Solidity Documentation

Release 0.4.25

Ethereum

Jul 17, 2018

Contents

Solidity is a contract-oriented, high-level language for implementing smart contracts. It was influenced by C++, Python and JavaScript and is designed to target the Ethereum Virtual Machine (EVM).

Solidity is statically typed, supports inheritance, libraries and complex user-defined types among other features.

As you will see, it is possible to create contracts for voting, crowdfunding, blind auctions, multi-signature wallets and more.

Note: The best way to try out Solidity right now is using Remix (it can take a while to load, please be patient). Remix is a web browser based IDE that allows you to write Solidity smart contracts, then deploy and run the smart contracts.

Warning: Since software is written by humans, it can have bugs. Thus, also smart contracts should be created following well-known best-practices in software development. This includes code review, testing, audits and correctness proofs. Also note that users are sometimes more confident in code than its authors. Finally, blockchains have their own things to watch out for, so please take a look at the section *Security Considerations*.

CHAPTER 1

Translations

This documentation is translated into several languages by community volunteers, but the English version stands as a reference.

- Simplified Chinese (in progress)
- Spanish
- Russian (rather outdated)
- Korean (in progress)

CHAPTER 2

Useful links

- Ethereum
- Changelog
- Story Backlog
- Source Code
- Ethereum Stackexchange
- Gitter Chat

CHAPTER 3

Available Solidity Integrations

- **Remix** Browser-based IDE with integrated compiler and Solidity runtime environment without server-side components.
- **IntelliJ IDEA plugin** Solidity plugin for IntelliJ IDEA (and all other JetBrains IDEs)
- **Visual Studio Extension** Solidity plugin for Microsoft Visual Studio that includes the Solidity compiler.
- **Package for SublimeText — Solidity language syntax** Solidity syntax highlighting for SublimeText editor.
- **Etheratom** Plugin for the Atom editor that features syntax highlighting, compilation and a runtime environment (Backend node & VM compatible).
- **Atom Solidity Linter** Plugin for the Atom editor that provides Solidity linting.
- **Atom Solium Linter** Configurable Solidty linter for Atom using Solium as a base.
- **Solium** Linter to identify and fix style and security issues in Solidity.
- **Solhint** Solidity linter that provides security, style guide and best practice rules for smart contract validation.
- **Visual Studio Code extension** Solidity plugin for Microsoft Visual Studio Code that includes syntax highlighting and the Solidity compiler.
- **Emacs Solidity** Plugin for the Emacs editor providing syntax highlighting and compilation error reporting.
- **Vim Solidity** Plugin for the Vim editor providing syntax highlighting.
- **Vim Syntastic** Plugin for the Vim editor providing compile checking.

Discontinued:

- **Mix IDE** Qt based IDE for designing, debugging and testing solidity smart contracts.
- **Ethereum Studio** Specialized web IDE that also provides shell access to a complete Ethereum environment.

CHAPTER 4

Solidity Tools

- **Dapp** Build tool, package manager, and deployment assistant for Solidity.
- **Solidity REPL** Try Solidity instantly with a command-line Solidity console.
- **solgraph** Visualize Solidity control flow and highlight potential security vulnerabilities.
- **evmdis** EVM Disassembler that performs static analysis on the bytecode to provide a higher level of abstraction than raw EVM operations.
- **Doxity** Documentation Generator for Solidity.

CHAPTER 5

Third-Party Solidity Parsers and Grammars

- **solidity-parser** Solidity parser for JavaScript
- **Solidity Grammar for ANTLR 4** Solidity grammar for the ANTLR 4 parser generator

CHAPTER 6

Language Documentation

On the next pages, we will first see a *simple smart contract* written in Solidity followed by the basics about *blockchains* and the *Ethereum Virtual Machine*.

The next section will explain several *features* of Solidity by giving useful *example contracts* Remember that you can always try out the contracts in your browser!

The last and most extensive section will cover all aspects of Solidity in depth.

If you still have questions, you can try searching or asking on the Ethereum Stackexchange site, or come to our gitter channel. Ideas for improving Solidity or this documentation are always welcome!

CHAPTER 7

Contents

Keyword Index, Search Page

7.1 Introduction to Smart Contracts

7.1.1 A Simple Smart Contract

Let us begin with the most basic example. It is fine if you do not understand everything right now, we will go into more detail later.

Storage

```
pragma solidity ^0.4.0;

contract SimpleStorage {
    uint storedData;

    function set(uint x) public {
        storedData = x;
    }

    function get() public view returns (uint) {
        return storedData;
    }
}
```

The first line simply tells that the source code is written for Solidity version 0.4.0 or anything newer that does not break functionality (up to, but not including, version 0.5.0). This is to ensure that the contract does not suddenly behave differently with a new compiler version. The keyword `pragma` is called that because, in general, pragmas are instructions for the compiler about how to treat the source code (e.g. pragma once).

A contract in the sense of Solidity is a collection of code (its *functions*) and data (its *state*) that resides at a specific address on the Ethereum blockchain. The line `uint storedData;` declares a state variable called `storedData` of type `uint` (unsigned integer of 256 bits). You can think of it as a single slot in a database that can be queried and altered by calling functions of the code that manages the database. In the case of Ethereum, this is always the owning contract. And in this case, the functions `set` and `get` can be used to modify or retrieve the value of the variable.

To access a state variable, you do not need the prefix `this.` as is common in other languages.

This contract does not do much yet (due to the infrastructure built by Ethereum) apart from allowing anyone to store a single number that is accessible by anyone in the world without a (feasible) way to prevent you from publishing this number. Of course, anyone could just call `set` again with a different value and overwrite your number, but the number will still be stored in the history of the blockchain. Later, we will see how you can impose access restrictions so that only you can alter the number.

Note: All identifiers (contract names, function names and variable names) are restricted to the ASCII character set. It is possible to store UTF-8 encoded data in string variables.

Warning: Be careful with using Unicode text, as similar looking (or even identical) characters can have different code points and as such will be encoded as a different byte array.

Subcurrency Example

The following contract will implement the simplest form of a cryptocurrency. It is possible to generate coins out of thin air, but only the person that created the contract will be able to do that (it is trivial to implement a different issuance scheme). Furthermore, anyone can send coins to each other without any need for registering with username and password — all you need is an Ethereum keypair.

```
pragma solidity >0.4.24;

contract Coin {
    // The keyword "public" makes those variables
    // readable from outside.
    address public minter;
    mapping (address => uint) public balances;

    // Events allow light clients to react to
    // changes efficiently.
    event Sent(address from, address to, uint amount);

    // This is the constructor whose code is
    // run only when the contract is created.
    constructor() public {
        minter = msg.sender;
    }

    function mint(address receiver, uint amount) public {
        if (msg.sender != minter) return;
        balances[receiver] += amount;
    }

    function send(address receiver, uint amount) public {
        if (balances[msg.sender] < amount) return;
```

(continues on next page)

(continued from previous page)

```
        balances[msg.sender] -= amount;
        balances[receiver] += amount;
        emit Sent(msg.sender, receiver, amount);
    }
}
```

This contract introduces some new concepts, let us go through them one by one.

The line `address public minter;` declares a state variable of type address that is publicly accessible. The `address` type is a 160-bit value that does not allow any arithmetic operations. It is suitable for storing addresses of contracts or keypairs belonging to external persons. The keyword `public` automatically generates a function that allows you to access the current value of the state variable from outside of the contract. Without this keyword, other contracts have no way to access the variable. The code of the function generated by the compiler is roughly equivalent to the following:

```
function minter() returns (address) { return minter; }
```

Of course, adding a function exactly like that will not work because we would have a function and a state variable with the same name, but hopefully, you get the idea - the compiler figures that out for you.

The next line, `mapping (address => uint) public balances;` also creates a public state variable, but it is a more complex datatype. The type maps addresses to unsigned integers. Mappings can be seen as hash tables which are virtually initialized such that every possible key exists and is mapped to a value whose byte-representation is all zeros. This analogy does not go too far, though, as it is neither possible to obtain a list of all keys of a mapping, nor a list of all values. So either keep in mind (or better, keep a list or use a more advanced data type) what you added to the mapping or use it in a context where this is not needed, like this one. The *getter function* created by the `public` keyword is a bit more complex in this case. It roughly looks like the following:

```
function balances(address _account) public view returns (uint) {
    return balances[_account];
}
```

As you see, you can use this function to easily query the balance of a single account.

The line `event Sent(address from, address to, uint amount);` declares a so-called "event" which is emitted in the last line of the function `send`. User interfaces (as well as server applications of course) can listen for those events being emitted on the blockchain without much cost. As soon as it is emitted, the listener will also receive the arguments `from`, `to` and `amount`, which makes it easy to track transactions. In order to listen for this event, you would use

```
Coin.Sent().watch({}, '', function(error, result) {
    if (!error) {
        console.log("Coin transfer: " + result.args.amount +
            " coins were sent from " + result.args.from +
            " to " + result.args.to + ".");
        console.log("Balances now:\n" +
            "Sender: " + Coin.balances.call(result.args.from) +
            "Receiver: " + Coin.balances.call(result.args.to));
    }
})
```

Note how the automatically generated function `balances` is called from the user interface.

The special function `Coin` is the constructor which is run during creation of the contract and cannot be called afterwards. It permanently stores the address of the person creating the contract: `msg` (together with `tx` and `block`) is a magic global variable that contains some properties which allow access to the blockchain. `msg.sender` is always the address where the current (external) function call came from.

Finally, the functions that will actually end up with the contract and can be called by users and contracts alike are `mint` and `send`. If `mint` is called by anyone except the account that created the contract, nothing will happen. On the other hand, `send` can be used by anyone (who already has some of these coins) to send coins to anyone else. Note that if you use this contract to send coins to an address, you will not see anything when you look at that address on a blockchain explorer, because the fact that you sent coins and the changed balances are only stored in the data storage of this particular coin contract. By the use of events it is relatively easy to create a "blockchain explorer" that tracks transactions and balances of your new coin.

7.1.2 Blockchain Basics

Blockchains as a concept are not too hard to understand for programmers. The reason is that most of the complications (mining, hashing, elliptic-curve cryptography, peer-to-peer networks, etc.) are just there to provide a certain set of features and promises. Once you accept these features as given, you do not have to worry about the underlying technology - or do you have to know how Amazon's AWS works internally in order to use it?

Transactions

A blockchain is a globally shared, transactional database. This means that everyone can read entries in the database just by participating in the network. If you want to change something in the database, you have to create a so-called transaction which has to be accepted by all others. The word transaction implies that the change you want to make (assume you want to change two values at the same time) is either not done at all or completely applied. Furthermore, while your transaction is applied to the database, no other transaction can alter it.

As an example, imagine a table that lists the balances of all accounts in an electronic currency. If a transfer from one account to another is requested, the transactional nature of the database ensures that if the amount is subtracted from one account, it is always added to the other account. If due to whatever reason, adding the amount to the target account is not possible, the source account is also not modified.

Furthermore, a transaction is always cryptographically signed by the sender (creator). This makes it straightforward to guard access to specific modifications of the database. In the example of the electronic currency, a simple check ensures that only the person holding the keys to the account can transfer money from it.

Blocks

One major obstacle to overcome is what, in Bitcoin terms, is called a "double-spend attack": What happens if two transactions exist in the network that both want to empty an account, a so-called conflict?

The abstract answer to this is that you do not have to care. An order of the transactions will be selected for you, the transactions will be bundled into what is called a "block" and then they will be executed and distributed among all participating nodes. If two transactions contradict each other, the one that ends up being second will be rejected and not become part of the block.

These blocks form a linear sequence in time and that is where the word "blockchain" derives from. Blocks are added to the chain in rather regular intervals - for Ethereum this is roughly every 17 seconds.

As part of the "order selection mechanism" (which is called "mining") it may happen that blocks are reverted from time to time, but only at the "tip" of the chain. The more blocks that are added on top, the less likely it is. So it might be that your transactions are reverted and even removed from the blockchain, but the longer you wait, the less likely it will be.

7.1.3 The Ethereum Virtual Machine

Overview

The Ethereum Virtual Machine or EVM is the runtime environment for smart contracts in Ethereum. It is not only sandboxed but actually completely isolated, which means that code running inside the EVM has no access to network, filesystem or other processes. Smart contracts even have limited access to other smart contracts.

Accounts

There are two kinds of accounts in Ethereum which share the same address space: **External accounts** that are controlled by public-private key pairs (i.e. humans) and **contract accounts** which are controlled by the code stored together with the account.

The address of an external account is determined from the public key while the address of a contract is determined at the time the contract is created (it is derived from the creator address and the number of transactions sent from that address, the so-called "nonce").

Regardless of whether or not the account stores code, the two types are treated equally by the EVM.

Every account has a persistent key-value store mapping 256-bit words to 256-bit words called **storage**.

Furthermore, every account has a **balance** in Ether (in "Wei" to be exact) which can be modified by sending transactions that include Ether.

Transactions

A transaction is a message that is sent from one account to another account (which might be the same or the special zero-account, see below). It can include binary data (its payload) and Ether.

If the target account contains code, that code is executed and the payload is provided as input data.

If the target account is the zero-account (the account with the address `0`), the transaction creates a **new contract**. As already mentioned, the address of that contract is not the zero address but an address derived from the sender and its number of transactions sent (the "nonce"). The payload of such a contract creation transaction is taken to be EVM bytecode and executed. The output of this execution is permanently stored as the code of the contract. This means that in order to create a contract, you do not send the actual code of the contract, but in fact code that returns that code when executed.

Note: While a contract is being created, its code is still empty. Because of that, you should not call back into the contract under construction until its constructor has finished executing.

Gas

Upon creation, each transaction is charged with a certain amount of **gas**, whose purpose is to limit the amount of work that is needed to execute the transaction and to pay for this execution. While the EVM executes the transaction, the gas is gradually depleted according to specific rules.

The **gas price** is a value set by the creator of the transaction, who has to pay `gas_price * gas` up front from the sending account. If some gas is left after the execution, it is refunded in the same way.

If the gas is used up at any point (i.e. it is negative), an out-of-gas exception is triggered, which reverts all modifications made to the state in the current call frame.

Storage, Memory and the Stack

Each account has a persistent memory area which is called **storage**. Storage is a key-value store that maps 256-bit words to 256-bit words. It is not possible to enumerate storage from within a contract and it is comparatively costly to read and even more so, to modify storage. A contract can neither read nor write to any storage apart from its own.

The second memory area is called **memory**, of which a contract obtains a freshly cleared instance for each message call. Memory is linear and can be addressed at byte level, but reads are limited to a width of 256 bits, while writes can be either 8 bits or 256 bits wide. Memory is expanded by a word (256-bit), when accessing (either reading or writing) a previously untouched memory word (ie. any offset within a word). At the time of expansion, the cost in gas must be paid. Memory is more costly the larger it grows (it scales quadratically).

The EVM is not a register machine but a stack machine, so all computations are performed on an area called the **stack**. It has a maximum size of 1024 elements and contains words of 256 bits. Access to the stack is limited to the top end in the following way: It is possible to copy one of the topmost 16 elements to the top of the stack or swap the topmost element with one of the 16 elements below it. All other operations take the topmost two (or one, or more, depending on the operation) elements from the stack and push the result onto the stack. Of course it is possible to move stack elements to storage or memory, but it is not possible to just access arbitrary elements deeper in the stack without first removing the top of the stack.

Instruction Set

The instruction set of the EVM is kept minimal in order to avoid incorrect implementations which could cause consensus problems. All instructions operate on the basic data type, 256-bit words. The usual arithmetic, bit, logical and comparison operations are present. Conditional and unconditional jumps are possible. Furthermore, contracts can access relevant properties of the current block like its number and timestamp.

Message Calls

Contracts can call other contracts or send Ether to non-contract accounts by the means of message calls. Message calls are similar to transactions, in that they have a source, a target, data payload, Ether, gas and return data. In fact, every transaction consists of a top-level message call which in turn can create further message calls.

A contract can decide how much of its remaining **gas** should be sent with the inner message call and how much it wants to retain. If an out-of-gas exception happens in the inner call (or any other exception), this will be signalled by an error value put onto the stack. In this case, only the gas sent together with the call is used up. In Solidity, the calling contract causes a manual exception by default in such situations, so that exceptions "bubble up" the call stack.

As already said, the called contract (which can be the same as the caller) will receive a freshly cleared instance of memory and has access to the call payload - which will be provided in a separate area called the **calldata**. After it has finished execution, it can return data which will be stored at a location in the caller's memory preallocated by the caller.

Calls are **limited** to a depth of 1024, which means that for more complex operations, loops should be preferred over recursive calls.

Delegatecall / Callcode and Libraries

There exists a special variant of a message call, named **delegatecall** which is identical to a message call apart from the fact that the code at the target address is executed in the context of the calling contract and `msg.sender` and `msg.value` do not change their values.

This means that a contract can dynamically load code from a different address at runtime. Storage, current address and balance still refer to the calling contract, only the code is taken from the called address.

This makes it possible to implement the "library" feature in Solidity: Reusable library code that can be applied to a contract's storage, e.g. in order to implement a complex data structure.

Logs

It is possible to store data in a specially indexed data structure that maps all the way up to the block level. This feature called **logs** is used by Solidity in order to implement **events**. Contracts cannot access log data after it has been created, but they can be efficiently accessed from outside the blockchain. Since some part of the log data is stored in bloom filters, it is possible to search for this data in an efficient and cryptographically secure way, so network peers that do not download the whole blockchain ("light clients") can still find these logs.

Create

Contracts can even create other contracts using a special opcode (i.e. they do not simply call the zero address). The only difference between these **create calls** and normal message calls is that the payload data is executed and the result stored as code and the caller / creator receives the address of the new contract on the stack.

Self-destruct

The only possibility that code is removed from the blockchain is when a contract at that address performs the `selfdestruct` operation. The remaining Ether stored at that address is sent to a designated target and then the storage and code is removed from the state.

Warning: Even if a contract's code does not contain a call to `selfdestruct`, it can still perform that operation using `delegatecall` or `callcode`.

Note: The pruning of old contracts may or may not be implemented by Ethereum clients. Additionally, archive nodes could choose to keep the contract storage and code indefinitely.

Note: Currently **external accounts** cannot be removed from the state.

7.2 Installing the Solidity Compiler

7.2.1 Versioning

Solidity versions follow semantic versioning and in addition to releases, **nightly development builds** are also made available. The nightly builds are not guaranteed to be working and despite best efforts they might contain undocumented and/or broken changes. We recommend using the latest release. Package installers below will use the latest release.

7.2.2 Remix

We recommend Remix for small contracts and for quickly learning Solidity.

Access Remix online, you don't need to install anything. If you want to use it without connection to the Internet, go to https://github.com/ethereum/browser-solidity/tree/gh-pages and download the .ZIP file as explained on that page.

Further options on this page detail installing commandline Solidity compiler software on your computer. Choose a commandline compiler if you are working on a larger contract or if you require more compilation options.

7.2.3 npm / Node.js

Use *npm* for a convenient and portable way to install *solcjs*, a Solidity compiler. The *solcjs* program has fewer features than all options further down this page. Our *Using the Commandline Compiler* documentation assumes you are using the full-featured compiler, *solc*. So if you install *solcjs* from *npm* then you will stop reading the documentation here and then continue to solc-js.

Note: The solc-js project is derived from the C++ *solc* by using Emscripten. *solc-js* can be used in JavaScript projects directly (such as Remix). Please refer to the solc-js repository for instructions.

```
npm install -g solc
```

Note: The commandline is named *solcjs*.

The comandline options of *solcjs* are not compatible with *solc* and tools (such as *geth*) expecting the behaviour of *solc* will not work with *solcjs*.

7.2.4 Docker

We provide up to date docker builds for the compiler. The `stable` repository contains released versions while the `nightly` repository contains potentially unstable changes in the develop branch.

```
docker run ethereum/solc:stable --version
```

Currently, the docker image only contains the compiler executable, so you have to do some additional work to link in the source and output directories.

7.2.5 Binary Packages

Binary packages of Solidity are available at solidity/releases.

We also have PPAs for Ubuntu. For the latest stable version.

```
sudo add-apt-repository ppa:ethereum/ethereum
sudo apt-get update
sudo apt-get install solc
```

If you want to use the cutting edge developer version:

```
sudo add-apt-repository ppa:ethereum/ethereum
sudo add-apt-repository ppa:ethereum/ethereum-dev
sudo apt-get update
sudo apt-get install solc
```

We are also releasing a snap package, which is installable in all the supported Linux distros. To install the latest stable version of solc:

```
sudo snap install solc
```

Or if you want to help testing the unstable solc with the most recent changes from the development branch:

```
sudo snap install solc --edge
```

Arch Linux also has packages, albeit limited to the latest development version:

```
pacman -S solidity
```

Homebrew is missing pre-built bottles at the time of writing, following a Jenkins to TravisCI migration, but Homebrew should still work just fine as a means to build-from-source. We will re-add the pre-built bottles soon.

```
brew update
brew upgrade
brew tap ethereum/ethereum
brew install solidity
```

If you need a specific version of Solidity you can install a Homebrew formula directly from Github.

View solidity.rb commits on Github.

Follow the history links until you have a raw file link of a specific commit of `solidity.rb`.

Install it using `brew`:

```
brew unlink solidity
# Install 0.4.8
brew install https://raw.githubusercontent.com/ethereum/homebrew-ethereum/
↪77cce03da9f289e5a3ffe579840d3c5dc0a62717/solidity.rb
```

Gentoo Linux also provides a solidity package that can be installed using `emerge`:

```
emerge dev-lang/solidity
```

7.2.6 Building from Source

Clone the Repository

To clone the source code, execute the following command:

```
git clone --recursive https://github.com/ethereum/solidity.git
cd solidity
```

If you want to help developing Solidity, you should fork Solidity and add your personal fork as a second remote:

```
cd solidity
git remote add personal git@github.com:[username]/solidity.git
```

Solidity has git submodules. Ensure they are properly loaded:

```
git submodule update --init --recursive
```

Prerequisites - macOS

For macOS, ensure that you have the latest version of Xcode installed. This contains the Clang C++ compiler, the Xcode IDE and other Apple development tools which are required for building C++ applications on OS X. If you are installing Xcode for the first time, or have just installed a new version then you will need to agree to the license before you can do command-line builds:

```
sudo xcodebuild -license accept
```

Our OS X builds require you to install the Homebrew package manager for installing external dependencies. Here's how to uninstall Homebrew, if you ever want to start again from scratch.

Prerequisites - Windows

You will need to install the following dependencies for Windows builds of Solidity:

Software	Notes
Git for Windows	Command-line tool for retrieving source from Github.
CMake	Cross-platform build file generator.
Visual Studio 2017 Build Tools	C++ compiler
Visual Studio 2017 (Optional)	C++ compiler and dev environment.

If you've already had one IDE and only need compiler and libraries, you could install Visual Studio 2017 Build Tools.

Visual Studio 2017 provides both IDE and necessary compiler and libraries. So if you have not got an IDE and prefer to develop solidity, Visual Studio 2017 may be an choice for you to get everything setup easily.

Here is the list of components that should be installed in Visual Studio 2017 Build Tools or Visual Studio 2017:

- Visual Studio C++ core features
- VC++ 2017 v141 toolset (x86,x64)
- Windows Universal CRT SDK
- Windows 8.1 SDK
- C++/CLI support

External Dependencies

We now have a "one button" script which installs all required external dependencies on macOS, Windows and on numerous Linux distros. This used to be a multi-step manual process, but is now a one-liner:

```
./scripts/install_deps.sh
```

Or, on Windows:

```
scripts\install_deps.bat
```

Command-Line Build

Be sure to install External Dependencies (see above) before build.

Solidity project uses CMake to configure the build. Building Solidity is quite similar on Linux, macOS and other Unices:

```
mkdir build
cd build
cmake .. && make
```

or even easier:

```
#note: this will install binaries solc and soltest at usr/local/bin
./scripts/build.sh
```

And even for Windows:

```
mkdir build
cd build
cmake -G "Visual Studio 15 2017 Win64" ..
```

This latter set of instructions should result in the creation of **solidity.sln** in that build directory. Double-clicking on that file should result in Visual Studio firing up. We suggest building **RelWithDebugInfo** configuration, but all others work.

Alternatively, you can build for Windows on the command-line, like so:

```
cmake --build . --config RelWithDebInfo
```

7.2.7 CMake options

If you are interested what CMake options are available run `cmake .. -LH`.

7.2.8 The version string in detail

The Solidity version string contains four parts:

- the version number
- pre-release tag, usually set to `develop.YYYY.MM.DD` or `nightly.YYYY.MM.DD`
- commit in the format of `commit.GITHASH`
- platform has arbitrary number of items, containing details about the platform and compiler

If there are local modifications, the commit will be postfixed with `.mod`.

These parts are combined as required by Semver, where the Solidity pre-release tag equals to the Semver pre-release and the Solidity commit and platform combined make up the Semver build metadata.

A release example: `0.4.8+commit.60cc1668.Emscripten.clang`.

A pre-release example: `0.4.9-nightly.2017.1.17+commit.6ecb4aa3.Emscripten.clang`

7.2.9 Important information about versioning

After a release is made, the patch version level is bumped, because we assume that only patch level changes follow. When changes are merged, the version should be bumped according to semver and the severity of the change. Finally, a release is always made with the version of the current nightly build, but without the `prerelease` specifier.

Example:

0. the 0.4.0 release is made

1. nightly build has a version of 0.4.1 from now on
2. non-breaking changes are introduced - no change in version
3. a breaking change is introduced - version is bumped to 0.5.0
4. the 0.5.0 release is made

This behaviour works well with the *version pragma*.

7.3 Solidity by Example

7.3.1 Voting

The following contract is quite complex, but showcases a lot of Solidity's features. It implements a voting contract. Of course, the main problems of electronic voting is how to assign voting rights to the correct persons and how to prevent manipulation. We will not solve all problems here, but at least we will show how delegated voting can be done so that vote counting is **automatic and completely transparent** at the same time.

The idea is to create one contract per ballot, providing a short name for each option. Then the creator of the contract who serves as chairperson will give the right to vote to each address individually.

The persons behind the addresses can then choose to either vote themselves or to delegate their vote to a person they trust.

At the end of the voting time, `winningProposal()` will return the proposal with the largest number of votes.

```
pragma solidity ^0.4.22;

/// @title Voting with delegation.
contract Ballot {
    // This declares a new complex type which will
    // be used for variables later.
    // It will represent a single voter.
    struct Voter {
        uint weight; // weight is accumulated by delegation
        bool voted;  // if true, that person already voted
        address delegate; // person delegated to
        uint vote;   // index of the voted proposal
    }

    // This is a type for a single proposal.
    struct Proposal {
        bytes32 name;   // short name (up to 32 bytes)
        uint voteCount; // number of accumulated votes
    }

    address public chairperson;

    // This declares a state variable that
    // stores a `Voter` struct for each possible address.
    mapping(address => Voter) public voters;

    // A dynamically-sized array of `Proposal` structs.
    Proposal[] public proposals;

    /// Create a new ballot to choose one of `proposalNames`.
```

(continues on next page)

(continued from previous page)

```
    constructor(bytes32[] memory proposalNames) public {
        chairperson = msg.sender;
        voters[chairperson].weight = 1;

        // For each of the provided proposal names,
        // create a new proposal object and add it
        // to the end of the array.
        for (uint i = 0; i < proposalNames.length; i++) {
            // `Proposal({...})` creates a temporary
            // Proposal object and `proposals.push(...)`
            // appends it to the end of `proposals`.
            proposals.push(Proposal({
                name: proposalNames[i],
                voteCount: 0
            }));
        }
    }

    // Give `voter` the right to vote on this ballot.
    // May only be called by `chairperson`.
    function giveRightToVote(address voter) public {
        // If the first argument of `require` evaluates
        // to `false`, execution terminates and all
        // changes to the state and to Ether balances
        // are reverted.
        // This used to consume all gas in old EVM versions, but
        // not anymore.
        // It is often a good idea to use `require` to check if
        // functions are called correctly.
        // As a second argument, you can also provide an
        // explanation about what went wrong.
        require(
            msg.sender == chairperson,
            "Only chairperson can give right to vote."
        );
        require(
            !voters[voter].voted,
            "The voter already voted."
        );
        require(voters[voter].weight == 0);
        voters[voter].weight = 1;
    }

    /// Delegate your vote to the voter `to`.
    function delegate(address to) public {
        // assigns reference
        Voter storage sender = voters[msg.sender];
        require(!sender.voted, "You already voted.");

        require(to != msg.sender, "Self-delegation is disallowed.");

        // Forward the delegation as long as
        // `to` also delegated.
        // In general, such loops are very dangerous,
        // because if they run too long, they might
        // need more gas than is available in a block.
        // In this case, the delegation will not be executed,
```

(continues on next page)

(continued from previous page)

```
        // but in other situations, such loops might
        // cause a contract to get "stuck" completely.
        while (voters[to].delegate != address(0)) {
            to = voters[to].delegate;

            // We found a loop in the delegation, not allowed.
            require(to != msg.sender, "Found loop in delegation.");
        }

        // Since `sender` is a reference, this
        // modifies `voters[msg.sender].voted`
        sender.voted = true;
        sender.delegate = to;
        Voter storage delegate_ = voters[to];
        if (delegate_.voted) {
            // If the delegate already voted,
            // directly add to the number of votes
            proposals[delegate_.vote].voteCount += sender.weight;
        } else {
            // If the delegate did not vote yet,
            // add to her weight.
            delegate_.weight += sender.weight;
        }
    }

    /// Give your vote (including votes delegated to you)
    /// to proposal `proposals[proposal].name`.
    function vote(uint proposal) public {
        Voter storage sender = voters[msg.sender];
        require(!sender.voted, "Already voted.");
        sender.voted = true;
        sender.vote = proposal;

        // If `proposal` is out of the range of the array,
        // this will throw automatically and revert all
        // changes.
        proposals[proposal].voteCount += sender.weight;
    }

    /// @dev Computes the winning proposal taking all
    /// previous votes into account.
    function winningProposal() public view
            returns (uint winningProposal_)
    {
        uint winningVoteCount = 0;
        for (uint p = 0; p < proposals.length; p++) {
            if (proposals[p].voteCount > winningVoteCount) {
                winningVoteCount = proposals[p].voteCount;
                winningProposal_ = p;
            }
        }
    }

    // Calls winningProposal() function to get the index
    // of the winner contained in the proposals array and then
    // returns the name of the winner
    function winnerName() public view
```

(continues on next page)

(continued from previous page)

```
            returns (bytes32 winnerName_)
    {
        winnerName_ = proposals[winningProposal()].name;
    }
}
```

Possible Improvements

Currently, many transactions are needed to assign the rights to vote to all participants. Can you think of a better way?

7.3.2 Blind Auction

In this section, we will show how easy it is to create a completely blind auction contract on Ethereum. We will start with an open auction where everyone can see the bids that are made and then extend this contract into a blind auction where it is not possible to see the actual bid until the bidding period ends.

Simple Open Auction

The general idea of the following simple auction contract is that everyone can send their bids during a bidding period. The bids already include sending money / ether in order to bind the bidders to their bid. If the highest bid is raised, the previously highest bidder gets her money back. After the end of the bidding period, the contract has to be called manually for the beneficiary to receive his money - contracts cannot activate themselves.

```
pragma solidity ^0.4.22;

contract SimpleAuction {
    // Parameters of the auction. Times are either
    // absolute unix timestamps (seconds since 1970-01-01)
    // or time periods in seconds.
    address public beneficiary;
    uint public auctionEnd;

    // Current state of the auction.
    address public highestBidder;
    uint public highestBid;

    // Allowed withdrawals of previous bids
    mapping(address => uint) pendingReturns;

    // Set to true at the end, disallows any change
    bool ended;

    // Events that will be fired on changes.
    event HighestBidIncreased(address bidder, uint amount);
    event AuctionEnded(address winner, uint amount);

    // The following is a so-called natspec comment,
    // recognizable by the three slashes.
    // It will be shown when the user is asked to
    // confirm a transaction.

    /// Create a simple auction with `_biddingTime`
```

(continues on next page)

(continued from previous page)

```
    /// seconds bidding time on behalf of the
    /// beneficiary address `_beneficiary`.
    constructor(
        uint _biddingTime,
        address _beneficiary
    ) public {
        beneficiary = _beneficiary;
        auctionEnd = now + _biddingTime;
    }

    /// Bid on the auction with the value sent
    /// together with this transaction.
    /// The value will only be refunded if the
    /// auction is not won.
    function bid() public payable {
        // No arguments are necessary, all
        // information is already part of
        // the transaction. The keyword payable
        // is required for the function to
        // be able to receive Ether.

        // Revert the call if the bidding
        // period is over.
        require(
            now <= auctionEnd,
            "Auction already ended."
        );

        // If the bid is not higher, send the
        // money back.
        require(
            msg.value > highestBid,
            "There already is a higher bid."
        );

        if (highestBid != 0) {
            // Sending back the money by simply using
            // highestBidder.send(highestBid) is a security risk
            // because it could execute an untrusted contract.
            // It is always safer to let the recipients
            // withdraw their money themselves.
            pendingReturns[highestBidder] += highestBid;
        }
        highestBidder = msg.sender;
        highestBid = msg.value;
        emit HighestBidIncreased(msg.sender, msg.value);
    }

    /// Withdraw a bid that was overbid.
    function withdraw() public returns (bool) {
        uint amount = pendingReturns[msg.sender];
        if (amount > 0) {
            // It is important to set this to zero because the recipient
            // can call this function again as part of the receiving call
            // before `send` returns.
            pendingReturns[msg.sender] = 0;
```

(continues on next page)

(continued from previous page)

```
            if (!msg.sender.send(amount)) {
                // No need to call throw here, just reset the amount owing
                pendingReturns[msg.sender] = amount;
                return false;
            }
        }
        return true;
    }

    /// End the auction and send the highest bid
    /// to the beneficiary.
    function auctionEnd() public {
        // It is a good guideline to structure functions that interact
        // with other contracts (i.e. they call functions or send Ether)
        // into three phases:
        // 1. checking conditions
        // 2. performing actions (potentially changing conditions)
        // 3. interacting with other contracts
        // If these phases are mixed up, the other contract could call
        // back into the current contract and modify the state or cause
        // effects (ether payout) to be performed multiple times.
        // If functions called internally include interaction with external
        // contracts, they also have to be considered interaction with
        // external contracts.

        // 1. Conditions
        require(now >= auctionEnd, "Auction not yet ended.");
        require(!ended, "auctionEnd has already been called.");

        // 2. Effects
        ended = true;
        emit AuctionEnded(highestBidder, highestBid);

        // 3. Interaction
        beneficiary.transfer(highestBid);
    }
}
```

Blind Auction

The previous open auction is extended to a blind auction in the following. The advantage of a blind auction is that there is no time pressure towards the end of the bidding period. Creating a blind auction on a transparent computing platform might sound like a contradiction, but cryptography comes to the rescue.

