On Verifying a File System Implementation
Konstantine Arkoudas, Karen Zee, Viktor Kuncak, and Martin Rinard
On Verifying a File System Implementation

Konstantine Arkoudas, Karen Zee, Viktor Kuncak, Martin Rinard

MIT Computer Science and AI Lab
{arkoudas,kz,vkuncak,rinard}@lcs.mit.edu
MIT CSAIL Technical Report 946

Abstract. We present a correctness proof for a basic file system implementation. This implementation contains key elements of standard Unix file systems such as inodes and fixed-size disk blocks. We prove the implementation correct by establishing a simulation relation between the specification of the file system (which models the file system as an abstract map from file names to sequences of bytes) and its implementation (which uses fixed-size disk blocks to store the contents of the files).

We used the Athena proof checker to represent and validate our proof. Our experience indicates that Athena’s use of block-structured natural deduction, support for structural induction and proof abstraction, and seamless connection with high-performance automated theorem provers were essential to our ability to successfully manage a proof of this size.

Table of Contents

1 Introduction .................................................. 2
2 A Simple File System ........................................ 3
3 Abstract specification of the file system .................. 4
   3.1 Specification of the abstract read operation ........... 4
   3.2 Specification of the abstract write operation ........... 5
4 File system implementation .................................... 5
   4.1 Definition of the concrete read operation ............. 6
   4.2 Definition of the concrete write operation ............. 7
5 The correctness proof ......................................... 10
   5.1 State abstraction and homomorphic simulation ....... 10
   5.2 Proof outline .......................................... 11
6 Reachability invariants ....................................... 12
   6.1 Proving invariants ...................................... 16
7 Proof automation with tactics ............................... 17
8 A sample lemma proof ........................................ 19
9 Related work .................................................. 23
10 Conclusions .................................................. 24
A Some standard Athena libraries ............................. 26
   A.1 Options ............................................... 26
   A.2 Finite maps ........................................... 27
   A.3 Resizable arrays ....................................... 27
   A.4 Natural numbers ....................................... 28
1 Introduction

In this paper we explore the challenges of verifying the core operations of a standard Unix file system [20, 16]. We formalize the specification of the file system as a map from file names to sequences of bytes, then formalize an implementation that uses such standard file system data structures as inodes and fixed-sized disk blocks. We verify the correctness of the implementation by proving the existence of a simulation relation between the specification and the implementation.

The proof is expressed and checked in Athena, an interactive theorem-proving environment based on denotational proof languages (DPLs [3]) for first-order logic with sorts and polymorphism. Athena uses a Fitch-style natural deduction calculus, formalized via the abstraction of assumption bases. High-level idioms that are frequently encountered in common mathematical reasoning (such as "pick any x and y ..." or "assume P in ...") are directly available to the user. Athena also includes a higher-order functional language in the style of Scheme and ML and offers flexible mechanisms for expressing proof-search algorithms in a trusted manner (akin to the ‘tactics’ and ‘tacticals’ of LCF-like systems such as HOL [11]).

The proof comprises 283 lemmas and theorems, and took 1.5 person-months of full-time work to complete. It consists of roughly 5,000 lines of Athena code, for an average of about 18 lines per lemma. It takes about 9 minutes to check on a high-end Pentium, for an average of 1.9 seconds per lemma. Athena seamlessly integrates cutting-edge automated theorem provers (ATPs) such as Vampire [21] and Spass [22] to mechanically prove tedious steps, leaving the user to focus on the interesting parts of the proof. Athena invokes Vampire and Spass over 2,000 times during the course of the proof. That the proof is still several thousand lines long reflects the sheer size of the problem. For instance, we needed to prove 12 invariants and there are 10 state-transforming operations, which translates to 120 lemmas for each invariant/operation pair \((I, f)\), each guaranteeing that \(f\) preserves \(I\). Most of these lemmas are non-trivial; many require induction, and several require a number of other auxiliary lemmas. Further complicating matters is the fact that we can show that some of these invariants are preserved only if we assume that certain other invariants hold. In these cases we must consider simultaneously the conjunction of several invariants. The resulting formulas are several pages long and have dozens of quantified variables. We believe that Athena’s combination of natural deduction, versatile mechanisms for proof abstraction, and seamless incorporation of very efficient ATPs were crucial to our ability to successfully complete a proof effort of this scale.

To place our results in a broader context, consider that organizations rely on storage systems in general and file systems in particular to store critical persistent data. Because errors can cause the file system to lose this data, it is important for the implementation to be correct. The standard wisdom is that core system components such as file systems will always remain beyond the reach of full correctness proofs, leaving extensive testing—and the possibility of undetected residual errors—as the only option. Our results, however, suggest that correctness proofs for crucial system components (especially for the key
algorithms and data structures at the heart of such components) may very well be within reach.

The remainder of the paper is structured as follows. Section 2 informally describes a simplified file system. Section 3 presents an abstract specification of the file system. This specification hides the complexity of implementation-specific data structures such as inodes and data blocks by representing files simply as indexable sequences of bytes. Section 4 presents our model of the implementation of the file system. This implementation contains many more details, e.g., the mapping from file names to inodes, as well as the representation of file contents using sequences of non-contiguous data blocks that are dynamically allocated on the disk. Section 5 presents the statement of the correctness criterion. This criterion uses an abstraction function [15] that maps the state of the implementation to the state of the specification. Section 5 also sketches out the overall strategy of the proof. Section 6 and Section 7 address the key role that invariants and proof tactics played in this project. Section 8 gives a flavor of our correctness proof by presenting a proof of a frame-condition lemma. Section 9 presents related work, and Section 10 concludes. The Appendix contains a description of the relevant parts of certain Athena libraries that were used in this project.

2 A Simple File System

In this section we describe the high-level structure of a simple file system. In Section 4 we present a formal model of such a file system.

In our file system the physical media is divided into blocks containing a fixed number of bytes. The contents of a file are divided into block-sized segments, and stored in a series of blocks that are not necessarily consecutive.

The file system associates each file with an inode, which is a data structure that contains information about the file, including the file size and which blocks contain the file data. Unlike actual UNIX file systems, the inodes in our system do not contain other information such as access privileges and time stamps.

There is only one directory, the root directory, which maps file names to inode numbers. No two file names can refer to the same file, so no two file identifiers can map to the same inode number. We also assume that the disk is unbounded—the file system has access to an infinite number of inodes and blocks.

To read a byte from a given file, the file system first looks up the file name in the root directory, and obtains the number of the corresponding inode. Assuming the file exists, the file system then looks up the inode. From the information in the inode, the file system determines if it is reading a byte that is within the bounds of the file size, and if so, which block contains the relevant byte. Finally, the file system reads the byte from that block and returns the value read.

A similar look-up process occurs when writing a byte in a file. In this case, if the file system is writing a byte that is within the bounds of the existing file size, it simply stores the new value to the appropriate byte. Otherwise, the file system extends the file up to the index of the byte it is writing. It then stores
the appropriate value to the byte it is writing, and a default pad value to the bytes in between.

Our formalization consists of a set of axioms in first-order logic with sorts, polymorphism, and structural induction. We use generic Athena libraries that contain axiomatizations of natural numbers, value options, finite maps, and re- sizable arrays; see the Appendix for a brief description of those libraries.

3 Abstract specification of the file system

Our specification is an abstract model of the file system that hides the complexity of data structures such as inodes and data blocks by representing files as indexable sequences of bytes.

The specification uses the following sorts (the first two are introduced as new primitive domains, while the latter two are defined as sort abbreviations):

- **sorts** Byte, FileID
- **define** File = RSArrayOf(Byte)
- **define** AbState = FMap(FileID, File)

The sort Byte is an abstract type whose values represent the units of file content. FileID is also an abstract type; its values represent file identifiers. We define File as a resizable array of Byte. The abstract state of the file system, AbState, is represented as a finite map from file identifiers (FileID) to file contents (File). We also introduce a distinguished element of Byte, called fillByte, which is used to pad a file in the case of an attempt to write at a position exceeding the file size: **declare** fillByte : Byte.

3.1 Specification of the abstract read operation

We begin by giving the signature of the abstract read operation, absRead:

**declare** absRead : FileID × Nat × AbState → ReadResult

Thus absRead takes a file identifier fid, an index i in the file, and an abstract file system state s; and returns an element of ReadResult. The latter is defined as the following datatype:

**datatype** ReadResult = EOF
| Ok(Byte) |
| FileNotFound |

Therefore, the result of any absRead operation is one of three things: EOF; if the index is out of bounds; FileNotFound, if the file does not exist; or, if all goes well, a value of the form Ok(v) for some byte v, representing the content of file fid at position i. More precisely, the semantics of absRead are given by the following three axioms:
\[ \forall \text{fid} \ i \ s . \ \text{lookUp}(\text{fid}, s) = \text{NONE} \Rightarrow \text{read}(\text{fid}, i, s) = \text{FileNotFoundException} \]

\[ \forall \text{fid} \ i \ s \ \text{file} . \left[ \text{lookUp}(\text{fid}, s) = \text{SOME}(\text{file}) \land \text{arrayLen(file)} \leq i \right] \Rightarrow \]

\[ \text{read}(\text{fid}, i, s) = \text{EOF} \]

\[ \forall \text{fid} \ i \ s \ \text{file} . \left[ \text{lookUp}(\text{fid}, s) = \text{SOME}(\text{file}) \land \text{arrayRead(file, i)} = \text{SOME}(v) \right] \Rightarrow \text{read}(\text{fid}, i, s) = \text{Ok}(v) \]

Using the equality conditions for finite maps and resizable arrays, we are able to prove the following extensionality theorem for abstract states:

\[ \forall s_1 \ s_2 , \ s_1 = s_2 \iff \left( \forall \text{fid} \ i . \ \text{read}(\text{fid}, i, s_1) = \text{read}(\text{fid}, i, s_2) \right) \]

