the model, we formally state and prove several essential properties of a valid blockchain – transactions are unique, each coin can be spent at most once and the new value is only created through block rewards. The model and the proofs are largely independent of Bitcoin specific details and easily generalize to any cryptocurrency blockchain based on the Unspent Transaction Output (UTXO) paradigm.

We mechanize all the results using the Coq proof assistant.

2012 ACM Subject Classification Theory of computation → Logic and verification

Keywords and phrases blockchain, Bitcoin, program verification, Coq

Digital Object Identifier 10.4230/OASIcs.FMBC.2020.7

Supplementary Material The Coq artifacts are available as as an open-source project at `https://github.com/krupic2402/bitcoin-blockchain-validation-formal-model`.

Funding *Ante Derek*: This research has been partially supported by the European Regional Development Fund under the grant KK.01.1.1.01.0009 (DATACROSS).

1 Introduction

In the past decade, due to the popularity of Bitcoin [18] and other cryptocurrencies, as well as new applications such as smart contracts [10], blockchain systems have attracted significant attention from the scientific community. The blockchain systems implement *distributed ledgers* where the data and transaction integrity is enforced using cryptography and consensus mechanisms.

Despite the openness of the Bitcoin system, serious design and implementation flaws have been discovered over the years. For example, a simple design flaw made it possible to include two different *coinbase* transactions with the same transaction identifier (TXID) into the blockchain [2]. The flaw was subsequently fixed in two *Bitcoin Improvement Proposals*: BIP 30 [2] made the older of the two transactions unspendable and included explicit checks

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler; Article No. 7; pp. 7:1–7:14
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

for uniqueness of TXID's, BIP 34 [3] mandated that coinbase transactions must include block height information, thereby fixing the design flaw. More recently (and more seriously), an implementation error in the transaction and block verification logic of the official Bitcoin client [5] made it possible for malicious miners to launch double-spending attacks.

In this paper, we build a formal model of Bitcoin's blockchain validation logic and we fully mechanize it using the Coq proof assistant [24]. We use the model to verify essential properties of a valid blockchain including the absence of both flaws described above.

We use the formal model of Bitcoin transactions by Atzei et al. [9] as the reference point for our formalization and mechanization efforts. We extend the simple "blockchain" model (i.e. a simple list of transactions) of [9] by adding an explicit *blockchain* data structure containing blocks of transactions linked by hash pointers. Our model also includes the complete treatment of coinbase transactions, the block height information as mandated by BIP34, transaction fees and block rewards. Finally, we model the blockchain validation procedures by formally specifying the sanity and validity checks performed by Bitcoin clients when adding new blocks; we define the blockchain to be *valid* if it passes the said validation procedures.

Contributions

Contributions of this paper are as follows:

1. We propose a fully mechanized model for Bitcoin transaction and the blockchain data structures. While simplified, the model includes many important details such as multi-signatures, segregated witnesses, absolute and relative locktimes, coinbase transactions, transaction fees and block rewards.
2. We define a denotational semantics for symbolic typed variant of Bitcoin Script language.
3. We define the sanity and validity checks performed by clients when adding new blocks to the blockchain.
4. We demonstrate the utility of the model by giving machine-verified proofs for three essential properties of a valid blockchain – same coin cannot be spent twice, transactions are unique, the total value of unspent coins is equal to the total value of block rewards.
5. We mechanize all the above results using the Coq proof assistant.

We make a number of simplifying assumptions in our work. First, we use the Dolev-Yao [15] model of cryptography where hash functions and digital signatures are abstract operations with perfect security properties. We simplify the Blockchain data structure by ignoring the Merkle trees that are normally used to include transactions and witnesses in block headers. Instead of a stack-based Script language and the corresponding execution model, we formalize the output scripts using an expression language with typed denotational semantics. Finally, many important aspects of the Bitcoin system such as the proof-of-work consensus mechanism, peer-to-peer network protocol, transaction and block discovery methods, etc. are out of scope of this work. Note that, there are efforts underway to mechanize those aspects of the Bitcoin system [22] – we view them as complementary to results presented in this paper.

We assume the reader is familiar with the Bitcoin system in general as well as the details of transaction and blockchain data structures including the notions of *inputs*, *outputs*, *witness* scripts and *coinbase transactions*. Due to space constraints, we defer details for the several aspects of the formal model (e.g., the semantics of the script language expressions) as well as proofs to the Coq artifacts.

$$\texttt{Satoshi}, \texttt{Index}, \texttt{Time} \triangleq \mathbb{N} \tag{1}$$
$$\texttt{PK}, \texttt{SK} \triangleq \mathbb{N} \tag{2}$$
$$is\_key\_pair : \texttt{PK} \to \texttt{SK} \to \texttt{bool} \tag{3}$$
$$\texttt{Modifier} \triangleq \{aa, an, as, sa, sn, ss\} \tag{4}$$

Figure 1 Basic definitions, key pairs and hash flags.

Outline

Section 2 presents our model of Bitcoin transactions formalized using the Coq proof assistant. In Section 3 we give the formal model of the blockchain data structure. In Section 4 we use the model to provide machine-verified proofs for the essential properties of a valid blockchain. In Section 5 we discuss the limitations of our model. We address related work in Section 6 and conclude in Section 7.

2 Formal Model of Bitcoin Transactions and Blockchain

We present a Coq model of the Bitcoin blockchain and the Bitcoin Script language. For now, we are primarily interested in transaction and blockchain validity.

Notation

For some type τ we use τ^* to denote the type of lists of elements of type τ. We denote the empty list as $[]$ and the singleton list containing some element x by $[x]$. We use '+' to denote list concatenation, $|\cdot|$ to denote list length, and $\in$ to denote list membership. Dot notation is used to denote access to individual members of structures. For example, we write $T.wit(i)$ to access the i-th index of the witness field of some transaction T. We will abbreviate $T.stub.inputs$ with $T.inputs$ (and similarly with other fields of transaction stubs). These notations might differ slightly from our Coq code but correspond to it in a one-to-one fashion.

2.1 The Transaction Model

We start out with a model of transactions and transaction histories, i.e., lists of transactions ordered by logical time. We model the Bitcoin Script language in order to provide an end-to-end model of transaction verification, although the proofs of various properties of our model could be made parametric with respect to a choice of the script language with relative ease, since their details tend to not affect higher-level properties.

As mentioned in the introduction, we use the symbolic approach when modeling cryptographic primitives. This allows us to simplify hashes of objects to only the objects themselves equipped with a decidable equality predicate, making the hash function essentially be the identity function which is injective and therefore also collision-resistant in a trivial way.

We begin by listing the basic definitions (Figure 1) which we will use throughout the rest of the formalization. Amounts of money (Satoshis, the name of the smallest Bitcoin denomination) and logical time in the system are both modeled as natural numbers for simplicity (1). Next, we define key pairs (2) for public-key digital signatures as trivial inductive types wrapping a value with decidable equality (in particular, a natural number) and we define a public and secret key to belong to the same pair if and only if they wrap equal values (3). We also define modifiers (4) corresponding to SIGHASH flags used in transaction signing [9].

Next, we need to define transactions (22). A regular transaction definition should consist of at least the following: a list of transaction inputs (16); a list of transaction outputs (17); a list of witness data associated with the inputs (24). Since we model SegWit [4], in our model we will distinguish between transactions and transactions paired with their respective witnesses depending on the context. The model of transactions also includes the absolute lock time (18) (nLockTime), which is a constraint on the earliest time the transaction can appear in a valid blockchain. While Bitcoin allows this to be either a block height or a UNIX timestamp depending on the range of the value [1], we only model some abstract, logical time. Extending the model to allow for block height or UNIX time should be fairly trivial. Unlike Atzei et al. [9], we also model coinbase transactions explicitly. They contain outputs but no inputs. These outputs represent the reward for mining of blocks and should be the sole supply of money in the system. They also contain their *block height*, i.e., the number of the block they are contained in in order to make them distinct as in BIP 34 [3].

Inputs (16) are references to outputs of other transactions, i.e., pairs of the referenced transaction and an index into its output list, along with a relative lock time which is another temporal constraint used in transaction verification. Unlike a Bitcoin implementation, this reference contains referenced transactions themselves instead of their hashes. Therefore, we require a decidable equality predicate on transactions, as well as an induction principle for its proof of correctness; we write an induction principle for transactions and their mutually inductive types manually. This is due to the fact that our inductive datatypes contain lists of the datatypes themselves – this creates an implicit mutual induction with lists which needs an induction principle more involved than ones Coq can automatically generate. This could have been avoided had we inlined lists (i.e. made our own datatypes using constructors analogous to *cons* and *nil*), however we would lose access to various existing theorems about lists contained in the standard library.

A transaction output (17) consists of its value in Satoshis and a script (5) for the verification of attempts to redeem the output. The Bitcoin Script language is a stack-based language that is used to write output scripts that verify that the conditions for redeeming the output are met. A script takes a fixed number of inputs which depends on the commands used; these inputs are called the witness and a redeeming transaction must provide them. Following Atzei et al. [9], we model the script language as an expression-based language instead as that allows us to easily specify denotational semantics for the scripts.

In a Bitcoin implementation all script values are just byte vectors at most 520 bytes long and their interpretation is made by the stack commands as either numbers, truth values, signatures, hashes etc. As we model hashes and signatures symbolically, we need our script input value type `StackValue` (9) to represent those possibilities as well, so we choose to impose a rudimentary type system on the values and their denotations that allows for integers (10), booleans (11), transaction signatures (13), and hashes of any type of value (12). As a transaction signature (26) is simply a wrapper for a secret key and a transaction "hash", a value will possibly contain transactions as well, making `StackValue` mutually inductive with transactions in our model (Figure 2).

The output script expression language is relatively simple. Most notable expression types are variables (6), constants of any `StackValue` (7), a multi-signature verification primitive (8) and several other arithmetic and comparison operations. We model it with an inductive type `Exp` (5) mutually inductive with `StackValue` and `TxStub` due to the fact that arbitrary

$$\texttt{Exp} : \texttt{Set} ::= \tag{5}$$
$$e\_var : \texttt{string} \to \texttt{Exp} \tag{6}$$
$$e\_const : \texttt{StackValue} \to \texttt{Exp} \tag{7}$$
$$e\_plus : \texttt{Exp} \to \texttt{Exp} \to \texttt{Exp}$$
$$e\_minus : \texttt{Exp} \to \texttt{Exp} \to \texttt{Exp}$$
$$e\_equal : \texttt{Exp} \to \texttt{Exp} \to \texttt{Exp}$$
$$e\_less : \texttt{Exp} \to \texttt{Exp} \to \texttt{Exp}$$
$$e\_if : \texttt{Exp} \to \texttt{Exp} \to \texttt{Exp} \to \texttt{Exp}$$
$$e\_length : \texttt{Exp} \to \texttt{Exp}$$
$$e\_hash : \texttt{Exp} \to \texttt{Exp}$$
$$e\_versig : \texttt{PK}^* \to \texttt{Exp}^* \to \texttt{Exp} \tag{8}$$
$$e\_abs\_after : \texttt{Time} \to \texttt{Exp} \to \texttt{Exp}$$
$$e\_rel\_after : \texttt{Time} \to \texttt{Exp} \to \texttt{Exp}$$

$$\texttt{StackValue} : \texttt{Set} ::= \tag{9}$$
$$sv\_int : \mathbb{Z} \to \texttt{StackValue} \tag{10}$$
$$sv\_bool : \texttt{bool} \to \texttt{StackValue} \tag{11}$$
$$sv\_hash : \texttt{StackValue} \to \texttt{StackValue} \tag{12}$$
$$sv\_sig : \texttt{TxStub} \to \texttt{SK} \to \texttt{Modifier} \to \texttt{Index} \to \texttt{StackValue} \tag{13}$$
$$\texttt{TxStub} : \texttt{Set} ::= \tag{14}$$
$$tx\_stub \ \{ \tag{15}$$
$$inputs : (\texttt{TxStub} \times \texttt{Index} \times \texttt{Time})^*; \tag{16}$$
$$outputs : (\texttt{Exp} \times \texttt{Satoshi})^*; \tag{17}$$
$$absLock : \texttt{Time} \ \} \tag{18}$$
$$coinbase \ \{ \tag{19}$$
$$block\_height : \mathbb{N} \ ; \tag{20}$$
$$outputs : (\texttt{Exp} \times \texttt{Satoshi})^* \ \} \tag{21}$$

$$\texttt{Tx} : \texttt{Set} ::= tx \ \{ \tag{22}$$
$$stub : \texttt{TxStub}; \tag{23}$$
$$witnesses : (\texttt{StackValue}^*)^* \ \} \tag{24}$$

Figure 2 Mutually inductive transaction, witness value and script definition.

`StackValue`s can be contained as constants in the expressions, which is made necessary by our imposed type system in order to meaningfully define arithmetic and comparison operations. The final result are three mutually inductive types (Figure 2) together with their mutual induction principle.

The witnesses (24) are data associated with each input. When verifying a redeeming attempt, they are used as the initial stack value in the output scripts of their associated inputs. Note that it is impossible to sign the witnesses along with the rest of transaction due to the fact that usually the witness data needs to contain the transaction signature itself. Not signing the witnesses implies that they can be changed before being included in a block, changing the hash of the transaction with witnesses included, a problem known as transaction malleability. This was resolved by the implementation of a protocol upgrade called SegWit (Segregated Witness) introduced by BIP141 [4]. We account for these subtleties in our model by separating the witnesses from input data in our model as well. In implementations of SegWit the witnesses are moved outside transaction data structures into their own Merkle tree stored in the containing block's coinbase transaction. To be able to talk about transaction history validity, we will sometimes have to associate transactions with their corresponding witnesses regardless of SegWit; to achieve this, we separate the transaction model into two layers of inductive types: the type `TxStub` (14) containing the transaction data save for the witnesses, and full transaction `Tx` (22) containing its stub and a list of witnesses (24). The transaction hash for input referencing purposes (TXID) is modeled by the the `TxStub` type.

$$\texttt{TxStubHash} ::= tx\_hash : \texttt{TxStub} \to \texttt{Modifier} \to \texttt{Index} \to \texttt{TxStubHash} \tag{25}$$
$$\texttt{Sig} ::= sig : \texttt{SK} \to \texttt{Modifier} \to \texttt{Index} \to \texttt{TxStub} \to \texttt{Sig} \tag{26}$$
$$ver : \texttt{PK} \to \texttt{Sig} \times \texttt{Modifier} \to \texttt{TxStub} \to \texttt{Index} \to \texttt{bool} \tag{27}$$
$$multi\_ver : \texttt{PK}^* \to (\texttt{Sig} \times \texttt{Modifier})^* \to \texttt{TxStub} \to \texttt{Index} \to \texttt{bool} \tag{28}$$

Figure 3 Signatures and routines for their verification.

2.2 Signature Verification and Output Redeeming

We now define our model of transaction signatures and their verification (Figure 3). A transaction signature is the SK-signed hash of a transaction with some fields disregarded in a way controlled by SIGHASH flags; in particular, some of the inputs are disregarded depending on the exact flags. We model hashes computed in this manner with the inductive type `TxStubHash` (25) wrapping the hashed transaction and the hashing flags, along with a comparison predicate which is based on transaction stub equality modulo hash flags. Signatures are represented by the inductive type `Sig` (26) wrapping everything a `TxStubHash` wraps, as well as the secret key. A signature needs to be paired with the hash flags used to compute it as they affect the result and are required for checking; this is implemented in Bitcoin by appending a byte denoting the hash flags to the signature. We model this explicitly by using `Sig` $\times$ `Modifier` even though we could introspect our inductive wrappers for their value.

We proceed to define single (27) and multiple (28) signature verification routines. We model successful signature verification with a public key using a simple check for pairedness of the given public key with the wrapped secret key with the function *is_key_pair* (3), and a check for hash equality by comparing both `TxStubHash` and the hash flags for equality; the verification succeeds if all comparisons do. Multiple signature verification tries to verify a list of signatures, in order, using an ordered list of public keys. The procedure repeatedly calls the single signature verification routine for each signature with successive public keys from the list until success, or until all public keys have been exhausted and no matching keys have been found; the whole routine succeeds if all signatures have been successfully verified and fails otherwise.

We define a straightforward denotational semantics for the script language based on Atzei et al. [9]. We impose a type system onto the values appearing in the script language (which are untyped in Bitcoin), consisting of the same types as `StackValue`, as well as a bottom type denoting failed computations or invalid types. We define the context of a witness $make\_context\ e\ T.wit(i)$ to be the mapping from variables ($free\_vars\ e$) to the values in the witness. The order of the variables is determined by a preorder traversal of the expression's syntax tree. The denotation of a script expression depends on the redeeming transaction, the index of the redeeming input and the context constructed from the corresponding witness. In the definition below, *den_bool* is a constructor for denotational values which wraps a boolean value. We refer the reader to the Coq development for details due to space constraints.

▶ **Definition 1** (Script verification). *We say a transaction T's i-th input verifies a script e if:*

$$verifies(T, i, e) \triangleq |free\_vars\ e| = |T.wit(i)| \wedge [\![e]\!]_{T,i,make\_context\ e\ T.wit(i)} = den\_bool\ \texttt{true}.$$

$$\texttt{TxHistory} \triangleq (\texttt{Tx} \times \texttt{Time})^* \quad (29)$$

$$UTXO : \texttt{TxHistory} \to (\texttt{TxStub} \times \texttt{Index})^* \quad (30)$$

$$STXO : \texttt{TxHistory} \to (\texttt{TxStub} \times \texttt{Index})^* \quad (31)$$

$$sum\_inputs : \texttt{TxStub} \to \texttt{Satoshi} \quad (32)$$

$$sum\_outputs : \texttt{TxStub} \to \texttt{Satoshi} \quad (33)$$

$$UTXO\_value : \texttt{TxHistory} \to \texttt{Satoshi} \quad (34)$$

$$coinbase\_value : \texttt{TxHistory} \to \texttt{Satoshi} \quad (35)$$

$$coinbase\_height : \texttt{TxHistory} \to \mathbb{N} \quad (36)$$

Figure 4 Transaction history model.

▶ **Definition 2** (Output redeeming). *We say the j-th input of transaction T_2 at logical time t_2* redeems *the i-th output of transaction T_1 at logical time t_1 for a value of v Satoshis if:*

$$\begin{aligned}
&redeems(T_1, i, t_1, v, T_2, j, t_2) \triangleq \\
(i) \quad &\exists\ relLock\ e, T_2.inputs(j) = (T_1, i, relLock) \ \wedge\ T_1.outputs(i) = (e, v) \ \wedge \\
(ii) \quad &T_2.absLock \le t_2 \ \wedge\ t_1 + relLock \le\ t_2 \ \wedge \\
(iii) \quad &verifies(T_2, j, e)
\end{aligned}$$

3 Blockchain Model and Validity

We begin our model of the Bitcoin blockchain by first considering transaction histories and their validity. We then define our model of the blockchain and its validity by requiring that the transaction history encoded by the blockchain be valid, among other things.