During the **bidding period**, a bidder does not actually send her bid, but only a hashed version of it. Since it is currently considered practically impossible to find two (sufficiently long) values whose hash values are equal, the bidder commits to the bid by that. After the end of the bidding period, the bidders have to reveal their bids: They send their values unencrypted and the contract checks that the hash value is the same as the one provided during the bidding period.

Another challenge is how to make the auction **binding and blind** at the same time: The only way to prevent the bidder from just not sending the money after he won the auction is to make her send it together with the bid. Since value transfers cannot be blinded in Ethereum, anyone can see the value.

The following contract solves this problem by accepting any value that is larger than the highest bid. Since this can of course only be checked during the reveal phase, some bids might be **invalid**, and this is on purpose (it even provides

an explicit flag to place invalid bids with high value transfers): Bidders can confuse competition by placing several high or low invalid bids.

```
pragma solidity >0.4.23 <0.5.0;

contract BlindAuction {
    struct Bid {
        bytes32 blindedBid;
        uint deposit;
    }

    address public beneficiary;
    uint public biddingEnd;
    uint public revealEnd;
    bool public ended;

    mapping(address => Bid[]) public bids;

    address public highestBidder;
    uint public highestBid;

    // Allowed withdrawals of previous bids
    mapping(address => uint) pendingReturns;

    event AuctionEnded(address winner, uint highestBid);

    /// Modifiers are a convenient way to validate inputs to
    /// functions. `onlyBefore` is applied to `bid` below:
    /// The new function body is the modifier's body where
    /// `_` is replaced by the old function body.
    modifier onlyBefore(uint _time) { require(now < _time); _; }
    modifier onlyAfter(uint _time) { require(now > _time); _; }

    constructor(
        uint _biddingTime,
        uint _revealTime,
        address _beneficiary
    ) public {
        beneficiary = _beneficiary;
        biddingEnd = now + _biddingTime;
        revealEnd = biddingEnd + _revealTime;
    }

    /// Place a blinded bid with `_blindedBid` =
    /// keccak256(abi.encodePacked(value, fake, secret)).
    /// The sent ether is only refunded if the bid is correctly
    /// revealed in the revealing phase. The bid is valid if the
    /// ether sent together with the bid is at least "value" and
    /// "fake" is not true. Setting "fake" to true and sending
    /// not the exact amount are ways to hide the real bid but
    /// still make the required deposit. The same address can
    /// place multiple bids.
    function bid(bytes32 _blindedBid)
        public
        payable
        onlyBefore(biddingEnd)
    {
        bids[msg.sender].push(Bid({
```

(continues on next page)

(continued from previous page)

```
            blindedBid: _blindedBid,
            deposit: msg.value
        }));
    }

    /// Reveal your blinded bids. You will get a refund for all
    /// correctly blinded invalid bids and for all bids except for
    /// the totally highest.
    function reveal(
        uint[] memory _values,
        bool[] memory _fake,
        bytes32[] memory _secret
    )
        public
        onlyAfter(biddingEnd)
        onlyBefore(revealEnd)
    {
        uint length = bids[msg.sender].length;
        require(_values.length == length);
        require(_fake.length == length);
        require(_secret.length == length);

        uint refund;
        for (uint i = 0; i < length; i++) {
            Bid storage bid = bids[msg.sender][i];
            (uint value, bool fake, bytes32 secret) =
                    (_values[i], _fake[i], _secret[i]);
            if (bid.blindedBid != keccak256(abi.encodePacked(value, fake, secret))) {
                // Bid was not actually revealed.
                // Do not refund deposit.
                continue;
            }
            refund += bid.deposit;
            if (!fake && bid.deposit >= value) {
                if (placeBid(msg.sender, value))
                    refund -= value;
            }
            // Make it impossible for the sender to re-claim
            // the same deposit.
            bid.blindedBid = bytes32(0);
        }
        msg.sender.transfer(refund);
    }

    // This is an "internal" function which means that it
    // can only be called from the contract itself (or from
    // derived contracts).
    function placeBid(address bidder, uint value) internal
            returns (bool success)
    {
        if (value <= highestBid) {
            return false;
        }
        if (highestBidder != address(0)) {
            // Refund the previously highest bidder.
            pendingReturns[highestBidder] += highestBid;
        }
```

(continues on next page)

(continued from previous page)

```
        highestBid = value;
        highestBidder = bidder;
        return true;
    }

    /// Withdraw a bid that was overbid.
    function withdraw() public {
        uint amount = pendingReturns[msg.sender];
        if (amount > 0) {
            // It is important to set this to zero because the recipient
            // can call this function again as part of the receiving call
            // before `transfer` returns (see the remark above about
            // conditions -> effects -> interaction).
            pendingReturns[msg.sender] = 0;

            msg.sender.transfer(amount);
        }
    }

    /// End the auction and send the highest bid
    /// to the beneficiary.
    function auctionEnd()
        public
        onlyAfter(revealEnd)
    {
        require(!ended);
        emit AuctionEnded(highestBidder, highestBid);
        ended = true;
        beneficiary.transfer(highestBid);
    }
}
```

7.3.3 Safe Remote Purchase

```
pragma solidity ^0.4.22;

contract Purchase {
    uint public value;
    address public seller;
    address public buyer;
    enum State { Created, Locked, Inactive }
    State public state;

    // Ensure that `msg.value` is an even number.
    // Division will truncate if it is an odd number.
    // Check via multiplication that it wasn't an odd number.
    constructor() public payable {
        seller = msg.sender;
        value = msg.value / 2;
        require((2 * value) == msg.value, "Value has to be even.");
    }

    modifier condition(bool _condition) {
        require(_condition);
        _;
```

(continues on next page)

(continued from previous page)

```
    }

    modifier onlyBuyer() {
        require(
            msg.sender == buyer,
            "Only buyer can call this."
        );
        _;
    }

    modifier onlySeller() {
        require(
            msg.sender == seller,
            "Only seller can call this."
        );
        _;
    }

    modifier inState(State _state) {
        require(
            state == _state,
            "Invalid state."
        );
        _;
    }

    event Aborted();
    event PurchaseConfirmed();
    event ItemReceived();

    /// Abort the purchase and reclaim the ether.
    /// Can only be called by the seller before
    /// the contract is locked.
    function abort()
        public
        onlySeller
        inState(State.Created)
    {
        emit Aborted();
        state = State.Inactive;
        seller.transfer(address(this).balance);
    }

    /// Confirm the purchase as buyer.
    /// Transaction has to include `2 * value` ether.
    /// The ether will be locked until confirmReceived
    /// is called.
    function confirmPurchase()
        public
        inState(State.Created)
        condition(msg.value == (2 * value))
        payable
    {
        emit PurchaseConfirmed();
        buyer = msg.sender;
        state = State.Locked;
    }
```

(continues on next page)

(continued from previous page)

```
    /// Confirm that you (the buyer) received the item.
    /// This will release the locked ether.
    function confirmReceived()
        public
        onlyBuyer
        inState(State.Locked)
    {
        emit ItemReceived();
        // It is important to change the state first because
        // otherwise, the contracts called using `send` below
        // can call in again here.
        state = State.Inactive;

        // NOTE: This actually allows both the buyer and the seller to
        // block the refund - the withdraw pattern should be used.

        buyer.transfer(value);
        seller.transfer(address(this).balance);
    }
}
```

7.3.4 Micropayment Channel

In this section we will learn how to build a simple implementation of a payment channel. It use cryptographics signatures to make repeated transfers of Ether between the same parties secure, instantaneous, and without transaction fees. To do it we need to understand how to sign and verify signatures, and setup the payment channel.

Creating and verifying signatures

Imagine Alice wants to send a quantity of Ether to Bob, i.e. Alice is the sender and the Bob is the recipient. Alice only needs to send cryptographically signed messages off-chain (e.g. via email) to Bob and it will be very similar to writing checks.

Signatures are used to authorize transactions, and they are a general tool that is available to smart contracts. Alice will build a simple smart contract that lets her transmit Ether, but in a unusual way, instead of calling a function herself to initiate a payment, she will let Bob do that, and therefore pay the transaction fee. The contract will work as follows:

1. Alice deploys the `ReceiverPays` contract, attaching enough Ether to cover the payments that will be made.
2. Alice authorizes a payment by signing a message with their private key.
3. Alice sends the cryptographically signed message to Bob. The message does not need to be kept secret (you will understand it later), and the mechanism for sending it does not matter.
4. Bob claims their payment by presenting the signed message to the smart contract, it verifies the authenticity of the message and then releases the funds.

Creating the signature

Alice does not need to interact with Ethereum network to sign the transaction, the process is completely offline. In this tutorial, we will sign messages in the browser using `web3.js` and `MetaMask`. In particular, we will use the standard way described in EIP-762, as it provides a number of other security benefits.

```
/// Hashing first makes a few things easier
var hash = web3.sha3("message to sign");
web3.personal.sign(hash, web3.eth.defaultAccount, function () {...});
```

Note that the `web3.personal.sign` prepends the length of the message to the signed data. Since we hash first, the message will always be exactly 32 bytes long, and thus this length prefix is always the same, making everything easier.

What to Sign

For a contract that fulfills payments, the signed message must include:

1. The recipient's address
2. The amount to be transferred
3. Protection against replay attacks

A replay attack is when a signed message is reused to claim authorization for a second action. To avoid replay attacks we will use the same as in Ethereum transactions themselves, a so-called nonce, which is the number of transactions sent by an account. The smart contract will check if a nonce is used multiple times.

There is another type of replay attacks, it occurs when the owner deploys a `ReceiverPays` smart contract, performs some payments, and then destroy the contract. Later, she decides to deploy the `RecipientPays` smart contract again, but the new contract does not know the nonces used in the previous deployment, so the attacker can use the old messages again.

Alice can protect against it including the contract's address in the message, and only messages containing contract's address itself will be accepted. This functionality can be found in the first two lines of the `claimPayment()` function in the full contract at the end of this chapter.

Packing arguments

Now that we have identified what information to include in the signed message, we are ready to put the message together, hash it, and sign it. For simplicity, we just concatenate the data. The ethereumjs-abi library provides a function called `soliditySHA3` that mimics the behavior of Solidity's `keccak256` function applied to arguments encoded using `abi.encodePacked`. Putting it all together, here is a JavaScript function that creates the proper signature for the `ReceiverPays` example:

```
// recipient is the address that should be paid.
// amount, in wei, specifies how much ether should be sent.
// nonce can be any unique number to prevent replay attacks
// contractAddress is used to prevent cross-contract replay attacks
function signPayment(recipient, amount, nonce, contractAddress, callback) {
    var hash = "0x" + ethereumjs.ABI.soliditySHA3(
        ["address", "uint256", "uint256", "address"],
        [recipient, amount, nonce, contractAddress]
    ).toString("hex");

    web3.personal.sign(hash, web3.eth.defaultAccount, callback);
}
```

Recovering the Message Signer in Solidity

In general, ECDSA signatures consist of two parameters, `r` and `s`. Signatures in Ethereum include a third parameter called `v`, that can be used to recover which account's private key was used to sign in the message, the transaction's sender. Solidity provides a built-in function ecrecover that accepts a message along with the `r`, `s` and `v` parameters and returns the address that was used to sign the message.

Extracting the Signature Parameters

Signatures produced by web3.js are the concatenation of `r`, `s` and `v`, so the first step is splitting those parameters back out. It can be done on the client, but doing it inside the smart contract means only one signature parameter needs to be sent rather than three. Splitting apart a byte array into component parts is a little messy. We will use inline assembly to do the job in the `splitSignature` function (the third function in the full contract at the end of this chapter).

Computing the Message Hash

The smart contract needs to know exactly what parameters were signed, and so it must recreate the message from the parameters and use that for signature verification. The functions `prefixed` and `recoverSigner` do this and their use can be found in the `claimPayment` function.

The full contract

```
pragma solidity ^0.4.24;

contract ReceiverPays {
    address owner = msg.sender;

    mapping(uint256 => bool) usedNonces;

    constructor() public payable {}

    function claimPayment(uint256 amount, uint256 nonce, bytes signature) public {
        require(!usedNonces[nonce]);
        usedNonces[nonce] = true;

        // this recreates the message that was signed on the client
        bytes32 message = prefixed(keccak256(abi.encodePacked(msg.sender, amount,␣
↪nonce, this)));

        require(recoverSigner(message, signature) == owner);

        msg.sender.transfer(amount);
    }

    /// destroy the contract and reclaim the leftover funds.
    function kill() public {
        require(msg.sender == owner);
        selfdestruct(msg.sender);
    }

    /// signature methods.
    function splitSignature(bytes sig)
```

(continues on next page)

(continued from previous page)

```
        internal
        pure
        returns (uint8 v, bytes32 r, bytes32 s)
    {
        require(sig.length == 65);

        assembly {
            // first 32 bytes, after the length prefix.
            r := mload(add(sig, 32))
            // second 32 bytes.
            s := mload(add(sig, 64))
            // final byte (first byte of the next 32 bytes).
            v := byte(0, mload(add(sig, 96)))
        }

        return (v, r, s);
    }

    function recoverSigner(bytes32 message, bytes sig)
        internal
        pure
        returns (address)
    {
        (uint8 v, bytes32 r, bytes32 s) = splitSignature(sig);

        return ecrecover(message, v, r, s);
    }

    /// builds a prefixed hash to mimic the behavior of eth_sign.
    function prefixed(bytes32 hash) internal pure returns (bytes32) {
        return keccak256(abi.encodePacked("\x19Ethereum Signed Message:\n32", hash));
    }
}
```

Writing a Simple Payment Channel

Alice will now build a simple but complete implementation of a payment channel. Payment channels use cryptographic signatures to make repeated transfers of Ether securely, instantaneously, and without transaction fees.

What is a Payment Channel?

Payment channels allow participants to make repeated transfers of Ether without using transactions. This means that the delays and fees associated with transactions can be avoided. We are going to explore a simple unidirectional payment channel between two parties (Alice and Bob). Using it involves three steps:

1. Alice funds a smart contract with Ether. This "opens" the payment channel.
2. Alice signs messages that specify how much of that Ether is owed to the recipient. This step is repeated for each payment.
3. Bob "closes" the payment channel, withdrawing their portion of the Ether and sending the remainder back to the sender.

Not ethat only steps 1 and 3 require Ethereum transactions, step 2 means that the sender transmits a cryptographically signed message to the recipient via off chain ways (e.g. email). This means only two transactions are required to

support any number of transfers.

Bob is guaranteed to receive their funds because the smart contract escrows the Ether and honors a valid signed message. The smart contract also enforces a timeout, so Alice is guaranteed to eventually recover their funds even if the recipient refuses to close the channel. It is up to the participants in a payment channel to decide how long to keep it open. For a short-lived transaction, such as paying an internet cafe for each minute of network access, or for a longer relationship, such as paying an employee an hourly wage, a payment could last for months or years.

Opening the Payment Channel

To open the payment channel, Alice deploys the smart contract, attaching the Ether to be escrowed and specifying the intendend recipient and a maximum duration for the channel to exist. It is the function `SimplePaymentChannel` in the contract, that is at the end of this chapter.

Making Payments

Alice makes payments by sending signed messages to Bob. This step is performed entirely outside of the Ethereum network. Messages are cryptographically signed by the sender and then transmitted directly to the recipient.

Each message includes the following information:

- The smart contract's address, used to prevent cross-contract replay attacks.
- The total amount of Ether that is owed the recipient so far.

A payment channel is closed just once, at the of a series of transfers. Because of this, only one of the messages sent will be redeemed. This is why each message specifies a cumulative total amount of Ether owed, rather than the amount of the individual micropayment. The recipient will naturally choose to redeem the most recent message because that is the one with the highest total. The nonce per-message is not needed anymore, because the smart contract will only honor a single message. The address of the smart contract is still used to prevent a message intended for one payment channel from being used for a different channel.

Here is the modified javascript code to cryptographically sign a message from the previous chapter:

```
function constructPaymentMessage(contractAddress, amount) {
    return ethereumjs.ABI.soliditySHA3(
        ["address", "uint256"],
        [contractAddress, amount]
    );
}

function signMessage(message, callback) {
    web3.personal.sign(
        "0x" + message.toString("hex"),
        web3.eth.defaultAccount,
        callback
    );
}

// contractAddress is used to prevent cross-contract replay attacks.
// amount, in wei, specifies how much Ether should be sent.

function signPayment(contractAddress, amount, callback) {
    var message = constructPaymentMessage(contractAddress, amount);
    signMessage(message, callback);
}
```

Closing the Payment Channel

When Bob is ready to receive their funds, it is time to close the payment channel by calling a `close` function on the smart contract. Closing the channel pays the recipient the Ether they are owed and destroys the contract, sending any remaining Ether back to Alice. To close the channel, Bob needs to provide a message signed by Alice.

The smart contract must verify that the message contains a valid signature from the sender. The process for doing this verification is the same as the process the recipient uses. The Solidity functions `isValidSignature` and `recoverSigner` work just like their JavaScript counterparts in the previous section. The latter is borrowed from the `ReceiverPays` contract in the previous chapter.

The `close` function can only be called by the payment channel recipient, who will naturally pass the most recent payment message because that message carries the highest total owed. If the sender were allowed to call this function, they could provide a message with a lower amount and cheat the recipient out of what they are owed.

The function verifies the signed message matches the given parameters. If everything checks out, the recipient is sent their portion of the Ether, and the sender is sent the rest via a `selfdestruct`. You can see the `close` function in the full contract.

Channel Expiration

Bob can close the payment channel at any time, but if they fail to do so, Alice needs a way to recover their escrowed funds. An *expiration* time was set at the time of contract deployment. Once that time is reached, Alice can call `claimTimeout` to recover their funds. You can see the `claimTimeout` function in the full contract.

After this function is called, Bob can no longer receive any Ether, so it is important that Bob closes the channel before the expiration is reached.

The full contract

```
pragma solidity ^0.4.24;

contract SimplePaymentChannel {
    address public sender;      // The account sending payments.
    address public recipient;   // The account receiving the payments.
    uint256 public expiration;  // Timeout in case the recipient never closes.

    constructor (address _recipient, uint256 duration)
        public
        payable
    {
        sender = msg.sender;
        recipient = _recipient;
        expiration = now + duration;
    }

    function isValidSignature(uint256 amount, bytes signature)
        internal
        view
        returns (bool)
    {
        bytes32 message = prefixed(keccak256(abi.encodePacked(this, amount)));

        // check that the signature is from the payment sender
        return recoverSigner(message, signature) == sender;
```

(continues on next page)

(continued from previous page)

```
    }

    /// the recipient can close the channel at any time by presenting a
    /// signed amount from the sender. the recipient will be sent that amount,
    /// and the remainder will go back to the sender
    function close(uint256 amount, bytes signature) public {
        require(msg.sender == recipient);
        require(isValidSignature(amount, signature));

        recipient.transfer(amount);
        selfdestruct(sender);
    }

    /// the sender can extend the expiration at any time
    function extend(uint256 newExpiration) public {
        require(msg.sender == sender);
        require(newExpiration > expiration);

        expiration = newExpiration;
    }

    /// if the timeout is reached without the recipient closing the channel,
    /// then the Ether is released back to the sender.
    function claimTimeout() public {
        require(now >= expiration);
        selfdestruct(sender);
    }

    /// All functions below this are just taken from the chapter
    /// 'creating and verifying signatures' chapter.

    function splitSignature(bytes sig)
        internal
        pure
        returns (uint8 v, bytes32 r, bytes32 s)
    {
        require(sig.length == 65);

        assembly {
            // first 32 bytes, after the length prefix
            r := mload(add(sig, 32))
            // second 32 bytes
            s := mload(add(sig, 64))
            // final byte (first byte of the next 32 bytes)
            v := byte(0, mload(add(sig, 96)))
        }

        return (v, r, s);
    }

    function recoverSigner(bytes32 message, bytes sig)
        internal
        pure
        returns (address)
    {
        (uint8 v, bytes32 r, bytes32 s) = splitSignature(sig);
```

(continues on next page)

(continued from previous page)

```
        return ecrecover(message, v, r, s);
    }

    /// builds a prefixed hash to mimic the behavior of eth_sign.
    function prefixed(bytes32 hash) internal pure returns (bytes32) {
        return keccak256(abi.encodePacked("\x19Ethereum Signed Message:\n32", hash));
    }
}
```

Note: The function `splitSignature` is very simple and does not use all security checks. A real implementation should use a more rigorously tested library, such as openzepplin's version of this code.

Verifying Payments

Unlike in our previous chapter, messages in a payment channel aren't redeemed right away. The recipient keeps track of the latest message and redeems it when it's time to close the payment channel. This means it's critical that the recipient perform their own verification of each message. Otherwise there is no guarantee that the recipient will be able to get paid in the end.

The recipient should verify each message using the following process:

1. Verify that the contact address in the message matches the payment channel.
2. Verify that the new total is the expected amount.
3. Verify that the new total does not exceed the amount of Ether escrowed.
4. Verify that the signature is valid and comes from the payment channel sender.

We'll use the ethereumjs-util library to write this verifications. The final step can be done a number of ways, but if it's being done in **JavaScript**. The following code borrows the *constructMessage* function from the signing **JavaScript code** above:

```
// this mimics the prefixing behavior of the eth_sign JSON-RPC method.
function prefixed(hash) {
    return ethereumjs.ABI.soliditySHA3(
        ["string", "bytes32"],
        ["\x19Ethereum Signed Message:\n32", hash]
    );
}

function recoverSigner(message, signature) {
    var split = ethereumjs.Util.fromRpcSig(signature);
    var publicKey = ethereumjs.Util.ecrecover(message, split.v, split.r, split.s);
    var signer = ethereumjs.Util.pubToAddress(publicKey).toString("hex");
    return signer;
}

function isValidSignature(contractAddress, amount, signature, expectedSigner) {
    var message = prefixed(constructPaymentMessage(contractAddress, amount));
    var signer = recoverSigner(message, signature);
    return signer.toLowerCase() ==
        ethereumjs.Util.stripHexPrefix(expectedSigner).toLowerCase();
}
```

7.4 Solidity in Depth

This section should provide you with all you need to know about Solidity. If something is missing here, please contact us on Gitter or make a pull request on Github.

7.4.1 Layout of a Solidity Source File

Source files can contain an arbitrary number of contract definitions, include directives and pragma directives.

Version Pragma

Source files can (and should) be annotated with a so-called version pragma to reject being compiled with future compiler versions that might introduce incompatible changes. We try to keep such changes to an absolute minimum and especially introduce changes in a way that changes in semantics will also require changes in the syntax, but this is of course not always possible. Because of that, it is always a good idea to read through the changelog at least for releases that contain breaking changes, those releases will always have versions of the form `0.x.0` or `x.0.0`.

The version pragma is used as follows:

```
pragma solidity ^0.4.0;
```

Such a source file will not compile with a compiler earlier than version 0.4.0 and it will also not work on a compiler starting from version 0.5.0 (this second condition is added by using `^`). The idea behind this is that there will be no breaking changes until version `0.5.0`, so we can always be sure that our code will compile the way we intended it to. We do not fix the exact version of the compiler, so that bugfix releases are still possible.

It is possible to specify much more complex rules for the compiler version, the expression follows those used by npm.

Importing other Source Files

Syntax and Semantics

Solidity supports import statements that are very similar to those available in JavaScript (from ES6 on), although Solidity does not know the concept of a "default export".

At a global level, you can use import statements of the following form:

```
import "filename";
```

This statement imports all global symbols from "filename" (and symbols imported there) into the current global scope (different than in ES6 but backwards-compatible for Solidity).

```
import * as symbolName from "filename";
```

...creates a new global symbol `symbolName` whose members are all the global symbols from `"filename"`.

```
import {symbol1 as alias, symbol2} from "filename";
```

...creates new global symbols `alias` and `symbol2` which reference `symbol1` and `symbol2` from `"filename"`, respectively.

Another syntax is not part of ES6, but probably convenient:

```
import "filename" as symbolName;
```

which is equivalent to `import * as symbolName from "filename";`.

Paths

In the above, `filename` is always treated as a path with `/` as directory separator, `.` as the current and `..` as the parent directory. When `.` or `..` is followed by a character except `/`, it is not considered as the current or the parent directory. All path names are treated as absolute paths unless they start with the current `.` or the parent directory `..`.

To import a file `x` from the same directory as the current file, use `import "./x" as x;`. If you use `import "x" as x;` instead, a different file could be referenced (in a global "include directory").

It depends on the compiler (see below) how to actually resolve the paths. In general, the directory hierarchy does not need to strictly map onto your local filesystem, it can also map to resources discovered via e.g. ipfs, http or git.

Use in Actual Compilers

When the compiler is invoked, it is not only possible to specify how to discover the first element of a path, but it is possible to specify path prefix remappings so that e.g. `github.com/ethereum/dapp-bin/library` is remapped to `/usr/local/dapp-bin/library` and the compiler will read the files from there. If multiple remappings can be applied, the one with the longest key is tried first. This allows for a "fallback-remapping" with e.g. `""` maps to `"/usr/local/include/solidity"`. Furthermore, these remappings can depend on the context, which allows you to configure packages to import e.g. different versions of a library of the same name.

solc:

For solc (the commandline compiler), these remappings are provided as `context:prefix=target` arguments, where both the `context:` and the `=target` parts are optional (where target defaults to prefix in that case). All remapping values that are regular files are compiled (including their dependencies). This mechanism is completely backwards-compatible (as long as no filename contains = or :) and thus not a breaking change. All imports in files in or below the directory `context` that import a file that starts with `prefix` are redirected by replacing `prefix` by `target`.

So as an example, if you clone `github.com/ethereum/dapp-bin/` locally to `/usr/local/dapp-bin`, you can use the following in your source file:

```
import "github.com/ethereum/dapp-bin/library/iterable_mapping.sol" as it_mapping;
```

and then run the compiler as

```
solc github.com/ethereum/dapp-bin/=/usr/local/dapp-bin/ source.sol
```

As a more complex example, suppose you rely on some module that uses a very old version of dapp-bin. That old version of dapp-bin is checked out at `/usr/local/dapp-bin_old`, then you can use

```
solc module1:github.com/ethereum/dapp-bin/=/usr/local/dapp-bin/ \
     module2:github.com/ethereum/dapp-bin/=/usr/local/dapp-bin_old/ \
     source.sol
```

so that all imports in `module2` point to the old version but imports in `module1` get the new version.

Note that solc only allows you to include files from certain directories: They have to be in the directory (or subdirectory) of one of the explicitly specified source files or in the directory (or subdirectory) of a remapping target. If you want to allow direct absolute includes, just add the remapping `=/`.

If there are multiple remappings that lead to a valid file, the remapping with the longest common prefix is chosen.

Remix:

Remix provides an automatic remapping for github and will also automatically retrieve the file over the network: You can import the iterable mapping by e.g. `import "github.com/ethereum/dapp-bin/library/iterable_mapping.sol" as it_mapping;`.

Other source code providers may be added in the future.

Comments

Single-line comments (`//`) and multi-line comments (`/*...*/`) are possible.

```
// This is a single-line comment.

/*
This is a
multi-line comment.
*/
```

Additionally, there is another type of comment called a natspec comment, for which the documentation is not yet written. They are written with a triple slash (`///`) or a double asterisk block(`/** ... */`) and they should be used directly above function declarations or statements. You can use Doxygen-style tags inside these comments to document functions, annotate conditions for formal verification, and provide a **confirmation text** which is shown to users when they attempt to invoke a function.

In the following example we document the title of the contract, the explanation for the two input parameters and two returned values.

```
pragma solidity ^0.4.0;

/** @title Shape calculator. */
contract ShapeCalculator {
    /** @dev Calculates a rectangle's surface and perimeter.
      * @param w Width of the rectangle.
      * @param h Height of the rectangle.
      * @return s The calculated surface.
      * @return p The calculated perimeter.
      */
    function rectangle(uint w, uint h) returns (uint s, uint p) {
        s = w * h;
        p = 2 * (w + h);
    }
}
```

7.4.2 Structure of a Contract

Contracts in Solidity are similar to classes in object-oriented languages. Each contract can contain declarations of *State Variables*, *Functions*, *Function Modifiers*, *Events*, *Struct Types* and *Enum Types*. Furthermore, contracts can inherit from other contracts.

State Variables

State variables are values which are permanently stored in contract storage.

```
pragma solidity ^0.4.0;

contract SimpleStorage {
    uint storedData; // State variable
    // ...
}
```

See the *Types* section for valid state variable types and *Visibility and Getters* for possible choices for visibility.

Functions

Functions are the executable units of code within a contract.

```
pragma solidity ^0.4.0;

contract SimpleAuction {
    function bid() public payable { // Function
        // ...
    }
}
```

Function Calls can happen internally or externally and have different levels of visibility (*Visibility and Getters*) towards other contracts.

Function Modifiers

Function modifiers can be used to amend the semantics of functions in a declarative way (see *Function Modifiers* in contracts section).

```
pragma solidity ^0.4.22;

contract Purchase {
    address public seller;

    modifier onlySeller() { // Modifier
        require(
            msg.sender == seller,
            "Only seller can call this."
        );
        _;
    }

    function abort() public onlySeller { // Modifier usage
        // ...
    }
}
```

Events

Events are convenience interfaces with the EVM logging facilities.

```
pragma solidity ^0.4.21;

```

(continues on next page)

(continued from previous page)

```
contract SimpleAuction {
    event HighestBidIncreased(address bidder, uint amount); // Event

    function bid() public payable {
        // ...
        emit HighestBidIncreased(msg.sender, msg.value); // Triggering event
    }
}
```

See *Events* in contracts section for information on how events are declared and can be used from within a dapp.

Struct Types

Structs are custom defined types that can group several variables (see *Structs* in types section).

```
pragma solidity ^0.4.0;

contract Ballot {
    struct Voter { // Struct
        uint weight;
        bool voted;
        address delegate;
        uint vote;
    }
}
```

Enum Types

Enums can be used to create custom types with a finite set of 'constant values' (see *Enums* in types section).

```
pragma solidity ^0.4.0;

contract Purchase {
    enum State { Created, Locked, Inactive } // Enum
}
```

7.4.3 Types

Solidity is a statically typed language, which means that the type of each variable (state and local) needs to be specified. Solidity provides several elementary types which can be combined to form complex types.

In addition, types can interact with each other in expressions containing operators. For a quick reference of the various operators, see *Order of Precedence of Operators*.

Value Types

The following types are also called value types because variables of these types will always be passed by value, i.e. they are always copied when they are used as function arguments or in assignments.

Booleans

`bool`: The possible values are constants `true` and `false`.

Operators:

- `!` (logical negation)
- `&&` (logical conjunction, "and")
- `||` (logical disjunction, "or")
- == (equality)
- `!=` (inequality)

The operators `||` and `&&` apply the common short-circuiting rules. This means that in the expression `f(x) || g(y)`, if `f(x)` evaluates to `true`, `g(y)` will not be evaluated even if it may have side-effects.

Integers

`int` / `uint`: Signed and unsigned integers of various sizes. Keywords `uint8` to `uint256` in steps of `8` (unsigned of 8 up to 256 bits) and `int8` to `int256`. `uint` and `int` are aliases for `uint256` and `int256`, respectively.

Operators:

- Comparisons: <=, <, ==, `!=`, >=, > (evaluate to `bool`)
- Bit operators: `&`, `|`, `^` (bitwise exclusive or), ~ (bitwise negation)
- Arithmetic operators: +, −, unary −, unary +, `*`, `/`, `%` (remainder), `**` (exponentiation), << (left shift), >> (right shift)

Division always truncates (it is just compiled to the `DIV` opcode of the EVM), but it does not truncate if both operators are *literals* (or literal expressions).

Division by zero and modulus with zero throws a runtime exception.

The result of a shift operation is the type of the left operand. The expression `x << y` is equivalent to `x * 2**y`, and, for positive integers, `x >> y` is equivalent to `x / 2**y`. For negative `x`, `x >> y` is equivalent to dividing by a power of `2` while rounding down (towards negative infinity). Shifting by a negative amount throws a runtime exception.

> **Warning:** Before version `0.5.0` a right shift `x >> y` for negative `x` was equivalent to `x / 2**y`, i.e. right shifts used rounding towards zero instead of rounding towards negative infinity.

Fixed Point Numbers

> **Warning:** Fixed point numbers are not fully supported by Solidity yet. They can be declared, but cannot be assigned to or from.

`fixed` / `ufixed`: Signed and unsigned fixed point number of various sizes. Keywords `ufixedMxN` and `fixedMxN`, where `M` represents the number of bits taken by the type and `N` represents how many decimal points are available. `M` must be divisible by 8 and goes from 8 to 256 bits. `N` must be between 0 and 80, inclusive. `ufixed` and `fixed` are aliases for `ufixed128x18` and `fixed128x18`, respectively.

Operators:

- Comparisons: <=, <, ==, !=, >=, > (evaluate to `bool`)
- Arithmetic operators: +, -, unary -, unary +, *, /, % (remainder)

Note: The main difference between floating point (`float` and `double` in many languages, more precisely IEEE 754 numbers) and fixed point numbers is that the number of bits used for the integer and the fractional part (the part after the decimal dot) is flexible in the former, while it is strictly defined in the latter. Generally, in floating point almost the entire space is used to represent the number, while only a small number of bits define where the decimal point is.