(1)

### 3.2 Specification of the abstract write operation

The abstract write operation has the following signature:

\[ \text{declare write} : \text{FileID} \times \text{Nat} \times \text{Byte} \times \text{AbState} \rightarrow \text{AbState} \]

This is the operation that defines state transitions in our file system. It takes as arguments a file identifier fid, an index i indicating a file position, a byte v representing the value to be written, and a file system state s. The result is a new state where the contents of the file associated with fid have been updated by storing v at position i. Note that if i exceeds the length of the file in state s, then in the resulting state the file will be extended to size i+1 and all newly allocated positions below i will be padded with the fillByte value. Finally, if fid does not correspond to a file in s, then an empty file of size i+1 is first created and then the value v is written. More precisely, we introduce the following axioms:

\[ \forall \text{fid} \ i \ v \ s . \ \text{lookUp}(\text{fid}, s) = \text{NONE} \Rightarrow \]

\[ \text{write}(\text{fid}, i, v, s) = \text{update}(s, \text{fid}, \text{arrayWrite} \left( \text{makeArray}(\text{fillByte}, i + 1), i, v, \text{fillByte} \right)) \]

\[ \forall \text{fid} \ i \ v \ s \ \text{file} . \ \text{lookUp}(\text{fid}, s) = \text{SOME}(\text{file}) \Rightarrow \]

\[ \text{write}(\text{fid}, i, v, s) = \text{update}(s, \text{fid}, \text{arrayWrite} \left( \text{file}, i, v, \text{fillByte} \right)) \]

### 4 File system implementation

Standard Unix file systems store the contents of each file in separate disk blocks, and maintain a table of structures called inodes that index those blocks and store various types of information about the file. Our implementation operates directly on the inodes and disk blocks and therefore models the operations that the file system performs on the disk. We omit details such as file permissions, dates, links, multi-layered directories, and optimizations such as caching. Some of these (e.g., permissions and date stamps) are orthogonal to the verification obligation and could be included with minimal changes to our proof, while others (e.g., caching) would likely introduce additional complexity.

File data is organized in Block units. A Block is an array of blockSize bytes, where blockSize is a positive constant. Specifically, we model a Block as a finite map from natural numbers to Byte:

\[ \text{define } \text{Block} = \text{FMap} \left( \text{Nat}, \text{Byte} \right) \]
We also define a distinguished element of Block, called initialBlock, such that:

\[
\forall i, i < \text{blockSize} \Rightarrow \text{lookUp}(i, \text{initialBlock}) = \text{SOME}(\text{fillByte}) \\
\forall i, \text{blockSize} \leq i \Rightarrow \text{lookUp}(i, \text{initialBlock}) = \text{NONE}
\]

In other words, an initialBlock consists of blockSize copies of fillByte.

File metadata is stored in inodes:

```plaintext
datatype INode = inode(fileSize : Nat, blockCount : Nat, blockList : FMap(Nat, Nat))
```

An INode is a datatype consisting of the file size in bytes and in blocks, and a list of block numbers. The list of block numbers is an array of the block numbers that contain the file data. We model this array as a finite map from natural numbers (array indices) to natural numbers (block numbers).

The data type State represents the file system state:

```plaintext
datatype State = state(inodeCount : Nat, stateBlockCount : Nat, 
\quad inodes : FMap(Nat, INode), blocks : FMap(Nat, Block), root : FMap(FileID, Nat))
```

A State consists of a count of the inodes in use; a count of the blocks in use; an array of inodes; an array of blocks; and the root directory. We model the array of inodes as a finite map from natural numbers (array indices) to INode (inodes). Likewise, we model the array of blocks as a finite map from natural numbers (array indices) to Block (blocks). We model the root directory as a finite map from FileID (file identifiers) to natural numbers (inode numbers).

We also define initialState, a distinguished element of State, which describes the initial state of the file system. In the initial state, no inodes or blocks are in use, and the root directory is empty:

```plaintext
declare initialState : State
initialState = state(0, 0, empty-map, empty-map, empty-map)
```

### 4.1 Definition of the concrete read operation

The concrete read operation, \textit{read}, has the following signature:

```plaintext
declare read : FileID × Nat × State → ReadResult
```

The \textit{read} operation takes a file identifier \( \text{fid} \), an index \( i \) in the file, and a concrete file system state \( s \), and returns an element of ReadResult. It first determines if \( \text{fid} \) is present in the root directory of \( s \). If not, \textit{read} returns FileNotFound. Otherwise, it looks up the corresponding inode. If \( i \) is not less than the file size, \textit{read} returns EOF. Otherwise, \textit{read} looks up the block containing the data and returns the relevant byte. The following axioms capture these semantics (for ease of presentation, we omit universal quantifiers from now on; all variables can be assumed to be universally quantified):

\footnote{As a convention, we use bold italic font to indicate the abstract-state version of something; e.g., abstract \textit{read} vs. concrete \textit{read}, an abstract state \( s \) vs. a concrete state \( s \), etc.}
4.2 Definition of the concrete write operation

The concrete write operation, write, takes a file identifier fid, a byte index i, the byte value v to write, and a state s, and returns the updated state:

\[
\text{declare write : FileID } \times \text{Nat } \times \text{Byte} \times \text{State} \rightarrow \text{State}
\]

\[
\text{[CW1]} \quad \text{let } s' = \text{alloc_inode(fid, s)} \text{ in}
\]

\[
\text{[CW2]} \quad \text{lookUp(fid, root(s)) = SOME(n) } \Rightarrow \text{write(fid, i, v, s) = writeExisting(n, i, v, s)}
\]

\[
\text{lookUp(fid, root(s)) = NONE } \Rightarrow \text{write(fid, i, v, s) = writeExisting(n, i, v, s')}
\]

If the file associated with fid already exists, write delegates the write to the helper function writeExisting. If the file does not exist, write first invokes alloc_inode, which creates a new, empty file, then calls writeExisting with the inode number of the new file.

alloc_inode takes a file identifier fid and a state s, and returns an updated state:

\[
\text{declare alloc_inode : FileID } \times \text{State} \rightarrow \text{State}
\]

\[
\text{getNext_inode(s) = state(inc + 1, bc, imm, bm, root) } \Rightarrow \text{alloc_inode(fid, s) = state(inc + 1, bc, imm, bm, update(root, fid, inc))}
\]

alloc_inode creates a new inode by invoking getNext_inode, then associates fid with the new inode.

getNext_inode takes a state and returns an updated state. It allocates and initializes a new inode:

\[
\text{declare getNext_inode : State} \rightarrow \text{State}
\]

\[
\text{getNext_inode(state(inc, bc, imm, bm, root)) = state(inc + 1, bc, update(imm, inc, inode(0, 0, empty-map)), bm, root)}
\]

writeExisting takes an inode number n, a byte index i, the byte value v to write, and a state s, and returns the updated state:

\[
\text{declare writeExisting : Nat } \times \text{Nat } \times \text{Byte} \times \text{State} \rightarrow \text{State}
\]

\[
\text{[WE1]} \quad \text{lookUp(n, inodes(s)) = SOME(inode) } \land \ (i \div \text{blockSize}) < \text{blockCount(inode)} \land \ i < \text{filesize(inode)} \Rightarrow \text{writeExisting(n, i, v, s) = writeNoExtend(n, i, v, s)}
\]

\[
\text{[WE2]} \quad \text{lookUp(n, inodes(s)) = SOME(inode) } \land \ (i \div \text{blockSize}) < \text{blockCount(inode)} \land \ \text{filesize(inode)} \leq i \Rightarrow \text{writeExisting(n, i, v, s) = writeSmallExtend(n, i, v, s)}
\]

\[
\text{[WE3]} \quad \text{lookUp(n, inodes(s)) = SOME(inode) } \land \ \text{blockCount(inode)} \leq (i \div \text{blockSize}) \Rightarrow \text{writeExisting(n, i, v, s) = writeNoExtend(n, i, v, \text{extendFile(n, i, s)})}
\]
If \( i \) is less than the file size, \( \text{writeExisting} \) delegates the writing to \( \text{writeNoExtend} \), which stores the value \( v \) in the appropriate location. If \( i \) is not less than the file size but is located in the last block of the file, \( \text{writeExisting} \) delegates to \( \text{writeSmallExtend} \), which stores the value \( v \) in the appropriate position and updates the file size. Otherwise, \( \text{writeExisting} \) first invokes \( \text{extendFile} \), which extends the file by the appropriate number of blocks, and then calls \( \text{writeNoExtend} \) on the updated state.