For Bitcoin to function as a currency, it is crucial to control the way in which money is created. Only coinbase transactions should increase the total sum of money in the system. However, if a transaction output was to be spent more than once, it would essentially act as duplicated money. Therefore, it is necessary to ensure that transaction outputs can be spent at most once. Transactions attempting to spend an already spent output, or spend an unspent output multiple times at once must be disallowed in a valid transaction history. We provide a formal definition of the transaction history validity predicate that enforces this and certain other conditions necessary for a history to be considered valid. We later prove that this property indeed implies that no double spending of transaction outputs is happening within a valid history, as well as that the total sum of unspent transaction outputs never exceeds supply, i.e., the sum of coinbase outputs.

We define a transaction history (29) as a list of transactions with witnesses and the logical time at which they occur. We also define the notions of spent and unspent transaction outputs; an output at index i of a transaction T_1 in the blockchain is *unspent* in a history TH if there is no transaction anywhere in TH that has an input (T_1, i), whereas an output is *spent* in TH if such a transaction and input exist. We define functions $STXO$ and $UTXO$ (31, 30) on histories that compute respectively the list of spent and unspent outputs, with outputs represented as pair of the containing transaction and the output's index. We also formally prove the obvious fact that every output of every transaction in a blockchain is either spent or unspent. We define the sum of values of inputs (32) and outputs (33) of a transaction, as well as the sum of values of all UTXO-s (34) and all coinbase outputs (35) in a transaction history which should represent the total supply of money in a transaction history following some validity rules which we will define. We define *coinbase_height* to be the number of coinbase transactions in a transaction history; note that this is going to be equal to the block height, but is formalized independently.

Using the work of Atzei et al. [9] as a reference point, we define *transaction history validity* (4) inductively by requiring that each valid transaction history is formed by a sequence of *valid updates* (3) each extending the history by a single transaction in a way that enforces the necessary invariants.

▶ **Definition 3** (Valid update for transaction histories).

$$
\begin{aligned}
&is\_valid\_update(TH, T, t) \triangleq \\
&(i) \quad \exists\ block\_height\ outputs,\ T.stub = coinbase\ block\_height\ outputs\ \land \\
&(ii) \quad \forall\ TH'\ T'\ t', TH = TH' + [(T', t')] \implies t' \leq t\ \land \\
&(iii) \quad T.block\_height = coinbase\_height\ TH \\
&\lor \\
&(iv) \quad sum\_inputs(T) \geq sum\_outputs(T)\ \land \\
&(v) \quad \forall\ TH'\ T'\ t', TH = TH' + [(T', t')] \implies t' \leq t\ \land \\
&(vi) \quad T.inputs \neq []\ \land \forall\ i\ j\ T'_i\ o_i\ r_i\ T'_j\ o_j\ r_j, i \neq j\ \land \\
&\quad T.inputs(i) = (T'_i, o_i, r_i)\ \land\ T.inputs(j) = (T'_j, o_j, r_j) \implies (T'_i, o_i) \neq (T'_j, o_j)\ \land \\
&(vii) \quad \forall\ j\ T'\ o\ r\ t'\ s\ v, (T', t') \in TH\ \land T.inputs(j) = (T', i, r)\ \land\ T'.outputs(i) = (s, v) \\
&\quad \implies (T', i) \in UTXO(TH)\ \land\ redeems(T', i, t', v, T, j, t)
\end{aligned}
$$

▶ **Definition 4** (Transaction history validity).

$$
\begin{aligned}
&tx\_history\_valid(TH) ::= \\
&\quad bc\_empty : TH = [] \rightarrow tx\_history\_valid(TH) \\
&\quad\quad bc\_cons : \forall\ TH'\ T\ t,\ TH = TH' + [(T, t)] \rightarrow tx\_history\_valid(TH') \\
&\quad\quad\quad \rightarrow valid\_update(TH', T, t') \rightarrow tx\_history\_valid(TH)
\end{aligned}
$$

Now we define a *blockchain* (Figure 5, 37) as an inductive type. Hash pointers to blocks are, as before, represented by the blocks themselves. As we do not deal with proof-of-work or consensus, the only contents of a block are the pointer to the previous block (40), the transactions (41) and witnesses (42) of the block, and the block's timestamp (43). Transactions and witnesses are both represented as lists instead of Merkle trees, but are separated according to SegWit. We also define *block_height* (44) to be the number of blocks in the blockchain, and *bc_to_tx_history* (45) to be a function that flattens a blockchain into the transaction history it represents by concatenating lists of transactions paired with their respective witnesses. We define the *block reward* (47), a function from block height of the block to be minted to the base value to include in the block's coinbase transaction; and $transaction\_fees$ (46) to be the sum of the differences between input and output value for each transaction in a list.

The definitions of valid updates of blockchains by blocks and valid blockchains are analogous to the definitions for transaction histories.

▶ **Definition 5** (Valid update for blockchains). *A blockchain B is validly updated with a new block containing $(transactions, witnesses, timestamp)$ when*

1. *transactions list contains exactly one coinbase transaction CB as the first transaction*
2. $CB.block\_height = block\_height\ B$
3. $sum\_outputs\ CB = block\_reward\ (block\_height\ B) + transaction\_fees\ transactions$
4. $tx\_history\_valid\ (bc\_to\_tx\_history\ (Block\ B\ transactions\ witnesses\ timestamp))$

$$\begin{aligned}
&\texttt{Blockchain} ::= && (37)\\
&\quad \mathit{Empty} && (38)\\
&\quad \mathit{Block}\ \{ && (39)\\
&\quad \mathit{prevBlock} : \texttt{Blockchain}; && (40)\\
&\quad \mathit{transactions} : \texttt{TxStub}^*; && (41)\\
&\quad \mathit{witnesses} : ((\texttt{StackValue}^*)^*)^*; && (42)\\
&\quad \mathit{timestamp} : \texttt{Time}\ \} && (43)
\end{aligned}$$

$$\begin{aligned}
&\mathit{block_height} : \texttt{Blockchain} \to \mathbb{N} && (44)\\
&\mathit{bc_to_tx_history} : \texttt{Blockchain} \to \texttt{TxHistory} && (45)\\
&\mathit{transaction_fees} : \texttt{TxStub}^* \to \texttt{Satoshi} && (46)\\
&\mathit{block_reward} : \mathbb{N} \to \texttt{Satoshi} && (47)
\end{aligned}$$

Figure 5 Blockchain model.

▶ **Definition 6** (Blockchain validity). *We define the validity of a blockchain inductively.*

- *An Empty blockchain is valid.*
- *A blockchain B with a block appended is valid whenever B was valid, the length of the block's transactions and witnesses lists is equal, and the appended block validly updates B.*

4 Formally Verified Blockchain Properties

With all the definitions in place, we move on to state several important properties of valid transaction histories and blockchains, which we have proven in our Coq development. Here we list only a part of the development due to space constraints; it can be seen in full along with the proofs in the accompanying materials.

First, we prove that blockchain validity implies the validity of the transaction history it stores. This follows directly from the definition of valid updates with blocks applied to the last block in the chain, if any. This result allows us to reason about valid transaction histories instead of blockchains, which can be more convenient e.g., when proving the impossibility of double spending in a valid blockchain (and transaction history).

▶ **Lemma 7** (Blockchain validity implies transaction history validity). *Let B be a valid blockchain. Then $bc\_to\_tx\_history$ B is a valid transaction history.*

Note that the definition of a transaction history does not order the transactions according to output spending. In a valid transaction history, however, every transaction input refers to an output of a transaction earlier in the history, which we proved as a lemma.

The first key property of the blockchain we consider is the impossibility of spending the same output multiple times.

▶ **Theorem 8** (No double spending). *Let B be a valid blockchain, and TH be its (valid) transaction history. Let T_i and T_j be two transactions in TH at indices i, j respectively. Then*

$$\forall k_i\ k_j,\ (T_i.inputs(k_i) = (T_i', l_i, t_{rl,i}) \wedge T_j.inputs(k_j) = (T_j', l_j, t_{rl,j}) \wedge (i, k_i) \neq (j, k_j)) \implies (T_i', l_i) \neq (T_j', l_j).$$

Another key property of valid transaction histories is that transactions identifiers are unique. While in reality we have to allow for the noninjectivity of hashes, in our model transactions are wholly unique within valid histories.

▶ **Theorem 9** (Transaction uniqueness). *Let B be a valid blockchain, and TH be its (valid) transaction history. Let txs be the list of* `TxStubs` *in the history (i.e., TH with timestamps and witnesses removed). Let T_i and T_j be transactions at indices i, j in txs, respectively. If $T_i = T_j$, then $i = j$.*

In the remainder of this section we consider properties of the total supply of money in the system. This should be equal to the sum of all coinbase output values, however it is also allowed to be smaller than that due to the presence of transaction fees.

▶ **Theorem 10** (Coinbase value bounds UTXO value above). *Let TH be a valid transaction history. Then*

$$UTXO\_value\ TH \leq coinbase\_value\ TH.$$

The following theorem illustrates the fact that only UTXO-s may be used as transaction inputs quantitatively. The proof follows from the definition of valid updates.

▶ **Theorem 11** (UTXO value bounds input value sum). *Let $TH + [(T, t)]$ be a valid transaction history. Then*

$$sum\_inputs\ (stub\ T) \leq UTXO\_value\ TH.$$

The final theorem is a strengthening of (10) shows that supply is exactly controlled by block rewards. It boils down to proving that transaction fees are properly collected in the coinbase transaction outputs of each block.

▶ **Theorem 12** (Total block reward equals UTXO value). *Let B be a valid blockchain, and $TH = bc\_to\_tx\_history\ B$. Then:*

$$UTXO\_value\ TH = \sum_{b=0}^{block\_height\ B-1} block\_reward\ b.$$

5 Limitations

Here we briefly discuss the limitations of our model and compare it to the Bitcoin client.

Since we use the symbolic model for digital signatures and hash functions, we are unable to prove the desired properties in the computational model of cryptography. Of course, we are also unable to extract the code for a verified client. We can overcome the latter by reusing Coq models of the cryptographic primitives (e.g. [7] for the SHA256 hash function) As for the former, since we are not concerned with the proof-of-work verification, we only rely on hash functions for data integrity and only need their collision resistance property. In the computational model, we could verify the properties under the assumption no collision occurred anywhere in that blockchain. For modeling properties of digital signatures in Coq we could attempt to use the toolset of the Foundational Cryptography Framework [21]. Alternatively, we could try to model the system within the universally composable (UC) security framework and replace the digital signature implementation with an ideal functionality similarly to the approach taken in [12] to develop mechanized analysis of a key exchange protocol.

The blockchain verification procedures are currently modeled mostly as first order inductive predicates rather than decidable routines and, hence, cannot be used to extract verified code. We plan to address this by writing the missing decision routines that generate proofs or disproofs of our propositions as well as routines that parse our model from serialized data, which would give us verified extractable blockchain validation code.

Comparison with the official Bitcoin client

First, we do not attempt to model several important aspects of the validation logic, since we do not consider them to be relevant to the correctness properties we wished to tackle first. Most notably, we omit proof-of-work verification and the corresponding data fields from the model. Transactions and blocks in our model do not have version numbers accounting for protocol updates. We do not enforce block and transaction size limits, coinbase maturity and we do not reject transactions with absurdly high fees.

We make a number of technical choices that result in a simpler formal model and diverge from the Bitcoin client. For example, we have explicit coinbase transactions while in the Bitcoin client coinbase transactions are stored in the same data structure and are distinguished by a single input field with the zero hash pointer. We feel that addressing these differences is a technical matter, albeit tedious and time consuming.

In our model, only transactions with segregated witnesses are supported, while the Bitcoin client additionally supports legacy transactions where the witness is a part of the transaction's input field. There are several other examples of extensions where both current and legacy features are supported. Moreover, these are almost always implemented in a backward-compatible manner. From a consensus perspective it is desirable that the blockchain verification procedures are updated by a *soft-fork* – old nodes *must* recognize the new blocks as valid. Hence, new features often need to be *hacked* into the existing protocol in order to satisfy the old validation procedures (e.g., see the "segregated witness" implementation [4]).

We feel that the multitude of supported options along with backwards-compatible implementations present the most significant challenge for building a complete mechanized formal model that is faithful to the wire-level protocol. Hence, more research is needed to produce methods of building and using such models without the exploding complexity.

6 Related Work

Bitcoin and similar systems have received a lot of attention in the scientific community in recent years with many attempts to formally specify and verify various aspects of blockchain systems.

Formal treatment of the Bitcoin system

First, we give an overview of formal models aimed at specification and verification of various aspects of the Bitcoin system.

Atzei et al.[9] give a formal model of Bitcoin transactions that we use a starting point for our formalization and mechanization efforts. The model includes transaction and blockchain data structures, as well as the semantics for the Bitcoin Script language. The model is used to formally prove "well-formedness" properties of the Bitcoin blockchain including the impossibility of double-spending. In contrast to a simplified "linked list" model of [9], we fully model blocks and the blockchain including coinbase transactions in each block, block height information, block rewards and transaction fees. For transactions themselves, our models are different in several details where we try to be closer to the behavior of the Bitcoin client. Most notably, in [9] segregated witnesses are a part of the transaction structure, while we store them in blocks, independently of transactions. Mechanization using the Coq proof assistant forces us to carefully specify all the details of the model. For example, mutually inductive definitions have to be explicitly taken into account. Similarly, we need to explicitly state and prove many assumptions that are implicit in [9], such as the temporal properties of spent outputs and explicit encoding of witnesses and hash functions. We replicate the no-double-spent result given in [9] but in a more general setting (with a blockchain data structure) and with a proof that is machine-verified. More importantly, we prove two additional properties of a valid blockchain.

In [13], the authors present the Extended UTXO model (EUTXO), which aims to extend Bitcoin's UTXO model in order to allow more expressive validation scripts. The work determines the expressive power of the model by showing its equivalence with Constraint Emitting Machines, a variant of state machines which is unimplementable in Bitcoin's script. As part of the work, the authors mechanize a transaction model very similar to the one in [9] using the Agda proof assistant. While close to our work, the authors do not attempt to follow Bitcoin specifically and work with a more general UTXO model (i.e. they do not model SegWit), and the overlapping part of the work does not extend beyond [9]. Thus, the previously stated differences between [9] (other than the mechanization) and our work apply here as well.

In [11] an alternative model is given for the semantics of blockchain transactions by using directed acyclic graphs to abstract the interactions of an incoming transaction with the blockchain. They provide a general blockchain model which they instantiate to to Bitcoin, Ethereum and Hyperledger Fabric systems.

Formal models of the Bitcoin Script language have also been an area of active research. In [8], the model of [9] is applied to development of a high-level domain specific language which then compiles into Bitcoin Script language, with the goal of systematically analyzing actual smart contracts proposed by researchers and Bitcoin developers. In [17] authors formalize the Bitcoin Script language with the goal of automatically finding inputs that satisfy a given script.

Finally, formal pen-and-paper treatments of Bitcoin's consensus mechanism include [16] where the focus is on quantifying the *quality* of the blockchain system by determining how many adversarial blocks are expected on the blockchain; and [14] where the authors work out the probability of a successful double-spending attack (assuming some nodes are malicious) and use the UPPAAL model checker to verify the results.

Consensus mechanization

In [22], the authors focus on mechanizing protocols and data structures necessary for establishing distributed consensus in blockchain systems. They formally prove a form of eventual consistency in a network,while precisely characterizing all assumptions on implementations of underlying security primitives. In [23], authors build and mechanize a probabilistic model of blockchain consensus with the eventual goal of stating and proving probabilistic security properties in a Byzantine setting. Other efforts towards automated verification of blockchain consensus mechanisms include [19, 20] that focus on the proposed proof-of-stake mechanism for the Ethereum system. All above efforts use the Coq proof assistant.

7 Conclusions and Future Work

In this paper, we have presented a Coq formalization for the Bitcoin's blockchain validation procedures including the models of basic data structures of the Bitcoin blockchain system and the denotational semantics for the typed variant of the Bitcoin Script language. We have used the model to provide machine-verified proofs for three essential properties of a valid blockchain: impossibility of double-spending, uniqueness of transactions and that cryptocurrency value is created only through block rewards.

In the future, we are going to discharge a number of simplifying assumptions and attempt to further bridge the gap between the abstract model and the reference client. In particular, we plan to model Merkle trees and use them to store transactions and witnesses in blocks. We also plan to make segregated witnesses optional and investigate the interaction between different types of transactions. More generally, we wish to investigate the scenarios where validity checks are updated. This will enable us to formally model the notion of soft-forks and evaluate proposed changes to the Bitcoin protocol such as spending rules based on Taproot, Schnorr signatures, and Merkle branches [6].

References

1 Bitcoin documentation. `https://en.bitcoin.it/wiki/Protocol_documentation`, 2010.

2 Bitcoin improvement proposal 30: Duplicate transactions. bip-0030.mediawiki, 2012.

3 Bitcoin improvement proposal 34: Block v2, height in coinbase. bip-0034.mediawiki, 2012.

4 Bitcoin improvement proposal 141: Segregated witness (consensus layer). bip-0141.mediawiki, 2015.

5 Double spending in bitcoin clients. CVE-2018-17144, 2018.

6 Bitcoin improvement proposal 341: Segwit version 1 spending rules. bip-0341.mediawiki, 2020.

7 Andrew W. Appel. Verification of a cryptographic primitive: Sha-256. *ACM Trans. Program. Lang. Syst.*, 37(2), April 2015. `doi:10.1145/2701415`.

8 Nicola Atzei, Massimo Bartoletti, Tiziana Cimoli, Stefano Lande, and Roberto Zunino. Sok: Unraveling bitcoin smart contracts. In *Principles of Security and Trust*, pages 217–242, Cham, 2018. Springer International Publishing.

9 Nicola Atzei, Massimo Bartoletti, Stefano Lande, and Roberto Zunino. A formal model of bitcoin transactions. *IACR Cryptology ePrint Archive*, 2017:1124, 2017.

10 Vitalik Buterin. A next-generation smart contract and decentralized application platform. `https://github.com/ethereum/wiki/wiki/White-Paper`, 2014. White-paper.