Address

`address`: Holds a 20 byte value (size of an Ethereum address). Address types also have members and serve as a base for all contracts.

Operators:

- <=, <, ==, !=, >= and >

Note: Starting with version 0.5.0 contracts do not derive from the address type, but can still be explicitly converted to address.

Members of Addresses

- `balance` and `transfer`

For a quick reference, see *Address Related*.

It is possible to query the balance of an address using the property `balance` and to send Ether (in units of wei) to an address using the `transfer` function:

```
address x = 0x123;
address myAddress = this;
if (x.balance < 10 && myAddress.balance >= 10) x.transfer(10);
```

Note: If `x` is a contract address, its code (more specifically: its fallback function, if present) will be executed together with the `transfer` call (this is a feature of the EVM and cannot be prevented). If that execution runs out of gas or fails in any way, the Ether transfer will be reverted and the current contract will stop with an exception.

- `send`

Send is the low-level counterpart of `transfer`. If the execution fails, the current contract will not stop with an exception, but `send` will return `false`.

Warning: There are some dangers in using `send`: The transfer fails if the call stack depth is at 1024 (this can always be forced by the caller) and it also fails if the recipient runs out of gas. So in order to make safe Ether transfers, always check the return value of `send`, use `transfer` or even better: use a pattern where the recipient withdraws the money.

- `call`, `callcode` and `delegatecall`

Furthermore, to interface with contracts that do not adhere to the ABI, or to get more direct control over the encoding, the function `call` is provided which takes a single byte array as input. The functions `abi.encode`, `abi.encodePacked`, `abi.encodeWithSelector` and `abi.encodeWithSignature` can be used to encode structured data.

Warning: All these functions are low-level functions and should be used with care. Specifically, any unknown contract might be malicious and if you call it, you hand over control to that contract which could in turn call back into your contract, so be prepared for changes to your state variables when the call returns. The regular way to interact with other contracts is to call a function on a contract object (`x.f()`).

:: note:: Previous versions of Solidity allowed these functions to receive arbitrary arguments and would also handle a first argument of type `bytes4` differently. These edge cases were removed in version 0.5.0.

`call` returns a boolean indicating whether the invoked function terminated (`true`) or caused an EVM exception (`false`). It is not possible to access the actual data returned with plain Solidity. However, using inline assembly it is possible to make a raw `call` and access the actual data returned with the `returndatacopy` instruction.

It is possible to adjust the supplied gas with the `.gas()` modifier:

```
namReg.call.gas(1000000)(abi.encodeWithSignature("register(string)", "MyName"));
```

Similarly, the supplied Ether value can be controlled too:

```
nameReg.call.value(1 ether)(abi.encodeWithSignature("register(string)", "MyName"));
```

Lastly, these modifiers can be combined. Their order does not matter:

```
nameReg.call.gas(1000000).value(1 ether)(abi.encodeWithSignature("register(string)",
↪"MyName"));
```

Note: It is not yet possible to use the gas or value modifiers on overloaded functions.

A workaround is to introduce a special case for gas and value and just re-check whether they are present at the point of overload resolution.

In a similar way, the function `delegatecall` can be used: the difference is that only the code of the given address is used, all other aspects (storage, balance, ...) are taken from the current contract. The purpose of `delegatecall` is to use library code which is stored in another contract. The user has to ensure that the layout of storage in both contracts is suitable for delegatecall to be used. Prior to homestead, only a limited variant called `callcode` was available that did not provide access to the original `msg.sender` and `msg.value` values.

All three functions `call`, `delegatecall` and `callcode` are very low-level functions and should only be used as a *last resort* as they break the type-safety of Solidity.

The `.gas()` option is available on all three methods, while the `.value()` option is not supported for `delegatecall`.

Note: All contracts can be converted to `address` type, so it is possible to query the balance of the current contract using `address(this).balance`.

Note: The use of `callcode` is discouraged and will be removed in the future.

Fixed-size byte arrays

`bytes1`, `bytes2`, `bytes3`, ..., `bytes32`. `byte` is an alias for `bytes1`.

Operators:

- Comparisons: <=, <, ==, !=, >=, > (evaluate to `bool`)
- Bit operators: `&`, `|`, ^ (bitwise exclusive or), ~ (bitwise negation), << (left shift), >> (right shift)
- Index access: If `x` is of type `bytesI`, then `x[k]` for `0 <= k < I` returns the `k` th byte (read-only).

The shifting operator works with any integer type as right operand (but will return the type of the left operand), which denotes the number of bits to shift by. Shifting by a negative amount will cause a runtime exception.

Members:

- `.length` yields the fixed length of the byte array (read-only).

Note: It is possible to use an array of bytes as `byte[]`, but it is wasting a lot of space, 31 bytes every element, to be exact, when passing in calls. It is better to use `bytes`.

Dynamically-sized byte array

`bytes`: Dynamically-sized byte array, see *Arrays*. Not a value-type!

`string`: Dynamically-sized UTF-8-encoded string, see *Arrays*. Not a value-type!

Address Literals

Hexadecimal literals that pass the address checksum test, for example `0xdCad3a6d3569DF655070DEd06cb7A1b2Ccd1D3AF` are of `address` type. Hexadecimal literals that are between 39 and 41 digits long and do not pass the checksum test produce a warning and are treated as regular rational number literals.

Note: The mixed-case address checksum format is defined in EIP-55.

Rational and Integer Literals

Integer literals are formed from a sequence of numbers in the range 0-9. They are interpreted as decimals. For example, `69` means sixty nine. Octal literals do not exist in Solidity and leading zeros are invalid.

Decimal fraction literals are formed by a `.` with at least one number on one side. Examples include `1.`, `.1` and `1.3`.

Scientific notation is also supported, where the base can have fractions, while the exponent cannot. Examples include `2e10`, `-2e10`, `2e-10`, `2.5e1`.

Number literal expressions retain arbitrary precision until they are converted to a non-literal type (i.e. by using them together with a non-literal expression). This means that computations do not overflow and divisions do not truncate in number literal expressions.

For example, `(2**800 + 1) - 2**800` results in the constant `1` (of type `uint8`) although intermediate results would not even fit the machine word size. Furthermore, `.5 * 8` results in the integer `4` (although non-integers were used in between).

Any operator that can be applied to integers can also be applied to number literal expressions as long as the operands are integers. If any of the two is fractional, bit operations are disallowed and exponentiation is disallowed if the exponent is fractional (because that might result in a non-rational number).

Note: Solidity has a number literal type for each rational number. Integer literals and rational number literals belong to number literal types. Moreover, all number literal expressions (i.e. the expressions that contain only number literals and operators) belong to number literal types. So the number literal expressions `1 + 2` and `2 + 1` both belong to the same number literal type for the rational number three.

Warning: Division on integer literals used to truncate in earlier versions, but it will now convert into a rational number, i.e. `5 / 2` is not equal to `2`, but to `2.5`.

Note: Number literal expressions are converted into a non-literal type as soon as they are used with non-literal expressions. Even though we know that the value of the expression assigned to `b` in the following example evaluates to an integer, but the partial expression `2.5 + a` does not type check so the code does not compile

```
uint128 a = 1;
uint128 b = 2.5 + a + 0.5;
```

String Literals

String literals are written with either double or single-quotes (`"foo"` or `'bar'`). They do not imply trailing zeroes as in C; `"foo"` represents three bytes not four. As with integer literals, their type can vary, but they are implicitly convertible to `bytes1`, ..., `bytes32`, if they fit, to `bytes` and to `string`.

String literals support escape characters, such as `\n`, `\xNN` and `\uNNNN`. `\xNN` takes a hex value and inserts the appropriate byte, while `\uNNNN` takes a Unicode codepoint and inserts an UTF-8 sequence.

Hexadecimal Literals

Hexademical Literals are prefixed with the keyword `hex` and are enclosed in double or single-quotes (`hex"001122FF"`). Their content must be a hexadecimal string and their value will be the binary representation of those values.

Hexademical Literals behave like String Literals and have the same convertibility restrictions.

Enums

Enums are one way to create a user-defined type in Solidity. They are explicitly convertible to and from all integer types but implicit conversion is not allowed. The explicit conversions check the value ranges at runtime and a failure causes an exception. Enums needs at least one member.

```
pragma solidity ^0.4.16;

contract test {
    enum ActionChoices { GoLeft, GoRight, GoStraight, SitStill }
    ActionChoices choice;
```

(continues on next page)

(continued from previous page)

```
    ActionChoices constant defaultChoice = ActionChoices.GoStraight;

    function setGoStraight() public {
        choice = ActionChoices.GoStraight;
    }

    // Since enum types are not part of the ABI, the signature of "getChoice"
    // will automatically be changed to "getChoice() returns (uint8)"
    // for all matters external to Solidity. The integer type used is just
    // large enough to hold all enum values, i.e. if you have more values,
    // `uint16` will be used and so on.
    function getChoice() public view returns (ActionChoices) {
        return choice;
    }

    function getDefaultChoice() public pure returns (uint) {
        return uint(defaultChoice);
    }
}
```

Function Types

Function types are the types of functions. Variables of function type can be assigned from functions and function parameters of function type can be used to pass functions to and return functions from function calls. Function types come in two flavours - *internal* and *external* functions:

Internal functions can only be called inside the current contract (more specifically, inside the current code unit, which also includes internal library functions and inherited functions) because they cannot be executed outside of the context of the current contract. Calling an internal function is realized by jumping to its entry label, just like when calling a function of the current contract internally.

External functions consist of an address and a function signature and they can be passed via and returned from external function calls.

Function types are notated as follows:

```
function (<parameter types>) {internal|external} [pure|view|payable] [returns (
↪<return types>)]
```

In contrast to the parameter types, the return types cannot be empty - if the function type should not return anything, the whole `returns (<return types>)` part has to be omitted.

By default, function types are internal, so the `internal` keyword can be omitted. In contrast, contract functions themselves are public by default, only when used as the name of a type, the default is internal.

If a function type variable is not initialized, calling it will result in an exception. The same happens if you call a function after using `delete` on it.

If external function types are used outside of the context of Solidity, they are treated as the `function` type, which encodes the address followed by the function identifier together in a single `bytes24` type.

Note that public functions of the current contract can be used both as an internal and as an external function. To use `f` as an internal function, just use `f`, if you want to use its external form, use `this.f`.

Additionally, public (or external) functions also have a special member called `selector`, which returns the *ABI function selector*:

```
pragma solidity ^0.4.16;

contract Selector {
  function f() public view returns (bytes4) {
    return this.f.selector;
  }
}
```

Example that shows how to use internal function types:

```
pragma solidity ^0.4.16;

library ArrayUtils {
  // internal functions can be used in internal library functions because
  // they will be part of the same code context
  function map(uint[] memory self, function (uint) pure returns (uint) f)
    internal
    pure
    returns (uint[] memory r)
  {
    r = new uint[](self.length);
    for (uint i = 0; i < self.length; i++) {
      r[i] = f(self[i]);
    }
  }
  function reduce(
    uint[] memory self,
    function (uint, uint) pure returns (uint) f
  )
    internal
    pure
    returns (uint r)
  {
    r = self[0];
    for (uint i = 1; i < self.length; i++) {
      r = f(r, self[i]);
    }
  }
  function range(uint length) internal pure returns (uint[] memory r) {
    r = new uint[](length);
    for (uint i = 0; i < r.length; i++) {
      r[i] = i;
    }
  }
}

contract Pyramid {
  using ArrayUtils for *;
  function pyramid(uint l) public pure returns (uint) {
    return ArrayUtils.range(l).map(square).reduce(sum);
  }
  function square(uint x) internal pure returns (uint) {
    return x * x;
  }
  function sum(uint x, uint y) internal pure returns (uint) {
    return x + y;
  }
}
```

Another example that uses external function types:

```
pragma solidity ^0.4.22;

contract Oracle {
  struct Request {
    bytes data;
    function(bytes memory) external callback;
  }
  Request[] requests;
  event NewRequest(uint);
  function query(bytes memory data, function(bytes memory) external callback) public {
    requests.push(Request(data, callback));
    emit NewRequest(requests.length - 1);
  }
  function reply(uint requestID, bytes memory response) public {
    // Here goes the check that the reply comes from a trusted source
    requests[requestID].callback(response);
  }
}

contract OracleUser {
  Oracle constant oracle = Oracle(0x1234567); // known contract
  function buySomething() {
    oracle.query("USD", this.oracleResponse);
  }
  function oracleResponse(bytes memory response) public {
    require(
        msg.sender == address(oracle),
        "Only oracle can call this."
    );
    // Use the data
  }
}
```

Note: Lambda or inline functions are planned but not yet supported.

Reference Types

Complex types, i.e. types which do not always fit into 256 bits have to be handled more carefully than the value-types we have already seen. Since copying them can be quite expensive, we have to think about whether we want them to be stored in **memory** (which is not persisting) or **storage** (where the state variables are held).

Data location

Every complex type, i.e. *arrays* and *structs*, has an additional annotation, the "data location", about whether it is stored in memory or in storage. Depending on the context, there is always a default, but it can be overridden by appending either `storage` or `memory` to the type. The default for function parameters (including return parameters) is `memory`, the default for local variables is `storage` and the location is forced to `storage` for state variables (obviously).

There is also a third data location, `calldata`, which is a non-modifiable, non-persistent area where function arguments are stored. Function parameters (not return parameters) of external functions are forced to `calldata` and behave mostly like `memory`.

Data locations are important because they change how assignments behave: assignments between storage and memory and also to a state variable (even from other state variables) always create an independent copy. Assignments to local storage variables only assign a reference though, and this reference always points to the state variable even if the latter is changed in the meantime. On the other hand, assignments from a memory stored reference type to another memory-stored reference type do not create a copy.

```
pragma solidity ^0.4.0;

contract C {
    uint[] x; // the data location of x is storage

    // the data location of memoryArray is memory
    function f(uint[] memory memoryArray) public {
        x = memoryArray; // works, copies the whole array to storage
        uint[] storage y = x; // works, assigns a pointer, data location of y is␣
↪storage
        y[7]; // fine, returns the 8th element
        y.length = 2; // fine, modifies x through y
        delete x; // fine, clears the array, also modifies y
        // The following does not work; it would need to create a new temporary /
        // unnamed array in storage, but storage is "statically" allocated:
        // y = memoryArray;
        // This does not work either, since it would "reset" the pointer, but there
        // is no sensible location it could point to.
        // delete y;
        g(x); // calls g, handing over a reference to x
        h(x); // calls h and creates an independent, temporary copy in memory
    }

    function g(uint[] storage storageArray) internal {}
    function h(uint[] memory memoryArray) public {}
}
```

Summary

Forced data location:

- parameters (not return) of external functions: calldata
- state variables: storage

Default data location:

- parameters (also return) of functions: memory
- all other local variables: storage

Arrays

Arrays can have a compile-time fixed size or they can be dynamic. For storage arrays, the element type can be arbitrary (i.e. also other arrays, mappings or structs). For memory arrays, it cannot be a mapping and has to be an ABI type if it is an argument of a publicly-visible function.

An array of fixed size `k` and element type `T` is written as `T[k]`, an array of dynamic size as `T[]`. As an example, an array of 5 dynamic arrays of `uint` is `uint[][5]` (note that the notation is reversed when compared to some other languages). To access the second uint in the third dynamic array, you use `x[2][1]` (indices are zero-based and access works in the opposite way of the declaration, i.e. `x[2]` shaves off one level in the type from the right).

Variables of type `bytes` and `string` are special arrays. A `bytes` is similar to `byte[]`, but it is packed tightly in calldata. `string` is equal to `bytes` but does not allow length or index access (for now). So `bytes` should always be preferred over `byte[]` because it is cheaper. As a rule of thumb, use `bytes` for arbitrary-length raw byte data and `string` for arbitrary-length string (UTF-8) data. If you can limit the length to a certain number of bytes, always use one of `bytes1` to `bytes32` because they are much cheaper.

Note: If you want to access the byte-representation of a string `s`, use `bytes(s).length` / `bytes(s)[7] = 'x';`. Keep in mind that you are accessing the low-level bytes of the UTF-8 representation, and not the individual characters!

It is possible to mark arrays `public` and have Solidity create a *getter*. The numeric index will become a required parameter for the getter.

Allocating Memory Arrays

Creating arrays with variable length in memory can be done using the `new` keyword. As opposed to storage arrays, it is **not** possible to resize memory arrays by assigning to the `.length` member.

```
pragma solidity ^0.4.16;

contract C {
    function f(uint len) public pure {
        uint[] memory a = new uint[](7);
        bytes memory b = new bytes(len);
        // Here we have a.length == 7 and b.length == len
        a[6] = 8;
    }
}
```

Array Literals / Inline Arrays

Array literals are arrays that are written as an expression and are not assigned to a variable right away.

```
pragma solidity ^0.4.16;

contract C {
    function f() public pure {
        g([uint(1), 2, 3]);
    }
    function g(uint[3] memory _data) public pure {
        // ...
    }
}
```

The type of an array literal is a memory array of fixed size whose base type is the common type of the given elements. The type of `[1, 2, 3]` is `uint8[3] memory`, because the type of each of these constants is `uint8`. Because of that, it was necessary to convert the first element in the example above to `uint`. Note that currently, fixed size memory arrays cannot be assigned to dynamically-sized memory arrays, i.e. the following is not possible:

```
// This will not compile.

pragma solidity ^0.4.0;
```

(continues on next page)

(continued from previous page)

```
contract C {
    function f() public {
        // The next line creates a type error because uint[3] memory
        // cannot be converted to uint[] memory.
        uint[] memory x = [uint(1), 3, 4];
    }
}
```

It is planned to remove this restriction in the future but currently creates some complications because of how arrays are passed in the ABI.

Members

length: Arrays have a `length` member to hold their number of elements. Dynamic arrays can be resized in storage (not in memory) by changing the `.length` member. This does not happen automatically when attempting to access elements outside the current length. The size of memory arrays is fixed (but dynamic, i.e. it can depend on runtime parameters) once they are created.

push: Dynamic storage arrays and `bytes` (not `string`) have a member function called `push` that can be used to append an element at the end of the array. The function returns the new length.

pop: Dynamic storage arrays and `bytes` (not `string`) have a member function called `pop` that can be used to remove an element from the end of the array.

Warning: It is not yet possible to use arrays of arrays in external functions.

Warning: Due to limitations of the EVM, it is not possible to return dynamic content from external function calls. The function `f` in `contract C { function f() returns (uint[]) { ... } }` will return something if called from web3.js, but not if called from Solidity.

The only workaround for now is to use large statically-sized arrays.

```
pragma solidity ^0.4.16;

contract ArrayContract {
    uint[2**20] m_aLotOfIntegers;
    // Note that the following is not a pair of dynamic arrays but a
    // dynamic array of pairs (i.e. of fixed size arrays of length two).
    bool[2][] m_pairsOfFlags;
    // newPairs is stored in memory - the default for function arguments

    function setAllFlagPairs(bool[2][] memory newPairs) public {
        // assignment to a storage array replaces the complete array
        m_pairsOfFlags = newPairs;
    }

    function setFlagPair(uint index, bool flagA, bool flagB) public {
        // access to a non-existing index will throw an exception
        m_pairsOfFlags[index][0] = flagA;
        m_pairsOfFlags[index][1] = flagB;
```

(continues on next page)

(continued from previous page)

```
    }

    function changeFlagArraySize(uint newSize) public {
        // if the new size is smaller, removed array elements will be cleared
        m_pairsOfFlags.length = newSize;
    }

    function clear() public {
        // these clear the arrays completely
        delete m_pairsOfFlags;
        delete m_aLotOfIntegers;
        // identical effect here
        m_pairsOfFlags.length = 0;
    }

    bytes m_byteData;

    function byteArrays(bytes memory data) public {
        // byte arrays ("bytes") are different as they are stored without padding,
        // but can be treated identical to "uint8[]"
        m_byteData = data;
        m_byteData.length += 7;
        m_byteData[3] = byte(8);
        delete m_byteData[2];
    }

    function addFlag(bool[2] memory flag) public returns (uint) {
        return m_pairsOfFlags.push(flag);
    }

    function createMemoryArray(uint size) public pure returns (bytes memory) {
        // Dynamic memory arrays are created using `new`:
        uint[2][] memory arrayOfPairs = new uint[2][](size);
        // Create a dynamic byte array:
        bytes memory b = new bytes(200);
        for (uint i = 0; i < b.length; i++)
            b[i] = byte(uint8(i));
        return b;
    }
}
```

Structs

Solidity provides a way to define new types in the form of structs, which is shown in the following example:

```
pragma solidity ^0.4.11;

contract CrowdFunding {
    // Defines a new type with two fields.
    struct Funder {
        address addr;
        uint amount;
    }

    struct Campaign {
```

(continues on next page)

(continued from previous page)

```
        address beneficiary;
        uint fundingGoal;
        uint numFunders;
        uint amount;
        mapping (uint => Funder) funders;
    }

    uint numCampaigns;
    mapping (uint => Campaign) campaigns;

    function newCampaign(address beneficiary, uint goal) public returns (uint␣
↪campaignID) {
        campaignID = numCampaigns++; // campaignID is return variable
        // Creates new struct and saves in storage. We leave out the mapping type.
        campaigns[campaignID] = Campaign(beneficiary, goal, 0, 0);
    }

    function contribute(uint campaignID) public payable {
        Campaign storage c = campaigns[campaignID];
        // Creates a new temporary memory struct, initialised with the given values
        // and copies it over to storage.
        // Note that you can also use Funder(msg.sender, msg.value) to initialise.
        c.funders[c.numFunders++] = Funder({addr: msg.sender, amount: msg.value});
        c.amount += msg.value;
    }

    function checkGoalReached(uint campaignID) public returns (bool reached) {
        Campaign storage c = campaigns[campaignID];
        if (c.amount < c.fundingGoal)
            return false;
        uint amount = c.amount;
        c.amount = 0;
        c.beneficiary.transfer(amount);
        return true;
    }
}
```

The contract does not provide the full functionality of a crowdfunding contract, but it contains the basic concepts necessary to understand structs. Struct types can be used inside mappings and arrays and they can itself contain mappings and arrays.

It is not possible for a struct to contain a member of its own type, although the struct itself can be the value type of a mapping member. This restriction is necessary, as the size of the struct has to be finite.

Note how in all the functions, a struct type is assigned to a local variable (of the default storage data location). This does not copy the struct but only stores a reference so that assignments to members of the local variable actually write to the state.

Of course, you can also directly access the members of the struct without assigning it to a local variable, as in `campaigns[campaignID].amount = 0`.

Mappings

Mapping types are declared as `mapping(_KeyType => _ValueType)`. Here `_KeyType` can be almost any type except for a mapping, a dynamically sized array, a contract, a function, an enum and a struct. `_ValueType` can actually be any type, including mappings.

Mappings can be seen as hash tables which are virtually initialized such that every possible key exists and is mapped to a value whose byte-representation is all zeros: a type's *default value*. The similarity ends here, though: The key data is not actually stored in a mapping, only its `keccak256` hash used to look up the value.

Because of this, mappings do not have a length or a concept of a key or value being "set".

Mappings are only allowed for state variables (or as storage reference types in internal functions).

It is possible to mark mappings `public` and have Solidity create a *getter*. The `_KeyType` will become a required parameter for the getter and it will return `_ValueType`.

The `_ValueType` can be a mapping too. The getter will have one parameter for each `_KeyType`, recursively.

```
pragma solidity ^0.4.0;

contract MappingExample {
    mapping(address => uint) public balances;

    function update(uint newBalance) public {
        balances[msg.sender] = newBalance;
    }
}

contract MappingUser {
    function f() public returns (uint) {
        MappingExample m = new MappingExample();
        m.update(100);
        return m.balances(this);
    }
}
```

Note: Mappings are not iterable, but it is possible to implement a data structure on top of them. For an example, see iterable mapping.

Operators Involving LValues

If `a` is an LValue (i.e. a variable or something that can be assigned to), the following operators are available as shorthands:

`a += e` is equivalent to `a = a + e`. The operators `-=`, `*=`, `/=`, `%=`, `|=`, `&=` and `^=` are defined accordingly. `a++` and `a--` are equivalent to `a += 1` / `a -= 1` but the expression itself still has the previous value of `a`. In contrast, `--a` and `++a` have the same effect on `a` but return the value after the change.

delete

`delete a` assigns the initial value for the type to `a`. I.e. for integers it is equivalent to `a = 0`, but it can also be used on arrays, where it assigns a dynamic array of length zero or a static array of the same length with all elements reset. For structs, it assigns a struct with all members reset.

`delete` has no effect on whole mappings (as the keys of mappings may be arbitrary and are generally unknown). So if you delete a struct, it will reset all members that are not mappings and also recurse into the members unless they are mappings. However, individual keys and what they map to can be deleted.

It is important to note that `delete a` really behaves like an assignment to `a`, i.e. it stores a new object in `a`.

```
pragma solidity ^0.4.0;

contract DeleteExample {
    uint data;
    uint[] dataArray;

    function f() public {
        uint x = data;
        delete x; // sets x to 0, does not affect data
        delete data; // sets data to 0, does not affect x which still holds a copy
        uint[] storage y = dataArray;
        delete dataArray; // this sets dataArray.length to zero, but as uint[] is a␣
↪complex object, also
        // y is affected which is an alias to the storage object
        // On the other hand: "delete y" is not valid, as assignments to local␣
↪variables
        // referencing storage objects can only be made from existing storage␣
↪objects.
    }
}
```

Conversions between Elementary Types

Implicit Conversions

If an operator is applied to different types, the compiler tries to implicitly convert one of the operands to the type of the other (the same is true for assignments). In general, an implicit conversion between value-types is possible if it makes sense semantically and no information is lost: `uint8` is convertible to `uint16` and `int128` to `int256`, but `int8` is not convertible to `uint256` (because `uint256` cannot hold e.g. `-1`). Furthermore, unsigned integers can be converted to bytes of the same or larger size, but not vice-versa. Any type that can be converted to `uint160` can also be converted to `address`.

Explicit Conversions

If the compiler does not allow implicit conversion but you know what you are doing, an explicit type conversion is sometimes possible. Note that this may give you some unexpected behaviour so be sure to test to ensure that the result is what you want! Take the following example where you are converting a negative `int8` to a `uint`:

```
int8 y = -3;
uint x = uint(y);
```

At the end of this code snippet, `x` will have the value `0xfffff..fd` (64 hex characters), which is -3 in the two's complement representation of 256 bits.

If a type is explicitly converted to a smaller type, higher-order bits are cut off:

```
uint32 a = 0x12345678;
uint16 b = uint16(a); // b will be 0x5678 now
```

Since 0.5.0 explicit conversions between integers and fixed-size byte arrays are only allowed, if both have the same size. To convert between integers and fixed-size byte arrays of different size, they first have to be explicitly converted to a matching size. This makes alignment and padding explicit:

```
uint16 x = 0xffff;
bytes32(uint256(x)); // pad on the left
bytes32(bytes2(x)); // pad on the right
```

7.4.4 Units and Globally Available Variables

Ether Units

A literal number can take a suffix of `wei`, `finney`, `szabo` or `ether` to convert between the subdenominations of Ether, where Ether currency numbers without a postfix are assumed to be Wei, e.g. `2 ether == 2000 finney` evaluates to `true`.

Time Units

Suffixes like `seconds`, `minutes`, `hours`, `days`, `weeks` and `years` after literal numbers can be used to convert between units of time where seconds are the base unit and units are considered naively in the following way:

- `1 == 1 seconds`
- `1 minutes == 60 seconds`
- `1 hours == 60 minutes`
- `1 days == 24 hours`
- `1 weeks == 7 days`
- `1 years == 365 days`

Take care if you perform calendar calculations using these units, because not every year equals 365 days and not even every day has 24 hours because of leap seconds. Due to the fact that leap seconds cannot be predicted, an exact calendar library has to be updated by an external oracle.

Note: The suffix `years` has been deprecated due to the reasons above and cannot be used starting version 0.5.0.

These suffixes cannot be applied to variables. If you want to interpret some input variable in e.g. days, you can do it in the following way:

```
function f(uint start, uint daysAfter) public {
    if (now >= start + daysAfter * 1 days) {
      // ...
    }
}
```

Special Variables and Functions

There are special variables and functions which always exist in the global namespace and are mainly used to provide information about the blockchain or are general-use utility functions.

Block and Transaction Properties

- `block.blockhash(uint blockNumber) returns (bytes32)`: hash of the given block - only works for 256 most recent, excluding current, blocks - deprecated in version 0.4.22 and replaced by `blockhash(uint blockNumber)`.
- `block.coinbase` (`address`): current block miner's address
- `block.difficulty` (`uint`): current block difficulty
- `block.gaslimit` (`uint`): current block gaslimit
- `block.number` (`uint`): current block number
- `block.timestamp` (`uint`): current block timestamp as seconds since unix epoch
- `gasleft() returns (uint256)`: remaining gas
- `msg.data` (`bytes`): complete calldata
- `msg.gas` (`uint`): remaining gas - deprecated in version 0.4.21 and to be replaced by `gasleft()`
- `msg.sender` (`address`): sender of the message (current call)
- `msg.sig` (`bytes4`): first four bytes of the calldata (i.e. function identifier)
- `msg.value` (`uint`): number of wei sent with the message
- `now` (`uint`): current block timestamp (alias for `block.timestamp`)
- `tx.gasprice` (`uint`): gas price of the transaction
- `tx.origin` (`address`): sender of the transaction (full call chain)

Note: The values of all members of `msg`, including `msg.sender` and `msg.value` can change for every **external** function call. This includes calls to library functions.

Note: Do not rely on `block.timestamp`, `now` and `blockhash` as a source of randomness, unless you know what you are doing.

Both the timestamp and the block hash can be influenced by miners to some degree. Bad actors in the mining community can for example run a casino payout function on a chosen hash and just retry a different hash if they did not receive any money.

The current block timestamp must be strictly larger than the timestamp of the last block, but the only guarantee is that it will be somewhere between the timestamps of two consecutive blocks in the canonical chain.

Note: The block hashes are not available for all blocks for scalability reasons. You can only access the hashes of the most recent 256 blocks, all other values will be zero.

ABI Encoding Functions

- `abi.encode(...) returns (bytes)`: ABI-encodes the given arguments
- `abi.encodePacked(...) returns (bytes)`: Performs *packed encoding* of the given arguments

- `abi.encodeWithSelector(bytes4 selector, ...) returns (bytes)`: ABI-encodes the given arguments starting from the second and prepends the given four-byte selector
- `abi.encodeWithSignature(string signature, ...) returns (bytes)`: Equivalent to `abi.encodeWithSelector(bytes4(keccak256(bytes(signature)), ...)`

Note: These encoding functions can be used to craft data for function calls without actually calling a function. Furthermore, `keccak256(abi.encodePacked(a, b))` is a way to compute the hash of structured data (although be aware that it is possible to craft a "hash collision" using different inputs types).

See the documentation about the *ABI* and the *tightly packed encoding* for details about the encoding.

Error Handling

`assert(bool condition)`: invalidates the transaction if the condition is not met - to be used for internal errors.

`require(bool condition)`: reverts if the condition is not met - to be used for errors in inputs or external components.

`require(bool condition, string message)`: reverts if the condition is not met - to be used for errors in inputs or external components. Also provides an error message.

`revert()`: abort execution and revert state changes

`revert(string reason)`: abort execution and revert state changes, providing an explanatory string

Mathematical and Cryptographic Functions

`addmod(uint x, uint y, uint k) returns (uint)`: compute `(x + y) % k` where the addition is performed with arbitrary precision and does not wrap around at `2**256`. Assert that `k != 0` starting from version 0.5.0.

`mulmod(uint x, uint y, uint k) returns (uint)`: compute `(x * y) % k` where the multiplication is performed with arbitrary precision and does not wrap around at `2**256`. Assert that `k != 0` starting from version 0.5.0.

`keccak256(bytes memory) returns (bytes32)`: compute the Ethereum-SHA-3 (Keccak-256) hash of the input

`sha256(bytes memory) returns (bytes32)`: compute the SHA-256 hash of the input

`sha3(bytes memory) returns (bytes32)`: alias to `keccak256`

`ripemd160(bytes memory) returns (bytes20)`: compute RIPEMD-160 hash of the input

`ecrecover(bytes32 hash, uint8 v, bytes32 r, bytes32 s) returns (address)`: recover the address associated with the public key from elliptic curve signature or return zero on error (example usage)

It might be that you run into Out-of-Gas for `sha256`, `ripemd160` or `ecrecover` on a *private blockchain*. The reason for this is that those are implemented as so-called precompiled contracts and these contracts only really exist after they received the first message (although their contract code is hardcoded). Messages to non-existing contracts are more expensive and thus the execution runs into an Out-of-Gas error. A workaround for this problem is to first send e.g. 1 Wei to each of the contracts before you use them in your actual contracts. This is not an issue on the official or test net.

Address Related

<address>.balance (uint256): balance of the *Address* in Wei

<address>.transfer(uint256 amount): send given amount of Wei to *Address*, throws on failure, forwards 2300 gas stipend, not adjustable

<address>.send(uint256 amount) returns (bool): send given amount of Wei to *Address*, returns `false` on failure, forwards 2300 gas stipend, not adjustable

<address>.call(bytes memory) returns (bool): issue low-level `CALL` with the given payload, returns `false` on failure, forwards all available gas, adjustable

<address>.callcode(bytes memory) returns (bool): issue low-level `CALLCODE` with the given payload, returns `false` on failure, forwards all available gas, adjustable

<address>.delegatecall(bytes memory) returns (bool): issue low-level `DELEGATECALL` with the given payload, returns `false` on failure, forwards all available gas, adjustable

For more information, see the section on *Address*.