\( \text{writeNoExtend} \) takes an inode number \( n \), a byte index \( i \), the byte value \( v \) to write, and a state \( s \), and returns the updated state after writing \( v \) at index \( i \):

\[
\text{declare writeNoExtend} : \text{Nat} \times \text{Nat} \times \text{Byte} \times \text{State} \rightarrow \text{State}
\]
\[
\text{lookUp}(n, \text{mode}(s)) = \text{SOME}(\text{inode}) \wedge
\text{lookUp}(i \text{ div blockSize}, \text{mode}(s)) = \text{SOME}(\text{bn}) \wedge
\text{lookUp}(\text{bn}, \text{blocks}(s)) = \text{SOME}(\text{block}) \Rightarrow
\text{writeNoExtend}(n, i, v, s) = \text{updateStateBM}(s, \text{bn}, \text{update}(\text{block}, i \text{ mod blockSize}, v))
\]

\( \text{writeNoExtend} \) uses the helper function \( \text{updateStateBM} \). The function \( \text{updateStateBM} \) takes the state, the block number \( \text{bn} \), and the block \( \text{block} \), and returns an updated state where \( \text{bn} \) maps to \( \text{block} \):

\[
\text{declare updateStateBM} : \text{State} \times \text{Nat} \times \text{Block} \rightarrow \text{State}
\]
\[
\text{updateStateBM}(\text{state} = (\text{inc}, \text{bc}, \text{imm}, \text{bn}, \text{root}), \text{bn}, \text{block}) =
\text{state}(\text{inc}, \text{bc}, \text{imm}, \text{update}(\text{bn}, \text{bn}, \text{block}), \text{root})
\]

\( \text{writeSmallExtend} \) takes an inode number \( n \), a byte index \( i \), the byte value \( v \) to write, and a state. It updates the file size and writes the byte value \( v \) at byte index \( i \) for the file associated with the inode number \( n \), and returns the updated state:

\[
\text{declare writeSmallExtend} : \text{Nat} \times \text{Nat} \times \text{Byte} \times \text{State} \rightarrow \text{State}
\]
\[
\text{lookUp}(n, \text{mode}(s)) = \text{SOME}(\text{inode}) \wedge
\text{lookUp}(i \text{ div blockSize}, \text{bn}) = \text{SOME}(\text{bn}) \wedge
\text{lookUp}(\text{bn}, \text{blocks}(s)) = \text{SOME}(\text{block}) \wedge i \leq \text{bn} \Rightarrow
\text{writeSmallExtend}(n, i, v, \text{state} = (\text{inc}, \text{bc}, \text{imm}, \text{bn}, \text{root})) =
\text{state}(\text{inc}, \text{bc}, \text{imm}, \text{update}(\text{bn}, n, \text{mode}(i + 1, \text{bc}, \text{bn})),
\text{update}(\text{bn}, \text{bn}, \text{update}(\text{block}, i \text{ mod blockSize}, v)), \text{root})
\]

\( \text{extendFile} \) takes an inode number \( n \), the byte index of the write, and the state \( s \). It delegates the task of allocating the necessary blocks to \( \text{allocBlocks} \):

\[
\text{declare extendFile} : \text{Nat} \times \text{Nat} \times \text{State} \rightarrow \text{State}
\]
\[
\text{lookUp}(n, \text{mode}(s)) = \text{SOME}(\text{inode}) \wedge \text{blockCount}(\text{inode}) \leq (j \text{ div blockSize}) \Rightarrow
\text{extendFile}(n, j, s) = \text{allocBlocks}(n, (j \text{ div blockSize}) - \text{blockCount}(\text{inode}) + 1, j, s)
\]

\( \text{allocBlocks} \) takes an inode number \( n \), the number of blocks to allocate, the byte index \( j \), and the state \( s \). We define it by primitive recursion:

\[
\text{declare allocBlocks} : \text{Nat} \times \text{Nat} \times \text{Nat} \times \text{State} \rightarrow \text{State}
\]
\[
[AB_1] \quad \text{allocBlocks}(n, 0, j, s) = s
\]
\[
[AB_2] \quad \text{allocBlocks}(n, k, j, s) = \text{state}(\text{inc}, \text{bc} + 1, \text{imm}, \text{bn}, \text{root}) \wedge
\text{lookUp}(n, \text{imm}) = \text{SOME}(\text{inode}(\text{bc}, \text{inc}, \text{bn})) \Rightarrow
\text{allocBlocks}(n, k + 1, j, s) = \text{allocBlocks}(n, k, j, \text{state}(\text{inc}, \text{bc} + 1,
\text{update}(\text{imm}, n, \text{mode}(j + 1, \text{bc} + 1, \text{update}(\text{bn}, \text{inc}, \text{bc})), \text{bn}, \text{root})))
\]
Fig. 1. The call graph of `write`.

`allocBlocks` uses the helper function `getNextBlock`, which takes the state `s`, allocates and initializes the next free block, and returns the updated state:

\[
\text{declare } \text{getNextBlock} : \text{State} \rightarrow \text{State} \\
\text{getNextBlock}(\text{state}(\text{inc}, \text{bc}, \text{imm}, \text{bm}, \text{root})) = \text{state}(\text{inc}, \text{bc} + 1, \text{imm}, \text{update}(\text{bm}, \text{bc}, \text{initialBlock}), \text{root})
\]

The call graph summarizing the `write` operation is shown in Figure 1. This call graph largely determines the auxiliary lemmas that need to be established every time we wish to prove a result about `write`. That is, whenever we need to prove a result `L` about `write`, we prove appropriate lemmas \(L_1\) and \(L_2\) about `allocINode` and `writeExisting`. In turn, \(L_1\) will rely on a lemma \(L_{11}\) about `getNextINode` and \(L_2\) will reference lemmas \(L_{21}, L_{22},\) and \(L_{23}\) about `writeNoExtend`, `writeSmallExtend`, and `extendFile`, respectively; and so on. In this way we obtain a lemma dependency graph for `L` whose structure mirrors that of the call graph for `write`.

In what follows we will restrict our attention to `reachable` states, those that can be obtained from the initial state by some finite sequence of `write` operations. Specifically, we define a predicate `reachableN` ("reachable in \(n\) steps") via two axioms: \(\text{reachableN}(s, 0) \equiv s = \text{initialState}\), and

\[
\text{reachableN}(s, n + 1) \Leftrightarrow \exists s' \ f i v. \text{reachableN}(s', n) \land s = \text{write}(f i v, s')
\]

We then set `reachable(s) \equiv \exists n. \text{reachableN}(s, n)`. We will write `\text{State}` for the set of all reachable states, and we will use the symbol \(\Downarrow\) to denote a reachable state. Propositions of the form \(\forall \cdots s \cdots P(\cdots \tilde{s} \cdots)\) and \(\exists \cdots s \cdots P(\cdots \tilde{s} \cdots)\) should be taken as abbreviations for \(\forall \cdots s \cdots, \text{reachable}(s) \Rightarrow P(\cdots \tilde{s} \cdots)\) and \(\exists \cdots s \cdots, \text{reachable}(s) \land P(\cdots \tilde{s} \cdots)\), respectively.
5 The correctness proof

5.1 State abstraction and homomorphic simulation

This section presents a correctness criterion for the implementation. The correctness criterion is specified using an abstraction function [15] that maps the state of the implementation to the state of the specification.

Consider the following binary relation $A$ from concrete to abstract states:

$$\forall \ s \ s'. A(s, s') \iff [\forall \ f d : i . \text{read}(f d, i, s) = \text{read}(f d, i, s')]$$

It follows directly from the extensionality principle on abstract states (1) that $A$ is functional:

$$\forall \ s \ s_1 \ s_2. A(s, s_1) \land A(s, s_2) \Rightarrow s_1 = s_2.$$ 

Accordingly, we postulate the existence of an abstraction function $\alpha : \text{State} \rightarrow \text{AbState}$ such that:

$$\forall \ s \ s'. \alpha(s) = s' \iff A(s, s').$$

That is, an abstracted state $\alpha(s)$ has the exact same contents as $s$: reading any position of a file in one state yields the same result as reading that position of the file in the other state.

\[\text{FileID} \times \text{Nat} \times \text{State} \xrightarrow{\text{read}} \text{ReadResult} \]
\[\text{FileID} \times \text{Nat} \times \text{AbState} \xrightarrow{\text{read}} \text{ReadResult} \]
\[\text{FileID} \times \text{Nat} \times \text{Byte} \times \text{State} \xrightarrow{\text{write}} \text{State} \]
\[\text{FileID} \times \text{Nat} \times \text{Byte} \times \text{AbState} \xrightarrow{\text{write}} \text{AbState} \]

Fig. 2. Commuting diagrams for the read and write operations.

A standard way of formalizing the requirement that an implementation $I$ is faithful to a specification $S$ is to express $I$ and $S$ as many-sorted algebras and establish a homomorphism from one to the other. In our case the two algebras are $I = (\text{FileID}, \text{Nat}, \text{Byte}, \text{State}; \text{read}, \text{write})$ and

$$S = (\text{FileID}, \text{Nat}, \text{Byte}, \text{AbState}; \text{read}, \text{write})$$

The embeddings from $I$ to $S$ for the carriers $\text{FileID}$, $\text{Nat}$, and $\text{Byte}$ are simply the identity functions on these domains; while the embedding from $\text{State}$ to
\( AbState \) is the abstraction mapping \( \alpha \). In order to prove that this translation yields a homomorphism we need to show that the two diagrams shown in Figure 2 commute. Symbolically, we need to prove the following:

\[
\forall \text{fid } i \in \mathcal{S}, \text{read}(\text{fid}, i, \mathcal{S}) = \text{read}(\text{fid}, i, \alpha(\mathcal{S})) \tag{2}
\]

and

\[
\forall \text{fid } i \in \mathcal{S}, \alpha(\text{write}(\text{fid}, i, v, \mathcal{S})) = \text{write}(\text{fid}, i, v, \alpha(\mathcal{S})) \tag{3}
\]

5.2 Proof outline

Goal (2) follows immediately from the definition of the abstraction function \( \alpha \). For (3), since the consequent is equality between two abstract states and we have already proven that two abstract states \( \mathcal{S}_1 \) and \( \mathcal{S}_2 \) are equal iff any abstract read operation yields identical results on \( \mathcal{S}_1 \) and \( \mathcal{S}_2 \), we transform (3) into the following:

\[
\forall \text{fid } i \in \mathcal{S}, \text{read}(\text{fid}, j, \alpha(\text{write}(\text{fid}, i, v, \mathcal{S}))) = \text{read}(\text{fid}, j, \text{write}(\text{fid}, i, v, \alpha(\mathcal{S})))
\]

Finally, using (2) on the above gives:

\[
\forall \text{fid } i \in \mathcal{S}, \text{read}(\text{fid}, j, \text{write}(\text{fid}, i, v, \mathcal{S})) = \text{read}(\text{fid}, j, \text{write}(\text{fid}, i, v, \alpha(\mathcal{S})))
\]

Therefore, choosing arbitrary \( \text{fid}, \text{fid}', j, i, v \), and \( \mathcal{S} \), we need to show \( L = R \), where \( L = \text{read}(\text{fid}, i, \text{write}(\text{fid}, j, v, \mathcal{S})) \) and

\[
R = \text{read}(\text{fid}', i, \text{write}(\text{fid}, j, v, \alpha(\mathcal{S})))
\]

Showing \( L = R \) is the main goal of the proof. We proceed by a case analysis as shown in Fig. 3. The decision tree of Fig. 3 has the following property: if the conditions that appear on a path from the root of the tree to an internal node \( u \) are all true, then the conditions at the children of \( u \) are mutually exclusive and jointly exhaustive (given that certain invariants hold, as discussed in Section 6). There are ultimately eight distinct cases to be considered, \( C_1 \) through \( C_8 \), appearing at the leaves of the tree. Exactly one of those eight cases must be true for any given \( \text{fid}, \text{fid}', j, i, v \), and \( \mathcal{S} \). We prove that \( L = R \) in all eight cases.