11 Christian Cachin, Angelo De Caro, Pedro Moreno-Sanchez, Björn Tackmann, and Marko Vukolic. The transaction graph for modeling blockchain semantics. *IACR Cryptology ePrint Archive*, 2017:1070, 2017.

12 R. Canetti, A. Stoughton, and M. Varia. Easyuc: Using easycrypt to mechanize proofs of universally composable security. In *2019 IEEE 32nd Computer Security Foundations Symposium (CSF)*, pages 167–16716, June 2019.

13 Manuel Chakravarty, James Chapman, Kenneth MacKenzie, Orestis Melkonian, Michael Peyton Jones, and Philip Wadler. The extended utxo model. In *Workshop on Trusted Smart Contracts*, 2020.

14 Kaylash Chaudhary, Ansgar Fehnker, Jan Cornelis van de Pol, and Mariëlle Ida Antoinette Stoelinga. Modeling and verification of the bitcoin protocol. In *Proceedings of the Workshop on Models for Formal Analysis of Real Systems (MARS 2015)*, Electronic Proceedings in Theoretical Computer Science, pages 46–60, Australia, November 2015. Open Publishing Association.

15 D. Dolev and A. Yao. On the security of public key protocols. *IEEE Transactions on Information Theory*, 29(2):198–208, March 1983.

16 Juan Garay, Aggelos Kiayias, and Nikos Leonardos. The bitcoin backbone protocol: Analysis and applications. In *Advances in Cryptology - EUROCRYPT 2015*, pages 281–310, Berlin, Heidelberg, 2015. Springer Berlin Heidelberg.

17 Rick Klomp and Andrea Bracciali. On symbolic verification of bitcoin's SCRIPT language. In *Data Privacy Management, Cryptocurrencies and Blockchain Technology*, pages 38–56, Cham, 2018. Springer International Publishing.

18 Satoshi Nakamoto. Bitcoin: a peer-to-peer electronic cash system. `https://bitcoin.org/bitcoin.pdf`, 2008.

19 Karl Palmskog, Milos Gligoric, Lucas Pena, Brandon Moore, and Grigore Rosu. Verification of casper in the coq proof assistant. `https://github.com/runtimeverification/casper-proofs`, 2018. Technical report.

20 Karl Palmskog, Milos Gligoric, Lucas Pena, Brandon Moore, and Grigore Rosu. Verifying finality for blockchain systems. In *CoqPL'19*, 2019.

21 Adam Petcher and Greg Morrisett. The foundational cryptography framework. In *Principles of Security and Trust*, pages 53–72, Berlin, Heidelberg, 2015. Springer Berlin Heidelberg.

22 George Pirlea and Ilya Sergey. Mechanising blockchain concensus. In *7th ACM SIGPLAN International Conference on Certified Programs and Proofs*. ACM New York, 2018.

23 Ilya Sergey and Kiran Gopinathan. Towards mechanising probabilistic properties of a blockchain. In *CoqPL'19*, 2019.

24 The Coq Development Team. The coq proof assistant, version 8.10.0, October 2019. `doi:10.5281/zenodo.3476303`.

Towards Verifying the Bitcoin-S Library

Ramon Boss
Bern University of Applied Sciences, Switzerland
ramon.boss@outlook.com

Kai Brünnler
Bern University of Applied Sciences, Switzerland
kai.bruennler@bfh.ch

Anna Doukmak
Bern University of Applied Sciences, Switzerland
anna.doukmak@gmail.com

Abstract

We try to verify properties of the Bitcoin-S library, a Scala implementation of parts of the Bitcoin protocol. We use the Stainless verifier which supports programs in a fragment of Scala called *Pure Scala*. Since Bitcoin-S is not written in this fragment, we extract the relevant code from it and rewrite it until we arrive at code that we successfully verify. In that process we find and fix two bugs in Bitcoin-S.

2012 ACM Subject Classification Theory of computation → Logic and verification

Keywords and phrases Bitcoin, Scala, Bitcoin-S, Stainless

Digital Object Identifier 10.4230/OASIcs.FMBC.2020.8

Category Short Paper

Supplementary Material The original Bitcoin-S code we started from, the extracted code, and the finally verified code are available in our GitHub repository [6]: `https://github.com/kaibr/bitcoin-s-verification`.

1 Introduction

For software handling cryptocurrency, correctness is clearly crucial. However, even in very well-tested software such as Bitcoin Core, serious bugs occur. The most recent example is the bug found in September 2018 [9] which essentially allowed to arbitrarily create new coins. Such software is thus a worthwhile target for formal verification. In this work, we set out to verify properties of the Bitcoin-S library with the Stainless verifier. So this is a case study in applying the Stainless verifier to existing real-world code.

The Bitcoin-S Library. The Bitcoin-S library is an implementation of parts of the Bitcoin protocol in Scala [10, 11]. In particular, it allows to serialize, deserialize, sign and validate Bitcoin transactions. The library uses immutable data structures and algebraic data types but is not specifically written with formal verification in mind. According to the website, the library is used in production, handling significant amounts of cryptocurrency each day [10].

The Stainless Verifier. Stainless is the successor of the Leon verifier and is developed at EPF Lausanne [2, 13, 1]. A distinguishing feature of Stainless is that it accepts specifications written in the programming language itself (Scala). Also, it focusses on counterexample finding in addition to proving correctness. Counterexamples are immediately useful to programmers, which can not be said about correctness proofs.

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler; Article No. 8; pp. 8:1–8:9
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

```scala
def factorial(n: Int): Int = {
    require(n >= 0)
    if (n == 0) {
      1
    } else {
      n * factorial(n - 1)
    }
} ensuring(res => res >= 0)
```

Figure 1 Factorial function with specification.

```
[  Info  ]  - Now solving 'postcondition' VC for factorial @10:3...
[  Info  ]  - Result for 'postcondition' VC for factorial @10:3:
[Warning ]  => INVALID
[Warning ] Found counter-example:
[Warning ]   n: Int -> 17
[  Info  ]   ┌───────────────────┐
[  Info  ] ╔═╡ stainless summary ╞══════════════════════════════════════════════════════════════════════════╗
[  Info  ] ║ └───────────────────┘                                                                          ║
[  Info  ] ║ factorial  postcondition                  valid from cache          src/TestFactorial.scala:10:3   1.055 ║
[  Info  ] ║ factorial  postcondition                  invalid         U:smt-z3  src/TestFactorial.scala:10:3   7.861 ║
[  Info  ] ║ factorial  precond. (call factorial(n - 1))  valid from cache       src/TestFactorial.scala:15:11  1.054 ║
[  Info  ] ╟┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄┄╢
[  Info  ] ║ total: 3    valid: 2    (2 from cache) invalid: 1    unknown: 0    time:   9.970                ║
[  Info  ] ╚════════════════════════════════════════════════════════════════════════════════════════════════╝
```

Figure 2 Stainless output for the factorial function.

The example in Figure 1 is adapted from the Stainless documentation [7] and shows how the verifier is used. Note how a precondition is specified using `require` and a postcondition using `ensuring`. Our function does not satisfy the specification. An overflow in the 32-bit integer type leads to a negative result for the input 17, as Stainless reports in Figure 2. Changing the type `Int` to `BigInt` will result in a successful verification.

The Pure Scala Fragment. The Scala fragment supported by Stainless comprises algebraic data types in the form of abstract classes, case classes and case objects, objects for grouping classes and functions, boolean expressions with short-circuit interpretation, generics with invariant type parameters, pattern matching, local and anonymous classes and more. In addition to Pure Scala Stainless also supports some imperative features, such as while loops and using a (mutable) variable in a local scope of a function. They turn out not to be relevant for our current work.

What will turn out to be more relevant for us are the Scala features which Stainless does not support, such as: inheritance by objects, abstract type members, and inner classes in case objects. Also, Stainless has its own library of some core data types and functions which are mapped to corresponding data types and functions inside of the SMT solver that Stainless ultimately relies on. Those data types in general do not have all the methods of the Scala data types. For example, the `BigInt` type in Scala has methods for bitwise operations while the `BigInt` type in Stainless does not.

Outline and Properties to Verify. In the next section we try to verify the property that a regular (non-coinbase) transaction can not generate new coins. We call it the *No-Inflation Property*. Trying to verify it, we uncover and fix a bug in the Bitcoin-S library. We then find that there is too much code involved that lies outside of the supported fragment to currently make this verification feasible. So we turn to a simpler property to verify. The simplest

```
1  def checkTransaction(transaction: Transaction): Boolean = {
2    val inputOutputsNotZero =
3      !(transaction.inputs.isEmpty || transaction.outputs.isEmpty)
4    val txNotLargerThanBlock =
5      transaction.bytes.size < Consensus.maxBlockSize
6    val outputsSpendValidAmountsOfMoney =
7      !transaction.outputs.exists(o =>
8        o.value < CurrencyUnits.zero || o.value > Consensus.maxMoney)
9
10   val outputValues = transaction.outputs.map(_.value)
11   val totalSpentByOutputs: CurrencyUnit =
12     outputValues.fold(CurrencyUnits.zero)(_ + _)
13   val allOutputsValidMoneyRange =
14     validMoneyRange(totalSpentByOutputs)
15   val prevOutputTxIds = transaction.inputs.map(_.previousOutput.txId)
16   val noDuplicateInputs =
17     prevOutputTxIds.distinct.size == prevOutputTxIds.size
18
19   val isValidScriptSigForCoinbaseTx = transaction.isCoinbase match {
20     case true =>
21       transaction.inputs.head.scriptSignature.asmBytes.size >= 2 &&
22         transaction.inputs.head.scriptSignature.asmBytes.size <= 100
23     case false =>
24       !transaction.inputs.exists(
25         _.previousOutput == EmptyTransactionOutPoint)
26   }
27   inputOutputsNotZero && txNotLargerThanBlock &&
28   outputsSpendValidAmountsOfMoney && noDuplicateInputs &&
29   allOutputsValidMoneyRange && noDuplicateInputs &&
30   isValidScriptSigForCoinbaseTx
31 }
```

Figure 3 The `checkTransaction` function.

possible property we can think of is the fact that adding zero satoshis to a given amount of satoshis yields the given amount of satoshis. We call it the *Addition-With-Zero Property* and we try to verify it in Section 3. Here as well we see that a significant part of the code lies outside of the supported fragment. We rewrite it until we arrive at code that we successfully verify. In that process we find and fix a second bug in Bitcoin-S.

2 The No-Inflation Property

An Attempt at Verification. Naively trying Stainless on the entire Bitcoin-S codebase results in many errors – as was to be expected. We tried to extract only the code relevant to the No-Inflation Property and to verify that. However, even the extracted code has more than 1500 lines and liberally uses Scala features outside of the supported fragment. We started to rewrite the code in the supported fragment, but quickly realized that a better approach is to first verify a simpler property depending on less code and later come back to the No-Inflation Property with more experience. However, during the process of trying to rewrite the code, we found a bug in the `checkTransaction` function shown in Figure 3.

A Bug in the checkTransaction Function. Given a transaction the function returns true if some basic checks succeed, otherwise false. For example, one of those checks is that both the list of inputs and list of outputs need to be non-empty.

Note particularly lines 15-17. Here, the value `prevOutputTxIds` gathers a list of all transaction identifiers referenced by the inputs of the current transaction. If the size of this list is the same as the size of this list with duplicates removed, we know that no transaction

```
15   val prevOutputs = transaction.inputs.map(_.previousOutput)
16   val noDuplicateInputs =
17     prevOutputs.distinct.size == prevOutputs.size
```

Figure 4 Bug Fix.

has been referenced twice. This prevents a transaction from spending two different outputs of the same previous transaction. The check is too strict: `checkTransaction` returns false for valid transactions.

The fix is simple: we perform the duplicate check on the `TransactionOutPoint` instances instead of on their transaction identifiers. Note that `TransactionOutPoint` is a case class and thus its notion of equality is just what we need: equality of of both the transaction identifier and the output index.

Specifically, we replace lines 15-17 as shown in Figure 4. We submitted this fix together with a corresponding unit test to the Bitcoin-S project in a pull request, which has been merged [5].

We now turn to the much simpler Addition-With-Zero Property.

3 The Addition-With-Zero Property

It is of course a crucial property we are verifying here: if zero satoshis were credited to your account, you would not want your balance to change! It is also the simplest meaningful property to verify that we can think of. However, the code involved in performing the addition of two satoshi amounts in Bitcoin-S is non-trivial. The reason for that is a peculiarity of consensus code: agreement with the majority is the only relevant notion of correctness. The most widely used bitcoin implementation by far is the reference implementation Bitcoin Core, written in C++. For consensus code, Bitcoin-S thus has little choice but to be in strict agreement with the reference implementation. To achieve that, it implements C-like data types in Scala and then implements functionality using those C-like data types. For example, the Satoshis class, which represents an amount of satoshis, is implemented using the class `Int64` which aims to represent the C-type `int64_t`.

Extracting the Relevant Code. The relevant code for the addition of satoshis is in two files: CurrencyUnits.scala and NumberType.scala. From those files we removed the majority of the code because it is not needed for the verification of our property. For example, we removed all number types except for `Int64` (so `Int32`, `UInt64`, etc.) because they are not used. We also removed the superclasses `Factory` and `NetworkElement` of `CurrencyUnit` and `Number`, respectively, because the inherited members are not used. We further removed all binary operations on `Number` that are not used, like subtraction and multiplication. The extracted code is shown in Figure 5 and Figure 6.

A Bug in the checkResult Function. Note the `checkResult` function on line 12 and the value `andMask` on line 23 of NumberType.scala. The function is intended to catch overflows by performing a bitwise conjunction of its argument with `andMask` and comparing the result with the argument. However, because of the way Java BigIntegers are represented [14] and because bitwise operations implicitly perform a sign extension [8] on the shorter operand, the function does not actually catch overflows.

```scala
package extracted.number

sealed abstract class Number[T <: Number[T]] {
  type A = BigInt
  protected def underlying: A
  def toLong: Long = toBigInt.bigInteger.longValueExact()
  def toBigInt: BigInt = underlying
  def andMask: BigInt
  def apply: A => T
  def +(num: T): T = apply(checkResult(underlying + num.underlying))

  private def checkResult(result: BigInt): A = {
    require((result & andMask) == result,
      "Result was out of bounds, got: " + result)
    result
  }
}

sealed abstract class SignedNumber[T <: Number[T]] extends Number[T]

sealed abstract class Int64 extends SignedNumber[Int64] {
  override def apply: A => Int64 = Int64(_)
  override def andMask = 0xffffffffffffffffL
}

trait BaseNumbers[T] {
  def zero: T
}

object Int64 extends BaseNumbers[Int64] {
  private case class Int64Impl(underlying: BigInt) extends Int64 {
    require(underlying >= -9223372036854775808L,
      "Number was too small for a int64, got: " + underlying)
    require(underlying <= 9223372036854775807L,
      "Number was too big for a int64, got: " + underlying)
  }

  lazy val zero = Int64(0)
  def apply(long: Long): Int64 = Int64(BigInt(long))
  def apply(bigInt: BigInt): Int64 = Int64Impl(bigInt)
}
```

Figure 5 Extracted Code from NumberType.scala.

```scala
1  package extracted.currency
2
3  import extracted.number.{BaseNumbers, Int64}
4
5  sealed abstract class CurrencyUnit {
6    type A
7    def satoshis: Satoshis
8    def ==(c: CurrencyUnit): Boolean = satoshis == c.satoshis
9    def +(c: CurrencyUnit): CurrencyUnit = {
10     Satoshis(satoshis.underlying + c.satoshis.underlying)
11   }
12   protected def underlying: A
13 }
14
15 sealed abstract class Satoshis extends CurrencyUnit {
16   override type A = Int64
17   override def satoshis: Satoshis = this
18   def toBigInt: BigInt = BigInt(toLong)
19   def toLong: Long = underlying.toLong
20   def ==(satoshis: Satoshis): Boolean = underlying == satoshis.underlying
21 }
22
23 object Satoshis extends BaseNumbers[Satoshis] {
24   val zero = Satoshis(Int64.zero)
25   def apply(int64: Int64): Satoshis = SatoshisImpl(int64)
26   private case class SatoshisImpl(underlying: Int64) extends Satoshis
27 }
```

Figure 6 Extracted Code from CurrencyUnits.scala.

While this is a potentially serious bug, it turns out that `checkResult` is only ever called inside a constructor call for a number type which contains the intended range check, see lines 32-35. The `checkResult` function thus can, and should, be removed entirely. The Bitcoin-S developers have acknowledged the bug and we submitted a pull request to fix it, which has been merged [4].

For further development of Bitcoin-S, this raises a question. If the goal of the `Int64` type is to emulate `int64_t` then why does it prevent overflows? To achieve strict agreement with Bitcoin Core, a better approach might be to remove overflow checking from the data type and to add it in exactly those places where it happens in Bitcoin Core.

Rewriting the Code. We now turn to the list of Scala features used by the extracted code which are not supported by Stainless and how to rewrite the code in the supported fragment.

All code changes are *equivalent* in the (admittedly narrow) sense that if the Addition-With-Zero Property holds for the rewritten code, then it also holds for the original code.

Inheriting Objects. In both files we have objects extending the BaseNumbers trait, on lines 30 and 23 respectively, which Stainless does not support. We simply turn those objects into case objects. That code is equivalent: case objects have various additional properties (for example, being serializable) but none of our code depends on the absence of those.

Abstract Type Members. In CurrencyUnits.scala on line 6 there is an abstract type that is not supported. Note that we can not simply replace it with a (supported) type parameter since the CurrencyUnit class uses one of its implementing classes: Satoshis. Since the Satoshis class overrides `A` with `Int64` anyway, we just remove the abstract type declaration and replace `A` by `Int64` everywhere.

Non-Literal BigInt Constructor Argument. In CurrencyUnits.scala on line 18 the BigInt constructor is called with a non-literal argument. As described before, the types in the Stainless library are more restricted than their Scala library counterparts. In particular, the Stainless BigInt constructor is restricted to literal arguments. So we simply replace `toLong` by `underlying.toBigInt`: instead of converting the underlying `Int64` (which in turn has an underlying `BigInt`) to `Long` and then back to `BigInt` we simply directly return the `BigInt`. This is an equivalent transformation: the only thing that might go wrong in the detour via `Long` is that the underlying `BigInt` does not fit into a `Long`. However, the only constructor of `Int64Impl` ensures exactly that and all functions producing `Int64` do so via this constructor.