Warning: There are some dangers in using `send`: The transfer fails if the call stack depth is at 1024 (this can always be forced by the caller) and it also fails if the recipient runs out of gas. So in order to make safe Ether transfers, always check the return value of `send`, use `transfer` or even better: Use a pattern where the recipient withdraws the money.

Note: Prior to version 0.5.0, Solidity allowed address members to be accessed by a contract instance, for example `this.balance`. This is now forbidden and an explicit conversion to address must be done: `address(this).balance`.

Note: If storage variables are accessed via a low-level delegatecall, the storage layout of the two contracts must align in order for the called contract to correctly access the storage variables of the calling contract by name. This is of course not the case if storage pointers are passed as function arguments as in the case for the high-level libraries.

Note: The use of `callcode` is discouraged and will be removed in the future.

Contract Related

this (current contract's type): the current contract, explicitly convertible to *Address*

selfdestruct(address recipient): destroy the current contract, sending its funds to the given *Address*

suicide(address recipient): deprecated alias to `selfdestruct`

Furthermore, all functions of the current contract are callable directly including the current function.

7.4.5 Expressions and Control Structures

Input Parameters and Output Parameters

As in Javascript, functions may take parameters as input; unlike in Javascript and C, they may also return arbitrary number of parameters as output.

Input Parameters

The input parameters are declared the same way as variables are. As an exception, unused parameters can omit the variable name. For example, suppose we want our contract to accept one kind of external calls with two integers, we would write something like:

```
pragma solidity ^0.4.16;

contract Simple {
    function taker(uint _a, uint _b) public pure {
        // do something with _a and _b.
    }
}
```

Output Parameters

The output parameters can be declared with the same syntax after the `returns` keyword. For example, suppose we wished to return two results: the sum and the product of the two given integers, then we would write:

```
pragma solidity ^0.4.16;

contract Simple {
    function arithmetic(uint _a, uint _b)
        public
        pure
        returns (uint o_sum, uint o_product)
    {
        o_sum = _a + _b;
        o_product = _a * _b;
    }
}
```

The names of output parameters can be omitted. The output values can also be specified using `return` statements. The `return` statements are also capable of returning multiple values, see *Returning Multiple Values*. Return parameters are initialized to zero; if they are not explicitly set, they stay to be zero.

Input parameters and output parameters can be used as expressions in the function body. There, they are also usable in the left-hand side of assignment.

Control Structures

Most of the control structures from JavaScript are available in Solidity except for `switch` and `goto`. So there is: `if`, `else`, `while`, `do`, `for`, `break`, `continue`, `return`, `? :`, with the usual semantics known from C or JavaScript.

Parentheses can *not* be omitted for conditionals, but curly brances can be omitted around single-statement bodies.

Note that there is no type conversion from non-boolean to boolean types as there is in C and JavaScript, so `if (1) { ... }` is *not* valid Solidity.

Returning Multiple Values

When a function has multiple output parameters, `return (v0, v1, ..., vn)` can return multiple values. The number of components must be the same as the number of output parameters.

Function Calls

Internal Function Calls

Functions of the current contract can be called directly ("internally"), also recursively, as seen in this nonsensical example:

```
pragma solidity ^0.4.16;

contract C {
    function g(uint a) public pure returns (uint ret) { return f(); }
    function f() internal pure returns (uint ret) { return g(7) + f(); }
}
```

These function calls are translated into simple jumps inside the EVM. This has the effect that the current memory is not cleared, i.e. passing memory references to internally-called functions is very efficient. Only functions of the same contract can be called internally.

External Function Calls

The expressions `this.g(8);` and `c.g(2);` (where `c` is a contract instance) are also valid function calls, but this time, the function will be called "externally", via a message call and not directly via jumps. Please note that function calls on `this` cannot be used in the constructor, as the actual contract has not been created yet.

Functions of other contracts have to be called externally. For an external call, all function arguments have to be copied to memory.

When calling functions of other contracts, the amount of Wei sent with the call and the gas can be specified with special options `.value()` and `.gas()`, respectively:

```
pragma solidity ^0.4.0;

contract InfoFeed {
    function info() public payable returns (uint ret) { return 42; }
}

contract Consumer {
    InfoFeed feed;
    function setFeed(address addr) public { feed = InfoFeed(addr); }
    function callFeed() public { feed.info.value(10).gas(800)(); }
}
```

The modifier `payable` has to be used for `info`, because otherwise, the *.value()* option would not be available.

Note that the expression `InfoFeed(addr)` performs an explicit type conversion stating that "we know that the type of the contract at the given address is `InfoFeed`" and this does not execute a constructor. Explicit type conversions have to be handled with extreme caution. Never call a function on a contract where you are not sure about its type.

We could also have used `function setFeed(InfoFeed _feed) { feed = _feed; }` directly. Be careful about the fact that `feed.info.value(10).gas(800)` only (locally) sets the value and amount of gas sent with the function call and only the parentheses at the end perform the actual call.

Function calls cause exceptions if the called contract does not exist (in the sense that the account does not contain code) or if the called contract itself throws an exception or goes out of gas.

Warning: Any interaction with another contract imposes a potential danger, especially if the source code of the contract is not known in advance. The current contract hands over control to the called contract and that may potentially do just about anything. Even if the called contract inherits from a known parent contract, the inheriting contract is only required to have a correct interface. The implementation of the contract, however, can be completely arbitrary and thus, pose a danger. In addition, be prepared in case it calls into other contracts of your system or even back into the calling contract before the first call returns. This means that the called contract can change state variables of the calling contract via its functions. Write your functions in a way that, for example, calls to external functions happen after any changes to state variables in your contract so your contract is not vulnerable to a reentrancy exploit.

Named Calls and Anonymous Function Parameters

Function call arguments can also be given by name, in any order, if they are enclosed in `{ }` as can be seen in the following example. The argument list has to coincide by name with the list of parameters from the function declaration, but can be in arbitrary order.

```
pragma solidity ^0.4.0;

contract C {
    function f(uint key, uint value) public {
        // ...
    }

    function g() public {
        // named arguments
        f({value: 2, key: 3});
    }
}
```

Omitted Function Parameter Names

The names of unused parameters (especially return parameters) can be omitted. Those parameters will still be present on the stack, but they are inaccessible.

```
pragma solidity ^0.4.16;

contract C {
    // omitted name for parameter
    function func(uint k, uint) public pure returns(uint) {
        return k;
    }
}
```

Creating Contracts via `new`

A contract can create a new contract using the `new` keyword. The full code of the contract being created has to be known in advance, so recursive creation-dependencies are not possible.

```
pragma solidity >0.4.24;

contract D {
    uint x;
    constructor(uint a) public payable {
        x = a;
    }
}

contract C {
    D d = new D(4); // will be executed as part of C's constructor

    function createD(uint arg) public {
        D newD = new D(arg);
    }

    function createAndEndowD(uint arg, uint amount) public payable {
        // Send ether along with the creation
        D newD = (new D).value(amount)(arg);
    }
}
```

As seen in the example, it is possible to forward Ether while creating an instance of `D` using the `.value()` option, but it is not possible to limit the amount of gas. If the creation fails (due to out-of-stack, not enough balance or other problems), an exception is thrown.

Order of Evaluation of Expressions

The evaluation order of expressions is not specified (more formally, the order in which the children of one node in the expression tree are evaluated is not specified, but they are of course evaluated before the node itself). It is only guaranteed that statements are executed in order and short-circuiting for boolean expressions is done. See *Order of Precedence of Operators* for more information.

Assignment

Destructuring Assignments and Returning Multiple Values

Solidity internally allows tuple types, i.e. a list of objects of potentially different types whose size is a constant at compile-time. Those tuples can be used to return multiple values at the same time. These can then either be assigned to newly declared variables or to pre-existing variables (or LValues in general):

```
pragma solidity >0.4.23 <0.5.0;

contract C {
    uint[] data;

    function f() public pure returns (uint, bool, uint) {
        return (7, true, 2);
    }
```

(continues on next page)

(continued from previous page)

```
    function g() public {
        // Variables declared with type and assigned from the returned tuple.
        (uint x, bool b, uint y) = f();
        // Common trick to swap values -- does not work for non-value storage types.
        (x, y) = (y, x);
        // Components can be left out (also for variable declarations).
        (data.length,,) = f(); // Sets the length to 7
    }
}
```

Note: Prior to version 0.5.0 it was possible to assign to tuples of smaller size, either filling up on the left or on the right side (which ever was empty). This is now disallowed, so both sides have to have the same number of components.

Complications for Arrays and Structs

The semantics of assignment are a bit more complicated for non-value types like arrays and structs. Assigning *to* a state variable always creates an independent copy. On the other hand, assigning to a local variable creates an independent copy only for elementary types, i.e. static types that fit into 32 bytes. If structs or arrays (including `bytes` and `string`) are assigned from a state variable to a local variable, the local variable holds a reference to the original state variable. A second assignment to the local variable does not modify the state but only changes the reference. Assignments to members (or elements) of the local variable *do* change the state.

Scoping and Declarations

A variable which is declared will have an initial default value whose byte-representation is all zeros. The "default values" of variables are the typical "zero-state" of whatever the type is. For example, the default value for a `bool` is `false`. The default value for the `uint` or `int` types is `0`. For statically-sized arrays and `bytes1` to `bytes32`, each individual element will be initialized to the default value corresponding to its type. Finally, for dynamically-sized arrays, `bytes` and `string`, the default value is an empty array or string.

Scoping in Solidity follows the widespread scoping rules of C99 (and many other languages): Variables are visible from the point right after their declaration until the end of a `{ }`-block. As an exception to this rule, variables declared in the initialization part of a for-loop are only visible until the end of the for-loop.

Variables and other items declared outside of a code block, for example functions, contracts, user-defined types, etc., do not change their scoping behaviour. This means you can use state variables before they are declared and call functions recursively.

As a consequence, the following examples will compile without warnings, since the two variables have the same name but disjoint scopes.

```
pragma solidity >0.4.24;
contract C {
    function minimalScoping() pure public {
        {
            uint same2 = 0;
        }

        {
            uint same2 = 0;
        }
```

(continues on next page)

(continued from previous page)

```
    }
}
```

As a special example of the C99 scoping rules, note that in the following, the first assignment to `x` will actually assign the outer and not the inner variable. In any case, you will get a warning about the outer variable being shadowed.

```
pragma solidity >0.4.24;
contract C {
    function f() pure public returns (uint) {
        uint x = 1;
        {
            x = 2; // this will assign to the outer variable
            uint x;
        }
        return x; // x has value 2
    }
}
```

Warning:

Before version 0.5.0 Solidity followed the same scoping rules as JavaScript, that is, a variable declared anywhere within a function would be in scope for the entire function, regardless where it was declared. Note that this is a breaking change. The following example shows a code snippet that used to compile but leads to an error starting from version 0.5.0.

```
// This will not compile

pragma solidity >0.4.24;
contract C {
    function f() pure public returns (uint) {
        x = 2;
        uint x;
        return x;
    }
}
```

Error handling: Assert, Require, Revert and Exceptions

Solidity uses state-reverting exceptions to handle errors. Such an exception will undo all changes made to the state in the current call (and all its sub-calls) and also flag an error to the caller. The convenience functions `assert` and `require` can be used to check for conditions and throw an exception if the condition is not met. The `assert` function should only be used to test for internal errors, and to check invariants. The `require` function should be used to ensure valid conditions, such as inputs, or contract state variables are met, or to validate return values from calls to external contracts. If used properly, analysis tools can evaluate your contract to identify the conditions and function calls which will reach a failing `assert`. Properly functioning code should never reach a failing assert statement; if this happens there is a bug in your contract which you should fix.

There are two other ways to trigger exceptions: The `revert` function can be used to flag an error and revert the current call. It is possible to provide a string message containing details about the error that will be passed back to the caller. The deprecated keyword `throw` can also be used as an alternative to `revert()` (but only without error message).

Note: From version 0.4.13 the `throw` keyword is deprecated and will be phased out in the future.

When exceptions happen in a sub-call, they "bubble up" (i.e. exceptions are rethrown) automatically. Exceptions to this rule are `send` and the low-level functions `call`, `delegatecall` and `callcode` – those return `false` in case of an exception instead of "bubbling up".

Warning: The low-level `call`, `delegatecall` and `callcode` will return success if the called account is non-existent, as part of the design of EVM. Existence must be checked prior to calling if desired.

Catching exceptions is not yet possible.

In the following example, you can see how `require` can be used to easily check conditions on inputs and how `assert` can be used for internal error checking. Note that you can optionally provide a message string for `require`, but not for `assert`.

```
pragma solidity >0.4.24;

contract Sharer {
    function sendHalf(address addr) public payable returns (uint balance) {
        require(msg.value % 2 == 0, "Even value required.");
        uint balanceBeforeTransfer = address(this).balance;
        addr.transfer(msg.value / 2);
        // Since transfer throws an exception on failure and
        // cannot call back here, there should be no way for us to
        // still have half of the money.
        assert(address(this).balance == balanceBeforeTransfer - msg.value / 2);
        return address(this).balance;
    }
}
```

An `assert`-style exception is generated in the following situations:

1. If you access an array at a too large or negative index (i.e. `x[i]` where `i >= x.length` or `i < 0`).
2. If you access a fixed-length `bytesN` at a too large or negative index.
3. If you divide or modulo by zero (e.g. `5 / 0` or `23 % 0`).
4. If you shift by a negative amount.
5. If you convert a value too big or negative into an enum type.
6. If you call a zero-initialized variable of internal function type.
7. If you call `assert` with an argument that evaluates to false.

A `require`-style exception is generated in the following situations:

1. Calling `throw`.
2. Calling `require` with an argument that evaluates to `false`.
3. If you call a function via a message call but it does not finish properly (i.e. it runs out of gas, has no matching function, or throws an exception itself), except when a low level operation `call`, `send`, `delegatecall` or `callcode` is used. The low level operations never throw exceptions but indicate failures by returning `false`.
4. If you create a contract using the `new` keyword but the contract creation does not finish properly (see above for the definition of "not finish properly").

5. If you perform an external function call targeting a contract that contains no code.
6. If your contract receives Ether via a public function without `payable` modifier (including the constructor and the fallback function).
7. If your contract receives Ether via a public getter function.
8. If a `.transfer()` fails.

Internally, Solidity performs a revert operation (instruction `0xfd`) for a `require`-style exception and executes an invalid operation (instruction `0xfe`) to throw an `assert`-style exception. In both cases, this causes the EVM to revert all changes made to the state. The reason for reverting is that there is no safe way to continue execution, because an expected effect did not occur. Because we want to retain the atomicity of transactions, the safest thing to do is to revert all changes and make the whole transaction (or at least call) without effect. Note that `assert`-style exceptions consume all gas available to the call, while `require`-style exceptions will not consume any gas starting from the Metropolis release.

The following example shows how an error string can be used together with revert and require:

```
pragma solidity ^0.4.22;

contract VendingMachine {
    function buy(uint amount) payable {
        if (amount > msg.value / 2 ether)
            revert("Not enough Ether provided.");
        // Alternative way to do it:
        require(
            amount <= msg.value / 2 ether,
            "Not enough Ether provided."
        );
        // Perform the purchase.
    }
}
```

The provided string will be *abi-encoded* as if it were a call to a function `Error(string)`. In the above example, `revert("Not enough Ether provided.");` will cause the following hexadecimal data be set as error return data:

```
0x08c379a0                                                         // Function␣
↪selector for Error(string)
0x0000000000000000000000000000000000000000000000000000000000000020 // Data offset
0x000000000000000000000000000000000000000000000000000000000000001a // String length
0x4e6f7420656e6f756768204574686572207072696465642e000000000000 // String data
```

7.4.6 Contracts

Contracts in Solidity are similar to classes in object-oriented languages. They contain persistent data in state variables and functions that can modify these variables. Calling a function on a different contract (instance) will perform an EVM function call and thus switch the context such that state variables are inaccessible.

Creating Contracts

Contracts can be created "from outside" via Ethereum transactions or from within Solidity contracts.

IDEs, such as Remix, make the creation process seamless using UI elements.

Creating contracts programmatically on Ethereum is best done via using the JavaScript API web3.js. As of today it has a method called web3.eth.Contract to facilitate contract creation.

When a contract is created, its constructor (a function declared with the `constructor` keyword) is executed once. A constructor is optional. Only one constructor is allowed, and this means overloading is not supported.

Internally, constructor arguments are passed *ABI encoded* after the code of the contract itself, but you do not have to care about this if you use `web3.js`.

If a contract wants to create another contract, the source code (and the binary) of the created contract has to be known to the creator. This means that cyclic creation dependencies are impossible.

```
pragma solidity ^0.4.22;

contract OwnedToken {
    // TokenCreator is a contract type that is defined below.
    // It is fine to reference it as long as it is not used
    // to create a new contract.
    TokenCreator creator;
    address owner;
    bytes32 name;

    // This is the constructor which registers the
    // creator and the assigned name.
    constructor(bytes32 _name) public {
        // State variables are accessed via their name
        // and not via e.g. this.owner. This also applies
        // to functions and especially in the constructors,
        // you can only call them like that ("internally"),
        // because the contract itself does not exist yet.
        owner = msg.sender;
        // We do an explicit type conversion from `address`
        // to `TokenCreator` and assume that the type of
        // the calling contract is TokenCreator, there is
        // no real way to check that.
        creator = TokenCreator(msg.sender);
        name = _name;
    }

    function changeName(bytes32 newName) public {
        // Only the creator can alter the name --
        // the comparison is possible since contracts
        // are implicitly convertible to addresses.
        if (msg.sender == address(creator))
            name = newName;
    }

    function transfer(address newOwner) public {
        // Only the current owner can transfer the token.
        if (msg.sender != owner) return;
        // We also want to ask the creator if the transfer
        // is fine. Note that this calls a function of the
        // contract defined below. If the call fails (e.g.
        // due to out-of-gas), the execution here stops
        // immediately.
        if (creator.isTokenTransferOK(owner, newOwner))
            owner = newOwner;
    }
}
```

(continues on next page)

(continued from previous page)

```
contract TokenCreator {
    function createToken(bytes32 name)
       public
       returns (OwnedToken tokenAddress)
    {
        // Create a new Token contract and return its address.
        // From the JavaScript side, the return type is simply
        // `address`, as this is the closest type available in
        // the ABI.
        return new OwnedToken(name);
    }

    function changeName(OwnedToken tokenAddress, bytes32 name)  public {
        // Again, the external type of `tokenAddress` is
        // simply `address`.
        tokenAddress.changeName(name);
    }

    function isTokenTransferOK(address currentOwner, address newOwner)
        public
        view
        returns (bool ok)
    {
        // Check some arbitrary condition.
        address tokenAddress = msg.sender;
        return (keccak256(abi.encodePacked(newOwner)) & 0xff) ==␣
↪(bytes20(tokenAddress) & 0xff);
    }
}
```

Visibility and Getters

Since Solidity knows two kinds of function calls (internal ones that do not create an actual EVM call (also called a "message call") and external ones that do), there are four types of visibilities for functions and state variables.

Functions can be specified as being `external`, `public`, `internal` or `private`, where the default is `public`. For state variables, `external` is not possible and the default is `internal`.

`external`: External functions are part of the contract interface, which means they can be called from other contracts and via transactions. An external function `f` cannot be called internally (i.e. `f()` does not work, but `this.f()` works). External functions are sometimes more efficient when they receive large arrays of data.

`public`: Public functions are part of the contract interface and can be either called internally or via messages. For public state variables, an automatic getter function (see below) is generated.

`internal`: Those functions and state variables can only be accessed internally (i.e. from within the current contract or contracts deriving from it), without using `this`.

`private`: Private functions and state variables are only visible for the contract they are defined in and not in derived contracts.

Note: Everything that is inside a contract is visible to all external observers. Making something `private` only prevents other contracts from accessing and modifying the information, but it will still be visible to the whole world outside of the blockchain.

The visibility specifier is given after the type for state variables and between parameter list and return parameter list for functions.

```
pragma solidity ^0.4.16;

contract C {
    function f(uint a) private pure returns (uint b) { return a + 1; }
    function setData(uint a) internal { data = a; }
    uint public data;
}
```

In the following example, `D`, can call `c.getData()` to retrieve the value of `data` in state storage, but is not able to call `f`. Contract `E` is derived from `C` and, thus, can call `compute`.

```
// This will not compile

pragma solidity ^0.4.0;

contract C {
    uint private data;

    function f(uint a) private pure returns(uint b) { return a + 1; }
    function setData(uint a) public { data = a; }
    function getData() public view returns(uint) { return data; }
    function compute(uint a, uint b) internal pure returns (uint) { return a + b; }
}

contract D {
    function readData() public {
        C c = new C();
        uint local = c.f(7); // error: member `f` is not visible
        c.setData(3);
        local = c.getData();
        local = c.compute(3, 5); // error: member `compute` is not visible
    }
}

contract E is C {
    function g() public {
        C c = new C();
        uint val = compute(3, 5); // access to internal member (from derived to␣
↪parent contract)
    }
}
```

Getter Functions

The compiler automatically creates getter functions for all **public** state variables. For the contract given below, the compiler will generate a function called `data` that does not take any arguments and returns a `uint`, the value of the state variable `data`. The initialization of state variables can be done at declaration.

```
pragma solidity ^0.4.0;

contract C {
    uint public data = 42;
}
```

(continues on next page)

(continued from previous page)

```
contract Caller {
    C c = new C();
    function f() public {
        uint local = c.data();
    }
}
```

The getter functions have external visibility. If the symbol is accessed internally (i.e. without `this.`), it is evaluated as a state variable. If it is accessed externally (i.e. with `this.`), it is evaluated as a function.

```
pragma solidity ^0.4.0;

contract C {
    uint public data;
    function x() public {
        data = 3; // internal access
        uint val = this.data(); // external access
    }
}
```

The next example is a bit more complex:

```
pragma solidity ^0.4.0;

contract Complex {
    struct Data {
        uint a;
        bytes3 b;
        mapping (uint => uint) map;
    }
    mapping (uint => mapping(bool => Data[])) public data;
}
```

It will generate a function of the following form:

```
function data(uint arg1, bool arg2, uint arg3) public returns (uint a, bytes3 b) {
    a = data[arg1][arg2][arg3].a;
    b = data[arg1][arg2][arg3].b;
}
```

Note that the mapping in the struct is omitted because there is no good way to provide the key for the mapping.

Function Modifiers

Modifiers can be used to easily change the behaviour of functions. For example, they can automatically check a condition prior to executing the function. Modifiers are inheritable properties of contracts and may be overridden by derived contracts.

```
pragma solidity >0.4.24;

contract owned {
    constructor() public { owner = msg.sender; }
    address owner;
```

(continues on next page)

(continued from previous page)

```
    // This contract only defines a modifier but does not use
    // it: it will be used in derived contracts.
    // The function body is inserted where the special symbol
    // `_;` in the definition of a modifier appears.
    // This means that if the owner calls this function, the
    // function is executed and otherwise, an exception is
    // thrown.
    modifier onlyOwner {
        require(
            msg.sender == owner,
            "Only owner can call this function."
        );
        _;
    }
}

contract mortal is owned {
    // This contract inherits the `onlyOwner` modifier from
    // `owned` and applies it to the `close` function, which
    // causes that calls to `close` only have an effect if
    // they are made by the stored owner.
    function close() public onlyOwner {
        selfdestruct(owner);
    }
}

contract priced {
    // Modifiers can receive arguments:
    modifier costs(uint price) {
        if (msg.value >= price) {
            _;
        }
    }
}

contract Register is priced, owned {
    mapping (address => bool) registeredAddresses;
    uint price;

    constructor(uint initialPrice) public { price = initialPrice; }

    // It is important to also provide the
    // `payable` keyword here, otherwise the function will
    // automatically reject all Ether sent to it.
    function register() public payable costs(price) {
        registeredAddresses[msg.sender] = true;
    }

    function changePrice(uint _price) public onlyOwner {
        price = _price;
    }
}

contract Mutex {
    bool locked;
    modifier noReentrancy() {
        require(
```

(continues on next page)

(continued from previous page)

```
            !locked,
            "Reentrant call."
        );
        locked = true;
        _;
        locked = false;
    }

    /// This function is protected by a mutex, which means that
    /// reentrant calls from within `msg.sender.call` cannot call `f` again.
    /// The `return 7` statement assigns 7 to the return value but still
    /// executes the statement `locked = false` in the modifier.
    function f() public noReentrancy returns (uint) {
        require(msg.sender.call(""));
        return 7;
    }
}
```

Multiple modifiers are applied to a function by specifying them in a whitespace-separated list and are evaluated in the order presented.

Warning: In an earlier version of Solidity, `return` statements in functions having modifiers behaved differently.

Explicit returns from a modifier or function body only leave the current modifier or function body. Return variables are assigned and control flow continues after the "_" in the preceding modifier.

Arbitrary expressions are allowed for modifier arguments and in this context, all symbols visible from the function are visible in the modifier. Symbols introduced in the modifier are not visible in the function (as they might change by overriding).

Constant State Variables

State variables can be declared as `constant`. In this case, they have to be assigned from an expression which is a constant at compile time. Any expression that accesses storage, blockchain data (e.g. `now`, `address(this).balance` or `block.number`) or execution data (`msg.value` or `gasleft()`) or make calls to external contracts are disallowed. Expressions that might have a side-effect on memory allocation are allowed, but those that might have a side-effect on other memory objects are not. The built-in functions `keccak256`, `sha256`, `ripemd160`, `ecrecover`, `addmod` and `mulmod` are allowed (even though they do call external contracts).

The reason behind allowing side-effects on the memory allocator is that it should be possible to construct complex objects like e.g. lookup-tables. This feature is not yet fully usable.

The compiler does not reserve a storage slot for these variables, and every occurrence is replaced by the respective constant expression (which might be computed to a single value by the optimizer).

Not all types for constants are implemented at this time. The only supported types are value types and strings.

```
pragma solidity ^0.4.0;

contract C {
    uint constant x = 32**22 + 8;
    string constant text = "abc";
    bytes32 constant myHash = keccak256("abc");
}
```

Functions

View Functions

Functions can be declared `view` in which case they promise not to modify the state.

Note: If the compiler's EVM target is Byzantium or newer (default) the opcode `STATICCALL` is used.

The following statements are considered modifying the state:

1. Writing to state variables.
2. *Emitting events*.
3. *Creating other contracts*.
4. Using `selfdestruct`.
5. Sending Ether via calls.
6. Calling any function not marked `view` or `pure`.
7. Using low-level calls.
8. Using inline assembly that contains certain opcodes.

```
pragma solidity >0.4.24;

contract C {
    function f(uint a, uint b) public view returns (uint) {
        return a * (b + 42) + now;
    }
}
```

Note: `constant` on functions used to be an alias to `view`, but this was dropped in version 0.5.0.

Note: Getter methods are marked `view`.

Note: Prior to version 0.5.0, the compiler did not use the `STATICCALL` opcode for `view` functions. This enabled state modifications in `view` functions through the use of invalid explicit type conversions. By using `STATICCALL` for `view` functions, modifications to the state are prevented on the level of the EVM.

Pure Functions

Functions can be declared `pure` in which case they promise not to read from or modify the state.

Note: If the compiler's EVM target is Byzantium or newer (default) the opcode `STATICCALL` is used.

In addition to the list of state modifying statements explained above, the following are considered reading from the state:

1. Reading from state variables.
2. Accessing `address(this).balance` or `<address>.balance`.
3. Accessing any of the members of `block`, `tx`, `msg` (with the exception of `msg.sig` and `msg.data`).
4. Calling any function not marked `pure`.
5. Using inline assembly that contains certain opcodes.

```
pragma solidity >0.4.24;

contract C {
    function f(uint a, uint b) public pure returns (uint) {
        return a * (b + 42);
    }
}
```

Note: Prior to version 0.5.0, the compiler did not use the `STATICCALL` opcode for `pure` functions. This enabled state modifications in `pure` functions through the use of invalid explicit type conversions. By using `STATICCALL` for `pure` functions, modifications to the state are prevented on the level of the EVM.

Warning: It is not possible to prevent functions from reading the state at the level of the EVM, it is only possible to prevent them from writing to the state (i.e. only `view` can be enforced at the EVM level, `pure` can not). It is a non-circumventable runtime checks done by the EVM.

Warning: Before version 0.4.17 the compiler did not enforce that `pure` is not reading the state. It is a compile-time type check, which can be circumvented doing invalid explicit conversions between contract types, because the compiler can verify that the type of the contract does not do state-changing operations, but it cannot check that the contract that will be called at runtime is actually of that type.

Warning: Before version 0.5.0 the compiler did not enforce that `view` is not writing the state.

Fallback Function

A contract can have exactly one unnamed function. This function cannot have arguments, cannot return anything and has to have `external` visibility. It is executed on a call to the contract if none of the other functions match the given function identifier (or if no data was supplied at all).

Furthermore, this function is executed whenever the contract receives plain Ether (without data). Additionally, in order to receive Ether, the fallback function must be marked `payable`. If no such function exists, the contract cannot receive Ether through regular transactions.

In the worst case, the fallback function can only rely on 2300 gas being available (for example when send or transfer is used), leaving not much room to perform other operations except basic logging. The following operations will consume more gas than the 2300 gas stipend:

- Writing to storage
- Creating a contract

- Calling an external function which consumes a large amount of gas
- Sending Ether

Like any function, the fallback function can execute complex operations as long as there is enough gas passed on to it.

Note: Even though the fallback function cannot have arguments, one can still use `msg.data` to retrieve any payload supplied with the call.

Warning: Contracts that receive Ether directly (without a function call, i.e. using `send` or `transfer`) but do not define a fallback function throw an exception, sending back the Ether (this was different before Solidity v0.4.0). So if you want your contract to receive Ether, you have to implement a payable fallback function.

Warning: A contract without a payable fallback function can receive Ether as a recipient of a *coinbase transaction* (aka *miner block reward*) or as a destination of a `selfdestruct`.

A contract cannot react to such Ether transfers and thus also cannot reject them. This is a design choice of the EVM and Solidity cannot work around it.

It also means that `address(this).balance` can be higher than the sum of some manual accounting implemented in a contract (i.e. having a counter updated in the fallback function).

```
pragma solidity >0.4.24;

contract Test {
    // This function is called for all messages sent to
    // this contract (there is no other function).
    // Sending Ether to this contract will cause an exception,
    // because the fallback function does not have the `payable`
    // modifier.
    function() external { x = 1; }
    uint x;
}


// This contract keeps all Ether sent to it with no way
// to get it back.
contract Sink {
    function() external payable { }
}

contract Caller {
    function callTest(Test test) public {
        address(test).call(abi.encodeWithSignature("nonExistingFunction()"));
        // results in test.x becoming == 1.

        // If someone sends ether to that contract,
        // the transaction will fail and reject the
        // Ether.
        address(test).send(2 ether);
    }
}
```

Function Overloading

A Contract can have multiple functions of the same name but with different arguments. This also applies to inherited functions. The following example shows overloading of the `f` function in the scope of contract `A`.

```
pragma solidity ^0.4.16;

contract A {
    function f(uint _in) public pure returns (uint out) {
        out = 1;
    }

    function f(uint _in, bytes32 _key) public pure returns (uint out) {
        out = 2;
    }
}
```

Overloaded functions are also present in the external interface. It is an error if two externally visible functions differ by their Solidity types but not by their external types.

```
// This will not compile
pragma solidity ^0.4.16;

contract A {
    function f(B _in) public pure returns (B out) {
        out = _in;
    }

    function f(address _in) public pure returns (address out) {
        out = _in;
    }
}

contract B {
}
```

Both `f` function overloads above end up accepting the address type for the ABI although they are considered different inside Solidity.

Overload resolution and Argument matching

Overloaded functions are selected by matching the function declarations in the current scope to the arguments supplied in the function call. Functions are selected as overload candidates if all arguments can be implicitly converted to the expected types. If there is not exactly one candidate, resolution fails.

Note: Return parameters are not taken into account for overload resolution.

```
pragma solidity ^0.4.16;

contract A {
    function f(uint8 _in) public pure returns (uint8 out) {
        out = _in;
    }

```

(continues on next page)

(continued from previous page)

```
    function f(uint256 _in) public pure returns (uint256 out) {
        out = _in;
    }
}
```

Calling `f(50)` would create a type error since `50` can be implicitly converted both to `uint8` and `uint256` types. On another hand `f(256)` would resolve to `f(uint256)` overload as `256` cannot be implicitly converted to `uint8`.

Events

Events allow the convenient usage of the EVM logging facilities, which in turn can be used to "call" JavaScript callbacks in the user interface of a dapp, which listen for these events.

Events are inheritable members of contracts. When they are called, they cause the arguments to be stored in the transaction's log - a special data structure in the blockchain. These logs are associated with the address of the contract and will be incorporated into the blockchain and stay there as long as a block is accessible (forever as of Frontier and Homestead, but this might change with Serenity). Log and event data is not accessible from within contracts (not even from the contract that created them).

SPV proofs for logs are possible, so if an external entity supplies a contract with such a proof, it can check that the log actually exists inside the blockchain. But be aware that block headers have to be supplied because the contract can only see the last 256 block hashes.

Up to three parameters can receive the attribute `indexed` which will cause the respective arguments to be stored in a special data structure as so-called "topics", which allows them to be searched for, for example when filtering a sequence of blocks for certain events. Events can always be filtered by the address of the contract that emitted the event. Also, the hash of the signature of the event is one of the topics except if you declared the event with `anonymous` specifier. This means that it is not possible to filter for specific anonymous events by name.

If arrays (including `string` and `bytes`) are used as indexed arguments, the Keccak-256 hash of it is stored as topic instead. This is because a topic can only hold a single word (32 bytes).