For each case \( C_i \), \( i = 1, \ldots, 8 \), we formulate and prove a pair of lemmas \( M_i \) and \( M_i' \) that facilitate the proof of the goal \( L = R \). Specifically, for each case \( C_i \) there are two possibilities:

1. \( L = R \) follows because both \( L \) and \( R \) reduce to a common term \( t \), with \( L = t \) following by virtue of lemma \( M_i \) and \( R = t \) following by virtue of lemma \( M_i' \):

   ![Diagram](image)

2. The desired identity follows because \( L \) and \( R \) respectively reduce to \( \text{read}(\text{fid}', i, \mathcal{S}) \) and \( \text{read}(\text{fid}', i, \alpha(\mathcal{S})) \), which are equal owing to (2). In this case, \( M_i \) is used to show \( L = \text{read}(\text{fid}', i, \mathcal{S}) \) and \( M_i' \) is used to show \( R = \text{read}(\text{fid}', i, \alpha(\mathcal{S})) \):
Fig. 3. Case analysis for proving the correctness of write.

The eight pairs of lemmas are shown in Figure 4. The “abstract-state” versions of the lemmas ([M_i], i = 1, ..., 8) are readily proved with the aid of Vampire from the axiomatizations of maps, resizable arrays, options, natural numbers, etc., and the specification axioms. The concrete lemmas M_i are much more challenging.

6 Reachability invariants

Reachable states have a number of properties that make them “well behaved.” For instance, if a file identifier is bound in the root of a state s to some inode number n, then we expect n to be bound in the mapping inodes(s). While this is not true for arbitrary states s, it is true for reachable states. In what follows, by a state invariant we will mean a unary predicate on states I(s) that is true for all reachable states, i.e., such that ∀ s. I(s).
There are 12 invariants \(inv_1, \ldots, inv_{11}\), that are of particular interest. The proof relies on them explicitly, i.e., at various points in the course of the argument we assume that all reachable states have these properties. Therefore, for the proof to be complete, we need to discharge these assumptions by proving that the properties in question are indeed invariants.

The process of guessing useful invariants—and then, more importantly, trying to prove them—was very helpful in strengthening our understanding of the implementation. More than once we conjectured false invariants, properties that appeared reasonable at first glance but later, when we tried to prove them,
turned out to be false. For instance, a seemingly sensible “size invariant” is that for every inode of size $fs$ and block count $bc$ we have

$$fs = \lfloor (bc - 1) \cdot \text{blockSize} \rfloor + (fs \mod \text{blockSize})$$

But this equality does not hold when the file size is a multiple of the block count. The proper invariant is \footnote{This invariant is equivalent to $bc = (fs + \text{blockSize} - 1) \div \text{blockSize}$.}

$$[fs \mod \text{blockSize} = 0 \Rightarrow fs = bc \cdot \text{blockSize}] \land [fs \mod \text{blockSize} \neq 0 \Rightarrow fs = ((bc - 1) \cdot \text{blockSize}) + (fs \mod \text{blockSize})]$$

where $\text{div}$ denotes integer division. For any inode of file size $fs$ and block count $bc$, we will write $\text{selinv}(fs,bc)$ to indicate that $fs$ and $bc$ are related as shown by the above formula.

Figure 5 presents the twelve reachability invariants for our file system implementation. In the sequel we focus on the first four invariants, $\text{inv}_0, \text{inv}_1, \text{inv}_2, \text{inv}_3$. These four invariants are fundamental and must be established before anything non-trivial can be proven about the system. They are also co-dependent, meaning that in order to prove that an operation preserves one of them, say $\text{inv}_j$, we often need to assume that the incoming state not only has $\text{inv}_j$ but also one or more of the other three invariants. For instance, we cannot prove that $\text{write}$ preserves $\text{inv}_0$, i.e., that

$$\forall i \ v \ s, \text{inv}_3(s) \Rightarrow \text{inv}_3(\text{write}(fid, i, v, s))$$

unless we also assume that $s$ has $\text{inv}_0$. Or suppose we want to prove that $\text{writeExisting}$ preserves any of the four invariants, say $\text{inv}_0$, so that our goal is to show $\text{inv}_0(\text{writeExisting}(n, i, v, s))$ on the assumptions

$$\text{lookUp}(n, \text{inodes}(s)) = \text{SOME}(\text{inode}(fs, bc, bl))$$  \hspace{1cm} (4)$$

and

$$\text{inv}_0(s)$$  \hspace{1cm} (5)$$

Consider the case

$$bc \leq i \ \text{div} \ \text{blockSize},$$

whereby $\text{writeExisting}(n, i, v, s)$ returns

$$\text{writeNoExtend}(n, i, v, \text{extendFile}(n, i, s)).$$

Since $\text{writeNoExtend}$ is conditionally defined, we need to show that its three preconditions are satisfied in the intermediate state $s_1 = \text{extendFile}(n, i, s)$. It is easy enough to show that the first precondition holds, i.e., that

$$\text{lookUp}(n, \text{inodes}(s_1)) = \text{SOME}(\text{inode}(fs_1, bc_1, bl_1))$$

for some $fs_1, bc_1$, and $bl_1$; this follows from (4) and an auxiliary lemma stating that $\text{extendFile}$ preserves the invariant $I(s) = \text{inDom}(n, \text{inodes}(s))$ (for
\( inv_0(s) : \) \[ \text{lookUp}( fid, \text{root}(s)) = \text{SOME}(n) \Rightarrow \text{inDom}(n, \text{inodes}(s)) \]

\( inv_1(s) : \) \[ \text{lookUp}( n, \text{inodes}(s)) = \text{SOME}(\text{inode}(fs, bc, bl)) \Rightarrow \]
\[ \text{inDom}(k, bl) \Leftrightarrow k < bc \]

\( inv_2(s) : \) \[ \text{lookUp}( n, \text{inodes}(s)) = \text{SOME}(\text{inode}) \land \]
\[ \text{lookUp}( bn, \text{blockList}(\text{inode})) = \text{SOME}(bn') \Rightarrow \]
\[ \text{inDom}(bn', \text{blocks}(s)) \]

\( inv_3(s) : \) \[ \text{lookUp}( n, \text{inodes}(s)) = \text{SOME}(\text{inode}(fs, bc, bl)) \Rightarrow \text{szInv}(fs, bc) \]

\( inv_4(s) : \) \[ \text{inDom}(bn, \text{blocks}(s)) \Leftrightarrow bn < \text{stateBlockCount}(s) \]

\( inv_5(s) : \) \[ \text{inDom}(\text{nodeNum}, \text{inodes}(s)) \Leftrightarrow \text{nodeNum} < \text{inodeCount}(s) \]

\( inv_6(s) : \) \[ \text{lookUp}(\text{nodeNum}, \text{inodes}(s)) = \text{SOME}(\text{inode}(fs, bc, bl)) \land bc = 0 \]
\[ \Rightarrow fs = 0 \]

\( inv_7(s) : \) \[ \text{sid}_1 \neq \text{sid}_2 \land \]
\[ \text{lookUp}(\text{sid}_1, \text{root}(s)) = \text{SOME}(\text{nodeNum}_1) \land \]
\[ \text{lookUp}(\text{sid}_2, \text{root}(s)) = \text{SOME}(\text{nodeNum}_2) \]
\[ \Rightarrow \text{nodeNum}_1 \neq \text{nodeNum}_2 \]

\( inv_8(s) : \) \[ \text{lookUp}(\text{nodeNum}, \text{inodes}(s)) = \text{SOME}(\text{inode}) \land \]
\[ \text{lookUp}(k, \text{blockList}(\text{inode})) = \text{SOME}(\text{bn}) \land \]
\[ \text{lookUp}(\text{bn}, \text{blocks}(s)) = \text{SOME}(\text{blocks}) \Rightarrow \]
\[ \text{inDom}(j, \text{block}) \Leftrightarrow j < \text{blockSize} \]

\( inv_9(s) : \) \[ \text{lookUp}(\text{nodeNum}_1, \text{inodes}(s)) = \text{SOME}(\text{node}_1) \land \]
\[ \text{lookUp}(\text{nodeNum}_2, \text{inodes}(s)) = \text{SOME}(\text{node}_2) \land \]
\[ \text{lookUp}(\text{k}_1, \text{blockList}(\text{node}_1)) = \text{SOME}(\text{bn}_1) \land \]
\[ \text{lookUp}(\text{k}_2, \text{blockList}(\text{node}_2)) = \text{SOME}(\text{bn}_2) \land \]
\[ \text{nodeNum}_1 \neq \text{nodeNum}_2 \]
\[ \Rightarrow \text{bn}_1 \neq \text{bn}_2 \]

\( inv_{10}(s) : \) \[ \text{lookUp}(\text{nodeNum}, \text{inodes}(s)) = \text{SOME}(\text{inode}) \land \]
\[ \text{lookUp}(\text{k}_1, \text{blockList}(\text{inode})) = \text{SOME}(\text{bn}_1) \land \]
\[ \text{lookUp}(\text{k}_2, \text{blockList}(\text{inode})) = \text{SOME}(\text{bn}_2) \land \]
\[ \text{k}_1 \neq \text{k}_2 \]
\[ \Rightarrow \text{bn}_1 \neq \text{bn}_2 \]

\( inv_{11}(s) : \) \[ \text{lookUp}(\text{nodeNum}, \text{inodes}(s)) = \text{SOME}(\text{inode}(fs, bc, bl)) \land \]
\[ i \text{ div } \text{blockSize} < bc \land fs \leq i \land \]
\[ \text{lookUp}(i \text{ div } \text{blockSize}, bl) = \text{SOME}(\text{bn}) \land \]
\[ \text{lookUp}(\text{bn}, \text{blocks}(s)) = \text{SOME}(\text{blocks}) \Rightarrow \]
\[ \text{lookUp}(i \text{ mod } \text{blockSize}, block) = \text{SOME}(	ext{fillByte}) \]

Fig. 5. Reachability Invariants
fixed inode number $n$). However, it is more challenging to show that the two remaining preconditions hold, i.e., that there exist $bn_1$ and $block_1$ such that $\text{lookUp}(i \div \text{BlockSize}, bn_1) = \text{SOME}(bn_1)$ and $\text{lookUp}(bn_1, blocks(s_1)) = \text{SOME}(block_1)$.