Self-Reference in Type Parameter Bound. In NumberTypes.scala both on lines 3 and 19 is a class with a type parameter and a type boundary that contains that type parameter itself. Stainless does not currently support such self-referential type boundaries. We opened an issue [3] on the Stainless repository and the developers have targeted version 0.4 to support self-referential type boundaries. Since our code only uses Number with type parameter `T` instantiated to `Int64`, we just remove the type parameter declaration and replace all its occurrences by `Int64`.

Missing Member `bigInteger` in BigInt. In NumberType on line 6 there is a reference to `bigInteger`. The Scala `BigInt` class is essentially a wrapper around `java.math.BigInteger`. `BigInt` has a member `bigInteger` which is the underlying instance of the Java class. The Java class has a method `longValueExact` which returns a `long` only if the `BigInteger` fits into a `long`, otherwise throws exception. Stainless does not support Java classes and in particular its `BigInt` has no member `bigInteger`. However, our code does not call `toLong` anymore, so we just remove it.

Type Members. In NumberType.scala there is a type member on line 4. Our version of Stainless (0.1) does not support type members. We just remove the declaration and replace all occurrences of `A` with `BigInt`, since `A` is never overwritten in an implementing class. Note that in the meantime Stainless has implemented support for type members [12]. Since version 0.2 verification should succeed without this change.

Missing Bitwise-And Method on BigInt. Contrary to Scala `BigInt`, the Stainless `BigInt` class does not support bitwise operations, in particular not the &-method used in NumberType.scala on line 13. However, as described above, the `checkResult` function is both broken and redundant, so we remove it and all calls to it.

Inner Class in Case Object. We have inner classes in NumberType.scala on line 31 and in CurrencyUnits.scala on line 26. Stainless does not support inner classes in a case object. We just move the inner classes out of the case objects. They do not interfere with any other code.

Message Parameter in Require. The calls of the require function on lines 32 and 34 in CurrencyUnits.scala have a second parameter: the error message. Stainless does not support the message parameter. We simply remove it.

```
  def +(c: CurrencyUnit): CurrencyUnit = {
    Satoshis(satoshis.underlying + c.satoshis.underlying)
  } ensuring (res =>
    (c == Satoshis.zero) ==> (res == this))
```

Figure 7 Addition function with specification.

```
[  Info  ]  - Now solving 'postcondition' VC for + @9:3...
[  Info  ]  - Result for 'postcondition' VC for + @9:3:
[  Info  ]  => VALID
[  Info  ]
[  Info  ]  ╔═ stainless summary ═══════════════════════════════════════════════════════════╗
[  Info  ]  ║                                                                               ║
[  Info  ]  ║ + postcondition   valid  U:smt-z3   verified/currency/CurrencyUnits.scala:9:3      1.451 ║
[  Info  ]  ║-------------------------------------------------------------------------------║
[  Info  ]  ║ total: 1    valid: 1    (0 from cache) invalid: 0    unknown: 0    time:   1.451 ║
[  Info  ]  ╚═══════════════════════════════════════════════════════════════════════════════╝
```

Figure 8 Stainless output for the rewritten code.

Missing Implicit Long to BigInt Conversion. The Scala `BigInt` class has implict conversions from `Long` which NumberType.scala uses on lines 32 and 34. They are missing in the Stainless `BigInt`. A `BigInt` constructor with a `Long` argument is also missing. We thus replace the `Long` literals by an explicit call to the `BigInt` constructor with a literal string argument, e.g. `BigInt("-9223...5808")`.

The Specification. Now that all our code is in the supported fragment, we can finally write our specification. We add a postcondition to the +-method of the `CurrencyUnit`-class (Figure 6, lines 9-11) resulting in Figure 7. We successfully verify it with Stainless, as the output in Figure 8 shows.

The original Bitcoin-S code we started from, the extracted code, and the finally verified code are available in our GitHub repository [6].

4 Conclusion and Future Work

We are happy to see some friendly green verifier output. However, apart from the bugs we found, the main conclusion of this work is that we had to non-trivially transform even a very small portion of the code (70 lines) in order to verify it. And that was true even though the code was purely functional to begin with. At the moment, it is probably unrealistic to routinely formally verify properties as part of the Bitcoin-S development process. However, Stainless development has already progressed (e.g. type members are supported in recent versions) and continues to do so (e.g. self-referential type bounds are on the roadmap). Some missing features that we identified are presumably very easy to support, like the message parameter in the require function. Some other features presumably require more substantial work, like bitwise operations on integer types.

On the other hand, Bitcoin-S uses features that might not be supported even by future Stainless versions, such as calls to Java code.

Given our experience, the best route towards integrating verification into the Bitcoin-S development process would be to re-implement parts of the library in Pure Scala. We would split the library into a *verified* and *non-verified* part, and use Stainless only on the *verified* part. It is then both a technical but also a political question how much code, if any, can be moved to the *verified* part. That is an interesting direction for future work.

References

1 Régis Blanc and Viktor Kuncak. Sound reasoning about integral data types with a reusable SMT solver interface. In Philipp Haller and Heather Miller, editors, *Proceedings of the 6th ACM SIGPLAN Symposium on Scala, Scala@PLDI 2015, Portland, OR, USA, June 15-17, 2015*, pages 35–40. ACM, 2015. `doi:10.1145/2774975.2774980`.

2 Régis Blanc, Viktor Kuncak, Etienne Kneuss, and Philippe Suter. An overview of the leon verification system: verification by translation to recursive functions. In *Proceedings of the 4th Workshop on Scala, SCALA@ECOOP 2013, Montpellier, France, July 2, 2013*, pages 1:1–1:10. ACM, 2013. `doi:10.1145/2489837.2489838`.

3 Ramon Boss. Issue 519: Unknown type parameter type T in self referencing generic. Accessed 2019-06-27. URL: `https://github.com/epfl-lara/stainless/issues/519`.

4 Ramon Boss. Remove redundant function checkresult. Accessed 2019-07-03. URL: `https://github.com/bitcoin-s/bitcoin-s/pull/565`.

5 Ramon Boss. Transaction can reference two different outputs of the same previous transaction. Accessed 2019-06-19. URL: `https://github.com/bitcoin-s/bitcoin-s/pull/435`.

6 Ramon Boss, Kai Brünnler, and Anna Doukmak. The bitcoin-s-verification repository. Accessed 2019-07-06. URL: `https://github.com/kaibr/bitcoin-s-verification`.

7 LARA Lab, École Polytechnique Fédérale de Lausanne. Stainless documentation. Accessed 2019-06-19. URL: `https://epfl-lara.github.io/stainless/`.

8 Oracle and/or its affiliates. Class BigInteger. Accessed 2019-07-03. URL: `https://docs.oracle.com/javase/8/docs/api/java/math/BigInteger.html`.

9 Jimmy Song. Bitcoin Core Bug CVE-2018–17144: An Analysis. Accessed 2019-06-20. URL: `https://hackernoon.com/bitcoin-core-bug-cve-2018-17144-an-analysis-f80d9d373362`.

10 Suredbits & the bitcoin-s developers. The bitcoin-s website. Accessed 2019-06-19. URL: `https://bitcoin-s.org`.

11 The bitcoin-s developers. The bitcoin-s repository. Accessed 2019-06-19. URL: `https://github.com/bitcoin-s`.

12 The Stainless developers. Type aliases, type members, and dependent function types. Accessed 2019-06-27. URL: `https://github.com/epfl-lara/stainless/pull/470`.

13 Nicolas Voirol, Etienne Kneuss, and Viktor Kuncak. Counter-example complete verification for higher-order functions. In Philipp Haller and Heather Miller, editors, *Proceedings of the 6th ACM SIGPLAN Symposium on Scala, Scala@PLDI 2015, Portland, OR, USA, June 15-17, 2015*, pages 18–29. ACM, 2015. `doi:10.1145/2774975.2774978`.

14 Wikipedia contributors. Two's complement. Accessed 2019-07-03. URL: `https://en.wikipedia.org/wiki/Two%27s_complement`.

On the Formal Verification of the Stellar Consensus Protocol

Giuliano Losa
Galois, Inc., Portland, Oregon, USA
giuliano@galois.com

Mike Dodds
Galois, Inc., Portland, Oregon, USA
miked@galois.com

Abstract

The Stellar Consensus Protocol (SCP) is a quorum-based BFT consensus protocol. However, instead of using threshold-based quorums, SCP is permissionless and its quorum system emerges from participants' self-declared trust relationships. In this paper, we describe the methodology we deploy to formally verify the safety and liveness of SCP for arbitrary but fixed configurations.

The proof uses a combination of Ivy and Isabelle/HOL. In Ivy, we model SCP in first-order logic, and we verify safety and liveness under eventual synchrony. In Isabelle/HOL, we prove the validity of our first-order encoding with respect to a more direct higher-order model. SCP is currently deployed in the Stellar Network, and we believe this is the first mechanized proof of both safety and liveness, specified in LTL, for a deployed BFT protocol.

2012 ACM Subject Classification Software and its engineering → Software verification; Networks → Protocol testing and verification

Keywords and phrases Consensus, Blockchains, First-Order Logic, Stellar, Ivy Prover, Decidability

Digital Object Identifier 10.4230/OASIcs.FMBC.2020.9

Funding This work was supported by the Stellar Development Foundation.

1 Introduction

Blockchains rely on Byzantine Fault-Tolerant (abbreviated BFT) consensus protocols to ensure that, despite the presence of malicious participants, the network of participants as a whole eventually reaches consensus on what block to append next to the blockchain. In many blockchains, the security of large amount of digital assets depend on the correctness of the blockchain's BFT consensus protocol, but designing BFT consensus protocols is notoriously difficult and serious flaws can remain undetected for years [1].

While formal verification can prevent many correctness issues in BFT consensus protocols, performing such verification is challenging for several reasons: BFT consensus protocols are designed to support an arbitrary number of participants; their executions and their reachable state-space are unbounded; they operate in asynchronous networks where the interleaving of messages is unpredictable; and finally, verifying termination is as important as verifying that participants never disagree.

In this paper, we summarize our approach to the formal verification of the main safety and liveness properties of the Stellar Consensus Protocol (abbreviated SCP) [7] using the Ivy methodology [10]. Both the safety and liveness proofs apply to a unique model of SCP. This model is parameterized by a fixed but arbitrary set of participants and denotes a set of infinite executions. To our knowledge, this is the first work that mechanically proves both safety and liveness, expressed in LTL, of a deployed BFT protocol under arbitrary configurations.

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler; Article No. 9; pp. 9:1–9:9
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

At a high level, verifying the safety of a protocol with Ivy entails 1) developing a set of axioms to express the protocol's underlying domain model as a first-order theory over uninterpreted sorts; 2) modeling the protocol in Ivy's procedural language; 3) developing an inductive invariant that implies the safety properties, while ensuring that verification conditions fall into the decidable first-order logic fragment EPR [5]. This is facilitated by Ivy's modular decomposition features [17].

For termination, or more generally liveness, Ivy provides a liveness to safety reduction [12] crafted specifically to help produce decidable verification conditions. Given a temporal property in First-Order Linear Temporal Logic (FO-LTL), Ivy automatically synthesises a transition system and an associated safety property such that if the synthesized system is safe, then the temporal property of the original system holds. The user can then verify that the synthesized transition system is safe using the safety verification methodology.

Producing EPR verification conditions ensures that Z3 can automatically and reliably determine their satisfiability. Compared to approaches that use automation but do not require decidability, Ivy's predictable automation greatly simplifies the mental model of the prover that the user must keep in mind when developing a proof. The user can thus stop worrying about the prover and instead focus on the properties of the protocol.

A key challenge in applying the Ivy methodology to SCP is to model SCP's permissionless Federated Byzantine Quorum Systems in first-order logic and in a way that is conductive to decidable reasoning in EPR. In SCP, every participant expresses agreement requirements with other nodes, and SCP relies on the properties of the resulting graph-like structure to solve consensus. At first sight, such a complex family of structures seem hard to axiomatize in first-order logic, let alone EPR.

The rest of the paper focuses on the first-order logic modeling of Federated Byzantine Quorum Systems. This model abstracts over significant aspects of SCP's quorum system. To provide evidence that the abstraction is sound, we verify some of its key properties with respect to a more concrete model in Isabelle/HOL.

With a first-order theory of Federated Byzantine Quorum Systems established, we verify that SCP's balloting protocol [7] satisfies its agreement property and that, under eventual synchrony, it satisfies its termination property (i.e. that every node eventually decides). This safety and liveness proof largely follows patterns identified previously during the verification of other consensus protocols in Ivy [14, 12, 3], and it is not described in this paper.

Our paper supplies evidence that BFT consensus protocol can be verified with decidable logics, which enables powerful yet stable automation, using the Ivy methodology. The proof is available online [8], and, for a more complete reference, we plan to publish an extended version of this paper with a detailed account of the safety and liveness proof that are omitted here.

2 Solving Consensus in a Federated Byzantine Quorum System

SCP must solve consensus, guaranteeing agreement and termination, in a permissionless system where nodes can join or leave without any synchronization and without the permission of any gatekeeper. There is thus no common notion of the set of all nodes. Moreover, the system is susceptible to Sybil attacks, in which attackers create a large number of identities to try to overwhelm the system. In such an environment, traditional threshold-based quorum systems, defined in terms of the total number of nodes, are thus of no use.

Other permissionless protocols like Bitcoin or Algorand use Proof-of-Work or Proof-of-Stake to defend against Sybil attacks. Stellar takes a different approach. The Stellar Network is intended as a platform to exchange digitized real-world assets (e.g. land parcels, retail coupons, national currencies, agricultural goods, etc.). Most participants are thus expected to engage with recognized identities and have real-world relationships with some (but not all) other participants in the network. SCP leverages these real-world relationships to defend against Sybil attacks, counting on real-world relationships to be difficult for an attacker to establish.

Concretely, each node in the Stellar Network is required to independently declare a set of *slices*, where each slice is a set of nodes. The intent is that a node n accepts some new information it hears on the network if and only if one of its slices unanimously agrees that the information is correct. Thus slices can be thought of as agreement requirements. Nodes advertise their slices throughout the network, and each node forms its own, personal notion of quorum based on its own slices and on the slices of other nodes it knows about, as follows. A quorum of n is defined as a set Q of nodes such that a) n has a slice included in Q and b) each member of Q has a slice included in Q. In other words, a quorum of n is a set that n satisfies the agreement requirement of n and of all its members. Formally, let $\text{slices}(n)$ denote the set of slices of node n. Then a set of nodes Q is a quorum of a node n if a) $\exists S \in \text{slices}(n).\ S \subseteq Q$ and b) $\forall m \in Q.\ \exists S \in \text{slices}(m).\ S \subseteq Q$. The resulting quorum system is called a Federated Byzantine Quorum System (abbreviated FBQS).

2.1 Intact and Intertwined Sets

With the notion of quorum in place, it seems possible to take a traditional threshold-based BFT consensus protocols, and only change how quorums are defined in order to obtain a consensus protocol for the Stellar Network. However, FBQS have some unusual properties that complicate the task. First, the notion of quorum is not global to the system; instead, each node has its own view of what a quorum is. Second, the quorums of a node depend on what slices other nodes declare; thus, Byzantine nodes can influence a well-behaved node's notion of what a quorum is. Third, it is possible that a subset of the nodes have quorums that intersect enough to guarantee safety, while some other subsets do not; thus, consensus may be solvable for only a strict subset of the system; there may even by two or more disjoint subsets of the system that form consensus islands that nevertheless diverge from each other.

What properties must a set of nodes satisfy in order for consensus to be solvable among its members? We do not know a precise answer to this question [9]. However, we can prove that SCP solves eventually synchronous consensus among sets of nodes called *intact sets*. A set I of well-behaved (non-Byzantine) nodes is intact when, regardless of what slices Byzantine nodes advertise: a) I enjoys quorum availability, i.e. the set I is a quorum for all its members, and b) I enjoys quorum intersection, i.e. if n_1 and n_2 are members of I, if Q_1 is a quorum of n_1, and if and Q_2 is a quorum of n_2, then the intersection of Q_1 and Q_2 contains a member of I. An important property of intact sets is that the union of two intact sects that intersect is also an intact set; thus maximal intact sets are disjoint and form consensus islands within the network.

Sets which have quorum intersection but which lack quorum availability are called *intertwined* sets. Precisely, a set S of nodes is intertwined when, if n_1 and n_2 are members of S, if Q_1 is a quorum of n_1, and if Q_2 is a quorum of n_2, then the intersection of Q_1 and Q_2 contains an intertwined member. SCP also guarantees that there will not be any disagreement among an *intertwined* set.

2.2 Termination and the Cascade Theorem

Thanks to the quorum intersection property, it is easy to guarantee agreement to an intertwined set. However, termination is more difficult to achieve. Traditional BFT consensus protocols often rely on eventual synchrony [4] to ensure termination. The idea is that, once the system becomes synchronous, the protocol can rely on all nodes having the same view of the system.

For example, suppose that, in a threshold quorum system, a quorum Q unanimously agrees on statement X. If the network is synchronous, then all nodes shortly notice that Q unanimously agrees on X. In this sense, they all form the same view of the fact "there is a quorum that is unanimous about X". Instead, if the quorum system is not a threshold quorum system but an FBQS, then no such common view arises because Q may be a quorum only of some nodes but not others.

SCP circumvents this problem using an epidemic propagation phenomenon that guarantees that, once the system is synchronous, if an intact node witnesses a unanimous quorum, then the knowledge that there is such a quorum soon propagates to the entire intact set, and Byzantine nodes cannot prevent propagation.

The epidemic propagation phenomenon is enabled by the Cascade Theorem. This theorem relies on the notion of slice-blocking set. A set B is a slice-blocking set for a node n when every slice of n intersects B. The cascading theorem states that if n is intact, Q is a quorum of n, and U is a superset of Q, then either all intact nodes belong to U, or U slice-blocks some intact node $m \notin U$.

Let us now get back to the example in which a quorum Q of an intact node unanimously agrees on statement X. We would like that all intact nodes learn the fact "there exists a quorum of an intact node that unanimously agrees on X". By the Cascade Theorem, either all intact nodes already know the fact, or there must be an intact node n that does not know it but that is slice-blocked by a set of intact nodes that know it. Thus, if we add the rule that n must accept a fact if slice-blocked by a set that already accepted the fact, then n newly accepts the fact. This process then repeats until the knowledge of the existence of Q propagates to the entire intact set.

Finally, we must also be sure that malicious nodes cannot use epidemic propagation to propagate forged facts. This is guaranteed because if n is intact and S slice-blocks n, then S contains an intact node.

3 Modeling Federated Byzantine Quorum Systems in EPR

In this section, we describe the first-order theory of Federated Byzantine Quorum System. This is the model we use in our proofs of safety and liveness (the proofs themselves are not described in detail in this paper).