All non-indexed arguments will be *ABI-encoded* into the data part of the log.

```
pragma solidity ^0.4.21;

contract ClientReceipt {
    event Deposit(
        address indexed _from,
        bytes32 indexed _id,
        uint _value
    );

    function deposit(bytes32 _id) public payable {
        // Events are emitted using `emit`, followed by
        // the name of the event and the arguments
        // (if any) in parentheses. Any such invocation
        // (even deeply nested) can be detected from
        // the JavaScript API by filtering for `Deposit`.
        emit Deposit(msg.sender, _id, msg.value);
    }
}
```

The use in the JavaScript API would be as follows:

```
var abi = /* abi as generated by the compiler */;
var ClientReceipt = web3.eth.contract(abi);
var clientReceipt = ClientReceipt.at("0x1234...ab67" /* address */);

var event = clientReceipt.Deposit();

// watch for changes
event.watch(function(error, result){
    // result will contain various information
    // including the arguments given to the `Deposit`
    // call.
    if (!error)
        console.log(result);
});

// Or pass a callback to start watching immediately
var event = clientReceipt.Deposit(function(error, result) {
    if (!error)
        console.log(result);
});
```

Low-Level Interface to Logs

It is also possible to access the low-level interface to the logging mechanism via the functions `log0`, `log1`, `log2`, `log3` and `log4`. `logi` takes `i + 1` parameter of type `bytes32`, where the first argument will be used for the data part of the log and the others as topics. The event call above can be performed in the same way as

```
pragma solidity ^0.4.10;

contract C {
    function f() public payable {
        bytes32 _id = 0x420042;
        log3(
            bytes32(msg.value),
            bytes32(0x50cb9fe53daa9737b786ab3646f04d0150dc50ef4e75f59509d83667ad5adb20),
            bytes32(uint256(msg.sender)),
            _id
        );
    }
}
```

where the long hexadecimal number is equal to `keccak256("Deposit(address,bytes32,uint256)")`, the signature of the event.

Additional Resources for Understanding Events

- Javascript documentation
- Example usage of events
- How to access them in js

Inheritance

Solidity supports multiple inheritance by copying code including polymorphism.

All function calls are virtual, which means that the most derived function is called, except when the contract name is explicitly given.

When a contract inherits from multiple contracts, only a single contract is created on the blockchain, and the code from all the base contracts is copied into the created contract.

The general inheritance system is very similar to Python's, especially concerning multiple inheritance.

Details are given in the following example.

```
pragma solidity ^0.4.22;

contract owned {
    constructor() { owner = msg.sender; }
    address owner;
}

// Use `is` to derive from another contract. Derived
// contracts can access all non-private members including
// internal functions and state variables. These cannot be
// accessed externally via `this`, though.
contract mortal is owned {
    function kill() {
        if (msg.sender == owner) selfdestruct(owner);
    }
}

// These abstract contracts are only provided to make the
// interface known to the compiler. Note the function
// without body. If a contract does not implement all
// functions it can only be used as an interface.
contract Config {
    function lookup(uint id) public returns (address adr);
}

contract NameReg {
    function register(bytes32 name) public;
    function unregister() public;
 }

// Multiple inheritance is possible. Note that `owned` is
// also a base class of `mortal`, yet there is only a single
// instance of `owned` (as for virtual inheritance in C++).
contract named is owned, mortal {
    constructor(bytes32 name) {
        Config config = Config(0xD5f9D8D94886E70b06E474c3fB14Fd43E2f23970);
        NameReg(config.lookup(1)).register(name);
    }

    // Functions can be overridden by another function with the same name and
    // the same number/types of inputs.  If the overriding function has different
    // types of output parameters, that causes an error.
    // Both local and message-based function calls take these overrides
    // into account.
    function kill() public {
```

(continues on next page)

(continued from previous page)

```
        if (msg.sender == owner) {
            Config config = Config(0xD5f9D8D94886E70b06E474c3fB14Fd43E2f23970);
            NameReg(config.lookup(1)).unregister();
            // It is still possible to call a specific
            // overridden function.
            mortal.kill();
        }
    }
}

// If a constructor takes an argument, it needs to be
// provided in the header (or modifier-invocation-style at
// the constructor of the derived contract (see below)).
contract PriceFeed is owned, mortal, named("GoldFeed") {
   function updateInfo(uint newInfo) public {
      if (msg.sender == owner) info = newInfo;
   }

   function get() public view returns(uint r) { return info; }

   uint info;
}
```

Note that above, we call `mortal.kill()` to "forward" the destruction request. The way this is done is problematic, as seen in the following example:

```
pragma solidity ^0.4.22;

contract owned {
    constructor() public { owner = msg.sender; }
    address owner;
}

contract mortal is owned {
    function kill() public {
        if (msg.sender == owner) selfdestruct(owner);
    }
}

contract Base1 is mortal {
    function kill() public { /* do cleanup 1 */ mortal.kill(); }
}

contract Base2 is mortal {
    function kill() public { /* do cleanup 2 */ mortal.kill(); }
}

contract Final is Base1, Base2 {
}
```

A call to `Final.kill()` will call `Base2.kill` as the most derived override, but this function will bypass `Base1.kill`, basically because it does not even know about `Base1`. The way around this is to use `super`:

```
pragma solidity ^0.4.22;

contract owned {
```

(continues on next page)

(continued from previous page)

```
    constructor() public { owner = msg.sender; }
    address owner;
}

contract mortal is owned {
    function kill() public {
        if (msg.sender == owner) selfdestruct(owner);
    }
}

contract Base1 is mortal {
    function kill() public { /* do cleanup 1 */ super.kill(); }
}


contract Base2 is mortal {
    function kill() public { /* do cleanup 2 */ super.kill(); }
}

contract Final is Base1, Base2 {
}
```

If `Base2` calls a function of `super`, it does not simply call this function on one of its base contracts. Rather, it calls this function on the next base contract in the final inheritance graph, so it will call `Base1.kill()` (note that the final inheritance sequence is – starting with the most derived contract: Final, Base2, Base1, mortal, owned). The actual function that is called when using super is not known in the context of the class where it is used, although its type is known. This is similar for ordinary virtual method lookup.

Constructors

A constructor is an optional function declared with the `constructor` keyword which is executed upon contract creation. Constructor functions can be either `public` or `internal`. If there is no constructor, the contract will assume the default constructor: `contructor() public {}`.

```
pragma solidity >0.4.24;

contract A {
    uint public a;

    constructor(uint _a) internal {
        a = _a;
    }
}

contract B is A(1) {
    constructor() public {}
}
```

A constructor set as `internal` causes the contract to be marked as *abstract*.

Warning: Prior to version 0.4.22, constructors were defined as functions with the same name as the contract. This syntax was deprecated and is not allowed anymore in version 0.5.0.

Arguments for Base Constructors

The constructors of all the base contracts will be called following the linearization rules explained below. If the base constructors have arguments, derived contracts need to specify all of them. This can be done in two ways:

```
pragma solidity ^0.4.22;

contract Base {
    uint x;
    constructor(uint _x) public { x = _x; }
}

contract Derived1 is Base(7) {
    constructor(uint _y) public {}
}

contract Derived2 is Base {
    constructor(uint _y) Base(_y * _y) public {}
}
```

One way is directly in the inheritance list (`is Base(7)`). The other is in the way a modifier would be invoked as part of the header of the derived constructor (`Base(_y * _y)`). The first way to do it is more convenient if the constructor argument is a constant and defines the behaviour of the contract or describes it. The second way has to be used if the constructor arguments of the base depend on those of the derived contract. Arguments have to be given either in the inheritance list or in modifier-style in the derived constructor. Specifying arguments in both places is an error.

If a derived contract doesn't specify the arguments to all of its base contracts' constructors, it will be abstract.

Multiple Inheritance and Linearization

Languages that allow multiple inheritance have to deal with several problems. One is the Diamond Problem. Solidity is similar to Python in that it uses "C3 Linearization" to force a specific order in the DAG of base classes. This results in the desirable property of monotonicity but disallows some inheritance graphs. Especially, the order in which the base classes are given in the `is` directive is important: You have to list the direct base contracts in the order from "most base-like" to "most derived". Note that this order is different from the one used in Python. In the following code, Solidity will give the error "Linearization of inheritance graph impossible".

```
// This will not compile

pragma solidity ^0.4.0;

contract X {}
contract A is X {}
contract C is A, X {}
```

The reason for this is that `C` requests `X` to override `A` (by specifying `A, X` in this order), but `A` itself requests to override `X`, which is a contradiction that cannot be resolved.

Inheriting Different Kinds of Members of the Same Name

When the inheritance results in a contract with a function and a modifier of the same name, it is considered as an error. This error is produced also by an event and a modifier of the same name, and a function and an event of the same name. As an exception, a state variable getter can override a public function.

Abstract Contracts

Contracts are marked as abstract when at least one of their functions lacks an implementation as in the following example (note that the function declaration header is terminated by `;`):

```
pragma solidity ^0.4.0;

contract Feline {
    function utterance() public returns (bytes32);
}
```

Such contracts cannot be compiled (even if they contain implemented functions alongside non-implemented functions), but they can be used as base contracts:

```
pragma solidity ^0.4.0;

contract Feline {
    function utterance() public returns (bytes32);
}

contract Cat is Feline {
    function utterance() public returns (bytes32) { return "miaow"; }
}
```

If a contract inherits from an abstract contract and does not implement all non-implemented functions by overriding, it will itself be abstract.

Note that a function without implementation is different from a *Function Type* even though their syntax looks very similar.

Example of function without implementation (a function declaration):

```
function foo(address) external returns (address);
```

Example of a Function Type (a variable declaration, where the variable is of type `function`):

```
function(address) external returns (address) foo;
```

Abstract contracts decouple the definition of a contract from its implementation providing better extensibility and self-documentation and facilitating patterns like the Template method and removing code duplication. Abstract contracts are useful in the same way that defining methods in an interface is useful. It is a way for the designer of the abstract contract to say "any child of mine must implement this method".

Interfaces

Interfaces are similar to abstract contracts, but they cannot have any functions implemented. There are further restrictions:

- Cannot inherit other contracts or interfaces.
- All declared functions must be external.
- Cannot define constructor.
- Cannot define variables.
- Cannot define structs.
- Cannot define enums.

Some of these restrictions might be lifted in the future.

Interfaces are basically limited to what the Contract ABI can represent, and the conversion between the ABI and an Interface should be possible without any information loss.

Interfaces are denoted by their own keyword:

```
pragma solidity ^0.4.11;

interface Token {
    function transfer(address recipient, uint amount) external;
}
```

Contracts can inherit interfaces as they would inherit other contracts.

Libraries

Libraries are similar to contracts, but their purpose is that they are deployed only once at a specific address and their code is reused using the `DELEGATECALL` (`CALLCODE` until Homestead) feature of the EVM. This means that if library functions are called, their code is executed in the context of the calling contract, i.e. `this` points to the calling contract, and especially the storage from the calling contract can be accessed. As a library is an isolated piece of source code, it can only access state variables of the calling contract if they are explicitly supplied (it would have no way to name them, otherwise). Library functions can only be called directly (i.e. without the use of `DELEGATECALL`) if they do not modify the state (i.e. if they are `view` or `pure` functions), because libraries are assumed to be stateless. In particular, it is not possible to destroy a library unless Solidity's type system is circumvented.

Libraries can be seen as implicit base contracts of the contracts that use them. They will not be explicitly visible in the inheritance hierarchy, but calls to library functions look just like calls to functions of explicit base contracts (`L.f()` if `L` is the name of the library). Furthermore, `internal` functions of libraries are visible in all contracts, just as if the library were a base contract. Of course, calls to internal functions use the internal calling convention, which means that all internal types can be passed and memory types will be passed by reference and not copied. To realize this in the EVM, code of internal library functions and all functions called from therein will at compile time be pulled into the calling contract, and a regular `JUMP` call will be used instead of a `DELEGATECALL`.

The following example illustrates how to use libraries (but be sure to check out *using for* for a more advanced example to implement a set).

```
pragma solidity ^0.4.22;

library Set {
  // We define a new struct datatype that will be used to
  // hold its data in the calling contract.
  struct Data { mapping(uint => bool) flags; }

  // Note that the first parameter is of type "storage
  // reference" and thus only its storage address and not
  // its contents is passed as part of the call.  This is a
  // special feature of library functions.  It is idiomatic
  // to call the first parameter `self`, if the function can
  // be seen as a method of that object.
  function insert(Data storage self, uint value)
      public
      returns (bool)
  {
      if (self.flags[value])
          return false; // already there
      self.flags[value] = true;
```

(continues on next page)

(continued from previous page)

```
        return true;
    }

    function remove(Data storage self, uint value)
        public
        returns (bool)
    {
        if (!self.flags[value])
            return false; // not there
        self.flags[value] = false;
        return true;
    }

    function contains(Data storage self, uint value)
        public
        view
        returns (bool)
    {
        return self.flags[value];
    }
}

contract C {
    Set.Data knownValues;

    function register(uint value) public {
        // The library functions can be called without a
        // specific instance of the library, since the
        // "instance" will be the current contract.
        require(Set.insert(knownValues, value));
    }
    // In this contract, we can also directly access knownValues.flags, if we want.
}
```

Of course, you do not have to follow this way to use libraries: they can also be used without defining struct data types. Functions also work without any storage reference parameters, and they can have multiple storage reference parameters and in any position.

The calls to `Set.contains`, `Set.insert` and `Set.remove` are all compiled as calls (`DELEGATECALL`) to an external contract/library. If you use libraries, take care that an actual external function call is performed. `msg.sender`, `msg.value` and `this` will retain their values in this call, though (prior to Homestead, because of the use of `CALLCODE`, `msg.sender` and `msg.value` changed, though).

The following example shows how to use memory types and internal functions in libraries in order to implement custom types without the overhead of external function calls:

```
pragma solidity ^0.4.16;

library BigInt {
    struct bigint {
        uint[] limbs;
    }

    function fromUint(uint x) internal pure returns (bigint memory r) {
        r.limbs = new uint[](1);
        r.limbs[0] = x;
    }
```

(continues on next page)

(continued from previous page)

```
    function add(bigint memory _a, bigint memory _b) internal pure returns (bigint␣
↪memory r) {
        r.limbs = new uint[](max(_a.limbs.length, _b.limbs.length));
        uint carry = 0;
        for (uint i = 0; i < r.limbs.length; ++i) {
            uint a = limb(_a, i);
            uint b = limb(_b, i);
            r.limbs[i] = a + b + carry;
            if (a + b < a || (a + b == uint(-1) && carry > 0))
                carry = 1;
            else
                carry = 0;
        }
        if (carry > 0) {
            // too bad, we have to add a limb
            uint[] memory newLimbs = new uint[](r.limbs.length + 1);
            uint i;
            for (i = 0; i < r.limbs.length; ++i)
                newLimbs[i] = r.limbs[i];
            newLimbs[i] = carry;
            r.limbs = newLimbs;
        }
    }

    function limb(bigint memory _a, uint _limb) internal pure returns (uint) {
        return _limb < _a.limbs.length ? _a.limbs[_limb] : 0;
    }

    function max(uint a, uint b) private pure returns (uint) {
        return a > b ? a : b;
    }
}

contract C {
    using BigInt for BigInt.bigint;

    function f() public pure {
        BigInt.bigint memory x = BigInt.fromUint(7);
        BigInt.bigint memory y = BigInt.fromUint(uint(-1));
        BigInt.bigint memory z = x.add(y);
    }
}
```

As the compiler cannot know where the library will be deployed at, these addresses have to be filled into the final bytecode by a linker (see *Using the Commandline Compiler* for how to use the commandline compiler for linking). If the addresses are not given as arguments to the compiler, the compiled hex code will contain placeholders of the form `__Set_______` (where `Set` is the name of the library). The address can be filled manually by replacing all those 40 symbols by the hex encoding of the address of the library contract.

Restrictions for libraries in comparison to contracts:

- No state variables
- Cannot inherit nor be inherited
- Cannot receive Ether

(These might be lifted at a later point.)

Call Protection For Libraries

As mentioned in the introduction, if a library's code is executed using a `CALL` instead of a `DELEGATECALL` or `CALLCODE`, it will revert unless a `view` or `pure` function is called.

The EVM does not provide a direct way for a contract to detect whether it was called using `CALL` or not, but a contract can use the `ADDRESS` opcode to find out "where" it is currently running. The generated code compares this address to the address used at construction time to determine the mode of calling.

More specifically, the runtime code of a library always starts with a push instruction, which is a zero of 20 bytes at compilation time. When the deploy code runs, this constant is replaced in memory by the current address and this modified code is stored in the contract. At runtime, this causes the deploy time address to be the first constant to be pushed onto the stack and the dispatcher code compares the current address against this constant for any non-view and non-pure function.

Using For

The directive `using A for B;` can be used to attach library functions (from the library `A`) to any type (`B`). These functions will receive the object they are called on as their first parameter (like the `self` variable in Python).

The effect of `using A for *;` is that the functions from the library `A` are attached to *any* type.

In both situations, *all* functions in the library are attached, even those where the type of the first parameter does not match the type of the object. The type is checked at the point the function is called and function overload resolution is performed.

The `using A for B;` directive is active only within the current contract, including within all of its functions, and has no effect outside of the contract in which it is used. The directive may only be used inside a contract, not inside any of its functions.

By including a library, its data types including library functions are available without having to add further code.

Let us rewrite the set example from the *Libraries* in this way:

```
pragma solidity ^0.4.16;

// This is the same code as before, just without comments
library Set {
  struct Data { mapping(uint => bool) flags; }

  function insert(Data storage self, uint value)
      public
      returns (bool)
  {
      if (self.flags[value])
        return false; // already there
      self.flags[value] = true;
      return true;
  }

  function remove(Data storage self, uint value)
      public
      returns (bool)
  {
      if (!self.flags[value])
          return false; // not there
      self.flags[value] = false;
      return true;
```

(continues on next page)

(continued from previous page)

```
  }

  function contains(Data storage self, uint value)
      public
      view
      returns (bool)
  {
      return self.flags[value];
  }
}

contract C {
    using Set for Set.Data; // this is the crucial change
    Set.Data knownValues;

    function register(uint value) public {
        // Here, all variables of type Set.Data have
        // corresponding member functions.
        // The following function call is identical to
        // `Set.insert(knownValues, value)`
        require(knownValues.insert(value));
    }
}
```

It is also possible to extend elementary types in that way:

```
pragma solidity ^0.4.16;

library Search {
    function indexOf(uint[] storage self, uint value)
        public
        view
        returns (uint)
    {
        for (uint i = 0; i < self.length; i++)
            if (self[i] == value) return i;
        return uint(-1);
    }
}

contract C {
    using Search for uint[];
    uint[] data;

    function append(uint value) public {
        data.push(value);
    }

    function replace(uint _old, uint _new) public {
        // This performs the library function call
        uint index = data.indexOf(_old);
        if (index == uint(-1))
            data.push(_new);
        else
            data[index] = _new;
    }
}
```

Note that all library calls are actual EVM function calls. This means that if you pass memory or value types, a copy will be performed, even of the `self` variable. The only situation where no copy will be performed is when storage reference variables are used.

7.4.7 Solidity Assembly

Solidity defines an assembly language that can also be used without Solidity. This assembly language can also be used as "inline assembly" inside Solidity source code. We start with describing how to use inline assembly and how it differs from standalone assembly and then specify assembly itself.

Inline Assembly

For more fine-grained control especially in order to enhance the language by writing libraries, it is possible to interleave Solidity statements with inline assembly in a language close to the one of the virtual machine. Due to the fact that the EVM is a stack machine, it is often hard to address the correct stack slot and provide arguments to opcodes at the correct point on the stack. Solidity's inline assembly tries to facilitate that and other issues arising when writing manual assembly by the following features:

- functional-style opcodes: `mul(1, add(2, 3))` instead of `push1 3 push1 2 add push1 1 mul`
- assembly-local variables: `let x := add(2, 3) let y := mload(0x40) x := add(x, y)`
- access to external variables: `function f(uint x) public { assembly { x := sub(x, 1) } }`
- labels: `let x := 10 repeat: x := sub(x, 1) jumpi(repeat, eq(x, 0))`
- loops: `for { let i := 0 } lt(i, x) { i := add(i, 1) } { y := mul(2, y) }`
- if statements: `if slt(x, 0) { x := sub(0, x) }`
- switch statements: `switch x case 0 { y := mul(x, 2) } default { y := 0 }`
- function calls: `function f(x) -> y { switch x case 0 { y := 1 } default { y := mul(x, f(sub(x, 1))) } }`

We now want to describe the inline assembly language in detail.

Warning: Inline assembly is a way to access the Ethereum Virtual Machine at a low level. This discards several important safety features of Solidity.

Note: TODO: Write about how scoping rules of inline assembly are a bit different and the complications that arise when for example using internal functions of libraries. Furthermore, write about the symbols defined by the compiler.

Example

The following example provides library code to access the code of another contract and load it into a `bytes` variable. This is not possible at all with "plain Solidity" and the idea is that assembly libraries will be used to enhance the language in such ways.

```
pragma solidity ^0.4.0;

library GetCode {
    function at(address _addr) public view returns (bytes memory o_code) {
        assembly {
            // retrieve the size of the code, this needs assembly
            let size := extcodesize(_addr)
            // allocate output byte array - this could also be done without assembly
            // by using o_code = new bytes(size)
            o_code := mload(0x40)
            // new "memory end" including padding
            mstore(0x40, add(o_code, and(add(add(size, 0x20), 0x1f), not(0x1f))))
            // store length in memory
            mstore(o_code, size)
            // actually retrieve the code, this needs assembly
            extcodecopy(_addr, add(o_code, 0x20), 0, size)
        }
    }
}
```

Inline assembly could also be beneficial in cases where the optimizer fails to produce efficient code. Please be aware that assembly is much more difficult to write because the compiler does not perform checks, so you should use it for complex things only if you really know what you are doing.

```
pragma solidity ^0.4.16;

library VectorSum {
    // This function is less efficient because the optimizer currently fails to
    // remove the bounds checks in array access.
    function sumSolidity(uint[] memory _data) public view returns (uint o_sum) {
        for (uint i = 0; i < _data.length; ++i)
            o_sum += _data[i];
    }

    // We know that we only access the array in bounds, so we can avoid the check.
    // 0x20 needs to be added to an array because the first slot contains the
    // array length.
    function sumAsm(uint[] memory _data) public view returns (uint o_sum) {
        for (uint i = 0; i < _data.length; ++i) {
            assembly {
                o_sum := add(o_sum, mload(add(add(_data, 0x20), mul(i, 0x20))))
            }
        }
    }

    // Same as above, but accomplish the entire code within inline assembly.
    function sumPureAsm(uint[] memory _data) public view returns (uint o_sum) {
        assembly {
           // Load the length (first 32 bytes)
           let len := mload(_data)

           // Skip over the length field.
           //
           // Keep temporary variable so it can be incremented in place.
           //
           // NOTE: incrementing _data would result in an unusable
           //       _data variable after this assembly block
```

(continues on next page)

(continued from previous page)

```
            let data := add(_data, 0x20)

            // Iterate until the bound is not met.
            for
                { let end := add(data, mul(len, 0x20)) }
                lt(data, end)
                { data := add(data, 0x20) }
            {
                o_sum := add(o_sum, mload(data))
            }
        }
    }
}
```

Syntax

Assembly parses comments, literals and identifiers exactly as Solidity, so you can use the usual `//` and `/* */` comments. Inline assembly is marked by `assembly { ... }` and inside these curly braces, the following can be used (see the later sections for more details)

- literals, i.e. `0x123`, `42` or `"abc"` (strings up to 32 characters)
- opcodes (in "instruction style"), e.g. `mload sload dup1 sstore`, for a list see below
- opcodes in functional style, e.g. `add(1, mlod(0))`
- labels, e.g. `name:`
- variable declarations, e.g. `let x := 7`, `let x := add(y, 3)` or `let x` (initial value of empty (0) is assigned)
- identifiers (labels or assembly-local variables and externals if used as inline assembly), e.g. `jump(name)`, `3 x add`
- assignments (in "instruction style"), e.g. `3 =: x`
- assignments in functional style, e.g. `x := add(y, 3)`
- blocks where local variables are scoped inside, e.g. `{ let x := 3 { let y := add(x, 1) } }`

Opcodes

This document does not want to be a full description of the Ethereum virtual machine, but the following list can be used as a reference of its opcodes.

If an opcode takes arguments (always from the top of the stack), they are given in parentheses. Note that the order of arguments can be seen to be reversed in non-functional style (explained below). Opcodes marked with `-` do not push an item onto the stack, those marked with `*` are special and all others push exactly one item onto the stack. Opcodes marked with `F`, `H`, `B` or `C` are present since Frontier, Homestead, Byzantium or Constantinople, respectively. Constantinople is still in planning and all instructions marked as such will result in an invalid instruction exception.

In the following, `mem[a...b)` signifies the bytes of memory starting at position `a` up to (excluding) position `b` and `storage[p]` signifies the storage contents at position `p`.

The opcodes `pushi` and `jumpdest` cannot be used directly.

In the grammar, opcodes are represented as pre-defined identifiers.

Instruction			Explanation
stop	-	F	stop execution, identical to return(0,0)
add(x, y)		F	x + y
sub(x, y)		F	x - y
mul(x, y)		F	x * y
div(x, y)		F	x / y
sdiv(x, y)		F	x / y, for signed numbers in two's complement
mod(x, y)		F	x % y
smod(x, y)		F	x % y, for signed numbers in two's complement
exp(x, y)		F	x to the power of y
not(x)		F	~x, every bit of x is negated
lt(x, y)		F	1 if x < y, 0 otherwise
gt(x, y)		F	1 if x > y, 0 otherwise
slt(x, y)		F	1 if x < y, 0 otherwise, for signed numbers in two's complement
sgt(x, y)		F	1 if x > y, 0 otherwise, for signed numbers in two's complement
eq(x, y)		F	1 if x == y, 0 otherwise
iszero(x)		F	1 if x == 0, 0 otherwise
and(x, y)		F	bitwise and of x and y
or(x, y)		F	bitwise or of x and y
xor(x, y)		F	bitwise xor of x and y
byte(n, x)		F	nth byte of x, where the most significant byte is the 0th byte
shl(x, y)		C	logical shift left y by x bits
shr(x, y)		C	logical shift right y by x bits
sar(x, y)		C	arithmetic shift right y by x bits
addmod(x, y, m)		F	(x + y) % m with arbitrary precision arithmetic
mulmod(x, y, m)		F	(x * y) % m with arbitrary precision arithmetic
signextend(i, x)		F	sign extend from (i*8+7)th bit counting from least significant
keccak256(p, n)		F	keccak(mem[p...(p+n)))
jump(label)	-	F	jump to label / code position
jumpi(label, cond)	-	F	jump to label if cond is nonzero
pc		F	current position in code
pop(x)	-	F	remove the element pushed by x
dup1 ... dup16		F	copy nth stack slot to the top (counting from top)
swap1 ... swap16	*	F	swap topmost and nth stack slot below it
mload(p)		F	mem[p...(p+32))
mstore(p, v)	-	F	mem[p...(p+32)) := v
mstore8(p, v)	-	F	mem[p] := v & 0xff (only modifies a single byte)
sload(p)		F	storage[p]
sstore(p, v)	-	F	storage[p] := v
msize		F	size of memory, i.e. largest accessed memory index
gas		F	gas still available to execution
address		F	address of the current contract / execution context
balance(a)		F	wei balance at address a
caller		F	call sender (excluding `delegatecall`)
callvalue		F	wei sent together with the current call
calldataload(p)		F	call data starting from position p (32 bytes)
calldatasize		F	size of call data in bytes
calldatacopy(t, f, s)	-	F	copy s bytes from calldata at position f to mem at position t
codesize		F	size of the code of the current contract / execution context

Table 1 – continued from previous p

Instruction			Explanation
codecopy(t, f, s)	-	F	copy s bytes from code at position f to mem at position t
extcodesize(a)		F	size of the code at address a
extcodecopy(a, t, f, s)	-	F	like codecopy(t, f, s) but take code at address a
returndatasize		B	size of the last returndata
returndatacopy(t, f, s)	-	B	copy s bytes from returndata at position f to mem at position t
create(v, p, s)		F	create new contract with code mem[p...(p+s)) and send v wei and return the new
create2(v, n, p, s)		C	create new contract with code mem[p...(p+s)) at address keccak256(<address> .
call(g, a, v, in, insize, out, outsize)		F	call contract at address a with input mem[in...(in+insize)) providing g gas and v
callcode(g, a, v, in, insize, out, outsize)		F	identical to `call` but only use the code from a and stay in the context of the curr
delegatecall(g, a, in, insize, out, outsize)		H	identical to `callcode` but also keep `caller` and `callvalue`
staticcall(g, a, in, insize, out, outsize)		B	identical to `call(g, a, 0, in, insize, out, outsize)` but do not
return(p, s)	-	F	end execution, return data mem[p...(p+s))
revert(p, s)	-	B	end execution, revert state changes, return data mem[p...(p+s))
selfdestruct(a)	-	F	end execution, destroy current contract and send funds to a
invalid	-	F	end execution with invalid instruction
log0(p, s)	-	F	log without topics and data mem[p...(p+s))
log1(p, s, t1)	-	F	log with topic t1 and data mem[p...(p+s))
log2(p, s, t1, t2)	-	F	log with topics t1, t2 and data mem[p...(p+s))
log3(p, s, t1, t2, t3)	-	F	log with topics t1, t2, t3 and data mem[p...(p+s))
log4(p, s, t1, t2, t3, t4)	-	F	log with topics t1, t2, t3, t4 and data mem[p...(p+s))
origin		F	transaction sender
gasprice		F	gas price of the transaction
blockhash(b)		F	hash of block nr b - only for last 256 blocks excluding current
coinbase		F	current mining beneficiary
timestamp		F	timestamp of the current block in seconds since the epoch
number		F	current block number
difficulty		F	difficulty of the current block
gaslimit		F	block gas limit of the current block

Literals

You can use integer constants by typing them in decimal or hexadecimal notation and an appropriate `PUSHi` instruction will automatically be generated. The following creates code to add 2 and 3 resulting in 5 and then computes the bitwise and with the string "abc". Strings are stored left-aligned and cannot be longer than 32 bytes.

```
assembly { 2 3 add "abc" and }
```

Functional Style

You can type opcode after opcode in the same way they will end up in bytecode. For example adding `3` to the contents in memory at position `0x80` would be

```
3 0x80 mload add 0x80 mstore
```

As it is often hard to see what the actual arguments for certain opcodes are, Solidity inline assembly also provides a "functional style" notation where the same code would be written as follows

```
mstore(0x80, add(mload(0x80), 3))
```

Functional style expressions cannot use instructional style internally, i.e. `1 2 mstore(0x80, add)` is not valid assembly, it has to be written as `mstore(0x80, add(2, 1))`. For opcodes that do not take arguments, the parentheses can be omitted.

Note that the order of arguments is reversed in functional-style as opposed to the instruction-style way. If you use functional-style, the first argument will end up on the stack top.

Access to External Variables and Functions

Solidity variables and other identifiers can be accessed by simply using their name. For memory variables, this will push the address and not the value onto the stack. Storage variables are different: Values in storage might not occupy a full storage slot, so their "address" is composed of a slot and a byte-offset inside that slot. To retrieve the slot pointed to by the variable `x`, you used `x_slot` and to retrieve the byte-offset you used `x_offset`.

In assignments (see below), we can even use local Solidity variables to assign to.

Functions external to inline assembly can also be accessed: The assembly will push their entry label (with virtual function resolution applied). The calling semantics in solidity are:

- the caller pushes `return label`, `arg1`, `arg2`, ..., `argn`
- the call returns with `ret1`, `ret2`, ..., `retm`

This feature is still a bit cumbersome to use, because the stack offset essentially changes during the call, and thus references to local variables will be wrong.

```
pragma solidity ^0.4.11;

contract C {
    uint b;
    function f(uint x) public returns (uint r) {
        assembly {
            r := mul(x, sload(b_slot)) // ignore the offset, we know it is zero
        }
    }
}
```

Note: If you access variables of a type that spans less than 256 bits (for example `uint64`, `address`, `bytes16` or `byte`), you cannot make any assumptions about bits not part of the encoding of the type. Especially, do not assume them to be zero. To be safe, always clear the data properly before you use it in a context where this is important: `uint32 x = f(); assembly { x := and(x, 0xffffffff) /* now use x */ }` To clean signed types, you can use the `signextend` opcode.

Labels

Note: Labels are deprecated. Please use functions, loops, if or switch statements instead.

Another problem in EVM assembly is that `jump` and `jumpi` use absolute addresses which can change easily. Solidity inline assembly provides labels to make the use of jumps easier. Note that labels are a low-level feature and it is possible to write efficient assembly without labels, just using assembly functions, loops, if and switch instructions (see below). The following code computes an element in the Fibonacci series.

```
{
    let n := calldataload(4)
    let a := 1
    let b := a
loop:
    jumpi(loopend, eq(n, 0))
    a add swap1
    n := sub(n, 1)
    jump(loop)
loopend:
    mstore(0, a)
    return(0, 0x20)
}
```

Please note that automatically accessing stack variables can only work if the assembler knows the current stack height. This fails to work if the jump source and target have different stack heights. It is still fine to use such jumps, but you should just not access any stack variables (even assembly variables) in that case.

Furthermore, the stack height analyser goes through the code opcode by opcode (and not according to control flow), so in the following case, the assembler will have a wrong impression about the stack height at label `two`:

```
{
    let x := 8
    jump(two)
    one:
        // Here the stack height is 2 (because we pushed x and 7),
        // but the assembler thinks it is 1 because it reads
        // from top to bottom.
        // Accessing the stack variable x here will lead to errors.
        x := 9
        jump(three)
    two:
        7 // push something onto the stack
        jump(one)
    three:
}
```

Declaring Assembly-Local Variables

You can use the `let` keyword to declare variables that are only visible in inline assembly and actually only in the current `{...}`-block. What happens is that the `let` instruction will create a new stack slot that is reserved for the variable and automatically removed again when the end of the block is reached. You need to provide an initial value for the variable which can be just `0`, but it can also be a complex functional-style expression.

```
pragma solidity ^0.4.16;

contract C {
    function f(uint x) public view returns (uint b) {
        assembly {
            let v := add(x, 1)
            mstore(0x80, v)
            {
                let y := add(sload(v), 1)
                b := y
            } // y is "deallocated" here
```

(continues on next page)

(continued from previous page)

```
            b := add(b, v)
        } // v is "deallocated" here
    }
}
```

Assignments

Assignments are possible to assembly-local variables and to function-local variables. Take care that when you assign to variables that point to memory or storage, you will only change the pointer and not the data.