But these would follow immediately if we could show that $s_1$ has $\text{inv}_1$ and $\text{inv}_2$, and that $i \div \text{BlockSize} < bc_1$. Showing that $s_1$ has $\text{inv}_1$ and $\text{inv}_2$ would also follow immediately if we strengthened our initial hypothesis (5) by additionally assuming that $s$ has $\text{inv}_1$ and $\text{inv}_2$, provided we have shown elsewhere that $\text{extendFile}$ preserves both of these invariants. However, showing $i \div \text{BlockSize} < bc_1$ presupposes that $s_1$ has $\text{inv}_3$. Consequently, we are led to assume that the original state $s$ has all four invariants. Provided we have already shown that $\text{extendFile}$ preserves each of the four invariants, it then follows that $s_1$ has all four of them, and hence that the preconditions of $\text{writeNoExtend}$ hold.

6.1 Proving invariants

Showing that a unary state property $I(s)$ is an invariant proceeds in two steps:

1. proving that $I$ holds for the initial state, $I(s_0)$; and
2. proving $\forall \text{fid} \ i \ v \ s. I(s) \Rightarrow I(\text{write}(\text{fid}, i, v, s))$.

Once both of these have been established, a routine induction on $n$ will show that

$$\forall n \ s. \text{reachableN}(s, n) \Rightarrow I(s).$$

It then follows directly by the definition of reachability that all reachable states have $I$.

Proving that the initial state has an invariant $inv_j$ is straightforward: in all twelve cases it is done automatically. The second step, proving that $\text{write}$ preserves $inv_j$, is more involved. Including $\text{write}$, the implementation comprises ten state-transforming operations, and control may flow from $\text{write}$ to any one of them. Accordingly, we need to show that all ten operations preserve the invariant under consideration. This means that for a total of ten operations $f_0, \ldots, f_9$ and twelve invariants $inv_0, \ldots, inv_{11}$, we need to prove 120 lemmas, each stating that $f_i$ preserves $inv_j$.

Most of the operations $f_i$ are defined conditionally, in the form

$$\forall \vec{x}_i \vec{y}_i. \text{PC}_i(\vec{x}_i, \vec{y}_i) \Rightarrow f_i(\vec{x}_i) = \cdots$$

where $\vec{x}_i, \vec{y}_i$ are lists of distinct variables; $\text{PC}_i(\vec{x}_i, \vec{y}_i)$, the “precondition” of $f_i$, is usually a conjunction of equations in the variables $\vec{x}_i$ and $\vec{y}_i$ (if $f_i$ is not defined conditionally then this can be regarded as the empty conjunction, i.e.,

---

3 By a “state-transforming operation” we mean one that takes a state as an argument and produces a state as output. There are ten such operations, nine of which are auxiliary functions (such as $\text{extendFile}$) invoked by $\text{write}$. 
---
as the constant \( \text{true} \). Therefore, each of the 120 invariant-preservation lemmas is of the form

\[
\forall x_i, y_i, s. [PC_i(x_i, y_i) \land I(s)] \Rightarrow inv_j(f_i(x_i))
\]

(6)

for \( i = 0, \ldots, 9 \) and \( j = 0, \ldots, 11 \), and where \( I(s) \) is of the form \( inv_j(s) \land inv_{i_r} \land \cdots \land inv_{i_k} \) where \( k \geq 0 \) and \( i_r \in \{0, 1, \ldots, 11\} \) for \( 1 \leq r \leq k \).

The large majority of the proof text (about 80% of it) is devoted to proving these lemmas. Some of them are surprisingly tricky to prove, and even those that are not particularly conceptually demanding can be challenging to manipulate, if for no other reason simply because of their volume. Given the size of the function preconditions and the size of the invariants (especially in those cases where we need to consider the conjunction of several invariants at once), an invariance lemma can span multiple pages of text. Proof goals of that scale test the limits even of cutting-edge ATPs. For instance, in the case of a proposition \( P \) that was several pages long (which arose in the proof of one of the invariance lemmas), Spess took over 10 minutes to prove the trivial goal \( P \Rightarrow P' \), where \( P' \) was simply an alphabetically renamed copy of \( P \) (Vampire was not able to prove it at all, at least within 20 minutes). Heavily skolemizing the formula and blindly following the resolution procedure prevented these systems from recognizing the goal as trivial. By contrast, using Athena’s native inference rules, the goal was derived instantaneously via the two-line deduction \textbf{assume P in claim P'}, because Athena treats alphabetically equivalent propositions as identical and has an efficient implementation of proposition look-ups. This speaks to the need to have a variety of reasoning mechanisms available in a uniform, integrated framework.

There are many additional lemmas that were used in proving the invariants or in proving other results after all twelve invariants had already been proven. We mention two typical ones:

\textbf{Lemma 1.} If \( fid_1 \neq fid_2 \) and

\[
\text{lookUp}(fid_2, \text{root}(s)) = x
\]

then \( \text{lookUp}(fid_2, \text{root}(\text{write}(fid_1, i, v, s))) = x \).

\textbf{Lemma 2.} If \( \text{lookUp}(n, \text{inodes}(s)) = \text{SOME}(inode_2) \) and

\[
\text{lookUp}(n, \text{inodes}(\text{allocBlocks}(n, k, j, s))) = \text{SOME}(inode_2)
\]

then \( \text{blockCount}(inode_2) = \text{blockCount}(inode_2) + k \).

\textbf{7 Proof automation with tactics}

After proving a few invariance lemmas for some of the operations it became apparent that a large portion of the reasoning was the same in every case and could thus be factored away for reuse.
Athena makes it easy to abstract concrete proofs into natural-deduction proof algorithms called *methods*. For every state-transforming operation $f_i$ we wrote a "preservation" method $P_i$ that takes an arbitrary invariant $I$ as input (expressed as a unary function that takes a state and constructs an appropriate proposition) and attempts to prove the corresponding invariance lemma.

$$\forall x_i y_i s \cdot [PC_i(x_i, y_i) \land I(s)] \Rightarrow I(f_i(x_i))$$

(7)

$p_i$ encapsulates all the generic reasoning involved in proving invariants for $f_i$. If any non-generic reasoning (specific to $I$) is additionally required, it is packaged into a proof continuation $K$ and passed into $p_i$ as a higher-order method argument. $p_i$ can then invoke $K$ at appropriate points within its body as needed. Similar methods for other functions make the overall proof substantially shorter—and easier to develop and to debug—than it would have been otherwise.

Consider, for example, proving that allocBlocks preserves a certain property $I$. This is always done by induction on $k$, the number of blocks to be allocated. Performing the base inductive step automatically, managing the inductive hypothesis, proving that the relevant precondition involving getNextBlock is satisfied in the context in which allocBlocks is called, deriving useful consequences of that fact, etc., these are all standard tasks that are repetitively performed regardless of $I$; we have abstracted all of them away in a higher-order method that accepts the $I$-specific reasoning as an input method.

Proof programability was useful in streamlining several other recurring patterns of reasoning, apart from dealing with invariants. A typical example is this: given a reachable state $s$, an inode number $n$ such that lookUp(n, inodes(s)) = SOME(inode(fs, bc, bl)), and an index $i < fs$, we often need to prove the existence of $bn$ and block such that lookUp(i div blockSize, bl) = SOME(bm) and

$$\text{lookUp}(bn, \text{blocks}(s)) = \text{SOME}(\text{block})$$

The reasoning runs as follows: first, from the reachability of $s$, we infer that it has certain invariants, including $inv_0$, $inv_1$, $inv_2$, and $inv_3$. From these invariants, the assumption $i < fs$, and standard arithmetic laws we may deduce $(i \div \text{blockSize}) < bc$. From this, our initial assumptions, and $inv_1$, we conclude that $i \div \text{blockSize}$ is in the domain of the mapping $bl$. Thus the existence of an appropriate $bn$ is ensured, and along with it, owing to $inv_2$, the existence of an appropriate block. We packaged this reasoning in a method find-bn-block that takes all the relevant quantities as inputs, assumes that the appropriate hypotheses are in the assumption base, and performs the appropriate inferences. The method also accepts a proof continuation $K$ that is invoked once the goal has been successfully derived.