3.1 Enabling Decidable Reasoning

We craft the FBQS model to meet two constraints: on the one hand, the model must enable decidable automated reasoning in EPR; on the other hand, the model must accurately capture the properties that FBQSs have in practice. Our solution is to abstract over some important aspects of FBQSs to make decidable reasoning possible, while formally verifying that the model is sound with respect to a more concrete model developed in Isabelle/HOL. By doing so, we trade off a relatively small manual proof effort in Isabelle/HOL in exchange for decidable automated reasoning in Ivy.

To enable decidable reasoning with Ivy, we model FBQSs as a first-order theory consisting of: a) a set of uninterpreted sorts, b) constants, functions, and relations over those sorts, and c) first-order axioms that constrain the models of the theory to structures that have properties sufficient for the balloting protocol to be correct. Moreover, we must use quantifier alternations and functions carefully, as those will impact our ability to keep protocol verification conditions in EPR.

A verification condition is in EPR when its sorts are stratified: for every pair of sorts a and b, say that b depends on a if either a) an existential quantifier on sort b is in the scope of a universal quantifier on sort a, or b) there is a function symbol of type $a_1, \cdots, a_n \to b$ with $a = a_j$ for some $j \in 1 \cdots n$; sorts are stratified if the dependencies between sort do not form any loops or cycles. For example, in the formula $\forall x.\exists y.P(x, y)$ where x is of sort a and y is of sort b and P is a predicate symbol, sort b depends on sort a but the formula is stratified. However, if both x and y have the same sort, then there is a sort dependency loop and the formula is not stratified.

Protocol verification conditions are formulas of the form $A \wedge I \wedge T \wedge \neg I'$, where A is the conjunction of the FBQS theory axioms, I is a protocol invariant, T is the protocol's transition relation, and I' is the post-state version of I. Thus unstratified verification conditions can arise because of the interaction between axioms, invariants, their negation, and the protocol's transition relation. It is thus wise to minimize the use of function symbols and quantifier alternation when developing the EPR FBQS theory.

In our experience, stratification is nevertheless likely to become an issue during protocol verification. However, Ivy has modular decomposition features specifically designed to help keep verification conditions decidable. The process of structuring proof modularly to ensure decidability is explained in details by Taube et al. [17]. In the case of liveness proofs, prophecy variables also help keep verification conditions decidable [13].

3.2 The Unique Challenges Posed by FBQSs

Developing an EPR theory of FBQSs is challenging because the notions we presented in Section 2, such as intact sets, slice-blocking sets, or the cascading theorem, are naturally second-order concepts. I.e. they are naturally expressed by quantifying over sets. While we cannot precisely capture the higher-order theory of sets in first-order logic, we can approximate it by using a first-order uninterpreted sort `nset`, a membership relation `member(N:node, S:nset)`, and appropriate axioms.

The full first-order theory of FBQSs appears in Figure 1. We model an arbitrary but fixed configuration, i.e. an arbitrary set of nodes with arbitrary slices and we consider a fixed intact set I among those nodes as well as a superset S of I such that S is intertwined (note that an intact set is inherently intertwined, so there is no inconsistency here). Instead of modeling slices explicitly, we only model the notions of intact node, intertwined node, quorum, and slice-blocking set.

Formally, in Figure 1, we introduce an uninterpreted sort `node`, denoting the set of all nodes, and an uninterpreted sort `nset`, denoting the powerset of the set of nodes (lines 1 and 2). Well-behaved, intertwined, and intact nodes are identified by corresponding unary relations (lines 3 to 5), and quorums of a node are identified by the binary relation `quorum_of` (line 7). Finally, the binary relation `member` (line 6) denotes set membership, and the binary relation `slice_blocking` identifies the slice-blocking sets of a node (line 8).

Given those sorts and relations, we obtain the first-order theory of FBQSs using the following axioms. First, line 9, we assert that intact nodes are intertwined, and that intertwined nodes are well-behaved. In line 10 and 11, we assert that quorums of well-behaved nodes are not empty. Then, line 12 to 15, we define two predicates to identify quorums of intertwined nodes and quorums of intact nodes. Then, line 16 and 17, we assert the quorum intersection property of intertwined nodes. Similarly, line 18 and 19, we assert the quorum intersection property of intact nodes. Line 20 and 21, we assert that if `N` is intact and `S` slice-blocks `N`, then `S` contains an intact node. Finally, line 22, we assert that the set of intact nodes is a quorum.

The conjunction of all the axioms is an EPR formula because the associated quantifier-alternation graph has a single dependency: sort `node` depends on sort `nset`. For example, lines 16 and 17 in Figure 1, the quorum intersection axiom for intertwined sets creates a dependency from sort `node` to sort `nset`. As explained in Section 3.1, this dependency may create a quantifier-alternation cycle when the axioms are conjoined with other formulas in a verification condition, and it is the user's responsibility to make use of Ivy's modularity features to avoid such a cycle when verifying a protocol; this process is explained in [17].

The reader may notice that the Cascade Theorem is missing from the axioms, and instead is expressed as an axiom schema in Figure 2. The reason is that we could not satisfactorily express it in first-order logic. The theorem states that if `p` is a predicate on nodes (i.e. a set of nodes) and `Q` is a quorum of an intact node whose intact members unanimously satisfy `p`, then either a) all intact nodes satisfy `p` or b) there exists an intact node `N` that does not satisfy `p` but that is slice-blocked by a set `S` whose members are exclusively intact and unanimously satisfy `p`. While other axioms quantify over a restricted family of sets, such as quorums or slice-blocking sets, the Cascade Theorem quantifies over all predicates `p`. It is thus inherently second-order. Ivy allows to express it as an axiom schema, but Ivy's proof automation cannot reason about such a second-order formula. Instead, Ivy allows to manually instantiate it, substituting `p` for a concrete predicate, to prove particular invariants. We use this technique in the termination proof of SCP. Note that, also when instantiating the Cascade Theorem, we must be careful not to introduce quantifier-alternation cycles.

Together, the axioms appearing in Figure 1 and the axiom schema of Figure 2 form the first-order theory of Federated Byzantine Quorum Systems.

3.3 Validating the Model

Asserting axioms instead of proving them as properties from basic definitions can be dangerous: even benign-looking axioms can turn out to be contradictory, e.g. because of a typo, thereby making any proof relying on them vacuous. To avoid this situation, we ask Ivy to find a model of the axioms of Figure 1 conjoined with the instantiations of the `cascade_thm` axiom schema that we use in the proof. Ivy confirms the existence of a model, which rules out any contradiction.

Another risk is that, although the axioms are not contradictory, they do not accurately model FBQSs. For instance, the first-order model abstracts over slices and instead considers that a node's quorums are fixed. This is limiting because, in reality, nodes are expected to change their slices in response to observed failures or changes in how much they trust other nodes. It is nevertheless interesting to prove that, under the assumption that well-behaved nodes do not change their slices, SCP is safe and live.

Another issue is that, as we have noted in Section 2, FBQSs have the peculiar property that, by crafting the slices they advertise, malicious nodes can dynamically influence a well-behaved node's notion of quorum. But in our model, the quorums of a well-behaved node are fixed. This is in fact a form of abstraction. Given a FBQS where well-behaved nodes have fixed slices and Byzantine nodes can advertise arbitrary slices, we defined its fixed-quorums counterpart, where we assign to each well-behaved node the set of all quorums that could

```
type node # the type of nodes; this is an uninterpreted, arbitrary non-empty type
type nset # the type of node sets
relation well_behaved(N:node)
relation intertwined(N:node)
relation intact(N:node)
relation member(N:node, S:nset) # this is the set membership relation
relation quorum_of(N:node, Q:nset)
relation slice_blocking(S:nset, N:node)
axiom ∀ N. (intact(N) → intertwined(N)) ∧ (intertwined(N) → well_behaved(N))
axiom (∃ N . well_behaved(N) ∧ quorum_of(N,Q))
     → ∃ N . well_behaved(N) ∧ member(N,Q)
definition quorum_of_intertwined(Q) =
  (∃ N. intertwined(N) ∧ quorum_of(N,Q))
definition quorum_of_intact(Q) =
 (∃ N. intact(N) ∧ quorum_of(N,Q))
axiom ∀ Q1,Q2. quorum_of_intertwined(Q1) ∧ quorum_of_intertwined(Q2)
  → ∃ N. intertwined(N) ∧ member(N,Q1) ∧ member(N,Q2)
axiom ∀ Q1,Q2. quorum_of_intact(Q1) ∧ quorum_of_intact(Q2)
  → ∃ N. intact(N) ∧ member(N,Q1) ∧ member(N,Q2)
axiom ∀ S. (∃ N. intact(N) ∧ slice_blocking(S,N))
  → ∃ N2. member(N2,S) ∧ intact(N2)
axiom ∃ Q. ∀ N. member(N,Q) ↔ intact(N) ∧ quorum_of(N,Q)
```

Figure 1 A model of Federated Byzantine Quorum Systems in the EPR fragment of first-order logic.

possibly arise given arbitrary malicious behavior of Byzantine nodes. Intuitively, in this new fixed-quorums model, it is harder to achieve consensus because the Byzantine adversary has more choices of quorums to manipulate. Thus a consensus algorithm that works in the fixed-quorums model will work in the model in which quorums can be dynamically shaped by Byzantine nodes.

Finally, we prove in Isabelle/HOL that the fixed-quorums model satisfies all the properties axiomatized in the first-order model; thus the first-order model is an abstraction of the fixed-quorums model. We now describe the Isabelle/HOL fixed-quorums model.

The Isabelle/HOL model formalizes FBQSs from the notion of slice. It assumes that well-behaved nodes have fixed slices, but it accounts for the situation in which malicious nodes dynamically shape the quorums of well-behaved nodes. To do so, we define a quorum Q of a node n as a set of nodes such that a) n has a slice included in Q and b) every *well-behaved* member of Q has a slice included in Q. Note how this definition of quorum subtly differs from the one of Section 2. By placing requirements only on well-behaved nodes, we account for any possible slices that could be advertised by malicious nodes. We then prove in Isabelle/HOL that all the axioms of the first-order model (Figure 1) and the Cascade Theorem (Figure 2) hold. This Isabelle/HOL theory is purely definitional (i.e. it does not use axioms).

There is no mechanically-checked connection between Isabelle/HOL and Ivy, and thus the best we can do is to carefully check, by hand, that the Ivy axioms correspond to the properties proved in Isabelle/HOL. Fortunately, the syntax and semantics of first-order formulas in Isabelle/HOL is very close to that of Ivy. This can be seen by comparing the Ivy

```
axiom [cascade_thm] {
  function p(N:node):bool
  property (∃ Q . quorum_of_intact(Q) ∧ (∀ N . intact(N) ∧ member(N,Q) → p(N)))
    → ((∀ N . intact(N) → p(N))
        ∨ (∃ N,S . intact(N) ∧ ¬p(N) ∧ slice_blocking(S,N)
           ∧ (∀ N2 . member(N2,S) → (intact(N2) ∧ p(N2)))))
}
```

Figure 2 The second-order Cascade Theorem as an axiom schema in Ivy.

axiom schema of Figure 2 with its Isabelle/HOL counterpart appearing in Figure 3.

```
theorem cascade:
  fixes P
  assumes "∃ Q . ∃ n . intact n ∧ quorum_of n Q ∧ (∀ n ∈ Q . intact n ⟶ P n)"
  obtains "∀ n . intact n ⟶ P n" | "∃ n S . intact n ∧ ¬P n
    ∧ (∀ S1 ∈ slices n . S ∩ S1 ≠ {}) ∧ (∀ n ∈ S. intact n ∧ P n)"
```

Figure 3 The Cascade Theorem in Isabelle/HOL.

4 Related Work

Lokhava et al [7] discuss the Stellar Network in the broader context of global payments; they also describe at a high level the formal verification effort that is the subject of the present paper. The purpose of the present paper is to dig into the technical details necessary to apply this technique to future proofs of BFT protocols. Losa et al. [9] show that FBQSs are an instance of the more general Personal Byzantine Quorum System model, and we reuse some of the Isabelle/HOL theories developed for this work.

Other works verify safety properties of BFT consensus protocols using Dafny, Coq, or Isabelle/HOL. For example, Alturki et al. verify safety properties of Algorand in Coq [2]. Palmskog et al.[15] and Nakamura et al.[11] verify properties of Ethereum's Casper CBC in Coq and Isabelle/HOL, respectively. Rahli verifies safety properties of PBFT in the Velisarios framework [16], which is based on Coq. IronFleet [6] verifies safety and liveness of a crash-tolerant implementation of Multi-Paxos using Dafny.

Isabelle/HOL, Dafny, and Coq are not restricted by decidable logics, but they lack the specific features that allow Ivy users to restrict verification conditions to a decidable fragment and in turn benefit from reliable proof automation. A series of papers describe the different aspects of decidable reasoning about protocols in Ivy. [14] focuses on modeling and safety verification of consensus protocols at a high level of abstraction. [3] presents a tool to synthesize first-order axioms modeling threshold-based quorum systems.[17] present Ivy's modularity features, which enable decidable safety verification of more complex protocols and their implementations. Finally, Ivy's liveness-to-safety reduction [12] allows decidable reasoning about liveness properties expressed in LTL. Ivy's support for prophecy variables [13] offers an additional tool that helps preserve decidability. In an extended version of this paper, we plan to present the Ivy proofs of safety and liveness of SCP and compare with the works cited above.

References

1 Ittai Abraham, Guy Gueta, Dahlia Malkhi, Lorenzo Alvisi, Rama Kotla, and Jean-Philippe Martin. Revisiting fast practical byzantine fault tolerance. *arXiv preprint arXiv:1712.01367*, 2017.

2 Musab A Alturki, Jing Chen, Victor Luchangco, Brandon Moore, Karl Palmskog, Lucas Peña, and Grigore Roşu. Towards a verified model of the algorand consensus protocol in coq. *arXiv preprint arXiv:1907.05523*, 2019.

3 Idan Berkovits, Marijana Lazić, Giuliano Losa, Oded Padon, and Sharon Shoham. Verification of Threshold-Based Distributed Algorithms by Decomposition to Decidable Logics. In Isil Dillig and Serdar Tasiran, editors, *Computer Aided Verification*, Lecture Notes in Computer Science, pages 245–266, Cham, 2019. Springer International Publishing.

4 Cynthia Dwork, Nancy Lynch, and Larry Stockmeyer. Consensus in the presence of partial synchrony. *Journal of the ACM (JACM)*, 35(2):288–323, 1988.

5 Yeting Ge and Leonardo De Moura. Complete instantiation for quantified formulas in satisfiabiliby modulo theories. In *Computer Aided Verification*, pages 306–320. Springer, 2009.

6 Chris Hawblitzel, Jon Howell, Manos Kapritsos, Jacob R. Lorch, Bryan Parno, Michael L. Roberts, Srinath Setty, and Brian Zill. IronFleet: Proving Practical Distributed Systems Correct. In *Proceedings of the 25th Symposium on Operating Systems Principles*, SOSP '15, pages 1–17, New York, NY, USA, 2015. ACM.

7 Marta Lokhava, Giuliano Losa, David Mazières, Graydon Hoare, Nicolas Barry, Eli Gafni, Jonathan Jove, Rafał Malinowsky, and Jed McCaleb. Fast and secure global payments with Stellar. In *Proceedings of the 27th ACM Symposium on Operating Systems Principles*, SOSP '19, pages 80–96, New York, NY, USA, October 2019. Association for Computing Machinery.

8 Giuliano Losa. `https://github.com/stellar/scp-proofs`, 2019.

9 Giuliano Losa, Eli Gafni, and David Mazières. Stellar Consensus by Instantiation. In Jukka Suomela, editor, *33rd International Symposium on Distributed Computing (DISC 2019)*, volume 146 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 27:1–27:15, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

10 Kenneth L. McMillan and Oded Padon. Deductive verification in decidable fragments with Ivy. In *International Static Analysis Symposium*, pages 43–55. Springer, 2018.

11 Ryuya Nakamura, Takayuki Jimba, and Dominik Harz. Refinement and Verification of CBC Casper. In *2019 Crypto Valley Conference on Blockchain Technology (CVCBT)*, pages 26–38, June 2019.

12 Oded Padon, Jochen Hoenicke, Giuliano Losa, Andreas Podelski, Mooly Sagiv, and Sharon Shoham. Reducing Liveness to Safety in First-Order Logic. In *45th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL 2018)*, Los Angeles, 2018.

13 Oded Padon, Jochen Hoenicke, Kenneth L. McMillan, Andreas Podelski, Mooly Sagiv, and Sharon Shoham. Temporal prophecy for proving temporal properties of infinite-state systems. In *2018 Formal Methods in Computer Aided Design (FMCAD)*, pages 1–11. IEEE, 2018.

14 Oded Padon, Giuliano Losa, Mooly Sagiv, and Sharon Shoham. Paxos Made EPR: Decidable Reasoning About Distributed Protocols. *Proc. ACM Program. Lang.*, 1(OOPSLA):108:1–108:31, October 2017.

15 Karl Palmskog, Milos Gligoric, Brandon Moore, Lucas Peña, and Grigore Roşu. Verification of Casper in the Coq Proof Assistant, 2018. URL: `http://hdl.handle.net/2142/102075`.

16 Vincent Rahli, Ivana Vukotic, Marcus Völp, and Paulo Esteves-Verissimo. Velisarios: Byzantine Fault-Tolerant Protocols Powered by Coq. In Amal Ahmed, editor, *Programming Languages and Systems*, Lecture Notes in Computer Science, pages 619–650, Cham, 2018 . Springer International Publishing.

17 Marcelo Taube, Giuliano Losa, Kenneth L. McMillan, Oded Padon, Mooly Sagiv, Sharon Shoham, James R. Wilcox, and Doug Woos. Modularity for decidability of deductive verification with applications to distributed systems. In *Proceedings of the 39th ACM SIGPLAN Conference on Programming Language Design and Implementation*, PLDI 2018, pages 662–677, New York, NY, USA, June 2018. Association for Computing Machinery.

Formal Specification and Model Checking of the Tendermint Blockchain Synchronization Protocol

Sean Braithwaite
Informal Systems, Lausanne, Switzerland

Ethan Buchman
Informal Systems, Toronto, Canada

Igor Konnov
Informal Systems, Wien, Austria

Zarko Milosevic
Informal Systems, Lausanne, Switzerland

Ilina Stoilkovska
Informal Systems, Wien, Austria

Josef Widder
Informal Systems, Wien, Austria

Anca Zamfir
Informal Systems, Lausanne, Switzerland

Abstract

Blockchain synchronization is one of the core protocols of Tendermint blockchains. In this short paper, we discuss our recent efforts in formal specification of the protocol and its implementation, as well as some initial model checking results. We demonstrate that the protocol quality and understanding can be improved by writing specifications and model checking them.