There are two kinds of assignments: functional-style and instruction-style. For functional-style assignments (`variable := value`), you need to provide a value in a functional-style expression that results in exactly one stack value and for instruction-style (`=: variable`), the value is just taken from the stack top. For both ways, the colon points to the name of the variable. The assignment is performed by replacing the variable's value on the stack by the new value.

```
{
    let v := 0 // functional-style assignment as part of variable declaration
    let g := add(v, 2)
    sload(10)
    =: v // instruction style assignment, puts the result of sload(10) into v
}
```

Note: Instruction-style assignment is deprecated.

If

The if statement can be used for conditionally executing code. There is no "else" part, consider using "switch" (see below) if you need multiple alternatives.

```
{
    if eq(value, 0) { revert(0, 0) }
}
```

The curly braces for the body are required.

Switch

You can use a switch statement as a very basic version of "if/else". It takes the value of an expression and compares it to several constants. The branch corresponding to the matching constant is taken. Contrary to the error-prone behaviour of some programming languages, control flow does not continue from one case to the next. There can be a fallback or default case called `default`.

```
{
    let x := 0
    switch calldataload(4)
    case 0 {
        x := calldataload(0x24)
    }
```

(continues on next page)

(continued from previous page)

```
    default {
        x := calldataload(0x44)
    }
    sstore(0, div(x, 2))
}
```

The list of cases does not require curly braces, but the body of a case does require them.

Loops

Assembly supports a simple for-style loop. For-style loops have a header containing an initializing part, a condition and a post-iteration part. The condition has to be a functional-style expression, while the other two are blocks. If the initializing part declares any variables, the scope of these variables is extended into the body (including the condition and the post-iteration part).

The following example computes the sum of an area in memory.

```
{
    let x := 0
    for { let i := 0 } lt(i, 0x100) { i := add(i, 0x20) } {
        x := add(x, mload(i))
    }
}
```

For loops can also be written so that they behave like while loops: Simply leave the initialization and post-iteration parts empty.

```
{
    let x := 0
    let i := 0
    for { } lt(i, 0x100) { } {      // while(i < 0x100)
        x := add(x, mload(i))
        i := add(i, 0x20)
    }
}
```

Functions

Assembly allows the definition of low-level functions. These take their arguments (and a return PC) from the stack and also put the results onto the stack. Calling a function looks the same way as executing a functional-style opcode.

Functions can be defined anywhere and are visible in the block they are declared in. Inside a function, you cannot access local variables defined outside of that function. There is no explicit `return` statement.

If you call a function that returns multiple values, you have to assign them to a tuple using `a, b := f(x)` or `let a, b := f(x)`.

The following example implements the power function by square-and-multiply.

```
{
    function power(base, exponent) -> result {
        switch exponent
        case 0 { result := 1 }
        case 1 { result := base }
```

(continues on next page)

(continued from previous page)

```
        default {
            result := power(mul(base, base), div(exponent, 2))
            switch mod(exponent, 2)
                case 1 { result := mul(base, result) }
        }
    }
}
```

Things to Avoid

Inline assembly might have a quite high-level look, but it actually is extremely low-level. Function calls, loops, ifs and switches are converted by simple rewriting rules and after that, the only thing the assembler does for you is re-arranging functional-style opcodes, managing jump labels, counting stack height for variable access and removing stack slots for assembly-local variables when the end of their block is reached. Especially for those two last cases, it is important to know that the assembler only counts stack height from top to bottom, not necessarily following control flow. Furthermore, operations like swap will only swap the contents of the stack but not the location of variables.

Conventions in Solidity

In contrast to EVM assembly, Solidity knows types which are narrower than 256 bits, e.g. `uint24`. In order to make them more efficient, most arithmetic operations just treat them as 256-bit numbers and the higher-order bits are only cleaned at the point where it is necessary, i.e. just shortly before they are written to memory or before comparisons are performed. This means that if you access such a variable from within inline assembly, you might have to manually clean the higher order bits first.

Solidity manages memory in a very simple way: There is a "free memory pointer" at position `0x40` in memory. If you want to allocate memory, just use the memory from that point on and update the pointer accordingly.

The first 64 bytes of memory can be used as "scratch space" for short-term allocation. The 32 bytes after the free memory pointer (i.e. starting at `0x60`) is meant to be zero permanently and is used as the initial value for empty dynamic memory arrays.

Elements in memory arrays in Solidity always occupy multiples of 32 bytes (yes, this is even true for `byte[]`, but not for `bytes` and `string`). Multi-dimensional memory arrays are pointers to memory arrays. The length of a dynamic array is stored at the first slot of the array and then only the array elements follow.

Warning: Statically-sized memory arrays do not have a length field, but it will be added soon to allow better convertibility between statically- and dynamically-sized arrays, so please do not rely on that.

Standalone Assembly

The assembly language described as inline assembly above can also be used standalone and in fact, the plan is to use it as an intermediate language for the Solidity compiler. In this form, it tries to achieve several goals:

1. Programs written in it should be readable, even if the code is generated by a compiler from Solidity.
2. The translation from assembly to bytecode should contain as few "surprises" as possible.
3. Control flow should be easy to detect to help in formal verification and optimization.

In order to achieve the first and last goal, assembly provides high-level constructs like `for` loops, `if` and `switch` statements and function calls. It should be possible to write assembly programs that do not make use of explicit

`SWAP`, `DUP`, `JUMP` and `JUMPI` statements, because the first two obfuscate the data flow and the last two obfuscate control flow. Furthermore, functional statements of the form `mul(add(x, y), 7)` are preferred over pure opcode statements like `7 y x add mul` because in the first form, it is much easier to see which operand is used for which opcode.

The second goal is achieved by compiling the higher level constructs to bytecode in a very regular way. The only non-local operation performed by the assembler is name lookup of user-defined identifiers (functions, variables, ...), which follow very simple and regular scoping rules and cleanup of local variables from the stack.

Scoping: An identifier that is declared (label, variable, function, assembly) is only visible in the block where it was declared (including nested blocks inside the current block). It is not legal to access local variables across function borders, even if they would be in scope. Shadowing is not allowed. Local variables cannot be accessed before they were declared, but labels, functions and assemblies can. Assemblies are special blocks that are used for e.g. returning runtime code or creating contracts. No identifier from an outer assembly is visible in a sub-assembly.

If control flow passes over the end of a block, pop instructions are inserted that match the number of local variables declared in that block. Whenever a local variable is referenced, the code generator needs to know its current relative position in the stack and thus it needs to keep track of the current so-called stack height. Since all local variables are removed at the end of a block, the stack height before and after the block should be the same. If this is not the case, a warning is issued.

Using `switch`, `for` and functions, it should be possible to write complex code without using `jump` or `jumpi` manually. This makes it much easier to analyze the control flow, which allows for improved formal verification and optimization.

Furthermore, if manual jumps are allowed, computing the stack height is rather complicated. The position of all local variables on the stack needs to be known, otherwise neither references to local variables nor removing local variables automatically from the stack at the end of a block will work properly.

Example:

We will follow an example compilation from Solidity to assembly. We consider the runtime bytecode of the following Solidity program:

```
pragma solidity ^0.4.16;

contract C {
  function f(uint x) public pure returns (uint y) {
    y = 1;
    for (uint i = 0; i < x; i++)
      y = 2 * y;
  }
}
```

The following assembly will be generated:

```
{
  mstore(0x40, 0x60) // store the "free memory pointer"
  // function dispatcher
  switch div(calldataload(0), exp(2, 226))
  case 0xb3de648b {
    let r := f(calldataload(4))
    let ret := $allocate(0x20)
    mstore(ret, r)
    return(ret, 0x20)
  }
  default { revert(0, 0) }
  // memory allocator
  function $allocate(size) -> pos {
```

(continues on next page)

(continued from previous page)

```
    pos := mload(0x40)
    mstore(0x40, add(pos, size))
  }
  // the contract function
  function f(x) -> y {
    y := 1
    for { let i := 0 } lt(i, x) { i := add(i, 1) } {
      y := mul(2, y)
    }
  }
}
```

Assembly Grammar

The tasks of the parser are the following:

- Turn the byte stream into a token stream, discarding C++-style comments (a special comment exists for source references, but we will not explain it here).
- Turn the token stream into an AST according to the grammar below
- Register identifiers with the block they are defined in (annotation to the AST node) and note from which point on, variables can be accessed.

The assembly lexer follows the one defined by Solidity itself.

Whitespace is used to delimit tokens and it consists of the characters Space, Tab and Linefeed. Comments are regular JavaScript/C++ comments and are interpreted in the same way as Whitespace.

Grammar:

```
AssemblyBlock = '{' AssemblyItem* '}'
AssemblyItem =
    Identifier |
    AssemblyBlock |
    AssemblyExpression |
    AssemblyLocalDefinition |
    AssemblyAssignment |
    AssemblyStackAssignment |
    LabelDefinition |
    AssemblyIf |
    AssemblySwitch |
    AssemblyFunctionDefinition |
    AssemblyFor |
    'break' |
    'continue' |
    SubAssembly
AssemblyExpression = AssemblyCall | Identifier | AssemblyLiteral
AssemblyLiteral = NumberLiteral | StringLiteral | HexLiteral
Identifier = [a-zA-Z_$] [a-zA-Z_0-9]*
AssemblyCall = Identifier '(' ( AssemblyExpression ( ',' AssemblyExpression )* )? ')'
AssemblyLocalDefinition = 'let' IdentifierOrList ( ':=' AssemblyExpression )?
AssemblyAssignment = IdentifierOrList ':=' AssemblyExpression
IdentifierOrList = Identifier | '(' IdentifierList ')'
IdentifierList = Identifier ( ',' Identifier)*
AssemblyStackAssignment = '=:' Identifier
LabelDefinition = Identifier ':'
```

(continues on next page)

(continued from previous page)

```
AssemblyIf = 'if' AssemblyExpression AssemblyBlock
AssemblySwitch = 'switch' AssemblyExpression AssemblyCase*
    ( 'default' AssemblyBlock )?
AssemblyCase = 'case' AssemblyExpression AssemblyBlock
AssemblyFunctionDefinition = 'function' Identifier '(' IdentifierList? ')'
    ( '->' '(' IdentifierList ')' )? AssemblyBlock
AssemblyFor = 'for' ( AssemblyBlock | AssemblyExpression )
    AssemblyExpression ( AssemblyBlock | AssemblyExpression ) AssemblyBlock
SubAssembly = 'assembly' Identifier AssemblyBlock
NumberLiteral = HexNumber | DecimalNumber
HexLiteral = 'hex' ('"' ([0-9a-fA-F]{2})* '"' | '\'' ([0-9a-fA-F]{2})* '\'')
StringLiteral = '"' ([^"\r\n\\] | '\\' .)* '"'
HexNumber = '0x' [0-9a-fA-F]+
DecimalNumber = [0-9]+
```

7.4.8 Miscellaneous

Layout of State Variables in Storage

Statically-sized variables (everything except mapping and dynamically-sized array types) are laid out contiguously in storage starting from position `0`. Multiple items that need less than 32 bytes are packed into a single storage slot if possible, according to the following rules:

- The first item in a storage slot is stored lower-order aligned.
- Elementary types use only that many bytes that are necessary to store them.
- If an elementary type does not fit the remaining part of a storage slot, it is moved to the next storage slot.
- Structs and array data always start a new slot and occupy whole slots (but items inside a struct or array are packed tightly according to these rules).

Warning: When using elements that are smaller than 32 bytes, your contract's gas usage may be higher. This is because the EVM operates on 32 bytes at a time. Therefore, if the element is smaller than that, the EVM must use more operations in order to reduce the size of the element from 32 bytes to the desired size.

It is only beneficial to use reduced-size arguments if you are dealing with storage values because the compiler will pack multiple elements into one storage slot, and thus, combine multiple reads or writes into a single operation. When dealing with function arguments or memory values, there is no inherent benefit because the compiler does not pack these values.

Finally, in order to allow the EVM to optimize for this, ensure that you try to order your storage variables and `struct` members such that they can be packed tightly. For example, declaring your storage variables in the order of `uint128, uint128, uint256` instead of `uint128, uint256, uint128`, as the former will only take up two slots of storage whereas the latter will take up three.

The elements of structs and arrays are stored after each other, just as if they were given explicitly.

Due to their unpredictable size, mapping and dynamically-sized array types use a Keccak-256 hash computation to find the starting position of the value or the array data. These starting positions are always full stack slots.

The mapping or the dynamic array itself occupies an (unfilled) slot in storage at some position `p` according to the above rule (or by recursively applying this rule for mappings to mappings or arrays of arrays). For a dynamic array, this slot stores the number of elements in the array (byte arrays and strings are an exception here, see below). For a mapping, the slot is unused (but it is needed so that two equal mappings after each other will use a different hash

distribution). Array data is located at `keccak256(p)` and the value corresponding to a mapping key `k` is located at `keccak256(k . p)` where `.` is concatenation. If the value is again a non-elementary type, the positions are found by adding an offset of `keccak256(k . p)`.

`bytes` and `string` store their data in the same slot where also the length is stored if they are short. In particular: If the data is at most `31` bytes long, it is stored in the higher-order bytes (left aligned) and the lowest-order byte stores `length * 2`. If it is longer, the main slot stores `length * 2 + 1` and the data is stored as usual in `keccak256(slot)`.

So for the following contract snippet:

```
pragma solidity ^0.4.0;

contract C {
  struct s { uint a; uint b; }
  uint x;
  mapping(uint => mapping(uint => s)) data;
}
```

The position of `data[4][9].b` is at `keccak256(uint256(9) . keccak256(uint256(4) . uint256(1))) + 1`.

Layout in Memory

Solidity reserves four 32 byte slots:

- `0x00` - `0x3f`: scratch space for hashing methods
- `0x40` - `0x5f`: currently allocated memory size (aka. free memory pointer)
- `0x60` - `0x7f`: zero slot

Scratch space can be used between statements (ie. within inline assembly). The zero slot is used as initial value for dynamic memory arrays and should never be written to (the free memory pointer points to `0x80` initially).

Solidity always places new objects at the free memory pointer and memory is never freed (this might change in the future).

Warning: There are some operations in Solidity that need a temporary memory area larger than 64 bytes and therefore will not fit into the scratch space. They will be placed where the free memory points to, but given their short lifecycle, the pointer is not updated. The memory may or may not be zeroed out. Because of this, one shouldn't expect the free memory to be zeroed out.

While it may seem like a good idea to use `msize` to arrive at a definitely zeroed out memory area, using such a pointer non-temporarily without updating the free memory pointer can have adverse results.

Layout of Call Data

When a Solidity contract is deployed and when it is called from an account, the input data is assumed to be in the format in *the ABI specification*. The ABI specification requires arguments to be padded to multiples of 32 bytes. The internal function calls use a different convention.

Internals - Cleaning Up Variables

When a value is shorter than 256-bit, in some cases the remaining bits must be cleaned. The Solidity compiler is designed to clean such remaining bits before any operations that might be adversely affected by the potential garbage in the remaining bits. For example, before writing a value to the memory, the remaining bits need to be cleared because the memory contents can be used for computing hashes or sent as the data of a message call. Similarly, before storing a value in the storage, the remaining bits need to be cleaned because otherwise the garbled value can be observed.

On the other hand, we do not clean the bits if the immediately following operation is not affected. For instance, since any non-zero value is considered `true` by `JUMPI` instruction, we do not clean the boolean values before they are used as the condition for `JUMPI`.

In addition to the design principle above, the Solidity compiler cleans input data when it is loaded onto the stack.

Different types have different rules for cleaning up invalid values:

Type	Valid Values	Invalid Values Mean
enum of n members	0 until n - 1	exception
bool	0 or 1	1
signed integers	sign-extended word	currently silently wraps; in the future exceptions will be thrown
unsigned integers	higher bits zeroed	currently silently wraps; in the future exceptions will be thrown

Internals - The Optimizer

The Solidity optimizer operates on assembly, so it can be and also is used by other languages. It splits the sequence of instructions into basic blocks at `JUMPs` and `JUMPDESTs`. Inside these blocks, the instructions are analysed and every modification to the stack, to memory or storage is recorded as an expression which consists of an instruction and a list of arguments which are essentially pointers to other expressions. The main idea is now to find expressions that are always equal (on every input) and combine them into an expression class. The optimizer first tries to find each new expression in a list of already known expressions. If this does not work, the expression is simplified according to rules like `constant + constant = sum_of_constants` or `X * 1 = X`. Since this is done recursively, we can also apply the latter rule if the second factor is a more complex expression where we know that it will always evaluate to one. Modifications to storage and memory locations have to erase knowledge about storage and memory locations which are not known to be different: If we first write to location x and then to location y and both are input variables, the second could overwrite the first, so we actually do not know what is stored at x after we wrote to y. On the other hand, if a simplification of the expression x - y evaluates to a non-zero constant, we know that we can keep our knowledge about what is stored at x.

At the end of this process, we know which expressions have to be on the stack in the end and have a list of modifications to memory and storage. This information is stored together with the basic blocks and is used to link them. Furthermore, knowledge about the stack, storage and memory configuration is forwarded to the next block(s). If we know the targets of all `JUMP` and `JUMPI` instructions, we can build a complete control flow graph of the program. If there is only one target we do not know (this can happen as in principle, jump targets can be computed from inputs), we have to erase all knowledge about the input state of a block as it can be the target of the unknown `JUMP`. If a `JUMPI` is found whose condition evaluates to a constant, it is transformed to an unconditional jump.

As the last step, the code in each block is completely re-generated. A dependency graph is created from the expressions on the stack at the end of the block and every operation that is not part of this graph is essentially dropped. Now code is generated that applies the modifications to memory and storage in the order they were made in the original code (dropping modifications which were found not to be needed) and finally, generates all values that are required to be on the stack in the correct place.

These steps are applied to each basic block and the newly generated code is used as replacement if it is smaller. If a basic block is split at a `JUMPI` and during the analysis, the condition evaluates to a constant, the `JUMPI` is replaced depending on the value of the constant, and thus code like

```
uint x = 7;
data[7] = 9;
if (data[x] != x + 2)
  return 2;
else
  return 1;
```

is simplified to code which can also be compiled from

```
data[7] = 9;
return 1;
```

even though the instructions contained a jump in the beginning.

Source Mappings

As part of the AST output, the compiler provides the range of the source code that is represented by the respective node in the AST. This can be used for various purposes ranging from static analysis tools that report errors based on the AST and debugging tools that highlight local variables and their uses.

Furthermore, the compiler can also generate a mapping from the bytecode to the range in the source code that generated the instruction. This is again important for static analysis tools that operate on bytecode level and for displaying the current position in the source code inside a debugger or for breakpoint handling.

Both kinds of source mappings use integer identifiers to refer to source files. These are regular array indices into a list of source files usually called `"sourceList"`, which is part of the combined-json and the output of the json / npm compiler.

Note: In the case of instructions that are not associated with any particular source file, the source mapping assigns an integer identifier of `-1`. This may happen for bytecode sections stemming from compiler-generated inline assembly statements.

The source mappings inside the AST use the following notation:

`s:l:f`

Where `s` is the byte-offset to the start of the range in the source file, `l` is the length of the source range in bytes and `f` is the source index mentioned above.

The encoding in the source mapping for the bytecode is more complicated: It is a list of `s:l:f:j` separated by `;`. Each of these elements corresponds to an instruction, i.e. you cannot use the byte offset but have to use the instruction offset (push instructions are longer than a single byte). The fields `s`, `l` and `f` are as above and `j` can be either `i`, `o` or `-` signifying whether a jump instruction goes into a function, returns from a function or is a regular jump as part of e.g. a loop.

In order to compress these source mappings especially for bytecode, the following rules are used:

- If a field is empty, the value of the preceding element is used.
- If a `:` is missing, all following fields are considered empty.

This means the following source mappings represent the same information:

`1:2:1;1:9:1;2:1:2;2:1:2;2:1:2`

`1:2:1;:9;2:1:2;;`

Tips and Tricks

- Use `delete` on arrays to delete all its elements.
- Use shorter types for struct elements and sort them such that short types are grouped together. This can lower the gas costs as multiple `SSTORE` operations might be combined into a single (`SSTORE` costs 5000 or 20000 gas, so this is what you want to optimise). Use the gas price estimator (with optimiser enabled) to check!
- Make your state variables public - the compiler will create *getters* for you automatically.
- If you end up checking conditions on input or state a lot at the beginning of your functions, try using *Function Modifiers*.
- If your contract has a function called `send` but you want to use the built-in send-function, use `address(contractVariable).send(amount)`.
- Initialize storage structs with a single assignment: `x = MyStruct({a: 1, b: 2});`

Note: If the storage struct has tightly packed properties, initialize it with separate assignments: `x.a = 1; x.b = 2;`. In this way it will be easier for the optimizer to update storage in one go, thus making assignment cheaper.

Cheatsheet

Order of Precedence of Operators

The following is the order of precedence for operators, listed in order of evaluation.

Precedence	Description	Operator
1	Postfix increment and decrement	`++`, `--`
	New expression	`new <typename>`
	Array subscripting	`<array>[<index>]`
	Member access	`<object>.<member>`
	Function-like call	`<func>(<args...>)`
	Parentheses	`(<statement>)`
2	Prefix increment and decrement	`++`, `--`
	Unary plus and minus	`+`, `-`
	Unary operations	`delete`
	Logical NOT	`!`
	Bitwise NOT	`~`
3	Exponentiation	`**`
4	Multiplication, division and modulo	`*`, `/`, `%`
5	Addition and subtraction	`+`, `-`
6	Bitwise shift operators	`<<`, `>>`
7	Bitwise AND	`&`
8	Bitwise XOR	`^`
9	Bitwise OR	`|`
10	Inequality operators	`<`, `>`, `<=`, `>=`
11	Equality operators	`==`, `!=`
12	Logical AND	`&&`
13	Logical OR	`||`
14	Ternary operator	`<conditional> ? <if-true> : <if-false>`
15	Assignment operators	`=`, `|=`, `^=`, `&=`, `<<=`, `>>=`, `+=`, `-=`, `*=`, `/=`, `%=`
16	Comma operator	`,`

Global Variables

- `abi.encode(...) returns (bytes)`: *ABI*-encodes the given arguments
- `abi.encodePacked(...) returns (bytes)`: Performs *packed encoding* of the given arguments
- **`abi.encodeWithSelector(bytes4 selector, ...) returns (bytes)`: *ABI*-encodes the given arguments** starting from the second and prepends the given four-byte selector
- `abi.encodeWithSignature(string signature, ...) returns (bytes)`: Equivalent to `abi.encodeWithSelector(bytes4(keccak256(bytes(signature)), ...)` `
- `block.blockhash(uint blockNumber) returns (bytes32)`: hash of the given block - only works for 256 most recent, excluding current, blocks - deprecated in version 0.4.22 and replaced by `blockhash(uint blockNumber)`.
- `block.coinbase` (`address`): current block miner's address
- `block.difficulty` (`uint`): current block difficulty
- `block.gaslimit` (`uint`): current block gaslimit
- `block.number` (`uint`): current block number
- `block.timestamp` (`uint`): current block timestamp
- `gasleft() returns (uint256)`: remaining gas
- `msg.data` (`bytes`): complete calldata
- `msg.gas` (`uint`): remaining gas - deprecated in version 0.4.21 and to be replaced by `gasleft()`
- `msg.sender` (`address`): sender of the message (current call)
- `msg.value` (`uint`): number of wei sent with the message
- `now` (`uint`): current block timestamp (alias for `block.timestamp`)
- `tx.gasprice` (`uint`): gas price of the transaction
- `tx.origin` (`address`): sender of the transaction (full call chain)
- `assert(bool condition)`: abort execution and revert state changes if condition is `false` (use for internal error)
- `require(bool condition)`: abort execution and revert state changes if condition is `false` (use for malformed input or error in external component)
- `require(bool condition, string message)`: abort execution and revert state changes if condition is `false` (use for malformed input or error in external component). Also provide error message.
- `revert()`: abort execution and revert state changes
- `revert(string message)`: abort execution and revert state changes providing an explanatory string
- `blockhash(uint blockNumber) returns (bytes32)`: hash of the given block - only works for 256 most recent blocks
- `keccak256(bytes memory) returns (bytes32)`: compute the Ethereum-SHA-3 (Keccak-256) hash of the input
- `sha3(bytes memory) returns (bytes32)`: an alias to `keccak256`
- `sha256(bytes memory) returns (bytes32)`: compute the SHA-256 hash of the input
- `ripemd160(bytes memory) returns (bytes20)`: compute the RIPEMD-160 hash of the input

- `ecrecover(bytes32 hash, uint8 v, bytes32 r, bytes32 s) returns (address)`: recover address associated with the public key from elliptic curve signature, return zero on error
- `addmod(uint x, uint y, uint k) returns (uint)`: compute `(x + y) % k` where the addition is performed with arbitrary precision and does not wrap around at `2**256`. Assert that `k != 0` starting from version 0.5.0.
- `mulmod(uint x, uint y, uint k) returns (uint)`: compute `(x * y) % k` where the multiplication is performed with arbitrary precision and does not wrap around at `2**256`. Assert that `k != 0` starting from version 0.5.0.
- `this` (current contract's type): the current contract, explicitly convertible to `address`
- `super`: the contract one level higher in the inheritance hierarchy
- `selfdestruct(address recipient)`: destroy the current contract, sending its funds to the given address
- `suicide(address recipient)`: a deprecated alias to `selfdestruct`
- `<address>.balance` (`uint256`): balance of the *Address* in Wei
- `<address>.send(uint256 amount) returns (bool)`: send given amount of Wei to *Address*, returns `false` on failure
- `<address>.transfer(uint256 amount)`: send given amount of Wei to *Address*, throws on failure

Note: Do not rely on `block.timestamp`, `now` and `blockhash` as a source of randomness, unless you know what you are doing.

Both the timestamp and the block hash can be influenced by miners to some degree. Bad actors in the mining community can for example run a casino payout function on a chosen hash and just retry a different hash if they did not receive any money.

The current block timestamp must be strictly larger than the timestamp of the last block, but the only guarantee is that it will be somewhere between the timestamps of two consecutive blocks in the canonical chain.

Note: The block hashes are not available for all blocks for scalability reasons. You can only access the hashes of the most recent 256 blocks, all other values will be zero.

Function Visibility Specifiers

```
function myFunction() <visibility specifier> returns (bool) {
    return true;
}
```

- `public`: visible externally and internally (creates a *getter function* for storage/state variables)
- `private`: only visible in the current contract
- `external`: only visible externally (only for functions) - i.e. can only be message-called (via `this.func`)
- `internal`: only visible internally

Modifiers

- `pure` for functions: Disallows modification or access of state.
- `view` for functions: Disallows modification of state.
- `payable` for functions: Allows them to receive Ether together with a call.
- `constant` for state variables: Disallows assignment (except initialisation), does not occupy storage slot.
- `anonymous` for events: Does not store event signature as topic.
- `indexed` for event parameters: Stores the parameter as topic.

Reserved Keywords

These keywords are reserved in Solidity. They might become part of the syntax in the future:

`abstract`, `after`, `alias`, `apply`, `auto`, `case`, `catch`, `copyof`, `default`, `define`, `final`, `immutable`, `implements`, `in`, `inline`, `let`, `macro`, `match`, `mutable`, `null`, `of`, `override`, `partial`, `promise`, `reference`, `relocatable`, `sealed`, `sizeof`, `static`, `supports`, `switch`, `try`, `type`, `typedef`, `typeof`, `unchecked`.

Language Grammar

```
SourceUnit = (PragmaDirective | ImportDirective | ContractDefinition)*

// Pragma actually parses anything up to the trailing ';' to be fully forward-
↪compatible.
PragmaDirective = 'pragma' Identifier ([^;]+) ';'

ImportDirective = 'import' StringLiteral ('as' Identifier)? ';'
        | 'import' ('*' | Identifier) ('as' Identifier)? 'from' StringLiteral ';'
        | 'import' '{' Identifier ('as' Identifier)? ( ',' Identifier ('as'␣
↪Identifier)? )* '}' 'from' StringLiteral ';'

ContractDefinition = ( 'contract' | 'library' | 'interface' ) Identifier
                     ( 'is' InheritanceSpecifier (',' InheritanceSpecifier )* )?
                     '{' ContractPart* '}'

ContractPart = StateVariableDeclaration | UsingForDeclaration
             | StructDefinition | ModifierDefinition | FunctionDefinition |␣
↪EventDefinition | EnumDefinition

InheritanceSpecifier = UserDefinedTypeName ( '(' Expression ( ',' Expression )* ')' )?

StateVariableDeclaration = TypeName ( 'public' | 'internal' | 'private' | 'constant'␣
↪)* Identifier ('=' Expression)? ';'
UsingForDeclaration = 'using' Identifier 'for' ('*' | TypeName) ';'
StructDefinition = 'struct' Identifier '{'
                     ( VariableDeclaration ';' (VariableDeclaration ';')* ) '}'

ModifierDefinition = 'modifier' Identifier ParameterList? Block
ModifierInvocation = Identifier ( '(' ExpressionList? ')' )?

FunctionDefinition = 'function' Identifier? ParameterList
```

(continues on next page)

(continued from previous page)

```
                    ( ModifierInvocation | StateMutability | 'external' | 'public' |
↪'internal' | 'private' )*
                    ( 'returns' ParameterList )? ( ';' | Block )
EventDefinition = 'event' Identifier EventParameterList 'anonymous'? ';'

EnumValue = Identifier
EnumDefinition = 'enum' Identifier '{' EnumValue? (',' EnumValue)* '}'

ParameterList = '(' ( Parameter (',' Parameter)* )? ')'
Parameter = TypeName StorageLocation? Identifier?

EventParameterList = '(' ( EventParameter (',' EventParameter )* )? ')'
EventParameter = TypeName 'indexed'? Identifier?

FunctionTypeParameterList = '(' ( FunctionTypeParameter (',' FunctionTypeParameter )*␣
↪)? ')'
FunctionTypeParameter = TypeName StorageLocation?

// semantic restriction: mappings and structs (recursively) containing mappings
// are not allowed in argument lists
VariableDeclaration = TypeName StorageLocation? Identifier

TypeName = ElementaryTypeName
         | UserDefinedTypeName
         | Mapping
         | ArrayTypeName
         | FunctionTypeName

UserDefinedTypeName = Identifier ( '.' Identifier )*

Mapping = 'mapping' '(' ElementaryTypeName '=>' TypeName ')'
ArrayTypeName = TypeName '[' Expression? ']'
FunctionTypeName = 'function' FunctionTypeParameterList ( 'internal' | 'external' |␣
↪StateMutability )*
                   ( 'returns' FunctionTypeParameterList )?
StorageLocation = 'memory' | 'storage' | 'calldata'
StateMutability = 'pure' | 'view' | 'payable'

Block = '{' Statement* '}'
Statement = IfStatement | WhileStatement | ForStatement | Block |␣
↪InlineAssemblyStatement |
            ( DoWhileStatement | PlaceholderStatement | Continue | Break | Return |
              Throw | EmitStatement | SimpleStatement ) ';'

ExpressionStatement = Expression
IfStatement = 'if' '(' Expression ')' Statement ( 'else' Statement )?
WhileStatement = 'while' '(' Expression ')' Statement
PlaceholderStatement = '_'
SimpleStatement = VariableDefinition | ExpressionStatement
ForStatement = 'for' '(' (SimpleStatement)? ';' (Expression)? ';'␣
↪(ExpressionStatement)? ')' Statement
InlineAssemblyStatement = 'assembly' StringLiteral? InlineAssemblyBlock
DoWhileStatement = 'do' Statement 'while' '(' Expression ')'
Continue = 'continue'
Break = 'break'
Return = 'return' Expression?
Throw = 'throw'
```

(continues on next page)

(continued from previous page)

```
EmitStatement = 'emit' FunctionCall
VariableDefinition = (VariableDeclaration | '(' VariableDeclaration? (','␣
↪VariableDeclaration? )* ')' ) ( '=' Expression )?
IdentifierList = '(' ( Identifier? ',' )* Identifier? ')'

// Precedence by order (see github.com/ethereum/solidity/pull/732)
Expression
  = Expression ('++' | '--')
  | NewExpression
  | IndexAccess
  | MemberAccess
  | FunctionCall
  | '(' Expression ')'
  | ('!' | '~' | 'delete' | '++' | '--' | '+' | '-') Expression
  | Expression '**' Expression
  | Expression ('*' | '/' | '%') Expression
  | Expression ('+' | '-') Expression
  | Expression ('<<' | '>>') Expression
  | Expression '&' Expression
  | Expression '^' Expression
  | Expression '|' Expression
  | Expression ('<' | '>' | '<=' | '>=') Expression
  | Expression ('==' | '!=') Expression
  | Expression '&&' Expression
  | Expression '||' Expression
  | Expression '?' Expression ':' Expression
  | Expression ('=' | '|=' | '^=' | '&=' | '<<=' | '>>=' | '+=' | '-=' | '*=' | '/=' ␣
↪| '%=') Expression
  | PrimaryExpression

PrimaryExpression = BooleanLiteral
                  | NumberLiteral
                  | HexLiteral
                  | StringLiteral
                  | TupleExpression
                  | Identifier
                  | ElementaryTypeNameExpression

ExpressionList = Expression ( ',' Expression )*
NameValueList = Identifier ':' Expression ( ',' Identifier ':' Expression )*

FunctionCall = Expression '(' FunctionCallArguments ')'
FunctionCallArguments = '{' NameValueList? '}'
                      | ExpressionList?

NewExpression = 'new' TypeName
MemberAccess = Expression '.' Identifier
IndexAccess = Expression '[' Expression? ']'

BooleanLiteral = 'true' | 'false'
NumberLiteral = ( HexNumber | DecimalNumber ) (' ' NumberUnit)?
NumberUnit = 'wei' | 'szabo' | 'finney' | 'ether'
           | 'seconds' | 'minutes' | 'hours' | 'days' | 'weeks' | 'years'
HexLiteral = 'hex' ('"' ([0-9a-fA-F]{2})* '"' | '\'' ([0-9a-fA-F]{2})* '\'')
StringLiteral = '"' ([^"\r\n\\] | '\\' .)* '"'
Identifier = [a-zA-Z_$] [a-zA-Z_$0-9]*
```

(continues on next page)

(continued from previous page)

```
HexNumber = '0x' [0-9a-fA-F]+
DecimalNumber = [0-9]+ ( '.' [0-9]* )? ( [eE] [0-9]+ )?

TupleExpression = '(' ( Expression? ( ',' Expression? )*  )? ')'
                | '[' ( Expression  ( ',' Expression  )*  )? ']'

ElementaryTypeNameExpression = ElementaryTypeName

ElementaryTypeName = 'address' | 'bool' | 'string' | Int | Uint | Byte | Fixed |␣
↪Ufixed

Int = 'int' | 'int8' | 'int16' | 'int24' | 'int32' | 'int40' | 'int48' | 'int56' |
↪'int64' | 'int72' | 'int80' | 'int88' | 'int96' | 'int104' | 'int112' | 'int120' |
↪'int128' | 'int136' | 'int144' | 'int152' | 'int160' | 'int168' | 'int176' | 'int184
↪' | 'int192' | 'int200' | 'int208' | 'int216' | 'int224' | 'int232' | 'int240' |
↪'int248' | 'int256'

Uint = 'uint' | 'uint8' | 'uint16' | 'uint24' | 'uint32' | 'uint40' | 'uint48' |
↪'uint56' | 'uint64' | 'uint72' | 'uint80' | 'uint88' | 'uint96' | 'uint104' |
↪'uint112' | 'uint120' | 'uint128' | 'uint136' | 'uint144' | 'uint152' | 'uint160' |
↪'uint168' | 'uint176' | 'uint184' | 'uint192' | 'uint200' | 'uint208' | 'uint216' |
↪'uint224' | 'uint232' | 'uint240' | 'uint248' | 'uint256'

Byte = 'byte' | 'bytes' | 'bytes1' | 'bytes2' | 'bytes3' | 'bytes4' | 'bytes5' |
↪'bytes6' | 'bytes7' | 'bytes8' | 'bytes9' | 'bytes10' | 'bytes11' | 'bytes12' |
↪'bytes13' | 'bytes14' | 'bytes15' | 'bytes16' | 'bytes17' | 'bytes18' | 'bytes19' |
↪'bytes20' | 'bytes21' | 'bytes22' | 'bytes23' | 'bytes24' | 'bytes25' | 'bytes26' |
↪'bytes27' | 'bytes28' | 'bytes29' | 'bytes30' | 'bytes31' | 'bytes32'

Fixed = 'fixed' | ( 'fixed' [0-9]+ 'x' [0-9]+ )

Ufixed = 'ufixed' | ( 'ufixed' [0-9]+ 'x' [0-9]+ )

InlineAssemblyBlock = '{' AssemblyItem* '}'

AssemblyItem = Identifier | FunctionalAssemblyExpression | InlineAssemblyBlock |␣
↪AssemblyLocalBinding | AssemblyAssignment | AssemblyLabel | NumberLiteral |␣
↪StringLiteral | HexLiteral
AssemblyLocalBinding = 'let' Identifier ':=' FunctionalAssemblyExpression
AssemblyAssignment = ( Identifier ':=' FunctionalAssemblyExpression ) | ( '=:'␣
↪Identifier )
AssemblyLabel = Identifier ':'
FunctionalAssemblyExpression = Identifier '(' AssemblyItem? ( ',' AssemblyItem )* ')'
```

7.5 Security Considerations

While it is usually quite easy to build software that works as expected, it is much harder to check that nobody can use it in a way that was **not** anticipated.