Another example is a slight extension of this method, named find-bn-block-ival, that operates under the same assumptions but, in addition to a block number and the block itself, yields a value $v$ such that lookUp(i mod blockSize, block) = SOME(v), which is possible because $i \mod \text{blockSize} < \text{blockSize}$. Yet another example of a streamlined proof method is an inductive method showing that an invariant holds for all reachable states.
8 A sample lemma proof

In this section we will prove lemma $[M_8]$, which can be viewed as a frame condition: it asserts that performing a write operation on a given file leaves the contents of every other file unchanged. More specifically, let $fid_1$ refer to the file to be written, let $fid_2$ be any file identifier distinct from $fid_1$, let $s$ be any reachable state, and let $s'$ be the state obtained from $s$ by writing some value into some byte position of $fid_1$. Then $[M_8]$ says that reading any byte of $fid_2$ in $s'$ yields the same result as reading that byte in $s$.

The proof relies on four auxiliary lemmas about write, given below. Lemmas (8) and (9) handle the case when $fid_1$ (the file to be written) already exists in $s$, while (10) and (11) apply to the case when $fid_1$ is unbound in the root of $s$. As usual, all the variables are assumed to be universally quantified.

\[
\begin{align*}
[\text{lookUp} (n_1, \text{inode}(s)) = \text{SOME}(\text{inode}_1) \land n \neq n_1 \land \\
\text{lookUp} (bn', \text{blockList}(\text{inode}_1)) = \text{SOME}(bn') \land \\
\text{lookUp} (bn', \text{blocks}(s)) = \text{SOME}(\text{blocks}) \land \text{lookUp} (fid, \text{root}(s)) = \text{SOME}(n) ] \Rightarrow (8) \\
\text{lookUp} (n_1, \text{inode}(\text{write}(fid, i, v, s))) = \text{SOME}(\text{inode}_1) \land \\
\text{lookUp} (bn', \text{blocks}(\text{write}(fid, i, v, s))) = \text{SOME}(\text{blocks}) \\
\end{align*}
\]

\[
\begin{align*}
[\text{lookUp} (n_1, \text{inode}(s)) = \text{SOME}(\text{inode}_2) \land n \neq n_1 \land \\
\text{lookUp} (fid, \text{root}(s)) = \text{SOME}(n) ] \Rightarrow (9) \\
\text{lookUp} (n_1, \text{inode}(\text{write}(fid, i, v, s))) = \text{SOME}(\text{inode}_1) \\
\end{align*}
\]

\[
\begin{align*}
[\text{lookUp} (n_1, \text{inode}(s)) = \text{SOME}(\text{inode}_1) \land \\
\text{lookUp} (bn', \text{blockList}(\text{inode}_1)) = \text{SOME}(bn') \land \\
\text{lookUp} (bn', \text{blocks}(s)) = \text{SOME}(\text{blocks}) \land \text{lookUp} (fid, \text{root}(s)) = \text{NONE} ] \Rightarrow (10) \\
\text{lookUp} (n_1, \text{inode}(\text{write}(fid, i, v, s))) = \text{SOME}(\text{inode}_1) \land \\
\text{lookUp} (bn', \text{blocks}(\text{write}(fid, i, v, s))) = \text{SOME}(\text{blocks}) \\
\end{align*}
\]

\[
\begin{align*}
[\text{lookUp} (n_1, \text{inode}(s)) = \text{SOME}(\text{inode}_2) \land \text{lookUp} (fid, \text{root}(s)) = \text{NONE} ] \Rightarrow (11) \\
\text{lookUp} (n_1, \text{inode}(\text{write}(fid, i, v, s))) = \text{SOME}(\text{inode}_1) \\
\end{align*}
\]

In turn, each of the above four lemmas about write relies on a number of other lemmas about the various operations in the call graph of write (see the relevant remarks in Section 4). We will state those lemmas after we present the proof of $[M_8]$.

We next present a natural-deduction style proof of $[M_8]$ to give the reader an idea of the abstraction level at which Athena proofs are written. We believe that the said level is roughly equivalent to the level at which a formally trained computer scientist would communicate the proof to another computer scientist of a similar background. The proof is rigorous and thorough, but does not descend to the level of primitive inference rules (such as introduction and elimination rules for the logical connectives or congruence rules for equality); the applications of such rules are fairly tedious steps that are filled in by Vampire. The overall proof is guided by constructs such as “pick any ...”, “assume that such and such holds”, “we distinguish two cases”, “from $P_1$, $P_2$ and $P_3$ we infer $P'$”, and so on.

The proof of $[M_8]$ is given below in English, but the level of detail and the overall structure of the argument are isomorphic to those of the formal Athena
deduction (for instance, the formal Athena proof runs to 120 lines, whereas the English proof below is about 64 lines).

**Lemma 3 (M_s).** If \( \text{fid}_1 \neq \text{fid}_2 \) then \( \text{read}(\text{fid}_2, i, \text{write}(\text{fid}_1, j, v, \hat{s})) = \text{read}(\text{fid}_2, i, \hat{s}) \).

**Proof.** Pick arbitrary \( \text{fid}_1, \text{fid}_2, i, j, v \), and \( \hat{s} \), and suppose that
\[
\text{fid}_1 \neq \text{fid}_2. \tag{12}
\]

We will prove the goal
\[
\text{read}(\text{fid}_2, i, \text{write}(\text{fid}_1, j, v, \hat{s})) = \text{read}(\text{fid}_2, i, \hat{s}) \tag{13}
\]
by distinguishing two (mutually exclusive and jointly exhaustive) cases:
\[
\text{lookUp}(\text{fid}_2, \text{root}(\hat{s})) = \text{NONE} \tag{14}
\]
and
\[
\exists n_2, \text{lookUp}(\text{fid}_2, \text{root}(\hat{s})) = \text{SOME}(n_2). \tag{15}
\]
If \( \text{fid}_2 \) is unbound in \( \text{root}(\hat{s}) \) (case (14)), then, by the definition of \( \text{read} \), we have
\[
\text{read}(\text{fid}_2, i, \hat{s}) = \text{FileNotFoundException}. \tag{16}
\]
By Lemma 1, (12), (14), and the reachability of \( \hat{s} \) we conclude
\[
\text{lookUp}(\text{fid}_2, \text{root}(\text{write}(\text{fid}_1, j, v, \hat{s}))) = \text{NONE} \tag{17}
\]
and therefore again by the definition of \( \text{read} \) we infer
\[
\text{read}(\text{fid}_2, i, \text{write}(\text{fid}_1, j, v, \hat{s})) = \text{FileNotFoundException} \tag{18}
\]
and hence (13) follows from (16) and (18). We now consider case (15), whereby
\[
\text{lookUp}(\text{fid}_2, \text{root}(\hat{s})) = \text{SOME}(n_2) \tag{19}
\]
for some inode number \( n_2 \). Since \( \hat{s} \) is reachable, it has \( \text{inu}_0 \), so that
\[
\text{lookUp}(n_2, \text{inodes}(\hat{s})) = \text{SOME}(\text{inode}(fs_2, be_2, bl_2)) \tag{20}
\]
for some \( fs_2, be_2 \), and \( bl_2 \). Moreover, we note that by Lemma 1, (19), (12), and the reachability of \( \hat{s} \), we have
\[
\text{lookUp}(\text{fid}_2, \text{root}(\text{write}(\text{fid}_1, j, v, \hat{s}))) = \text{SOME}(n_2), \tag{21}
\]
We proceed by distinguishing two cases, \( i < fs_2 \) and \( fs_2 \leq i \). Suppose first that \( i < fs_2 \). In that case it becomes evident by inspection that all the preconditions of method \( \text{find-blk-block-val} \) are satisfied: \( \hat{s} \) has the required invariants because it is reachable; \( n_2 \) is mapped by the inode mapping of \( \hat{s} \) to the inode
comprising $fs_2$, $bc_2$, and $bl_2$; and $i < fs_2$. Therefore, we are able to prove that there exist $bn_2$, $block_2$, and $v_2$ such that

$$\text{lookUp}(i \text{ div blockSize}, bl_2) = \text{SOME}(bn_2)$$

(22)

$$\text{lookUp}(bn_2, \text{blocks}(\hat{s})) = \text{SOME}(block_2)$$

(23)

and

$$\text{lookUp}(i \text{ mod blockSize}, block_2) = \text{SOME}(v_2).$$

(24)

It now follows from (19), (20), the assumption $i < fs_2$, (22), (23), (24), and the definition of read that

$$\text{read}(fid_2, i, \hat{s}) = \text{Ok}(v_2)$$

(25)

and therefore our goal (13) becomes reduced to proving

$$\text{read}(fid_2, i, \text{write}(fid_1, j, v, \hat{s})) = \text{Ok}(v_2),$$

(26)

We establish (26) by considering two subcases. First, suppose that $fid_1$ is unbound in the root of $\hat{s}$, i.e.,

$$\text{lookUp}(fid_1, \text{root}(\hat{s})) = \text{NONE}. $$

(27)

Then by (27), (20), (22), (23), the reachability of $\hat{s}$ and Lemma (10), we conclude

$$\begin{align*}
\text{lookUp}(n_2, \text{inodes} & (\text{write}(fid_1, j, v, \hat{s}))) = \\
& \text{SOME}(\text{inode}(fs_2, bc_2, bl_2))
\end{align*}$$

(28)

and

$$\text{lookUp}(bn_2, \text{blocks}(\text{write}(fid_1, j, v, \hat{s}))) = \text{SOME}(block_2). $$

(29)

Accordingly, by the definition of read, (21), (28), the assumption $i < fs_2$, (22), (29), and (24), we obtain the desired (26).

Now suppose, by contrast, that

$$\text{lookUp}(fid_1, \text{root}(\hat{s})) = \text{SOME}(n_1) $$

(30)

for some inode number $n_1$. Since $\hat{s}$ is reachable, it has the invariant inv$_7$, so from (30), (19), and (12) we conclude

$$n_1 \neq n_2.$$ 

(31)

From (8), the reachability of $\hat{s}$, (20), (31), (22), (23), and (30) we can now again derive (28) and (29). Hence, by the definition of read, (21), (28), the assumption $i < fs_2$, (22), (29), and (24) we obtain (26).