2012 ACM Subject Classification Software and its engineering → Software verification; Software and its engineering → Software fault tolerance; Theory of computation → Logic and verification

Keywords and phrases Blockchain, Fault Tolerance, Byzantine Faults, Model Checking

Digital Object Identifier 10.4230/OASIcs.FMBC.2020.10

Category Short Paper

Supplementary Material `https://github.com/tendermint/spec/tree/master/rust-spec/fastsync`

Funding Supported by Interchain Foundation (Switzerland).

1 Introduction

Tendermint is a state-of-the art Byzantine-fault-tolerant state machine replication (BFT SMR) engine equipped with a flexible interface supporting arbitrary state machines written in any programming language [2]. Tendermint is particularly popular for proof-of-stake blockchains, and constitutes a core component of the Cosmos Project [3]. At the heart of the Cosmos Project is the InterBlockchain Communication (IBC) protocol for reliable communication between independent BFT SMs; what TCP is for computers, IBC aims to be for blockchains.

Multiple Tendermint-based blockchains currently run in production on the public Internet and have for over a year, with new ones launching regularly, carrying billions of dollars of cumulative value in the market capitalizations of their respective cryptocurrencies. One of the primary deployments, the so-called Cosmos Hub blockchain, is operated by a diverse set of 125 consensus forming nodes connected to one another over an open-membership gossip network consisting of hundreds of other nodes.

Tendermint was the first to apply traditional BFT consensus protocols to blockchain systems [9]. The core Tendermint BFT consensus protocol constitutes a modern implementation of the consensus algorithm for Byzantine faults with Authentication from [6] built on top of an efficient gossiping layer. The latest description of the consensus protocol can be found in the technical report [4]. Tendermint consensus has been a source of inspiration for a wide variety of blockchain systems that have followed [15, 5], though few, if any, have achieved its level of maturity in production.

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler; Article No. 10; pp. 10:1–10:8
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

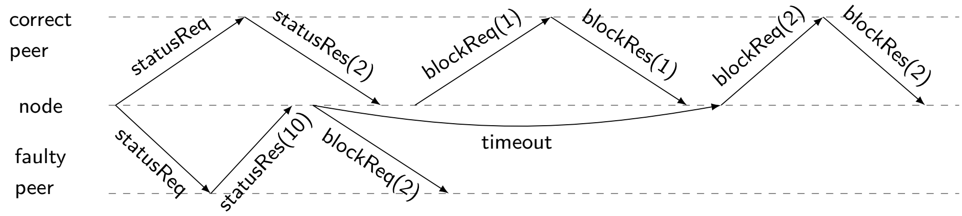

Figure 1 A Fastsync execution for a fully unsynchronized node of height 1.

The reference implementation of the Tendermint software is written in Go [14]. Under the hood, it consists of several fault-tolerant distributed protocols that interact to ensure efficient operation:

Consensus. Core BFT consensus protocol including the gossiping of proposals, blocks, and votes.

Evidence. To incentivize consensus participants to follow the consensus protocol (and not behave faulty), in the proof-of-stake systems, misbehavior is punished by destroying stake. This protocol gossips evidence of malicious behavior in the form of conflicting signatures.

Mempool. A protocol to gossip transactions, ensuring transactions eventually end up in a block are distributed to all participants.

Peer Exchange. Gossiping is based on communication only with a subset of the peers. Managing a list of available peers and selecting the peers based on performance metrics is done by this protocol.

Blockchain synchronization (Fastsync). If a peer gets disconnected by the network for some time, it might miss the most recent blocks in the blockchain. A node that recovers from such a disconnection uses the blockchain synchronization protocol called Fastsync to learn blocks without going through consensus.

We are conducting a project to formally specify and model check these protocols. The first protocol we considered was the blockchain synchronization protocol called *Fastsync*. Specifications can be found in English [10] and TLA$^+$ [11].

Fastsync. A full node that connects to a Tendermint blockchain needs to synchronize its state to the latest global state of the network. One way to achieve this is to update its local copy of the blockchain and replay all transactions, using Fastsync: initially, the node has a local copy of a blockchain prefix and the corresponding application state that may be out of date. The node queries its peers for the blocks that were decided on by the Tendermint blockchain since the time the full node was disconnected from the system. After receiving these blocks, the protocol executes the transactions that are stored in the blocks, in order to synchronize to the current height of the blockchain and the corresponding application state.

Figure 1 shows a typical execution of the Blockchain Synchronization protocol. In this execution, a new node connects to two full nodes: a correct peer and a faulty peer. The node requests the blockchain heights of the peers by issuing `statusReq`. Once a peer replies with its height, e.g., with `statusRes(10)`, the node can request for a block i by sending the message `blockReq(i)`. In our example, the correct peer receives the request `blockReq(1)` for block 1 and replies with the message `blockRes(1)` that contains the block. In a Tendermint blockchain, the commit for block (signed votes messages) `h` is contained in block `h+1`, and thus a node performing Fastsync must receive two sequential blocks before it can verify fully

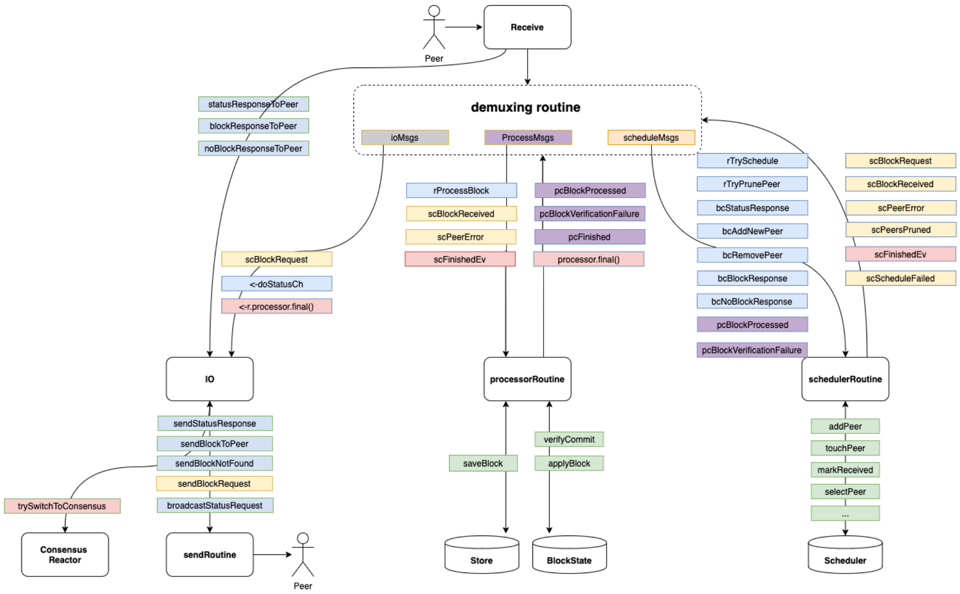

Figure 2 Concurrent threads of execution in the Fastsync implementation [13].

the first one. If verification succeeds, the first block is accepted; if it fails, both blocks are rejected, since it is not known which block was faulty. When the node rejects a block, it suspects the sending peer of being faulty and evicts this peer from the set of peers. The same happens when a peer does not reply within a predefined time interval. In our example, the faulty peer is evicted, and the node finishes synchronization with the correct peer.

The above example may produce an impression that it is easy to specify and verify correctness of Fastsync. (The authors of this paper thought so.) By writing several protocol specifications in English and TLA$^+$ and by running model checkers, we have found that the specifications, in particular in the presence of faulty peers, are intricate.

2 Architecture

The most recent implementation of the Fastsync protocol, called V2, is the result of significant refactoring to improve testability and determinism, as described in the Architectural Decision Record [13]. In the original design, a go-routine (thread of execution) was spawned for each block requested, and was responsible for both protocol logic and network IO. In the V2 design, protocol logic is decoupled from IO by using three concurrent threads of execution: a scheduler, a processor, and a demuxer, as per Figure 2. Rounded-corner rectangles represent concurrent threads that exchange the events that are depicted by rectangles on the edges between threads.

The demuxer acts as an internal bridge: it is responsible for translating between internal events and network IO messages, and for routing events between components. Both the scheduler and processor are structured as finite state machines with input and output events. Input events are received on an unbounded priority queue, with higher priority for error events. Output events are emitted on a blocking, bounded channel. Network IO is handled by the Tendermint p2p subsystem, where messages are sent in a non-blocking manner.

The IO component is responsible for exchanging (sending and receiving) Fastsync protocol messages with peers. There is one send and one receive routine per peer (denoted *Receive* and *sendRoutine* on Figure 2, respectively).

The scheduler contains the business logic for tracking peers and determining which block to request from whom. The scheduler receives relevant protocol messages from peers (for example *bcBlockResponse* and *bcStatusResponse*), but also internal events that are the result of the block processing in the processor (the events carry the information of whether a block was successfully processed or there was an error). The scheduler schedules block requests by emitting internal events (*scBlockRequest*) and also informs the processor about internal processing, for example, when block response is received (*scBlockReceived*) or if there is an error in peer behaviour (*scPeerError*).

The processor handles the computationally expensive block processing, including verification of consensus signatures and execution of all transactions. It manages the block store (denoted *Store* on Figure 2) and interacts with the Tendermint block execution component (denoted *BlockState*). Furthermore, it informs the scheduler whether a block processing was successful (*pcBlockProcessed*) or it has led to an error (*pcBlockVerificationFailure*).

Once the Fastsync protocol terminates, this is signaled to the Tendermint consensus component (denoted *ConsensusReactor*) with a *trySwitchToConsensus* event.

3 Specifications in English and TLA$^+$

Structured Specification in English. We start our formalization by a structured English specification [10], that consists of four parts:

1. *Blockchain.* Formalization of relevant properties of the blockchain and its blocks.
2. *Sequential problem statement.* Parts of the sequential safety specification read as follows:

   > "Let bh be the height of the blockchain at the time Fastsync starts. When the protocol terminates, it outputs a list of all blocks from its initial height to some height $terminationHeight \geq bh - 1$". (Fastsync cannot synchronize to the maximum height bh as in Tendermint, verification of block at height h requires the commit from the block at height $h + 1$.)

   This specification is sequential, as it ignores that the blockchain is implemented in a distributed system, in which validators may be faulty. Even if they are correct, they locally have prefixes of different lengths, so that bh is not clearly defined a priori.
3. *Distributed aspects.* Here we introduce the computational model and the refinement of the problem statement. For instance, the above translates to:

   > "Let $maxh$ be the maximum height of a correct peer to which the node is connected at the time Fastsync starts. If the protocol terminates successfully, it is at some height $terminationHeight \geq maxh - 1$."
4. *Distributed protocol.* Specification of the protocol, where we describe inputs, outputs, variables, and functions used by the protocol. We specify functions mainly by preconditions, postconditions, and error conditions. Further, we provide invariants over the protocol variables. These inform both the implementation and the verification effort.

Specifications in TLA$^+$. The structure of the English specification already highlighted interesting properties of the protocols and pointed to some issues. As it is written in natural language, the English specification is ambiguous. To eliminate the ambiguities, we have written three TLA$^+$ specifications, which focus on different aspects of the protocol and its architecture:

- *High-level specification (HLS).* This specification contains the minimal set of interactions in the synchronization protocol. Its primary purpose is to highlight safety and termination. HLS was mainly written by the researchers.
- *Low-level specification (LLS).* While HLS captures the distributed protocol, there was a significant gap between HLS and the implementation. For instance, the implementation uses additional messages and contains detailed error codes, which are missing in HLS. The low-level specification is much closer to the implementation, and it was mainly written by distributed system engineers.
- *Concurrency specification (CRS).* As discussed above, the V2 implementation utilizes several threads that communicate via queues. To formally capture this structure, we wrote a specification that models threads and message queues.

We discuss the modeling assumptions of HLS [11], which builds the basis for our model checking work discussed in Section 4. (1) The node starts with a finite set of peers, which can shrink, when the node suspects peers of being faulty. This set is partitioned in two subsets: correct and faulty. (2) The blockchain can grow up to a fixed height. By fixing the parameters of (1) and (2), we run finite-state model checking with TLC [8] and APALACHE [7, 1]. We model the distributed system as two components: the node and its peers. These two components communicate via two variables: `outMsg` that keeps an output message from the node to a peer, and `inMsg` that keeps an input message from a peer to the node; both variables may be set to `None`, indicating that there is no message. The components alternate their steps: The odd turns belong to the node, whereas the even turns belong to the peers.

This approach is simple but powerful. On one hand, it dramatically decreases the state space, as there are no queues, and alternation produces fewer states than disjunction. On the other hand, it does not decrease precision, as the peers consume and produce message at a high degree of non-determinism. Moreover, this approach allows us to easily formulate fairness in the system as weak fairness over the variable `turn`, which encodes the scheduled component.

Finally, V2 relies on several timeouts to guarantee termination. In TLA$^+$ specifications, we simply model timeouts with non-determinism and weak fairness.

4 Model Checking with TLC and Apalache

While developing TLA$^+$ specifications, we were using TLA$^+$ Toolbox and the TLC model checker [8]. We also checked the safety properties with the new symbolic model checker APALACHE [7, 1]. So far, we have checked the specifications for tiny parameters such as 1–3 peers and small Blockchain heights[1]. Table 1 summarizes the results and running times of TLC and APALACHE. A central property is the protocol's termination:

$$\mathrm{WF}_{turn}(\mathit{FlipTurn}) \Rightarrow \Diamond(\mathit{state} = \text{“}\mathit{finished}\text{”}) \qquad \text{(Termination)}$$

[1] The original protocol specification in TLA$^+$ is available in the main branch: `https://github.com/informalsystems/tendermint-rs/tree/master/docs/spec/fastsync`. The refined protocol specification was located in a pull request at the moment of writing (July 23, 2020): `https://github.com/informalsystems/tendermint-rs/pull/466`. After peer-review of the updates, the refined specification will be merged into the main branch.

In the early experiments, we did not find violations of this property, as TLC did not finish its exhaustive search. However, later, after refining the specification, TLC reported a counterexample at depth 7, which is indicated with $[\text{✗}]_{=7}$ in Table 1. The refined termination property looks as follows:

$$\Big(\mathrm{WF}_{turn}(FlipTurn) \\ \wedge \diamond(inMsg.type = \text{“}syncTimeout\text{”} \wedge blockPool.height \leq blockPool.syncHeight)\Big) \\ \Rightarrow \diamond(state = \text{“}finished\text{”}) \qquad (\mathsf{TerminationByTO})$$

A global timeout guarantees that the protocol terminates, no matter what happens. TLC did not report liveness violations of TerminationByTO up to depth 14, which is indicated with $[\text{✓}]_{<14}$ in Table 1. However, the TLC search was not exhaustive, as we have terminated the model checker after 24 hours. We did not check liveness with APALACHE, as it only supports safety at the moment. The more interesting property is "synchronization", whose intuitive meaning is that when Fastsync terminates, it has reached the height of the blockchain. Let us formalize this as Sync1: To see that our modeling is precise, let us start with a property we know to be slightly wrong, namely, when the protocol finishes, it reaches the maximum height among the heights of the correct peers, i.e.,

$$state = \text{“}finished\text{”} \Rightarrow blockPool.height \geq MaxCorrectPeerHeight(blockPool) \qquad (\mathsf{Sync1})$$

The model checkers report counterexamples. One reason is that to verify a block h, one needs the commit signatures from block $h+1$. We also observe that we do not require that the node that runs Fastsync needs to be connected to correct peers. Hence, we fix it in Sync2 by stating that height $MaxCorrectPeerHeight(blockPool) - 1$ should be reached when the node is connected to correct peers. This property also fails. This time we observe that a global timeout – that guarantees Termination – may terminate Fastsync before it has reached the maximal height. We add a precondition for "no timeout", and call the property Sync3. Neither TLC, nor APALACHE produce a counterexample up to computation depth 15 and 21, respectively.

The following two properties might appear to be intuitively correct, but the model checkers produce counterexamples. SyncFromCorrect states that the accepted blocks originate only from the correct processes. This property fails, as it does not consider that faulty peers may behave correct in an execution prefix (before they show faulty behavior). Thus, the initial intuition fails. CorrectNeverSuspected states that the correct peers are never removed from the peer set. This would be a desirable property, but the latest implementation V2 does not guarantee it.

5 Conclusions

We approach this work with a process-oriented goal in mind: By *Verification-Driven Development* [12] we understand a design process for distributed systems that makes it easier to test and verify the software. The re-design of the Fastsync protocol that resulted in a decomposition into state machines should be understood under this aspect. The design documents, namely the English and the TLA$^+$ specifications, are artifacts of this design process, and are means of communication between researchers, software engineers, and verification engineers. The structured English specification strikes a balance between mathematical rigor and readability. It serves as the base (i) for formal verification efforts in TLA$^+$ that

Table 1 Model checking results for TLC and APALACHE against the high-level specification for the following parameters: 1 correct peer, 1 faulty peer, 4 blocks (Apple MacBook Pro 2019). The symbols in "result" are: found a bug ✗, and no bug up to given length [✓].

Property	TLC (6 CPUs, 13 GB)			APALACHE (1 CPU)	
	result	time	#states	result	time
Sync1	$[✗]_{=5}$	28s	1.1M	$[✗]_{=5}$	16s
Sync2	$[✗]_{=5}$	28s	1.1M	$[✗]_{=5}$	16s
Sync3	$[✓]_{<15}$	TO	875M	$[✓]_{<21}$	1h25m
Termination	$[✗]_{=7}$	7m25s	2.9M	*not supported*	
TerminationByTO	$[✓]_{<14}$	TO	461M	*not supported*	
SyncFromCorrect	$[✗]_{=9}$	8m34s	4.5M	$[✗]_{=9}$	1m23s
CorrectNeverSuspected	$[✗]_{=3}$	4s	126K	$[✗]_{=3}$	9s

will give precise semantics, and (ii) for implementations. The annotations with invariants, preconditions, and postconditions are very helpful for the software engineers to guide the implementation.