In Solidity, this is even more important because you can use smart contracts to handle tokens or, possibly, even more valuable things. Furthermore, every execution of a smart contract happens in public and, in addition to that, the source code is often available.

Of course you always have to consider how much is at stake: You can compare a smart contract with a web service that is open to the public (and thus, also to malicious actors) and perhaps even open source. If you only store your

grocery list on that web service, you might not have to take too much care, but if you manage your bank account using that web service, you should be more careful.

This section will list some pitfalls and general security recommendations but can, of course, never be complete. Also, keep in mind that even if your smart contract code is bug-free, the compiler or the platform itself might have a bug. A list of some publicly known security-relevant bugs of the compiler can be found in the *list of known bugs*, which is also machine-readable. Note that there is a bug bounty program that covers the code generator of the Solidity compiler.

As always, with open source documentation, please help us extend this section (especially, some examples would not hurt)!

7.5.1 Pitfalls

Private Information and Randomness

Everything you use in a smart contract is publicly visible, even local variables and state variables marked `private`.

Using random numbers in smart contracts is quite tricky if you do not want miners to be able to cheat.

Re-Entrancy

Any interaction from a contract (A) with another contract (B) and any transfer of Ether hands over control to that contract (B). This makes it possible for B to call back into A before this interaction is completed. To give an example, the following code contains a bug (it is just a snippet and not a complete contract):

```
pragma solidity ^0.4.0;

// THIS CONTRACT CONTAINS A BUG - DO NOT USE
contract Fund {
    /// Mapping of ether shares of the contract.
    mapping(address => uint) shares;
    /// Withdraw your share.
    function withdraw() public {
        if (msg.sender.send(shares[msg.sender]))
            shares[msg.sender] = 0;
    }
}
```

The problem is not too serious here because of the limited gas as part of `send`, but it still exposes a weakness: Ether transfer can always include code execution, so the recipient could be a contract that calls back into `withdraw`. This would let it get multiple refunds and basically retrieve all the Ether in the contract. In particular, the following contract will allow an attacker to refund multiple times as it uses `call` which forwards all remaining gas by default:

```
pragma solidity ^0.4.0;

// THIS CONTRACT CONTAINS A BUG - DO NOT USE
contract Fund {
    /// Mapping of ether shares of the contract.
    mapping(address => uint) shares;
    /// Withdraw your share.
    function withdraw() public {
        if (msg.sender.call.value(shares[msg.sender])(""))
            shares[msg.sender] = 0;
    }
}
```

To avoid re-entrancy, you can use the Checks-Effects-Interactions pattern as outlined further below:

```
pragma solidity ^0.4.11;

contract Fund {
    /// Mapping of ether shares of the contract.
    mapping(address => uint) shares;
    /// Withdraw your share.
    function withdraw() public {
        uint share = shares[msg.sender];
        shares[msg.sender] = 0;
        msg.sender.transfer(share);
    }
}
```

Note that re-entrancy is not only an effect of Ether transfer but of any function call on another contract. Furthermore, you also have to take multi-contract situations into account. A called contract could modify the state of another contract you depend on.

Gas Limit and Loops

Loops that do not have a fixed number of iterations, for example, loops that depend on storage values, have to be used carefully: Due to the block gas limit, transactions can only consume a certain amount of gas. Either explicitly or just due to normal operation, the number of iterations in a loop can grow beyond the block gas limit which can cause the complete contract to be stalled at a certain point. This may not apply to `view` functions that are only executed to read data from the blockchain. Still, such functions may be called by other contracts as part of on-chain operations and stall those. Please be explicit about such cases in the documentation of your contracts.

Sending and Receiving Ether

- Neither contracts nor "external accounts" are currently able to prevent that someone sends them Ether. Contracts can react on and reject a regular transfer, but there are ways to move Ether without creating a message call. One way is to simply "mine to" the contract address and the second way is using `selfdestruct(x)`.
- If a contract receives Ether (without a function being called), the fallback function is executed. If it does not have a fallback function, the Ether will be rejected (by throwing an exception). During the execution of the fallback function, the contract can only rely on the "gas stipend" (2300 gas) being available to it at that time. This stipend is not enough to access storage in any way. To be sure that your contract can receive Ether in that way, check the gas requirements of the fallback function (for example in the "details" section in Remix).
- There is a way to forward more gas to the receiving contract using `addr.call.value(x)("")`. This is essentially the same as `addr.transfer(x)`, only that it forwards all remaining gas and opens up the ability for the recipient to perform more expensive actions (and it only returns a failure code and does not automatically propagate the error). This might include calling back into the sending contract or other state changes you might not have thought of. So it allows for great flexibility for honest users but also for malicious actors.
- If you want to send Ether using `address.transfer`, there are certain details to be aware of:
    1. If the recipient is a contract, it causes its fallback function to be executed which can, in turn, call back the sending contract.
    2. Sending Ether can fail due to the call depth going above 1024. Since the caller is in total control of the call depth, they can force the transfer to fail; take this possibility into account or use `send` and make sure to always check its return value. Better yet, write your contract using a pattern where the recipient can withdraw Ether instead.

3. Sending Ether can also fail because the execution of the recipient contract requires more than the allotted amount of gas (explicitly by using `require`, `assert`, `revert`, `throw` or because the operation is just too expensive) - it "runs out of gas" (OOG). If you use `transfer` or `send` with a return value check, this might provide a means for the recipient to block progress in the sending contract. Again, the best practice here is to use a *"withdraw" pattern instead of a "send" pattern*.

Callstack Depth

External function calls can fail any time because they exceed the maximum call stack of 1024. In such situations, Solidity throws an exception. Malicious actors might be able to force the call stack to a high value before they interact with your contract.

Note that `.send()` does **not** throw an exception if the call stack is depleted but rather returns `false` in that case. The low-level functions `.call()`, `.callcode()` and `.delegatecall()` behave in the same way.

tx.origin

Never use tx.origin for authorization. Let's say you have a wallet contract like this:

```
pragma solidity >0.4.24;

// THIS CONTRACT CONTAINS A BUG - DO NOT USE
contract TxUserWallet {
    address owner;

    constructor() public {
        owner = msg.sender;
    }

    function transferTo(address dest, uint amount) public {
        require(tx.origin == owner);
        dest.transfer(amount);
    }
}
```

Now someone tricks you into sending ether to the address of this attack wallet:

```
pragma solidity >0.4.24;

interface TxUserWallet {
    function transferTo(address dest, uint amount) external;
}

contract TxAttackWallet {
    address owner;

    constructor() public {
        owner = msg.sender;
    }

    function() external {
        TxUserWallet(msg.sender).transferTo(owner, msg.sender.balance);
    }
}
```

If your wallet had checked `msg.sender` for authorization, it would get the address of the attack wallet, instead of the owner address. But by checking `tx.origin`, it gets the original address that kicked off the transaction, which is still the owner address. The attack wallet instantly drains all your funds.

Minor Details

- Types that do not occupy the full 32 bytes might contain "dirty higher order bits". This is especially important if you access `msg.data` - it poses a malleability risk: You can craft transactions that call a function `f(uint8 x)` with a raw byte argument of `0xff000001` and with `0x00000001`. Both are fed to the contract and both will look like the number `1` as far as `x` is concerned, but `msg.data` will be different, so if you use `keccak256(msg.data)` for anything, you will get different results.

7.5.2 Recommendations

Take Warnings Seriously

If the compiler warns you about something, you should better change it. Even if you do not think that this particular warning has security implications, there might be another issue buried beneath it. Any compiler warning we issue can be silenced by slight changes to the code.

Also try to enable the "0.5.0" safety features as early as possible by adding `pragma experimental "v0.5.0";`. Note that in this case, the word `experimental` does not mean that the safety features are in any way risky, it is just a way to enable some features that are not yet part of the latest version of Solidity due to backwards compatibility.

Restrict the Amount of Ether

Restrict the amount of Ether (or other tokens) that can be stored in a smart contract. If your source code, the compiler or the platform has a bug, these funds may be lost. If you want to limit your loss, limit the amount of Ether.

Keep it Small and Modular

Keep your contracts small and easily understandable. Single out unrelated functionality in other contracts or into libraries. General recommendations about source code quality of course apply: Limit the amount of local variables, the length of functions and so on. Document your functions so that others can see what your intention was and whether it is different than what the code does.

Use the Checks-Effects-Interactions Pattern

Most functions will first perform some checks (who called the function, are the arguments in range, did they send enough Ether, does the person have tokens, etc.). These checks should be done first.

As the second step, if all checks passed, effects to the state variables of the current contract should be made. Interaction with other contracts should be the very last step in any function.

Early contracts delayed some effects and waited for external function calls to return in a non-error state. This is often a serious mistake because of the re-entrancy problem explained above.

Note that, also, calls to known contracts might in turn cause calls to unknown contracts, so it is probably better to just always apply this pattern.

Include a Fail-Safe Mode

While making your system fully decentralised will remove any intermediary, it might be a good idea, especially for new code, to include some kind of fail-safe mechanism:

You can add a function in your smart contract that performs some self-checks like "Has any Ether leaked?", "Is the sum of the tokens equal to the balance of the contract?" or similar things. Keep in mind that you cannot use too much gas for that, so help through off-chain computations might be needed there.

If the self-check fails, the contract automatically switches into some kind of "failsafe" mode, which, for example, disables most of the features, hands over control to a fixed and trusted third party or just converts the contract into a simple "give me back my money" contract.

7.5.3 Formal Verification

Using formal verification, it is possible to perform an automated mathematical proof that your source code fulfills a certain formal specification. The specification is still formal (just as the source code), but usually much simpler.

Note that formal verification itself can only help you understand the difference between what you did (the specification) and how you did it (the actual implementation). You still need to check whether the specification is what you wanted and that you did not miss any unintended effects of it.

7.6 Using the compiler

7.6.1 Using the Commandline Compiler

Note: This section doesn't apply to *solcjs*.

One of the build targets of the Solidity repository is `solc`, the solidity commandline compiler. Using `solc --help` provides you with an explanation of all options. The compiler can produce various outputs, ranging from simple binaries and assembly over an abstract syntax tree (parse tree) to estimations of gas usage. If you only want to compile a single file, you run it as `solc --bin sourceFile.sol` and it will print the binary. If you want to get some of the more advanced output variants of `solc`, it is probably better to tell it to output everything to separate files using `solc -o outputDirectory --bin --ast --asm sourceFile.sol`.

Before you deploy your contract, activate the optimizer while compiling using `solc --optimize --bin sourceFile.sol`. By default, the optimizer will optimize the contract for 200 runs. If you want to optimize for initial contract deployment and get the smallest output, set it to `--runs=1`. If you expect many transactions and don't care for higher deployment cost and output size, set `--runs` to a high number.

The commandline compiler will automatically read imported files from the filesystem, but it is also possible to provide path redirects using `prefix=path` in the following way:

```
solc github.com/ethereum/dapp-bin/=/usr/local/lib/dapp-bin/ =/usr/local/lib/fallback␣
↪file.sol
```

This essentially instructs the compiler to search for anything starting with `github.com/ethereum/dapp-bin/` under `/usr/local/lib/dapp-bin` and if it does not find the file there, it will look at `/usr/local/lib/fallback` (the empty prefix always matches). `solc` will not read files from the filesystem that lie outside of the remapping targets and outside of the directories where explicitly specified source files reside, so things like `import "/etc/passwd";` only work if you add `=/` as a remapping.

If there are multiple matches due to remappings, the one with the longest common prefix is selected.

For security reasons the compiler has restrictions what directories it can access. Paths (and their subdirectories) of source files specified on the commandline and paths defined by remappings are allowed for import statements, but everything else is rejected. Additional paths (and their subdirectories) can be allowed via the `--allow-paths /sample/path,/another/sample/path` switch.

If your contracts use *libraries*, you will notice that the bytecode contains substrings of the form `__LibraryName_____`. You can use `solc` as a linker meaning that it will insert the library addresses for you at those points:

Either add `--libraries "Math:0x1234567890123456789 0 Heap:0xabcdef0123456"` to your command to provide an address for each library or store the string in a file (one library per line) and run `solc` using `--libraries fileName`.

If `solc` is called with the option `--link`, all input files are interpreted to be unlinked binaries (hex-encoded) in the `__LibraryName_____`-format given above and are linked in-place (if the input is read from stdin, it is written to stdout). All options except `--libraries` are ignored (including `-o`) in this case.

If `solc` is called with the option `--standard-json`, it will expect a JSON input (as explained below) on the standard input, and return a JSON output on the standard output.

7.6.2 Compiler Input and Output JSON Description

These JSON formats are used by the compiler API as well as are available through `solc`. These are subject to change, some fields are optional (as noted), but it is aimed at to only make backwards compatible changes.

The compiler API expects a JSON formatted input and outputs the compilation result in a JSON formatted output.

Comments are of course not permitted and used here only for explanatory purposes.

Input Description

```
{
  // Required: Source code language, such as "Solidity", "serpent", "lll", "assembly",
↪ etc.
  language: "Solidity",
  // Required
  sources:
  {
    // The keys here are the "global" names of the source files,
    // imports can use other files via remappings (see below).
    "myFile.sol":
    {
      // Optional: keccak256 hash of the source file
      // It is used to verify the retrieved content if imported via URLs.
      "keccak256": "0x123...",
      // Required (unless "content" is used, see below): URL(s) to the source file.
      // URL(s) should be imported in this order and the result checked against the
      // keccak256 hash (if available). If the hash doesn't match or none of the
      // URL(s) result in success, an error should be raised.
      "urls":
      [
        "bzzr://56ab...",
        "ipfs://Qma...",
        "file:///tmp/path/to/file.sol"
      ]
    },
```

(continues on next page)

(continued from previous page)

```
    "mortal":
    {
      // Optional: keccak256 hash of the source file
      "keccak256": "0x234...",
      // Required (unless "urls" is used): literal contents of the source file
      "content": "contract mortal is owned { function kill() { if (msg.sender ==␣
↪owner) selfdestruct(owner); } }"
    }
  },
  // Optional
  settings:
  {
    // Optional: Sorted list of remappings
    remappings: [ ":g/dir" ],
    // Optional: Optimizer settings
    optimizer: {
      // disabled by default
      enabled: true,
      // Optimize for how many times you intend to run the code.
      // Lower values will optimize more for initial deployment cost, higher values␣
↪will optimize more for high-frequency usage.
      runs: 200
    },
    evmVersion: "byzantium", // Version of the EVM to compile for. Affects type␣
↪checking and code generation. Can be homestead, tangerineWhistle, spuriousDragon,␣
↪byzantium or constantinople
    // Metadata settings (optional)
    metadata: {
      // Use only literal content and not URLs (false by default)
      useLiteralContent: true
    },
    // Addresses of the libraries. If not all libraries are given here, it can result␣
↪in unlinked objects whose output data is different.
    libraries: {
      // The top level key is the the name of the source file where the library is␣
↪used.
      // If remappings are used, this source file should match the global path after␣
↪remappings were applied.
      // If this key is an empty string, that refers to a global level.
      "myFile.sol": {
        "MyLib": "0x123123..."
      }
    }
    // The following can be used to select desired outputs.
    // If this field is omitted, then the compiler loads and does type checking, but␣
↪will not generate any outputs apart from errors.
    // The first level key is the file name and the second is the contract name,␣
↪where empty contract name refers to the file itself,
    // while the star refers to all of the contracts.
    //
    // The available output types are as follows:
    //   abi - ABI
    //   ast - AST of all source files
    //   legacyAST - legacy AST of all source files
    //   devdoc - Developer documentation (natspec)
    //   userdoc - User documentation (natspec)
    //   metadata - Metadata
```

(continues on next page)

(continued from previous page)

```
    //   ir - New assembly format before desugaring
    //   evm.assembly - New assembly format after desugaring
    //   evm.legacyAssembly - Old-style assembly format in JSON
    //   evm.bytecode.object - Bytecode object
    //   evm.bytecode.opcodes - Opcodes list
    //   evm.bytecode.sourceMap - Source mapping (useful for debugging)
    //   evm.bytecode.linkReferences - Link references (if unlinked object)
    //   evm.deployedBytecode* - Deployed bytecode (has the same options as evm.
↪bytecode)
    //   evm.methodIdentifiers - The list of function hashes
    //   evm.gasEstimates - Function gas estimates
    //   ewasm.wast - eWASM S-expressions format (not supported atm)
    //   ewasm.wasm - eWASM binary format (not supported atm)
    //
    // Note that using a using `evm`, `evm.bytecode`, `ewasm`, etc. will select every
    // target part of that output. Additionally, `*` can be used as a wildcard to␣
↪request everything.
    //
    outputSelection: {
      // Enable the metadata and bytecode outputs of every single contract.
      "*": {
        "*": [ "metadata", "evm.bytecode" ]
      },
      // Enable the abi and opcodes output of MyContract defined in file def.
      "def": {
        "MyContract": [ "abi", "evm.bytecode.opcodes" ]
      },
      // Enable the source map output of every single contract.
      "*": {
        "*": [ "evm.bytecode.sourceMap" ]
      },
      // Enable the legacy AST output of every single file.
      "*": {
        "": [ "legacyAST" ]
      }
    }
  }
}
```

Output Description

```
{
  // Optional: not present if no errors/warnings were encountered
  errors: [
    {
      // Optional: Location within the source file.
      sourceLocation: {
        file: "sourceFile.sol",
        start: 0,
        end: 100
      },
      // Mandatory: Error type, such as "TypeError", "InternalCompilerError",
↪"Exception", etc.
      // See below for complete list of types.
      type: "TypeError",
```

(continues on next page)

(continued from previous page)

```
      // Mandatory: Component where the error originated, such as "general", "ewasm",␣
↪etc.
      component: "general",
      // Mandatory ("error" or "warning")
      severity: "error",
      // Mandatory
      message: "Invalid keyword"
      // Optional: the message formatted with source location
      formattedMessage: "sourceFile.sol:100: Invalid keyword"
    }
  ],
  // This contains the file-level outputs. In can be limited/filtered by the␣
↪outputSelection settings.
  sources: {
    "sourceFile.sol": {
      // Identifier (used in source maps)
      id: 1,
      // The AST object
      ast: {},
      // The legacy AST object
      legacyAST: {}
    }
  },
  // This contains the contract-level outputs. It can be limited/filtered by the␣
↪outputSelection settings.
  contracts: {
    "sourceFile.sol": {
      // If the language used has no contract names, this field should equal to an␣
↪empty string.
      "ContractName": {
        // The Ethereum Contract ABI. If empty, it is represented as an empty array.
        // See https://github.com/ethereum/wiki/wiki/Ethereum-Contract-ABI
        abi: [],
        // See the Metadata Output documentation (serialised JSON string)
        metadata: "{...}",
        // User documentation (natspec)
        userdoc: {},
        // Developer documentation (natspec)
        devdoc: {},
        // Intermediate representation (string)
        ir: "",
        // EVM-related outputs
        evm: {
          // Assembly (string)
          assembly: "",
          // Old-style assembly (object)
          legacyAssembly: {},
          // Bytecode and related details.
          bytecode: {
            // The bytecode as a hex string.
            object: "00fe",
            // Opcodes list (string)
            opcodes: "",
            // The source mapping as a string. See the source mapping definition.
            sourceMap: "",
            // If given, this is an unlinked object.
            linkReferences: {
```

(continues on next page)

(continued from previous page)

```
            "libraryFile.sol": {
              // Byte offsets into the bytecode. Linking replaces the 20 bytes␣
↪located there.
              "Library1": [
                { start: 0, length: 20 },
                { start: 200, length: 20 }
              ]
            }
          }
        },
        // The same layout as above.
        deployedBytecode: { },
        // The list of function hashes
        methodIdentifiers: {
          "delegate(address)": "5c19a95c"
        },
        // Function gas estimates
        gasEstimates: {
          creation: {
            codeDepositCost: "420000",
            executionCost: "infinite",
            totalCost: "infinite"
          },
          external: {
            "delegate(address)": "25000"
          },
          internal: {
            "heavyLifting()": "infinite"
          }
        }
      },
      // eWASM related outputs
      ewasm: {
        // S-expressions format
        wast: "",
        // Binary format (hex string)
        wasm: ""
      }
    }
  }
}
```

Error types

1. `JSONError`: JSON input doesn't conform to the required format, e.g. input is not a JSON object, the language is not supported, etc.
2. `IOError`: IO and import processing errors, such as unresolvable URL or hash mismatch in supplied sources.
3. `ParserError`: Source code doesn't conform to the language rules.
4. `DocstringParsingError`: The NatSpec tags in the comment block cannot be parsed.
5. `SyntaxError`: Syntactical error, such as `continue` is used outside of a `for` loop.
6. `DeclarationError`: Invalid, unresolvable or clashing identifier names. e.g. `Identifier not found`

7. `TypeError`: Error within the type system, such as invalid type conversions, invalid assignments, etc.
8. `UnimplementedFeatureError`: Feature is not supported by the compiler, but is expected to be supported in future versions.
9. `InternalCompilerError`: Internal bug triggered in the compiler - this should be reported as an issue.
10. `Exception`: Unknown failure during compilation - this should be reported as an issue.
11. `CompilerError`: Invalid use of the compiler stack - this should be reported as an issue.
12. `FatalError`: Fatal error not processed correctly - this should be reported as an issue.
13. `Warning`: A warning, which didn't stop the compilation, but should be addressed if possible.

7.7 Contract Metadata

The Solidity compiler automatically generates a JSON file, the contract metadata, that contains information about the current contract. It can be used to query the compiler version, the sources used, the ABI and NatSpec documentation in order to more safely interact with the contract and to verify its source code.

The compiler appends a Swarm hash of the metadata file to the end of the bytecode (for details, see below) of each contract, so that you can retrieve the file in an authenticated way without having to resort to a centralized data provider.

Of course, you have to publish the metadata file to Swarm (or some other service) so that others can access it. The file can be output by using `solc --metadata` and the file will be called `ContractName_meta.json`. It will contain Swarm references to the source code, so you have to upload all source files and the metadata file.

The metadata file has the following format. The example below is presented in a human-readable way. Properly formatted metadata should use quotes correctly, reduce whitespace to a minimum and sort the keys of all objects to arrive at a unique formatting. Comments are of course also not permitted and used here only for explanatory purposes.

```
{
  // Required: The version of the metadata format
  version: "1",
  // Required: Source code language, basically selects a "sub-version"
  // of the specification
  language: "Solidity",
  // Required: Details about the compiler, contents are specific
  // to the language.
  compiler: {
    // Required for Solidity: Version of the compiler
    version: "0.4.6+commit.2dabbdf0.Emscripten.clang",
    // Optional: Hash of the compiler binary which produced this output
    keccak256: "0x123..."
  },
  // Required: Compilation source files/source units, keys are file names
  sources:
  {
    "myFile.sol": {
      // Required: keccak256 hash of the source file
      "keccak256": "0x123...",
      // Required (unless "content" is used, see below): Sorted URL(s)
      // to the source file, protocol is more or less arbitrary, but a
      // Swarm URL is recommended
      "urls": [ "bzzr://56ab..." ]
    },
    "mortal": {
```

(continues on next page)

(continued from previous page)

```
      // Required: keccak256 hash of the source file
      "keccak256": "0x234...",
      // Required (unless "url" is used): literal contents of the source file
      "content": "contract mortal is owned { function kill() { if (msg.sender ==
↪owner) selfdestruct(owner); } }"
    }
  },
  // Required: Compiler settings
  settings:
  {
    // Required for Solidity: Sorted list of remappings
    remappings: [ ":g/dir" ],
    // Optional: Optimizer settings (enabled defaults to false)
    optimizer: {
      enabled: true,
      runs: 500
    },
    // Required for Solidity: File and name of the contract or library this
    // metadata is created for.
    compilationTarget: {
      "myFile.sol": "MyContract"
    },
    // Required for Solidity: Addresses for libraries used
    libraries: {
      "MyLib": "0x123123..."
    }
  },
  // Required: Generated information about the contract.
  output:
  {
    // Required: ABI definition of the contract
    abi: [ ... ],
    // Required: NatSpec user documentation of the contract
    userdoc: [ ... ],
    // Required: NatSpec developer documentation of the contract
    devdoc: [ ... ],
  }
}
```

Note: Note the ABI definition above has no fixed order. It can change with compiler versions.

Note: Since the bytecode of the resulting contract contains the metadata hash, any change to the metadata will result in a change of the bytecode. Furthermore, since the metadata includes a hash of all the sources used, a single whitespace change in any of the source codes will result in a different metadata, and subsequently a different bytecode.

7.7.1 Encoding of the Metadata Hash in the Bytecode

Because we might support other ways to retrieve the metadata file in the future, the mapping `{"bzzr0": <Swarm hash>}` is stored CBOR-encoded. Since the beginning of that encoding is not easy to find, its length is added in a two-byte big-endian encoding. The current version of the Solidity compiler thus adds the following to the end of the deployed bytecode:

```
0xa1 0x65 'b' 'z' 'z' 'r' '0' 0x58 0x20 <32 bytes swarm hash> 0x00 0x29
```

So in order to retrieve the data, the end of the deployed bytecode can be checked to match that pattern and use the Swarm hash to retrieve the file.

7.7.2 Usage for Automatic Interface Generation and NatSpec

The metadata is used in the following way: A component that wants to interact with a contract (e.g. Mist) retrieves the code of the contract, from that the Swarm hash of a file which is then retrieved. That file is JSON-decoded into a structure like above.

The component can then use the ABI to automatically generate a rudimentary user interface for the contract.

Furthermore, Mist can use the userdoc to display a confirmation message to the user whenever they interact with the contract.

Additional information about Ethereum Natural Specification (NatSpec) can be found here.

7.7.3 Usage for Source Code Verification

In order to verify the compilation, sources can be retrieved from Swarm via the link in the metadata file. The compiler of the correct version (which is checked to be part of the "official" compilers) is invoked on that input with the specified settings. The resulting bytecode is compared to the data of the creation transaction or `CREATE` opcode data. This automatically verifies the metadata since its hash is part of the bytecode. Excess data corresponds to the constructor input data, which should be decoded according to the interface and presented to the user.

7.8 Contract ABI Specification

7.8.1 Basic Design

The Contract Application Binary Interface (ABI) is the standard way to interact with contracts in the Ethereum ecosystem, both from outside the blockchain and for contract-to-contract interaction. Data is encoded according to its type, as described in this specification. The encoding is not self describing and thus requires a schema in order to decode.

We assume the interface functions of a contract are strongly typed, known at compilation time and static. No introspection mechanism will be provided. We assume that all contracts will have the interface definitions of any contracts they call available at compile-time.

This specification does not address contracts whose interface is dynamic or otherwise known only at run-time. Should these cases become important they can be adequately handled as facilities built within the Ethereum ecosystem.

7.8.2 Function Selector

The first four bytes of the call data for a function call specifies the function to be called. It is the first (left, high-order in big-endian) four bytes of the Keccak (SHA-3) hash of the signature of the function. The signature is defined as the canonical expression of the basic prototype, i.e. the function name with the parenthesised list of parameter types. Parameter types are split by a single comma - no spaces are used.

Note: The return type of a function is not part of this signature. In *Solidity's function overloading* return types are not considered. The reason is to keep function call resolution context-independent. The JSON description of the ABI however contains both inputs and outputs. See (the *JSON ABI*)

7.8.3 Argument Encoding

Starting from the fifth byte, the encoded arguments follow. This encoding is also used in other places, e.g. the return values and also event arguments are encoded in the same way, without the four bytes specifying the function.

7.8.4 Types

The following elementary types exist:

- `uint<M>`: unsigned integer type of `M` bits, `0 < M <= 256`, `M % 8 == 0`. e.g. `uint32`, `uint8`, `uint256`.
- `int<M>`: two's complement signed integer type of `M` bits, `0 < M <= 256`, `M % 8 == 0`.
- `address`: equivalent to `uint160`, except for the assumed interpretation and language typing. For computing the function selector, `address` is used.
- `uint`, `int`: synonyms for `uint256`, `int256` respectively. For computing the function selector, `uint256` and `int256` have to be used.
- `bool`: equivalent to `uint8` restricted to the values 0 and 1. For computing the function selector, `bool` is used.
- `fixed<M>x<N>`: signed fixed-point decimal number of `M` bits, `8 <= M <= 256`, `M % 8 ==0`, and `0 < N <= 80`, which denotes the value `v` as `v / (10 ** N)`.
- `ufixed<M>x<N>`: unsigned variant of `fixed<M>x<N>`.
- `fixed`, `ufixed`: synonyms for `fixed128x18`, `ufixed128x18` respectively. For computing the function selector, `fixed128x18` and `ufixed128x18` have to be used.
- `bytes<M>`: binary type of `M` bytes, `0 < M <= 32`.
- `function`: an address (20 bytes) followed by a function selector (4 bytes). Encoded identical to `bytes24`.

The following (fixed-size) array type exists:

- `<type>[M]`: a fixed-length array of `M` elements, `M >= 0`, of the given type.

The following non-fixed-size types exist:

- `bytes`: dynamic sized byte sequence.
- `string`: dynamic sized unicode string assumed to be UTF-8 encoded.
- `<type>[]`: a variable-length array of elements of the given type.

Types can be combined to a tuple by enclosing them inside parentheses, separated by commas:

- `(T1,T2,...,Tn)`: tuple consisting of the types `T1`, ..., `Tn`, `n >= 0`

It is possible to form tuples of tuples, arrays of tuples and so on. It is also possible to form zero-tuples (where `n == 0`).

Note: Solidity supports all the types presented above with the same names with the exception of tuples. The ABI tuple type is utilised for encoding Solidity `structs`.

7.8.5 Formal Specification of the Encoding

We will now formally specify the encoding, such that it will have the following properties, which are especially useful if some arguments are nested arrays:

Properties:

1. The number of reads necessary to access a value is at most the depth of the value inside the argument array structure, i.e. four reads are needed to retrieve `a_i[k][l][r]`. In a previous version of the ABI, the number of reads scaled linearly with the total number of dynamic parameters in the worst case.
2. The data of a variable or array element is not interleaved with other data and it is relocatable, i.e. it only uses relative "addresses"

We distinguish static and dynamic types. Static types are encoded in-place and dynamic types are encoded at a separately allocated location after the current block.