We finally consider the possibility $fs_2 \leq i$. In that case the definition of read in tandem with (19) and (20) entails

$$\text{read}(fid_2, i, \hat{s}) = \text{EOF}. $$

(32)

As before, we again distinguish two subcases, according to whether or not $fid_1$ is bound in the root of $\hat{s}$, and we use lemmas (9) and (11), respectively, to infer
(28). In combination with (21), it follows from the definition of \( \text{read} \) that in either case we have
\[
\text{read}(\text{fid}_2, i, \text{write}(\text{fid}_2, j, v, s)) = \text{EOF}
\tag{33}
\]
and the desired equality now follows from (32) and (33). This completes our case analysis and the proof. \( \square \)

Finally, we list below the remaining lemmas needed for lemmas (8), (9), (10), and (11).

\text{writeSmallExtendPreservesInodeAndBlockMaps}:
\[
\text{lookUp} (n_1, \text{inode}s(s)) = \text{SOME}(\text{inode}_1) \land n \neq n_1 \land \\
\text{lookUp} (b_1, \text{blockList}(\text{inode}_1)) = \text{SOME}(\text{block}_1) \land \\
\text{lookUp} (b_1, \text{blocks}(s)) = \text{SOME}(\text{block}_1) \land \text{inv}_0(s) \\
\text{lookUp} (n, \text{inode}s(s)) = \text{SOME}(\text{inode}_2, \text{bs}, \text{bl}) \land \\
\text{lookUp} (i \div \text{blockSize}, \text{bl}) = \text{SOME}(\text{bn}) \land \\
\text{lookUp} (b_1, \text{blocks}(s)) = \text{SOME}(\text{block}_1) \land \text{fs} \leq i \Rightarrow \\
\text{lookUp} (n_2, \text{inode}s(\text{writeSmallExtend}(n_2, i, v, s))) = \text{SOME}(\text{inode}_2) \land \\
\text{lookUp} (b_1, \text{blocks}(\text{writeSmallExtend}(n_2, i, v, s))) = \text{SOME}(\text{block}_1)
\]

\text{writeSmallExtendPreservesInodeMap}:
\[
\text{lookUp} (n_1, \text{inode}s(s)) = \text{SOME}(\text{inode}_1) \land n \neq n_1 \land \\
\text{lookUp} (n, \text{inode}s(s)) = \text{SOME}(\text{inode}_2, \text{bs}, \text{bl}) \land \\
\text{lookUp} (i \div \text{blockSize}, \text{bl}) = \text{SOME}(\text{bn}) \land \\
\text{lookUp} (b, \text{blocks}(s)) = \text{SOME}(\text{block}_1) \land \text{fs} \leq i \Rightarrow \\
\text{lookUp} (n_1, \text{inode}s(\text{writeSmallExtend}(n, i, v, s))) = \text{SOME}(\text{inode}_1)
\]

\text{writeNoExtendPreservesInodeAndBlockMaps}:
\[
\text{lookUp} (n_1, \text{inode}s(s)) = \text{SOME}(\text{inode}_1) \land n \neq n_1 \land \\
\text{lookUp} (b_1, \text{blockList}(\text{inode}_1)) = \text{SOME}(\text{block}_1) \land \\
\text{lookUp} (b_1, \text{blocks}(s)) = \text{SOME}(\text{block}_1) \land \text{inv}_0(s) \\
\text{lookUp} (n, \text{inode}s(s)) = \text{SOME}(\text{inode}) \land \\
\text{lookUp} (i \div \text{blockSize}, \text{blockList}(\text{inode})) = \text{SOME}(\text{bn}) \land \\
\text{lookUp} (b, \text{blocks}(s)) = \text{SOME}(\text{block}) \Rightarrow \\
\text{lookUp} (n_1, \text{inode}s(\text{writeNoExtend}(n, i, v, s))) = \text{SOME}(\text{inode}_1) \land \\
\text{lookUp} (b_1, \text{blocks}(\text{writeNoExtend}(n, i, v, s))) = \text{SOME}(\text{block}_1)
\]

\text{writeNoExtendPreservesInodeMap}:
\[
\text{lookUp} (n_1, \text{inode}s(s)) = \text{SOME}(\text{inode}_1) \land n \neq n_1 \land \\
\text{lookUp} (n, \text{inode}s(s)) = \text{SOME}(\text{inode}) \land \\
\text{lookUp} (i \div \text{blockSize}, \text{blockList}(\text{inode})) = \text{SOME}(\text{bn}) \land \\
\text{lookUp} (b, \text{blocks}(s)) = \text{SOME}(\text{block}) \Rightarrow \\
\text{lookUp} (n_1, \text{inode}s(\text{writeNoExtend}(n, i, v, s))) = \text{SOME}(\text{inode}_1)
\]
allocBlocksPreservesINodeAndBlockMaps:
lookUp(n₁, modes(s)) = SOME(inode₁) ∧ n ≠ n₁ ∧
inDom(n₁, modes(s)) ∧ inv₁(s) ∧
lookUp(bₙ₁, blockList(inode₁)) = SOME(bₙ₁') ∧
lookUp(bₙ₁', blocks(s)) = SOME(block₁) ⇒
lookUp(n₁, modes(allocBlocks(n₁, k₁, fs₁, s₁))) = SOME(inode₁) ∧
lookUp(bₙ₁', blocks(allocBlocks(n₁, k₁, fs₁, s₁))) = SOME(block₁)

allocBlocksPreservesINodeMap:
lookUp(n₁, modes(s)) = SOME(inode₁) ∧ n ≠ n₁ ∧
inDom(n₁, modes(s)) ⇒
lookUp(n₁, modes(allocBlocks(n₁, k₁, fs₁, s₁))) = SOME(inode₁)

extendFilePreservesINodeAndBlockMaps:
lookUp(n₁, modes(s)) = SOME(inode₁) ∧ n ≠ n₁ ∧
inDom(n₁, modes(s)) ∧ inv₁(s) ∧
lookUp(bₙ₁, blockList(inode₁)) = SOME(bₙ₁') ∧
lookUp(bₙ₁', blocks(s)) = SOME(block₁) ⇒
lookUp(n₁, modes(extendFile(n₁, i₁, s₁))) = SOME(inode₁) ∧
lookUp(bₙ₁', blocks(extendFile(n₁, i₁, s₁))) = SOME(block₁)

extendFilePreservesINodeMap:
lookUp(n₁, modes(s)) = SOME(inode₁) ∧ n ≠ n₁ ∧
inDom(n₁, modes(s)) ⇒
lookUp(n₁, modes(extendFile(n₁, i₁, s₁))) = SOME(inode₁)

writeExistingPreservesINodeAndBlockMaps:
inv₁(s) ∧ inv₂(s) ∧ invₚᵢ(s) ∧ inv₁ₗ(s) ∧ inv₁₀(s)
lookUp(n₁, modes(s)) = SOME(inode₁) ∧
n ≠ n₁ ∧ lookUp(bₙ₁, blockList(inode₁)) = SOME(bₙ₁') ∧
lookUp(bₙ₁', blocks(s)) = SOME(block₁) ∧
lookUp(n, modes(s)) = SOME(inode(fs, bc, bl)) ⇒
lookUp(n₁, modes(writeExisting(n₁, i₁, v₁, s₁))) = SOME(inode₁) ∧
lookUp(bₙ₁', blocks(writeExisting(n₁, i₁, v₁, s₁))) = SOME(block₁)

writeExistingPreservesINodeMap:
inv₁(s) ∧ inv₂(s) ∧ invₚᵢ(s) ∧ n ≠ n₁
lookUp(n₁, modes(s)) = SOME(inode₁)
lookUp(n, modes(s)) = SOME(inode(fs, bc, bl)) ⇒
lookUp(n₁, modes(writeExisting(n₁, i₁, v₁, s₁))) = SOME(inode₁)

9 Related work

Techniques for verifying the correct use of file system interfaces expressed as finite state machines are presented in [9, 10, 8, 2]. In this paper we have addressed the more difficult problem of showing that the file system implementation conforms to its specification. Consequently, our proof obligations are stronger and we have resorted to more general deductive verification. Static analysis techniques that handle more complex data structures include predicate abstraction and shape
analysis [19, 18, 14, 6]. These approaches are promising for automating proofs of program properties, but have not been used so far to show full functional correctness, as we do here. Security properties of a Unix file system are studied in [23, Chapter 10]; these properties are orthogonal to the correct functioning of a file system for storing and reading data. A sample specification of a widely used file system is [1]. Simple abstract models of file systems have also been developed in Z [24, Chapter 15].

Alloy [12] is a specification language based on a first-order relational calculus that has been used to describe the directory structure of a file system (but without modelling read and write operations). The Alloy Analyzer is a model finder for Alloy specifications that can be used to check structural properties of file systems in finite scope. The use of Alloy is complementary to proofs [4]. Alloy is useful for debugging, whereas our proofs ensure that the refinement relation holds for any number of files, any file sizes, and all sequences of operations. In addition, readable, high-level proofs can be viewed as explanations of why the file system implementation is correct, and therefore provide guidance to developers on how to modify the system in the future while preserving its correctness.

It is interesting to consider whether the verification burden would be lighter with a system such as PVS [17] or ACL2 [13] that makes heavy use of automatic decision procedures for combinations of first-order theories such as arrays, lists, linear arithmetic, etc. We note that our use of high-performance off-the-shelf ATPs already provides a considerable degree of automation. In our experience, both Vampire and Spass have proven quite effective in non-inductive reasoning about lists, arrays, etc., simply on the basis of first-order axiomatizations of these domains. Our experience supports a recent benchmark study by Armando et al. [5], which showed that a state-of-the-art paramodulation-based prover with a fair search strategy compares favorably with CVC [7] in reasoning about arrays with extensionality.