The formalization also led to a better understanding of the liveness properties that we expect and want from blockchain synchronization protocols. It also improved the understanding of the discrepancies between the current implementations (Fastsync V0, V1, and V2). We have found several liveness issues that come from unexpected behavior of faulty peers. For instance, rather than reporting bad blocks, faulty peers may be very slow in reporting good blocks. If they report them slower than the blockchain grows, but fast enough to not lead to a timeout at the node, V2 may never terminate. This highlights that a vital requirement had not been explicitly captured before, namely, a relationship between timeout duration, block generation rate, and message end-to-end delays. We made these requirements explicit as part of the English specification; they constitute timing assumptions upon which the protocol is based. As this is closely related to real-time, we are not able to capture and reproduce this with TLA$^+$. However, TLA$^+$ counterexamples and the English specifications helped us in isolating this scenario.

For safety verification, we can replace a timeout by a non-deterministic event that may occur at any time. For liveness we have to treat the relation of timeouts to message delays and processing times precisely. The extensive use of timeouts in the current implementation poses a challenge to liveness verification. Some of our current research challenges are therefore timeouts, and we are interested in answering the following questions: How to limit timeouts in the implementation? What is the most effective way to use timeouts in the implementation in order to stay precise in the verification? How can we capture the relation of the (local) timeouts to (global) message delays in model checking? We keep these challenges for future work.

Model checkers and the produced counterexamples were quite helpful in understanding and refining the protocol properties. After refining the protocol, which results in larger state space, we found that TLC could not reach error states within the reasonable time frame of one hour. However, APALACHE was still finding errors within 10 minutes, which was still interactive enough for us. As future work, we also plan to find an inductive invariant and prove its correctness with APALACHE (for fixed but larger parameters).

References

1 APALACHE: a symbolic model checker for TLA$^+$. https://github.com/konnov/apalache/, 2020. Last accessed: May 20, 2020.

2 Ethan Buchman. Tendermint: Byzantine fault tolerance in the age of Blockchains. Master's thesis, University of Guelph, 2016. URL: http://hdl.handle.net/10214/9769.

3 Ethan Buchman and Jae Kwon. Cosmos whitepaper: a network of distributed ledgers, 2016. URL: https://cosmos.network/resources/whitepaper.

4 Ethan Buchman, Jae Kwon, and Zarko Milosevic. The latest gossip on BFT consensus. *arXiv preprint arXiv:1807.04938*, 2018. URL: https://arxiv.org/abs/1807.04938.

5 Vitalik Buterin and Virgil Griffith. Casper the friendly finality gadget. *arXiv preprint arXiv:1710.09437*, 2017.

6 Cynthia Dwork, Nancy Lynch, and Larry Stockmeyer. Consensus in the presence of partial synchrony. *J.ACM*, 35(2):288–323, 1988.

7 Igor Konnov, Jure Kukovec, and Thanh-Hai Tran. TLA+ model checking made symbolic. *PACMPL*, 3(OOPSLA):123:1–123:30, 2019.

8 Markus Alexander Kuppe, Leslie Lamport, and Daniel Ricketts. The TLA+ toolbox. In *F-IDE@FM 2019*, pages 50–62, 2019.

9 Jae Kwon. Tendermint: Consensus without mining. *Draft v. 0.6, fall*, 1(11), 2014.

10 Informal Systems. Fastsync – English specification, 2020. URL: https://github.com/informalsystems/tendermint-rs/blob/master/docs/spec/fastsync/fastsync.md.

11 Informal Systems. Fastsync – TLA$^+$ specification, 2020. URL: https://github.com/informalsystems/tendermint-rs/blob/master/docs/spec/fastsync/fastsync.tla.

12 Informal Systems. Verification-Driven Development: An Informal Guide, 2020. URL: https://github.com/informalsystems/VDD/blob/master/guide/guide.md.

13 ADR 043: Blockhchain reactor riri-org, 2020. URL: https://github.com/tendermint/tendermint/blob/master/docs/architecture/adr-043-blockchain-riri-org.md.

14 Tendermint core, reference implementation in Go, 2020. URL: https://github.com/tendermint/tendermint.

15 Maofan Yin, Dahlia Malkhi, Michael K Reiter, Guy Golan Gueta, and Ittai Abraham. Hotstuff: BFT consensus with linearity and responsiveness. In *PODC*, pages 347–356, 2019.

Inter-Blockchain Protocols with the Isabelle Infrastructure Framework

Florian Kammüller
Middlesex University London, UK
Technische Universität Berlin, Germany
f.kammueller@mdx.ac.uk

Uwe Nestmann
Technische Universität Berlin, Germany
firstname.secondname@tu-berlin.de

Abstract

The main incentives of blockchain technology are distribution and distributed change, consistency, and consensus. Beyond just being a distributed ledger for digital currency, smart contracts add transaction protocols to blockchains to execute terms of a contract in a blockchain network. Inter-blockchain (IBC) protocols define and control exchanges between different blockchains.

The Isabelle Infrastructure framework has been designed to serve security and privacy for IoT architectures by formal specification and stepwise attack analysis and refinement[1]. A major case study of this framework is a distributed health care scenario for data consistency for GDPR compliance. This application led to the development of an abstract system specification of blockchains for IoT infrastructures.

In this paper, we first give a summary of the concept of IBC. We then introduce an instantiation of the Isabelle Infrastructure framework to model blockchains. Based on this we extend this model to instantiate different blockchains and formalize IBC protocols. We prove the concept by defining the generic property of global consistency and prove it in Isabelle.

2012 ACM Subject Classification Networks → Peer-to-peer protocols; Networks → Security protocols; Software and its engineering → Software verification and validation

Keywords and phrases Blockchain, smart contracts, interactive theorem proving, inter-blockchain protocols

Digital Object Identifier 10.4230/OASIcs.FMBC.2020.11

Funding *Florian Kammüller*: Research for this paper has been supported by CHIST-ERA grant 101112, SUCCESS.

1 Introduction

Inter-blockchain (IBC) protocols is a concept driven by industry. It serves to provide "reliable and secure communication between deterministic processes" [24] that run on independent blockchains or distributed ledgers. Practical application of IBC are for example the Cosmos Hub [5] "the first of thousands of interconnected blockchains" with the purpose of facilitating transfers between blockchains.

A formal specification of IBC within a Higher Order Logic theorem prover like Isabelle has the advantage that it provides a very rigorous model of the IBC concepts enabling mechanically verified properties. In principle, from such a formalization, executable code into many standard programming languages like Haskell or Scala can be generated. However, such code generation would always be understood to provide only reference implementations. Moreover, the major insights from specifying a practice oriented concept like IBC is that

[1] In this paper we do neither illustrate attack tree analysis nor security refinement.

2nd Workshop on Formal Methods for Blockchains (FMBC 2020).
Editors: Bruno Bernardo and Diego Marmsoler; Article No. 11; pp. 11:1–11:12
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

the formal specification is mainly useful to provide a more abstract yet more precise model that carefully picks out the central concepts used within the application, here IBC. In doing this, the used methodology, here Isabelle, can provide as a framework existing work to immediately support the IBC specification. We rely heavily on the Isabelle Infrastructure framework [15] as an existing instantiation of Isabelle/HOL (which we will simply refer to as Isabelle within this paper). This framework offers a range of predefined concepts like Kripke structures and CTL, as well as state transition relations, actors, and policies that can be readily instantiated to the current application of IBC. Besides extracting a more abstract but precise specification of IBC, the resulting scientific advantage is to show that as a product of this process it becomes feasible to lay open crucial basic properties that result from the application domain (blockchain security). As the main result of this kind, we formally establish a global consistency property, define it formally on our IBC model and prove a consistency preservation theorem that shows the safety of our formal IBC semantics.

The contributions of this paper are

- summarizing the main features of IBC into a logical conceptual model,
- building a formal model of IBC in Isabelle as an instance of the Isabelle Infrastructure framework but extending it with sets of infrastructures,
- illustrating the feasibility of the formal model by expressing a global consistency property and formally proving it in Isabelle.

The last point seems to suggest that IBC can be seen as a "blockchain of blockchains".

1.1 Inter-blockchain protocols (IBC)

In this section, we summarize the main concepts of the IBC following the practice-oriented description [24]: we refer to the relevant section of the principal documentation[24], giving precise reference to section numbers. Figure 1 is a copy an overview architectural sketch provided by the main specification [24].

```
+-------------------------------+                             +-------------------------------+
| Distributed Ledger A          |                             | Distributed Ledger B          |
|                               |                             |                               |
| +--------------------------+  |                             | +--------------------------+  |
| | State Machine            |  |                             | | State Machine            |  |
| |                          |  |                             | |                          |  |
| | +----------+     +-----+ |  |        +---------+          | | +-----+     +----------+ |  |
| | | Module A | <-> | IBC | |  | <----> | Relayer | <---->   | | | IBC | <-> | Module B | |  |
| | +----------+     +-----+ |  |        +---------+          | | +-----+     +----------+ |  |
| +--------------------------+  |                             | +--------------------------+  |
+-------------------------------+                             +-------------------------------+
```

Figure 1 Architecture of IBC[24].

One of the main abstractions used in IBC comprising its architectural description is the *actor* [24, Section 1.1.1] which is the same as a *user*. Instances given to exemplify this are: a human end user, a module or smart contract running on a blockchain, or an off-chain *relayer* process. This relayer process represents the logical core of the IBC. It is a process that is outside any of the blockchains ("off-chain" [24]) that is responsible for "relaying" IBC data packets between blockchains. It can scan their states and submit data.

The notion of state machine is very central in IBC: the terms *machine*, *chain*, *blockchain*, or *ledger* are used interchangeably [24, Section 1.1.2] to denote a state machine that implements part or all of the IBC. In using the Isabelle Infrastructure framework – whose core part is the formal definition of a state machine semantics through a state transition relation – we follow this important architectural spirit.

Consensus is not explicitly defined but somewhat implicitly by the notion of consensus algorithm "the protocol used by the set of processes operating a distributed ledger to come to agreement on the same state" [24, 1.1.5] where "Consensus state" is defined next as information about the "state of a consensus algorithm" [24, 1.1.6]. We can safely understand consensus to mean the agreement of the actors on the next state with respect to the state transition relation.

1.2 Isabelle Infrastructure framework

The Isabelle Infrastructure is built in the interactive generic theorem prover Isabelle/HOL [19]. As a framework, it supports formalization and proof of systems with actors and policies. It originally emerged from verification of insider threat scenarios but it soon became clear that the theoretical concepts, like temporal logic combined with Kripke structures and a generic notion of state transitions were very suitable to be combined with attack trees into a formal security engineering process [3] and framework [10].

Figure 2 gives an overview of the Isabelle Infrastructure framework with its layers of object-logics – each level below embeds the one above showing the novel contribution of this paper in blue on the top. The formal model of IBC in Isabelle uses the Isabelle

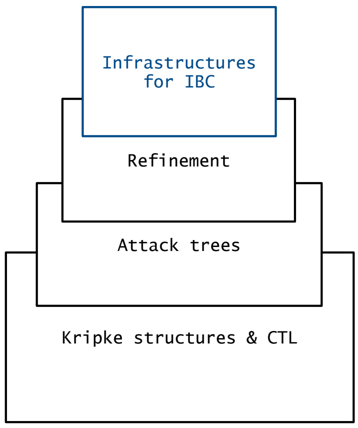

Figure 2 Generic Isabelle Infrastructure framework applied to Inter-blockchain protocols (IBC).

Infrastructure framework instantiating it by reusing its concept of *actors* for users, processes running on blockchains, or relayers running off-chain. Technically, an Isabelle theory file `IBC.thy` builds on top of the theories for Kripke structures and CTL (`MC.thy`), attack trees (`AT.thy`), and security refinement (`Refinement.thy`). Thus all these concepts can be used to specify the formal model for IBC, express relevant and interesting properties and conduct interactive proofs (with the full support of the powerful and highly automated proof support of Isabelle). The IBC theory itself is an adaptation of the Infrastructure theory of the Isabelle Infrastructure framework and reuses (or slightly adapts) existing concepts. In the remainder of this paper, we introduce the model that we conceived for IBC. All Isabelle sources are available online [12].

2 IBC in Isabelle

2.1 Overview

In the following, we give a detailed description of the central parts of the formal Isabelle theory of IBC, pointing out and motivating special design decisions. In addition to the short general intro to the Isabelle Infrastructure framework of the previous section, we provide explanations of all used Isabelle specific specification concepts on the fly.

The IBC is supposed to work for any type of blockchain, for example, Bitcoin or Ethereum, therefore the formal model abstracts from specific details of a specific blockchain. Similar to the IBC specification [24], the Isabelle formalization focuses on the central IBC concepts as depicted in Figure 1: ledgers, actors or modules, respectively, and the relayer process interacting via the IBC protocol with the modules within the distributed ledgers. In our formal model based on the Isabelle Infrastructure framework, we represent each blockchain as an infrastructure containing nodes on which the modules (actors) are running. Data items are assigned to actors. The ledgers of each infrastructure keep control over the data items. That is, a ledger is a unique assignment that controls the access to a data item and keeps a record of where the data item resides within this and other blockchains. The IBC enables just that: a unified view over a whole range of heterogeneous blockchains that exchange data consistently. Therefore, our formal model goes beyond the usual application of the Isabelle Infrastructure framework, e.g. [8], and considers sets of infrastructures (representing different blockchains).

2.2 Ledgers

Actors are a general concept provided by the Isabelle Infrastructure framework and can be used directly to represent the *actor* concept in IBC.

```
typedecl actor
type_synonym identity = string
consts Actor :: string ⇒ actor
```

Similar to the general Infrastructure framework, actors can perform actions. However, in this instantiation to IBC we redefine the actions representing the central activities of the relayer scanning each blockchain's state and submitting transactions (see Section 1.1).

```
datatype action = scan | submit
```

The Decentralized Label Model (DLM) [17] allows labeling data with owners and readers. We also adopt this definition of security labeled data as already formalised in [10]. Labeled data is given by the type `dlm × data` where `data` can be any data type.

```
type_synonym data = string
type_synonym dlm = identity × identity set
```

One major achievement of a blockchain is that it acts like a distributed ledger, that is, a global accounting book. A distributed ledger is a unique consistent transcript keeping track of protected data across a distributed system. In our application, the ledger must mainly keep track of where the data resides for any labeled data item. To express the system requirement that processing may not change the security and privacy labels of data, we introduce a type of security and privacy preserving functions.

```
typedef label_fun = {f :: dlm × data ⇒ dlm × data.
                      ∀ x. fst x = fst (f x)}
```

We formalize a ledger thus as a type of partial functions that maps a data item to a pair of the data's label and the set of locations where the data item is registered. Since all function in HOL are total, we use a standard Isabelle way of representing partial functions using the type constructor `option`. This type constructor lifts every type α to the type α `option` which consists of the unique constant `None` and the range of elements `Some x` for all `x`$\in \alpha$.

```
type_synonym ledger = data ⇒ (dlm × node set)option
```

Since the type `ledger` is a function type, it automatically constrains each data item `d` in its domain to have at most one range element `Some(l,N)`, that is, at most one valid data label `l` of type `dlm` and a list of current blockchain nodes `N` at which this data item is transcribed.

```
lemma ledger_def_prop: ∀ lg:: ledger. ∀ d:: data.
     lg d = None | (∃! l. (∃! L. lg d = Some(l, L)))
```

In an earlier application of the Isabelle Infrastructure framework to IoT security and privacy[15], we established a formal notion of blockchain. However, there we used a more explicit logical characterization in an Isabelle type definition which creates additional proof effort and makes formulas more complex. The current representation of the ledger type as a partial function type is more concise and implicitly carries the requested uniqueness properties. Note that the defining property of the ledger type is now proved from the used type constructors by the above lemma instead of being specified into the type as in the earlier formalization [15].

2.3 Infrastructures as blockchains

The datatype `sc_fun` formalizes any action that is sent or received between different blockchains and may have effects on the labeled data. Therefore the inputs to the send and receive messages are two identities of sender and receiver as well as the dlm label and the concerned data.

```
datatype sc_fun = Send identity × identity × dlm × data
                | Receive identity × identity × dlm × data
```

In addition to specifying the potential types of smart contracts, we need to provide a way of keeping track of the transactions that are executed within a blockchain. To this end, we define the following type of `transaction_record` which is a list of all executed smart contracts.

```
type_synonym transaction_record = sc_fun list
```

The central component that builds the system state is an infrastructure. Since we use the Isabelle Infrastructure framework, we consider blockchains as infrastructures. The essential architecture of such an infrastructure is a simple graph of blockchain nodes on which the processes (actors) reside given as the first component `(node ×node)set` of the below datatype `igraph`. Besides this basic architecture, this infrastructure graph also stores the other components of the blockchain. The second input is a function that assigns a set of actor identities to each node in the graph representing the current location of the actors. The next input associates actors to a pair of string sets by a pair-valued function whose first range component is a set describing the credentials in the possession of an actor and the second component is a set defining the roles the actor can take on. An infrastructure graph also allows assigning a string to each location to represent some current state information of that location. Finally, the ledger is added as a separate component as well as the transaction record.

```
datatype igraph =
         Lgraph (node × node)set
                node ⇒ node set
                actor ⇒ (string set × string set)
                node ⇒ string
                ledger
                transaction_record
```

Corresponding projection functions for each of the components of an infrastructure graph are provided. They are omitted here for brevity but are available in the online version [12]); they are named `gra` for the actual set of pairs of locations, `agra` for the actor map, `cgra` for the credentials, and `lgra` for the state of a location `ledgra` for the ledger component in the graph and `trec` for the transaction record. Infrastructures contain an infrastructure graph and a policy given by a function that assigns local policies over a graph to all locations of the graph.

```
datatype infrastructure =
  Infrastructure  igraph
                  [igraph, location] ⇒ apolicy set
```

There are projection functions `graphI` and `delta` when applied to an infrastructure return the graph and the policy, respectively.

Policies specify the expected behaviour of actors of an infrastructure. We define the behaviour of actors using a predicate `enables`: within infrastructure `I`, at location `l`, an actor `h` is enabled to perform an action `a` if there is a pair `(p,e)` in the local policy of `l` – `delta I l` projects to the local policy – such that action `a` is in the action set `e` and the policy predicate `p` holds for actor `h`.

```
enables I l h a = ∃ (p,e) ∈ delta I l. a ∈ e ∧ p h
```

Compared to the applications of the Isabelle Infrastructure framework, e.g. [8], we do not make use of policies to model the constraints of our application. However different to previous applications, the IBC challenges the framework in other ways leading to slight extensions.

2.4 Relayer and set of blockchains

To model the relayer, we also use infrastructures: the relayer is a distinguished infrastructure. It could be thought of as another distributed application with various relayer processes to avoid bottlenecks but for simplicity, we assume that there is one specific actor `''relayer''` that resides on a specific node in the relayer infrastructure.