10 Conclusions

We have presented a correctness proof for the key operations (reading and writing) of a file system based on Unix implementations. We are not aware of any other file system verification attempts dealing with such strong properties as the simulation relation condition, for all possible sequences of file system operations and without a priori bounds on the number of files or their sizes. Despite the apparent simplicity of this particular specification and implementation, our proofs shed light on the general kinds of reasoning that would be required in establishing full functional correctness for any file system. Our results suggest that a combination of state-of-the-art formal methods techniques greatly facilitates the deductive verification of crucial software infrastructure components such as file systems.

We have found Athena to be a powerful framework for carrying out a complex verification effort. Polymorphic sorts and structures allow for natural data modelling; strong support for structural induction facilitates inductive reasoning.
over such datatypes; a block-structured natural deduction format helps to make proofs more readable and writable; a higher-order functional metalanguage and assumption base semantics allow for powerful trusted proof tactics; and the use of first-order logic allows for smooth integration with state-of-the-art first-order ATPs, keeping the proof steps at a high level of detail. Our use of these features was essential in dealing with the strong properties arising from the simulation relation condition, where most of the complexity stems from the details of unbounded data structures.

Acknowledgements. We thank Darko Marinov and Alexandru Salcianu for useful comments on an earlier version of this manuscript.

References

A Some standard Athena libraries

A.1 Options

Options in Athena are represented as follows:

\[
\text{datatype } Option(S) = \text{NONE} | \text{SOME}(S)
\]

Here \( S \) is a sort parameter. Thus \( Option \) can be viewed as a sort constructor that takes an arbitrary sort \( S \) and builds a new sort, \( Option(S) \).

Datatypes in Athena are free algebras with corresponding induction principles. For instance, the following axioms are automatically generated from the above definition:

\[
\forall x : Option(S) \cdot x = \text{NONE} \lor (\exists v : S \cdot x = \text{SOME}(v)) 
\]

(34)

\[
\forall v : S \cdot \text{NONE} \neq \text{SOME}(v)
\]

(35)

\[
\forall v_1 : S, v_2 : S \cdot \text{SOME}(v_1) = \text{SOME}(v_2) \Rightarrow v_1 = v_2
\]

(36)

Note that in the above axioms we annotated quantified variables with their sorts for readability purposes. In practice Athena uses a Hindley-Milner algorithm to infer the most general possible sorts of quantified variables, so such annotations are not necessary; we omit them in the remainder of this Appendix.

Structural induction may be performed on datatypes using a built-in syntax form that Athena offers for that purpose, and which automates much of the tedium associated with inductive proofs (e.g., managing inductive hypotheses in multiply nested inductive arguments).
A.2 Finite maps

Polymorphic finite maps are introduced in Athena as follows:

structure FMap(D, R) = empty-map | update(FMap(D, R), D, R)

Here D and R are sort parameters, representing the sorts of the domain and the range of the map, respectively. The declaration states that every finite map from D to R is either the empty-map or else it is of the form update(m, x, v), i.e., it is an update of some other map m, obtained by binding the argument x to the value v (potentially overwriting whatever assignment x might have had in m).

Like data types, structures are inductively generated: axioms of the form (34) are valid for structures, and induction may be performed on them. However, structures are not necessarily freely generated (elements are not “uniquely readable”), hence Athena does not generate axioms such as (36) for structures. We introduce two additional useful function symbols for finite maps:

\[ \text{lookUp} : D \times \text{FMap}(D, R) \to \text{Option}(D) \]
\[ \text{inDom} : D \times \text{FMap}(D, R) \to \text{Boolean} \]

whose semantics are captured by the following four axioms:

\[ [M_1] \forall x. \text{lookUp}(x, \text{empty-map}) = \text{NONE} \]
\[ [M_2] \forall x v m. \text{lookUp}(x, \text{update}(m, x, v)) = \text{SOME}(v) \]
\[ [M_3] \forall x y v m. x \neq y \Rightarrow \text{lookUp}(x, \text{update}(m, y, v)) = \text{lookUp}(x, m) \]
\[ [M_4] \forall x m. \text{inDom}(x, m) \Rightarrow \exists v. \text{lookUp}(x, m) = \text{SOME}(v) \]

We also have an extensionality axiom for finite maps:

\[ [\text{FMEExt}] \forall m_1 m_2. [\forall x. \text{lookUp}(x, m_1) = \text{lookUp}(x, m_2)] \Rightarrow m_1 = m_2 \]

A.3 Resizable arrays

Resizable arrays are inductively generated by the following structure:

structure RSAArray(S) = makeArray(S, Nat) | arrayWrite(RSAArray(S), Nat, S, S)

That is, a resizable array is either of the form makeArray(x, n), which is a freshly constructed array of length n with the element x in every location from 0 to n-1; or else it is of the form arrayWrite(A, i, x, f), i.e., obtained from an already existing array A by writing the value x into slot i. If i happens to be outside the bounds of A (i.e., arrayLen(A) ≤ i), then the length will increase to i + 1, the value x will be written into the i^{th} position of this extended array, and all the other newly allocated slots will be padded with the “fill” value f. This is made more clear in the axioms of Figure 6. Two additional useful functions are:

\[ \text{arrayLen} : \text{RSAArray}(S) \to \text{Nat} \]
\[ \text{arrayRead} : \text{RSAArray}(S) \times \text{Nat} \to \text{Option}(S) \]

Their semantics are captured by the nine axioms [A1]—[A9] shown in Figure 6. Finally, we have an extensionality axiom for arrays:

\[ [\text{RSAExt}] \forall A_1 A_2. [\forall i. \text{arrayRead}(A_1, i) = \text{arrayRead}(A_2, i)] \Rightarrow A_1 = A_2. \]
Fig. 6. The semantics of resizable arrays

A.4 Natural numbers

Numeric reasoning played an important role in this project. Although no deep number-theoretic results were needed, it was still necessary to introduce all the usual arithmetic operations, including the remainder operation, and derive many simple results for them. We start by introducing the natural numbers as an algebraic datatype:

\[
\text{datatype Nat = zero | succ(Nat)}
\]

This definition automatically generates the following axioms:

\[
\forall x. \text{zero} \neq \text{succ}(x)
\]
\[
\forall x, y. \text{succ}(x) = \text{succ}(y) \Rightarrow x = y
\]
\[
\forall x. x = \text{zero} \lor (\exists y. x = \text{succ}(y))
\]

which are then added to the assumption base.

Next, we introduce function symbols for the predecessor operation:

\[
\text{declare pred: Nat \rightarrow Nat}
\]
as well as for (binary) addition, subtraction, multiplication, division, and remainder:

\[
\text{declare +, -, *, div, mod : Nat \rightarrow Nat}
\]

We also introduce operators for numeric comparisons:

\[
\text{declare <, \leq: Nat \times Nat \rightarrow Boolean}
\]
The semantics of these symbols are given via equational axioms (possibly conditionally equational axioms) that capture the usual primitive recursive definitions of these operations. For example, predecessor is defined as a total function as follows:

\[ \text{pred}(\text{zero}) = \text{zero} \land \forall x. \text{pred}(\text{succ}(x)) = x \]

The definition of binary addition is given via the two axioms:

\[ \forall y. \text{zero} + y = y \]
\[ \forall x, y. \text{succ}(x) + y = \text{succ}(x + y) \]

The definitions of subtraction, multiplication, and numeric comparisons are given by the following axioms:

\[ \forall x. \text{zero} - x = \text{zero} \]
\[ \forall x. x - \text{zero} = x \]
\[ \forall x, y. \text{succ}(x) - \text{succ}(y) = x - y \]
\[ \forall y. \text{zero} \ast y = \text{zero} \]
\[ \forall x, y. \text{succ}(x) \ast y = y + (x \ast y) \]
\[ \forall x. (x < \text{zero}) = \text{false} \]
\[ \forall x. (\text{zero} < \text{succ}(x)) = \text{true} \]
\[ \forall x, y. (\text{succ}(x) < \text{succ}(y)) = x < y \]

The less-than-or-equal symbol is defined in terms of less-than:

\[ x \leq y \iff x = y \lor x < y \]

The definitions of quotient and remainder are as follows:

\[ \forall x. x \; \text{div} \; \text{zero} = \text{zero} \]
\[ \forall x, y. x < y \Rightarrow x \; \text{div} \; y = \text{zero} \]
\[ \forall x, y. (y \neq \text{zero}) \land \neg(x < y) \Rightarrow x \; \text{div} \; y = \text{succ}((x - y) \; \text{div} \; x) \]
\[ \forall x. x \; \text{mod} \; \text{zero} = x \]
\[ \forall x, y. x < y \Rightarrow x \; \text{mod} \; y = x \]
\[ \forall x, y. x < y \Rightarrow x \; \text{mod} \; y = x \]
\[ \forall x, y. (y \neq \text{zero}) \land \neg(x < y) \Rightarrow x \; \text{mod} \; y = (x - y) \; \text{mod} \; y \]

From the above definitions, a number of useful properties can be derived, e.g., that addition is commutative. Most of these properties are derivable only with the aid of a mathematical induction principle—in our case, structural induction on the datatype \text{Nat}. Structural induction in this case corresponds to conventional mathematical induction on the natural numbers. Occasionally it is very convenient to be able to use \textit{strong induction} instead, whereby one inductively assumes the truth of the statement \( P(n) \) for all \( m < n \). For instance, the so-called “division algorithm” result, which states

\[ 0 < b \Rightarrow [(a \; \text{div} \; b) \ast b] + [a \; \text{mod} \; b] = a \]
can be readily proved by strong induction but is much more tedious with conventional induction. In Athena, a strong induction principle on natural numbers is currently formulated as a primitive method. Figure 7 depicts some numeric results that were needed at various points in the project. Most of them were proved automatically by Athena methods that mechanize induction, but a few of them required more detailed proofs. The reader can refer to the file nat.ath in the source code listing for details.
Fig. 7. Useful results about the natural numbers.