We express protocols as traces of execution steps of IBC transaction steps, that is, lists of smart contracts `sc_fun` (see previous section). Using traces of execution steps to represent protocols, follows the classical method of the inductive approach to security protocol verification originally devised by Paulson [22] and already successfully used for the Isabelle Infrastructure framework, for example, [13] and more recently [11, 9].

```
datatype ibc_protocol = Protocol sc_fun list set
```

The datatype `blockchainset` puts together the IBC protocol as a triple: as the first element it includes the IBC protocol, the second element is the list of infrastructures where each element is one blockchain involved in the IBC, and the third element is a single distinguished infrastructure, the relayer.

```
datatype blockchainset = Infs ibc_protocol
                              infrastructure list
                              infrastructure
```

To round off these new datatypes, we provide additional projection functions and constructors. For a given blockchain `Il`, the projection `trcs Il` returns the `sc_fun list set` representing the protocol, the projection `the_Il` returns the list of infrastructures of all involved blockchains, and `relayer Il` gives the distinguished infrastructure, the third element, which is the relayer infrastructure. To facilitate handling of data transactions, we define some update functions: the function application `upd_ld d lN I` updates a ledger at the data point `d` to now contain the pair `lN` of a dlm label and a set of nodes of residences of the data. Scaling this up to the level of infrastructures, the function application `upd_Il d lN Il` updates all blockchains in the infrastructure list of the blockchainset `Il` using the former ledger update `upd_ld`. A function `replace` allows to replace an infrastructure `I` in a blockchainset `Il`. See the online resources [12] for technical details and implementations of these definitions.

2.5 Consensus

The consensus algorithm may be different for each blockchain employed in the IBC. Therefore, we cannot make any assumptions at the general specification level of the IBC about it. Yet, we still want to use it in the description of the IBC protocol semantics. Therefore, we apply a trick: we declare `Consensus` to be a constant at the level of the specification of the IBC.

```
consts Consensus :: infrastructure ⇒ blockchainset ⇒ blockchainset
```

In Isabelle this means that `Consensus` is a function mapping an infrastructure and a system state of type `blockchain` to `blockchain` but there is no semantics attached to this constant. The constant is part of the theory `IBC.thy` and can be used in it like any other defined element but it has no meaning. However, a semantics can be later attached to it in an application of the IBC theory to specific blockchains. This could be done in the current context for example using a definition in a locale [14].

```
locale ConsensusExample =
fixes cons_algo :: infrastructure ⇒ blockchainset ⇒ infrastructure
defines cons_algo_def: cons_algo I Il = ...
fixes Consensus :: infrastructure ⇒ blockchainset ⇒ blockchainset
defines Consensus_def: Consensus I Il = replace (cons_algo I Il) I Il
```

The predicate `Consensus` redefines the semantics within the locale `ConsensusExample`. The first locale definition is omitted here for simplicity. We could imagine that it is a description of a consensus algorithm that can depend on all the state constituents, like actors, nodes, and policies of the blockchain `I` but also of the surrounding blockchainset including the relayer state and the current protocol state. The definition of the constant `Consensus` lifts the algorithm to the blockchain by using the replace function defined as part of the infrastructure for blockchainsets (see Section 2.4 or refer to the Isabelle code [12]).

2.6 IBC state transition semantics

The semantics of the IBC state machines is defined by a state transition relation over blockchain sets. That is, we define a syntactic infix notation `Il → Il'` to denote that blockchain sets `Il` and `Il'` are in this relation.

```
inductive state_transition_in ::
       [blockchainset, blockchainset] ⇒ bool "(_ → _)"
```

The rules of the inductive definition `state_transition_in` allow the definition of the intended behaviour of the relayer scanning an arbitrary blockchain (see Section 1.1). The relayer stores the results in its own transaction record. The following rule `scan` is the first of two inductive definition rules defining the transition relation →: if an infrastructure `I` is in the blockchainset `Il`, the actor (process, module) resides at node `n` in the graph `G` of `I`; `R` is the relayer and thus enabled to scan. The follow up state `Il'` of `Il` is given by extending any current protocol trace `l` using the specially defined function `insertp` by the transaction `Send(a,b,(a,as), d)`. Also the relayer's trace record `trec R` is extended by the same transaction.

```
scan : inbc I Il ⟹ G = graphI I ⟹ a @G n ⟹ n ∈ nodes G ⟹
       R = graphI (relayer Il) ⟹ r @R n' ⟹ n' @R nodes R ⟹
       relrole (relayer Il) (Actor r) ⟹
       enables I n (Actor r) scan ⟹
       ledgra G d = Some ((a, as), N) ⟹ r ∈ as ⟹
       R' = Infrastructure
                 (Lgraph (gra R)(agra R)(cgra R)(lgra R)
                         ((ledgra R)(d := Some((a, as),N)))
                         (trec R))
                 (delta (relayer Il)) ⟹
       l ∈ trcs Il ⟹ Consensus I Il = Il' ⟹
       Il' = insertp ((Send(a,b,(a,as), d)) # l) (replrel R' Il)
       ⟹ Il → Il'
```

Additionally, the relayer can submit data onto an arbitrary blockchain (see Section 1.1). The second rule `submit` of → defines its semantics: between the infrastructures `I` and `J` which are both in the blockchain set `Il` the relayer `R` can submit data `d` from an owner `a` to an owner `b` if the ledger component `ledgra R` of the relayer's infrastructure `R` is updated to the new owner in both blockchains. The update is achieved using the function update `:=` of Isabelle's function theory updating the point `d` to the new value `Some((b, bs), N)`. In the construction of the next state blockchainset `Il'` the specially defined update operators mentioned in Section 2.4 are used: `replrel` for updating the relayer and `bc_upd` for the infrastructure list representing the "client" blockchains. Note the latter realizes the consistent update in both involved infrastructures `I` and `J`.

```
submit : G = graphI I ⟹ inbc I Il ⟹ a @G n ⟹ n ∈ nodes G ⟹
         ledgra G d = Some ((a, as), N) ⟹
         H = graphI J ⟹ inbc J Il ⟹ b @H n' ⟹ n' ∈ nodes H ⟹
         ledgra H d = Some ((a, as), N) ⟹
         R = graphI (relayer Il) ⟹ r @R n'' ⟹ n'' ∈ nodes R ⟹
         relrole (relayer Il) (Actor r) ⟹
         enables J n' (Actor r) submit ⟹
         r ∈ as ⟹
         R' = Infrastructure
                (Lgraph (gra R)(agra R)(cgra R)(lgra R)
                        ((ledgra R)(d := Some((b, bs),N)))
                        (trec R))
               (delta (relayer Il))  ⟹
         Il' = insertp (Receive(a,b,(a,as),d)# l)
                       (replrel R' (bc_upd d ((b,as), N) Il)) ⟹
```

```
          Consensus J Il = Il'
          ⟹ Il → Il'
```

The real advantage of the Isabelle Infrastructure framework comes into play when using the possibility of instantiation of axiomatic type classes provided by Isabelle. Since state transitions have been defined by an axiomatic type class in the framework within the theory for Kripke structures and CTL, we can now instantiate blockchainsets as state and thereby inherit the entire logic, constructors and theorems.

```
instantiation blockchainset :: state
```

3 Global consistency

To illustrate the use of the abstract formal model of IBC presented in this paper, we show that we can exhibit an important property: global consistency. That is, if the IBC scans and submits between blockchains it must not introduce inconsistencies.

Expressing this property alone represents a proof of concept since it shows that our IBC model is detailed enough to capture explicitly the notion of consistent data representation across different blockchains. *Proving* the property is a non-trivial contribution (see proof scripts [12]) that helped exhibiting a range of useful auxiliary definitions and lemmas as we will highlight in this section when discussing the global consistency theorem. The proofs were greatly helped by the recent advances in proof automation in Isabelle using sledgehammer [21]. The fact that the property is provable shows that the model and in particular its semantics conform to the intuition described in [24]. The formalization and proof also highlight the pros and cons of our model as discussed in the Conclusions in Section 4.

We first define global consistency as the property that the individual ledgers in each blockchain in an IBC blockchainset agree on the data, that is, they all hold consistent information about the access control of the data (the first part of type `dlm` of the `ledgra` output (see Section 2.2)) and where the data resides: the set of nodes that are the second component of the `ledgra` output.

```
Global_consistency Il = (∀ I I'. inbc I Il → inbc I' Il →
       (∀ d. (ledgra (graphI I') d) = (ledgra (graphI I) d)))
```

The theorem shows that if global consistency holds, then a step of the state transition does preserve it.

```
theorem consistency_preservation:
global_consistency Il ⟹ (Il → Il') ⟹ global_consistency Il'
```

Preservation of global consistency guarantees that any transaction happening within IBC preserves one consistent view over all data, their access control, and residence. If initially data is not visible on all blockchains, not all ledgers are equal. However, if eventually data has traveled across, all ledgers become the same: the blockchainset becomes like one blockchain: a "blockchain of blockchains".

4 Conclusions, related work, and outlook

In this paper, we have provided an abstract formal model of the Inter-blockchain protocol (IBC) [24] as an instantiation of the Isabelle Infrastructure framework. We have detailed the formal presentation in Isabelle and the extensions to the Isabelle Infrastructure framework, most notably by defining sets of (heterogeneous) blockchains including protocols and a distinguished relayer. The abstraction we conceived for this model has been first validated by a proof of concept by sketching how the abstract notion of Consensus can be instantiated by a locale (Section 2.5). Furthermore, we have defined a global consistency property over blockchainsets proving that our abstraction yields the desired expressivity (Section 3). We have proved a preservation theorem for global consistency in Isabelle. Summarizing, our model allows to prove meta-theoretical results but is not too abstract to allow instantiation onto concrete blockchains and their Consensus algorithms. As a more general thought, the dealings with global consistency seem to suggest that IBC creates a blockchain of blockchains.

4.1 Related Work

Relevant examples for the investigation of formal support for blockchains and smart contracts can be found in abundance in the proceedings of the first FMBC workshop [2]. We only discuss the few most closely related ones from there since others are either focusing on specific blockchains (unlike the generic IBC we consider) or are differing in the formal approach (not using theorem provers and thus not addressing the same level of expressivity and assurance).

A range of works formalizes smart contracts typical for the Ethereum virtual machine. For example, using the K framework [23], the Lem language [7], and F* [6]. We focus here on the work that has been performed in the K-framework [23]. The K-framework is a semantics framework enabling to produce executable operational semantics for programming languages. K also provides tools like parsers, interpreters, model-checkers and program verifiers. It has been applied to provide a verification environment for the Ethereum Virtual Machine EVM [20] which is useful for verifying programme modules within Ethereum's smart contract systems, for example, Ethereum's Name Service (ENS) [25].

In comparison to those dedicated verification environments for specific blockchains, like Ethereum, our formal model strongly abstracts from technical detail. This abstraction is necessary to accommodate a global view that allows to reason about the communication between a heterogeneous set of blockchains.

A few works use model checkers and SMT solvers, for example [4]. Deductive verification platforms like Why3 [11,13] have been also used for smart contracts. Interactive proof assistants (e.g. Isabelle/HOL or Coq) have been used before for modeling and proving properties about Ethereum and Tezos smart contracts [1].

Very related is the work by Nielsen and Spitters on Smart Contract Interactions in Coq [18]. The authors construct a model of smart contracts that allows for inter-contract communication generalizing over depth-first execution blockchains like Ethereum and breadth-first execution blockchains like Tezos. They use Coq's functional language Galina to express smart contracts. Besides the obvious difference of being a Coq development rather than an Isabelle development, we address the high level protocol language IBC instead of focusing on generalized smart contracts.

Maybe even more closely related is the work on the specification of the dedicated security framework Cap9 in Isabelle [16]. Compared to us it focuses again on the expression of smart contracts and does not have the inter-blockchain aspect like our IBC.

4.2 Outlook

The global consistency preservation theorem proves the concept of the IBC specification and also shows that the formalization in itself is a useful experiment: extracting a closed abstract model of the IBC from the technical specification [24] has immediately produced the consistency question. The abstraction allowed to define semantics in which a strong global consistency theorem could be proved within Isabelle in reasonably short time. It should be understood that these are first steps that mainly serve to prove the concept of using the Isabelle Infrastructure framework for advancing the IBC. A clear next step is to elaborate the sketched application example of Section 2.5 of a concrete blockchain and its consensus algorithm. A much more challenging next step is to refine the model by elaborating a more concrete IBC protocol example by instantiation of the `ibc_prot` component of the `blockchainset` type. This would be a fruitful future avenue for applied research in collaboration with the designers of IBC.

The notions of attack trees and security refinement have not been applied in this application of the Isabelle Infrastructure framework but can be seen in other applications, for example to auction protocols [13], GDPR [8], or IoT security [10]. Nevertheless, the current application has brought about much improvement on the formalization of the ledger datatype as well as instantiating the generic state of the framework to sets of infrastructures and defining their state transition.

The Isabelle Infrastructure framework subsumes the earlier Isabelle Insider framework, for example [13]. Thus there is the possibility to reason about malicious agents that are in the group of trusted participants. This could be used to reason about participants that do not comply to the IBC protocol and in terms of Consensus it would enable reasoning on Byzantine fault tolerance. Using attack tree analysis and security refinement in a security engineering cycle [15] could then be used to develop secure IBC solutions.

References

1. S. Amani, M. Bégel, M. Bortin, and M. Staples. Towards verifying ethereum smart contract bytecode in isabelle/hol. In *Proceedings of the 7th ACM SIGPLAN International Conference on Certified Programs and Proofs*, pages 66–77. ACM, 2018.
2. N. Catano, D. Marmsoler, and B. Bernardo, editors. *Pre-proceedings of the First Workshop on Formal Methods for Blockchains, FMBC*, 2019. Selected papers to appear in Springer LNCS. URL: `https://sites.google.com/view/fmbc`.
3. CHIST-ERA. Success: Secure accessibility for the internet of things, 2016. http://www.chistera.eu/projects/success.
4. Sylvain Conchon, Alexandrina Korneva1, and Fatiha Zaidi. Verifying smart contracts with cubicle, 2019. Selected papers to appear in Springer LNCS. URL: `https://sites.google.com/view/fmbc`.
5. Cosmos. Cosmos hub, 2020. accessed 23.1.2020. URL: `https://hub.cosmos.network/master/hub-overview/overview.html`.
6. I. Grishchenko, M. Maffei, and C. Schneidewind. A semantic framework for the security analysis of ethereum smart contracts. In L. Bauer and R. Ksters, editors, *Principles of Security and Trust*, Lecture Notes in Computer Science, pages 243–269. Springer, 2017.
7. Y. Hirai. Defining the ethereum virtual machine for interactive theorem provers. In M. Brenner, K. Rohloff, J. Bonneau, A. Miller, P. Y. Ryan, V. Teague, A. Braccialiand M. Sala, F. Pintore, and M. Jakobsson, editors, *Financial Cryptography and Data Security*, Lecture Notes in Computer Science, pages 520–535. Springer, 2017.
8. F. Kammüller. Formal modeling and analysis of data protection for gdpr compliance of iot healthcare systems. In *IEEE Systems, Man and Cybernetics, SMC2018.* IEEE, 2018.
9. F. Kammüller. Attack trees in isabelle extended with probabilities for quantum cryptography. *Computer & Security*, 87, 2019. URL: `//doi.org/10.1016/j.cose.2019.101572`.
10. F. Kammüller. Combining secure system design with risk assessment for iot healthcare systems. In *Workshop on Security, Privacy, and Trust in the IoT, SPTIoT'19, colocated with IEEE PerCom.* IEEE, 2019.

11 F. Kammüller. Qkd in isabelle – bayesian calculation. *arXiv*, cs.CR, 2019. URL: `https://arxiv.org/abs/1905.00325`.

12 F. Kammüller. Isabelle infrastructure framework for ibc, 2020. Isabelle sources for IBC formalisation. URL: `https://github.com/flokam/IsabelleSC`.

13 F. Kammüller, M. Kerber, and C.W. Probst. Towards formal analysis of insider threats for auctions. In *8th ACM CCS International Workshop on Managing Insider Security Threats, MIST'16*. ACM, 2016.

14 F. Kammüller, M. Wenzel, and L. C. Paulson. Locales – a sectioning concept for Isabelle. In Y. Bertot, G. Dowek, A. Hirschowitz, C. Paulin, and L. Thery, editors, *Theorem Proving in Higher Order Logics, 12th International Conference, TPHOLs'99*, volume 1690 of *LNCS*. Springer, 1999.

15 Florian Kammüller. A formal development cycle for security engineering in isabelle, 2020. `arXiv:2001.08983`.

16 Mikhail Mandrykin1, Jake O'Shannessy, Jacob Payne, and Ilya Shchepetkov. Formal specification of a security framework for smart contracts, 2019. Selected papers to appear in Springer LNCS. URL: `https://sites.google.com/view/fmbc`.

17 A. C. Myers and B. Liskov. Complete, safe information flow with decentralized labels. In *Proceedings of the IEEE Symposium on Security and Privacy*. IEEE, 1999.

18 J. Botsch Nielsen and B. Spitters. Smart contract interactions in coq, 2019. Selected papers to appear in Springer LNCS. URL: `https://sites.google.com/view/fmbc`.

19 T. Nipkow, L. C. Paulson, and M. Wenzel. *Isabelle/HOL – A Proof Assistant for Higher-Order Logic*, volume 2283 of *LNCS*. Springer-Verlag, 2002.

20 Daejun Park, Yi Zhang, Manasvi Saxena, Philip Daian, and Grigore Roundefinedu. A formal verification tool for ethereum vm bytecode. In *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, ESEC/FSE 2018, pages 912–915, New York, NY, USA, 2018. Association for Computing Machinery. `doi:10.1145/3236024.3264591`.

21 Lawrence Paulson and Jasmin Blanchette. Three years of experience with sledgehammer, a practical link between automatic and interactive theorem provers. *Proceedings of the 8th International Workshop on the Implementation of Logics*, February 2015. `doi:10.29007/tnfd`.

22 Lawrence C. Paulson. The inductive approach to verifying cryptographic protocols. *Journal of Computer Security*, 6(1-2):85–128, 1998. URL: `http://iospress.metapress.com/content/5wlu8p2am1du051d/`.

23 Grigore Rosu. Specifying languages and verifying programs with k. *2013 15th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*, pages 28–31, 2013.

24 The IBC Specification Team. The interblockchain communication protocol, 2020. 4th May 2020 — 1.0.0-rc5. URL: `https://github.com/cosmos/ics/blob/master/spec.pdf`.

25 Duy Minh Vo. *Verification of Smart Contracts using the K-framework*. PhD thesis, Technische Universität Berlin, 2018.