Definition: The following types are called "dynamic":

- `bytes`
- `string`
- `T[]` for any `T`
- `T[k]` for any dynamic `T` and any `k >= 0`
- `(T1,...,Tk)` if `Ti` is dynamic for some `1 <= i <= k`

All other types are called "static".

Definition: `len(a)` is the number of bytes in a binary string `a`. The type of `len(a)` is assumed to be `uint256`.

We define `enc`, the actual encoding, as a mapping of values of the ABI types to binary strings such that `len(enc(X))` depends on the value of `X` if and only if the type of `X` is dynamic.

Definition: For any ABI value `X`, we recursively define `enc(X)`, depending on the type of `X` being

- `(T1,...,Tk)` for `k >= 0` and any types `T1`, ..., `Tk`

  `enc(X) = head(X(1)) ... head(X(k)) tail(X(1)) ... tail(X(k))`

  where `X = (X(1), ..., X(k))` and `head` and `tail` are defined for `Ti` being a static type as

  `head(X(i)) = enc(X(i))` and `tail(X(i)) = ""` (the empty string)

  and as

  ```
  head(X(i)) = enc(len(head(X(1)) ... head(X(k)) tail(X(1)) ...
  tail(X(i-1)) )) tail(X(i)) = enc(X(i))
  ```

  otherwise, i.e. if `Ti` is a dynamic type.

  Note that in the dynamic case, `head(X(i))` is well-defined since the lengths of the head parts only depend on the types and not the values. Its value is the offset of the beginning of `tail(X(i))` relative to the start of `enc(X)`.

- `T[k]` for any `T` and `k`:

  `enc(X) = enc((X[0], ..., X[k-1]))`

  i.e. it is encoded as if it were a tuple with `k` elements of the same type.

- `T[]` where `X` has `k` elements (`k` is assumed to be of type `uint256`):

  `enc(X) = enc(k) enc([X[0], ..., X[k-1]])`

  i.e. it is encoded as if it were an array of static size `k`, prefixed with the number of elements.

- `bytes`, of length `k` (which is assumed to be of type `uint256`):

  `enc(X) = enc(k) pad_right(X)`, i.e. the number of bytes is encoded as a `uint256` followed by the actual value of `X` as a byte sequence, followed by the minimum number of zero-bytes such that `len(enc(X))` is a multiple of 32.

- `string`:

  `enc(X) = enc(enc_utf8(X))`, i.e. `X` is utf-8 encoded and this value is interpreted as of `bytes` type and encoded further. Note that the length used in this subsequent encoding is the number of bytes of the utf-8 encoded string, not its number of characters.

- `uint<M>`: `enc(X)` is the big-endian encoding of `X`, padded on the higher-order (left) side with zero-bytes such that the length is 32 bytes.

- `address`: as in the `uint160` case

- `int<M>`: `enc(X)` is the big-endian two's complement encoding of `X`, padded on the higher-order (left) side with `0xff` for negative `X` and with zero bytes for positive `X` such that the length is 32 bytes.

- `bool`: as in the `uint8` case, where `1` is used for `true` and `0` for `false`

- `fixed<M>x<N>`: `enc(X)` is `enc(X * 10**N)` where `X * 10**N` is interpreted as a `int256`.

- `fixed`: as in the `fixed128x18` case

- `ufixed<M>x<N>`: `enc(X)` is `enc(X * 10**N)` where `X * 10**N` is interpreted as a `uint256`.

- `ufixed`: as in the `ufixed128x18` case

- `bytes<M>`: `enc(X)` is the sequence of bytes in `X` padded with trailing zero-bytes to a length of 32 bytes.

Note that for any `X`, `len(enc(X))` is a multiple of 32.

7.8.6 Function Selector and Argument Encoding

All in all, a call to the function `f` with parameters `a_1, ..., a_n` is encoded as

```
function_selector(f) enc((a_1, ..., a_n))
```

and the return values `v_1, ..., v_k` of `f` are encoded as

```
enc((v_1, ..., v_k))
```

i.e. the values are combined into a tuple and encoded.

7.8.7 Examples

Given the contract:

```
pragma solidity ^0.4.16;

contract Foo {
  function bar(bytes3[2] memory) public pure {}
  function baz(uint32 x, bool y) public pure returns (bool r) { r = x > 32 || y; }
  function sam(bytes memory, bool, uint[] memory) public pure {}
}
```

Thus for our `Foo` example if we wanted to call `baz` with the parameters `69` and `true`, we would pass 68 bytes total, which can be broken down into:

- `0xcdcd77c0`: the Method ID. This is derived as the first 4 bytes of the Keccak hash of the ASCII form of the signature `baz(uint32,bool)`.
- `0x0000000000000000000000000000000000000000000000000000000000000045`: the first parameter, a uint32 value `69` padded to 32 bytes
- `0x0000000000000000000000000000000000000000000000000000000000000001`: the second parameter - boolean `true`, padded to 32 bytes

In total:

```
0xcdcd77c0000000000000000000000000000000000000000000000000000000000000004500000000000000000000000000000
```

It returns a single `bool`. If, for example, it were to return `false`, its output would be the single byte array `0x0000000000000000000000000000000000000000000000000000000000000000`, a single bool.

If we wanted to call `bar` with the argument `["abc", "def"]`, we would pass 68 bytes total, broken down into:

- `0xfce353f6`: the Method ID. This is derived from the signature `bar(bytes3[2])`.
- `0x6162630000000000000000000000000000000000000000000000000000000000`: the first part of the first parameter, a `bytes3` value `"abc"` (left-aligned).
- `0x6465660000000000000000000000000000000000000000000000000000000000`: the second part of the first parameter, a `bytes3` value `"def"` (left-aligned).

In total:

```
0xfce353f6616263000000000000000000000000000000000000000000000000000000006465660000000000000000000000000
```

If we wanted to call `sam` with the arguments `"dave"`, `true` and `[1,2,3]`, we would pass 292 bytes total, broken down into:

- `0xa5643bf2`: the Method ID. This is derived from the signature `sam(bytes,bool,uint256[])`. Note that `uint` is replaced with its canonical representation `uint256`.
- `0x0000000000000000000000000000000000000000000000000000000000000060`: the location of the data part of the first parameter (dynamic type), measured in bytes from the start of the arguments block. In this case, `0x60`.
- `0x0000000000000000000000000000000000000000000000000000000000000001`: the second parameter: boolean true.
- `0x00000000000000000000000000000000000000000000000000000000000000a0`: the location of the data part of the third parameter (dynamic type), measured in bytes. In this case, `0xa0`.
- `0x0000000000000000000000000000000000000000000000000000000000000004`: the data part of the first argument, it starts with the length of the byte array in elements, in this case, 4.

- `0x6461766500000000000000000000000000000000000000000000000000000000`: the contents of the first argument: the UTF-8 (equal to ASCII in this case) encoding of `"dave"`, padded on the right to 32 bytes.
- `0x0000000000000000000000000000000000000000000000000000000000000003`: the data part of the third argument, it starts with the length of the array in elements, in this case, 3.
- `0x0000000000000000000000000000000000000000000000000000000000000001`: the first entry of the third parameter.
- `0x0000000000000000000000000000000000000000000000000000000000000002`: the second entry of the third parameter.
- `0x0000000000000000000000000000000000000000000000000000000000000003`: the third entry of the third parameter.

In total:

```
0xa5643bf20000000000000000000000000000000000000000000000000000000000000060000000000000000000000000000000
```

7.8.8 Use of Dynamic Types

A call to a function with the signature `f(uint,uint32[],bytes10,bytes)` with values `(0x123, [0x456, 0x789], "1234567890", "Hello, world!")` is encoded in the following way:

We take the first four bytes of `sha3("f(uint256,uint32[],bytes10,bytes)")`, i.e. `0x8be65246`. Then we encode the head parts of all four arguments. For the static types `uint256` and `bytes10`, these are directly the values we want to pass, whereas for the dynamic types `uint32[]` and `bytes`, we use the offset in bytes to the start of their data area, measured from the start of the value encoding (i.e. not counting the first four bytes containing the hash of the function signature). These are:

- `0x0000000000000000000000000000000000000000000000000000000000000123` (`0x123` padded to 32 bytes)
- `0x0000000000000000000000000000000000000000000000000000000000000080` (offset to start of data part of second parameter, 4*32 bytes, exactly the size of the head part)
- `0x3132333435363738393000000000000000000000000000000000000000000000` (`"1234567890"` padded to 32 bytes on the right)
- `0x00000000000000000000000000000000000000000000000000000000000000e0` (offset to start of data part of fourth parameter = offset to start of data part of first dynamic parameter + size of data part of first dynamic parameter = 4*32 + 3*32 (see below))

After this, the data part of the first dynamic argument, `[0x456, 0x789]` follows:

- `0x0000000000000000000000000000000000000000000000000000000000000002` (number of elements of the array, 2)
- `0x0000000000000000000000000000000000000000000000000000000000000456` (first element)
- `0x0000000000000000000000000000000000000000000000000000000000000789` (second element)

Finally, we encode the data part of the second dynamic argument, `"Hello, world!"`:

- `0x000000000000000000000000000000000000000000000000000000000000000d` (number of elements (bytes in this case): 13)

- `0x48656c6c6f2c20776f726c6421000000000000000000000000000000000000000000` (`"Hello, world!"` padded to 32 bytes on the right)

All together, the encoding is (newline after function selector and each 32-bytes for clarity):

```
0x8be65246
  0000000000000000000000000000000000000000000000000000000000000123
  0000000000000000000000000000000000000000000000000000000000000080
  3132333435363738393000000000000000000000000000000000000000000000
  00000000000000000000000000000000000000000000000000000000000000e0
  0000000000000000000000000000000000000000000000000000000000000002
  0000000000000000000000000000000000000000000000000000000000000456
  0000000000000000000000000000000000000000000000000000000000000789
  000000000000000000000000000000000000000000000000000000000000000d
  48656c6c6f2c20776f726c6421000000000000000000000000000000000000000000
```

Let us apply the same principle to encode the data for a function with a signature `g(uint[][],string[])` with values `([[1, 2], [3]], ["one", "two", "three"])` but start from the most atomic parts of the encoding:

First we encode the length and data of the first embedded dynamic array `[1, 2]` of the first root array `[[1, 2], [3]]`:

- `0x0000000000000000000000000000000000000000000000000000000000000002` (number of elements in the first array, 2; the elements themselves are `1` and `2`)
- `0x0000000000000000000000000000000000000000000000000000000000000001` (first element)
- `0x0000000000000000000000000000000000000000000000000000000000000002` (second element)

Then we encode the length and data of the second embedded dynamic array `[3]` of the first root array `[[1, 2], [3]]`:

- `0x0000000000000000000000000000000000000000000000000000000000000001` (number of elements in the second array, 1; the element is `3`)
- `0x0000000000000000000000000000000000000000000000000000000000000003` (first element)

Then we need to find the offsets `a` and `b` for their respective dynamic arrays `[1, 2]` and `[3]`. To calculate the offsets we can take a look at the encoded data of the first root array `[[1, 2], [3]]` enumerating each line in the encoding:

```
0 - a                                                                - offset of [1,␣
↪2]
1 - b                                                                - offset of [3]
2 - 0000000000000000000000000000000000000000000000000000000000000002 - count for [1,␣
↪2]
3 - 0000000000000000000000000000000000000000000000000000000000000001 - encoding of 1
4 - 0000000000000000000000000000000000000000000000000000000000000002 - encoding of 2
5 - 0000000000000000000000000000000000000000000000000000000000000001 - count for [3]
6 - 0000000000000000000000000000000000000000000000000000000000000003 - encoding of 3
```

Offset `a` points to the start of the content of the array `[1, 2]` which is line 2 (64 bytes); thus `a = 0x0000000000000000000000000000000000000000000000000000000000000040`.

Offset `b` points to the start of the content of the array `[3]` which is line 5 (160 bytes); thus `b = 0x00000000000000000000000000000000000000000000000000000000000000a0`.

Then we encode the embedded strings of the second root array:

- `0x0000000000000000000000000000000000000000000000000000000000000003` (number of characters in word `"one"`)
- `0x6f6e650000000000000000000000000000000000000000000000000000000000` (utf8 representation of word `"one"`)
- `0x0000000000000000000000000000000000000000000000000000000000000003` (number of characters in word `"two"`)
- `0x74776f0000000000000000000000000000000000000000000000000000000000` (utf8 representation of word `"two"`)
- `0x0000000000000000000000000000000000000000000000000000000000000005` (number of characters in word `"three"`)
- `0x7468726565000000000000000000000000000000000000000000000000000000` (utf8 representation of word `"three"`)

In parallel to the first root array, since strings are dynamic elements we need to find their offsets `c`, `d` and `e`:

```
0 - c                                                                - offset for "one
↪"
1 - d                                                                - offset for "two
↪"
2 - e                                                                - offset for
↪"three"
3 - 0000000000000000000000000000000000000000000000000000000000000003 - count for "one"
4 - 6f6e650000000000000000000000000000000000000000000000000000000000 - encoding of
↪"one"
5 - 0000000000000000000000000000000000000000000000000000000000000003 - count for "two"
6 - 74776f0000000000000000000000000000000000000000000000000000000000 - encoding of
↪"two"
7 - 0000000000000000000000000000000000000000000000000000000000000005 - count for
↪"three"
8 - 7468726565000000000000000000000000000000000000000000000000000000 - encoding of
↪"three"
```

Offset `c` points to the start of the content of the string `"one"` which is line 3 (96 bytes); thus `c = 0x0000000000000000000000000000000000000000000000000000000000000060`.

Offset `d` points to the start of the content of the string `"two"` which is line 5 (160 bytes); thus `d = 0x00000000000000000000000000000000000000000000000000000000000000a0`.

Offset `e` points to the start of the content of the string `"three"` which is line 7 (224 bytes); thus `e = 0x00000000000000000000000000000000000000000000000000000000000000e0`.

Note that the encodings of the embedded elements of the root arrays are not dependent on each other and have the same encodings for a function with a signature `g(string[],uint[][])`.

Then we encode the length of the first root array:

- `0x0000000000000000000000000000000000000000000000000000000000000002` (number of elements in the first root array, 2; the elements themselves are `[1, 2]` and `[3]`)

Then we encode the length of the second root array:

- `0x0000000000000000000000000000000000000000000000000000000000000003` (number of strings in the second root array, 3; the strings themselves are `"one"`, `"two"` and `"three"`)

Finally we find the offsets `f` and `g` for their respective root dynamic arrays `[[1, 2], [3]]` and `["one", "two", "three"]`, and assemble parts in the correct order:

```
0x2289b18c                                                          - function␣
↪signature
 0 - f                                                              - offset of [[1,
↪ 2], [3]]
 1 - g                                                              - offset of [
↪"one", "two", "three"]
 2 - 0000000000000000000000000000000000000000000000000000000000000002 - count for [[1,
↪ 2], [3]]
 3 - 0000000000000000000000000000000000000000000000000000000000000040 - offset of [1,␣
↪2]
 4 - 00000000000000000000000000000000000000000000000000000000000000a0 - offset of [3]
 5 - 0000000000000000000000000000000000000000000000000000000000000002 - count for [1,␣
↪2]
 6 - 0000000000000000000000000000000000000000000000000000000000000001 - encoding of 1
 7 - 0000000000000000000000000000000000000000000000000000000000000002 - encoding of 2
 8 - 0000000000000000000000000000000000000000000000000000000000000001 - count for [3]
 9 - 0000000000000000000000000000000000000000000000000000000000000003 - encoding of 3
10 - 0000000000000000000000000000000000000000000000000000000000000003 - count for [
↪"one", "two", "three"]
11 - 0000000000000000000000000000000000000000000000000000000000000060 - offset for
↪"one"
12 - 00000000000000000000000000000000000000000000000000000000000000a0 - offset for
↪"two"
13 - 00000000000000000000000000000000000000000000000000000000000000e0 - offset for
↪"three"
14 - 0000000000000000000000000000000000000000000000000000000000000003 - count for "one
↪"
15 - 6f6e650000000000000000000000000000000000000000000000000000000000 - encoding of
↪"one"
16 - 0000000000000000000000000000000000000000000000000000000000000003 - count for "two
↪"
17 - 74776f0000000000000000000000000000000000000000000000000000000000 - encoding of
↪"two"
18 - 0000000000000000000000000000000000000000000000000000000000000005 - count for
↪"three"
19 - 7468726565000000000000000000000000000000000000000000000000000000 - encoding of
↪"three"
```

Offset `f` points to the start of the content of the array `[[1, 2], [3]]` which is line 2 (64 bytes); thus `f = 0x0000000000000000000000000000000000000000000000000000000000000040`.

Offset `g` points to the start of the content of the array `["one", "two", "three"]` which is line 10 (320 bytes); thus `g = 0x0000000000000000000000000000000000000000000000000000000000000140`.

7.8.9 Events

Events are an abstraction of the Ethereum logging/event-watching protocol. Log entries provide the contract's address, a series of up to four topics and some arbitrary length binary data. Events leverage the existing function ABI in order to interpret this (together with an interface spec) as a properly typed structure.

Given an event name and series of event parameters, we split them into two sub-series: those which are indexed and those which are not. Those which are indexed, which may number up to 3, are used alongside the Keccak hash of the event signature to form the topics of the log entry. Those which are not indexed form the byte array of the event.

In effect, a log entry using this ABI is described as:

- `address`: the address of the contract (intrinsically provided by Ethereum);

- `topics[0]: keccak(EVENT_NAME+"("+EVENT_ARGS.map(canonical_type_of).join(",")+")")` (`canonical_type_of` is a function that simply returns the canonical type of a given argument, e.g. for `uint indexed foo`, it would return `uint256`). If the event is declared as `anonymous` the `topics[0]` is not generated;
- `topics[n]`: `EVENT_INDEXED_ARGS[n - 1]` (`EVENT_INDEXED_ARGS` is the series of `EVENT_ARGS` that are indexed);
- `data`: `abi_serialise(EVENT_NON_INDEXED_ARGS)` (`EVENT_NON_INDEXED_ARGS` is the series of `EVENT_ARGS` that are not indexed, `abi_serialise` is the ABI serialisation function used for returning a series of typed values from a function, as described above).

For all fixed-length Solidity types, the `EVENT_INDEXED_ARGS` array contains the 32-byte encoded value directly. However, for *types of dynamic length*, which include `string`, `bytes`, and arrays, `EVENT_INDEXED_ARGS` will contain the *Keccak hash* of the encoded value, rather than the encoded value directly. This allows applications to efficiently query for values of dynamic-length types (by setting the hash of the encoded value as the topic), but leaves applications unable to decode indexed values they have not queried for. For dynamic-length types, application developers face a trade-off between fast search for predetermined values (if the argument is indexed) and legibility of arbitrary values (which requires that the arguments not be indexed). Developers may overcome this tradeoff and achieve both efficient search and arbitrary legibility by defining events with two arguments — one indexed, one not — intended to hold the same value.

7.8.10 JSON

The JSON format for a contract's interface is given by an array of function and/or event descriptions. A function description is a JSON object with the fields:

- `type`: `"function"`, `"constructor"`, or `"fallback"` (the *unnamed "default" function*);
- `name`: the name of the function;
- `inputs`: an array of objects, each of which contains:
  - `name`: the name of the parameter;
  - `type`: the canonical type of the parameter (more below).
  - `components`: used for tuple types (more below).
- `outputs`: an array of objects similar to `inputs`, can be omitted if function doesn't return anything;
- `payable`: `true` if function accepts ether, defaults to `false`;
- `stateMutability`: a string with one of the following values: `pure` (*specified to not read blockchain state*), `view` (*specified to not modify the blockchain state*), `nonpayable` and `payable` (same as `payable` above).
- `constant`: `true` if function is either `pure` or `view`

`type` can be omitted, defaulting to `"function"`.

Constructor and fallback function never have `name` or `outputs`. Fallback function doesn't have `inputs` either.

Sending non-zero ether to non-payable function will throw. Don't do it.

An event description is a JSON object with fairly similar fields:

- `type`: always `"event"`
- `name`: the name of the event;
- `inputs`: an array of objects, each of which contains:
  - `name`: the name of the parameter;

  - `type`: the canonical type of the parameter (more below).
  - `components`: used for tuple types (more below).
  - `indexed`: `true` if the field is part of the log's topics, `false` if it one of the log's data segment.
- `anonymous`: `true` if the event was declared as `anonymous`.

For example,

```
pragma solidity >0.4.24;

contract Test {
  constructor() public { b = 0x1234567890123456789012345678901234567890; }
  event Event(uint indexed a, bytes32 b);
  event Event2(uint indexed a, bytes32 b);
  function foo(uint a) public { emit Event(a, b); }
  bytes32 b;
}
```

would result in the JSON:

```
[{
"type":"event",
"inputs": [{"name":"a","type":"uint256","indexed":true},{"name":"b","type":"bytes32",
↪"indexed":false}],
"name":"Event"
}, {
"type":"event",
"inputs": [{"name":"a","type":"uint256","indexed":true},{"name":"b","type":"bytes32",
↪"indexed":false}],
"name":"Event2"
}, {
"type":"function",
"inputs": [{"name":"a","type":"uint256"}],
"name":"foo",
"outputs": []
}]
```

Handling tuple types

Despite that names are intentionally not part of the ABI encoding they do make a lot of sense to be included in the JSON to enable displaying it to the end user. The structure is nested in the following way:

An object with members `name`, `type` and potentially `components` describes a typed variable. The canonical type is determined until a tuple type is reached and the string description up to that point is stored in `type` prefix with the word `tuple`, i.e. it will be `tuple` followed by a sequence of `[]` and `[k]` with integers `k`. The components of the tuple are then stored in the member `components`, which is of array type and has the same structure as the top-level object except that `indexed` is not allowed there.

As an example, the code

```
pragma solidity ^0.4.19;
pragma experimental ABIEncoderV2;

contract Test {
  struct S { uint a; uint[] b; T[] c; }
  struct T { uint x; uint y; }
```

(continues on next page)

(continued from previous page)

```
  function f(S memory s, T memory t, uint a) public { }
  function g() public returns (S memory s, T memory t, uint a) {}
}
```

would result in the JSON:

```
[
  {
    "name": "f",
    "type": "function",
    "inputs": [
      {
        "name": "s",
        "type": "tuple",
        "components": [
          {
            "name": "a",
            "type": "uint256"
          },
          {
            "name": "b",
            "type": "uint256[]"
          },
          {
            "name": "c",
            "type": "tuple[]",
            "components": [
              {
                "name": "x",
                "type": "uint256"
              },
              {
                "name": "y",
                "type": "uint256"
              }
            ]
          }
        ]
      },
      {
        "name": "t",
        "type": "tuple",
        "components": [
          {
            "name": "x",
            "type": "uint256"
          },
          {
            "name": "y",
            "type": "uint256"
          }
        ]
      },
      {
        "name": "a",
        "type": "uint256"
      }
```

(continues on next page)

(continued from previous page)

```
    ],
    "outputs": []
  }
]
```

7.8.11 Non-standard Packed Mode

Through `abi.encodePacked()`, Solidity supports a non-standard packed mode where:

- no *function selector* is encoded,
- types shorter than 32 bytes are neither zero padded nor sign extended and
- dynamic types are encoded in-place and without the length.

As an example encoding `int1, bytes1, uint16, string` with values `-1, 0x42, 0x2424, "Hello, world!"` results in

```
0xff42242448656c6c6f2c20776f726c6421
  ^^                                 int1(-1)
    ^^                               bytes1(0x42)
      ^^^^                           uint16(0x2424)
          ^^^^^^^^^^^^^^^^^^^^^^^^^^ string("Hello, world!") without a length field
```

More specifically, each statically-sized type takes as many bytes as its range has and dynamically-sized types like `string`, `bytes` or `uint[]` are encoded without their length field. This means that the encoding is ambiguous as soon as there are two dynamically-sized elements.

Note that constants will be packed using the minimum number of bytes required to store them. This means that, for example, `abi.encodePacked(0) == abi.encodePacked(uint8(0)) == hex"00"` and `abi.encodePacked(0x12345678) == abi.encodePacked(uint32(0x12345678)) == hex"12345678"`.

If padding is needed, explicit type conversions can be used: `abi.encodePacked(uint16(0x12)) == hex"0012"`.

7.9 Yul

Yul (previously also called JULIA or IULIA) is an intermediate language that can compile to various different backends (EVM 1.0, EVM 1.5 and eWASM are planned). Because of that, it is designed to be a usable common denominator of all three platforms. It can already be used for "inline assembly" inside Solidity and future versions of the Solidity compiler will even use Yul as intermediate language. It should also be easy to build high-level optimizer stages for Yul.

Note: Note that the flavour used for "inline assembly" does not have types (everything is `u256`) and the built-in functions are identical to the EVM opcodes. Please resort to the inline assembly documentation for details.

The core components of Yul are functions, blocks, variables, literals, for-loops, if-statements, switch-statements, expressions and assignments to variables.

Yul is typed, both variables and literals must specify the type with postfix notation. The supported types are `bool`, `u8`, `s8`, `u32`, `s32`, `u64`, `s64`, `u128`, `s128`, `u256` and `s256`.

Yul in itself does not even provide operators. If the EVM is targeted, opcodes will be available as built-in functions, but they can be reimplemented if the backend changes. For a list of mandatory built-in functions, see the section below.

The following example program assumes that the EVM opcodes `mul`, `div` and `mod` are available either natively or as functions and computes exponentiation.

```
{
    function power(base:u256, exponent:u256) -> result:u256
    {
        switch exponent
        case 0:u256 { result := 1:u256 }
        case 1:u256 { result := base }
        default:
        {
            result := power(mul(base, base), div(exponent, 2:u256))
            switch mod(exponent, 2:u256)
                case 1:u256 { result := mul(base, result) }
        }
    }
}
```

It is also possible to implement the same function using a for-loop instead of with recursion. Here, we need the EVM opcodes `lt` (less-than) and `add` to be available.

```
{
    function power(base:u256, exponent:u256) -> result:u256
    {
        result := 1:u256
        for { let i := 0:u256 } lt(i, exponent) { i := add(i, 1:u256) }
        {
            result := mul(result, base)
        }
    }
}
```

7.9.1 Specification of Yul

This chapter describes Yul code. It is usually placed inside a Yul object, which is described in the following chapter.

Grammar:

```
Block = '{' Statement* '}'
Statement =
    Block |
    FunctionDefinition |
    VariableDeclaration |
    Assignment |
    Expression |
    Switch |
    ForLoop |
    BreakContinue
FunctionDefinition =
    'function' Identifier '(' TypedIdentifierList? ')'
    ( '->' TypedIdentifierList )? Block
VariableDeclaration =
    'let' TypedIdentifierList ( ':=' Expression )?
Assignment =
```

(continues on next page)

(continued from previous page)

```
    IdentifierList ':=' Expression
Expression =
    FunctionCall | Identifier | Literal
If =
    'if' Expression Block
Switch =
    'switch' Expression Case* ( 'default' Block )?
Case =
    'case' Literal Block
ForLoop =
    'for' Block Expression Block Block
BreakContinue =
    'break' | 'continue'
FunctionCall =
    Identifier '(' ( Expression ( ',' Expression )* )? ')'
Identifier = [a-zA-Z_$] [a-zA-Z_0-9]*
IdentifierList = Identifier ( ',' Identifier)*
TypeName = Identifier | BuiltinTypeName
BuiltinTypeName = 'bool' | [us] ( '8' | '32' | '64' | '128' | '256' )
TypedIdentifierList = Identifier ':' TypeName ( ',' Identifier ':' TypeName )*
Literal =
    (NumberLiteral | StringLiteral | HexLiteral | TrueLiteral | FalseLiteral) ':'␣
↪TypeName
NumberLiteral = HexNumber | DecimalNumber
HexLiteral = 'hex' ('"' ([0-9a-fA-F]{2})* '"' | '\'' ([0-9a-fA-F]{2})* '\'')
StringLiteral = '"' ([^"\r\n\\] | '\\' .)* '"'
TrueLiteral = 'true'
FalseLiteral = 'false'
HexNumber = '0x' [0-9a-fA-F]+
DecimalNumber = [0-9]+
```

Restrictions on the Grammar

Switches must have at least one case (including the default case). If all possible values of the expression is covered, the default case should not be allowed (i.e. a switch with a `bool` expression and having both a true and false case should not allow a default case).

Every expression evaluates to zero or more values. Identifiers and Literals evaluate to exactly one value and function calls evaluate to a number of values equal to the number of return values of the function called.

In variable declarations and assignments, the right-hand-side expression (if present) has to evaluate to a number of values equal to the number of variables on the left-hand-side. This is the only situation where an expression evaluating to more than one value is allowed.

Expressions that are also statements (i.e. at the block level) have to evaluate to zero values.

In all other situations, expressions have to evaluate to exactly one value.

The `continue` and `break` statements can only be used inside loop bodies and have to be in the same function as the loop (or both have to be at the top level). The condition part of the for-loop has to evaluate to exactly one value.

Literals cannot be larger than the their type. The largest type defined is 256-bit wide.

Scoping Rules

Scopes in Yul are tied to Blocks (exceptions are functions and the for loop as explained below) and all declarations (`FunctionDefinition`, `VariableDeclaration`) introduce new identifiers into these scopes.

Identifiers are visible in the block they are defined in (including all sub-nodes and sub-blocks). As an exception, identifiers defined in the "init" part of the for-loop (the first block) are visible in all other parts of the for-loop (but not outside of the loop). Identifiers declared in the other parts of the for loop respect the regular syntatical scoping rules. The parameters and return parameters of functions are visible in the function body and their names cannot overlap.

Variables can only be referenced after their declaration. In particular, variables cannot be referenced in the right hand side of their own variable declaration. Functions can be referenced already before their declaration (if they are visible).

Shadowing is disallowed, i.e. you cannot declare an identifier at a point where another identifier with the same name is also visible, even if it is not accessible.

Inside functions, it is not possible to access a variable that was declared outside of that function.

Formal Specification

We formally specify Yul by providing an evaluation function E overloaded on the various nodes of the AST. Any functions can have side effects, so E takes two state objects and the AST node and returns two new state objects and a variable number of other values. The two state objects are the global state object (which in the context of the EVM is the memory, storage and state of the blockchain) and the local state object (the state of local variables, i.e. a segment of the stack in the EVM). If the AST node is a statement, E returns the two state objects and a "mode", which is used for the `break` and `continue` statements. If the AST node is an expression, E returns the two state objects and as many values as the expression evaluates to.

The exact nature of the global state is unspecified for this high level description. The local state `L` is a mapping of identifiers `i` to values `v`, denoted as `L[i] = v`.

For an identifier `v`, let `$v` be the name of the identifier.

We will use a destructuring notation for the AST nodes.

```
E(G, L, <{St1, ..., Stn}>: Block) =
    let G1, L1, mode = E(G, L, St1, ..., Stn)
    let L2 be a restriction of L1 to the identifiers of L
    G1, L2, mode
E(G, L, St1, ..., Stn: Statement) =
    if n is zero:
        G, L, regular
    else:
        let G1, L1, mode = E(G, L, St1)
        if mode is regular then
            E(G1, L1, St2, ..., Stn)
        otherwise
            G1, L1, mode
E(G, L, FunctionDefinition) =
    G, L, regular
E(G, L, <let var1, ..., varn := rhs>: VariableDeclaration) =
    E(G, L, <var1, ..., varn := rhs>: Assignment)
E(G, L, <let var1, ..., varn>: VariableDeclaration) =
    let L1 be a copy of L where L1[$vari] = 0 for i = 1, ..., n
    G, L1, regular
E(G, L, <var1, ..., varn := rhs>: Assignment) =
    let G1, L1, v1, ..., vn = E(G, L, rhs)
    let L2 be a copy of L1 where L2[$vari] = vi for i = 1, ..., n
```

(continues on next page)

(continued from previous page)

```
    G, L2, regular
E(G, L, <for { i1, ..., in } condition post body>: ForLoop) =
    if n >= 1:
        let G1, L1, mode = E(G, L, i1, ..., in)
        // mode has to be regular due to the syntactic restrictions
        let G2, L2, mode = E(G1, L1, for {} condition post body)
        // mode has to be regular due to the syntactic restrictions
        let L3 be the restriction of L2 to only variables of L
        G2, L3, regular
    else:
        let G1, L1, v = E(G, L, condition)
        if v is false:
            G1, L1, regular
        else:
            let G2, L2, mode = E(G1, L, body)
            if mode is break:
                G2, L2, regular
            else:
                G3, L3, mode = E(G2, L2, post)
                E(G3, L3, for {} condition post body)
E(G, L, break: BreakContinue) =
    G, L, break
E(G, L, continue: BreakContinue) =
    G, L, continue
E(G, L, <if condition body>: If) =
    let G0, L0, v = E(G, L, condition)
    if v is true:
        E(G0, L0, body)
    else:
        G0, L0, regular
E(G, L, <switch condition case l1:t1 st1 ... case ln:tn stn>: Switch) =
    E(G, L, switch condition case l1:t1 st1 ... case ln:tn stn default {})
E(G, L, <switch condition case l1:t1 st1 ... case ln:tn stn default st'>: Switch) =
    let G0, L0, v = E(G, L, condition)
    // i = 1 .. n
    // Evaluate literals, context doesn't matter
    let _, _, v1 = E(G0, L0, l1)
    ...
    let _, _, vn = E(G0, L0, ln)
    if there exists smallest i such that vi = v:
        E(G0, L0, sti)
    else:
        E(G0, L0, st')

E(G, L, <name>: Identifier) =
    G, L, L[$name]
E(G, L, <fname(arg1, ..., argn)>: FunctionCall) =
    G1, L1, vn = E(G, L, argn)
    ...
    G(n-1), L(n-1), v2 = E(G(n-2), L(n-2), arg2)
    Gn, Ln, v1 = E(G(n-1), L(n-1), arg1)
    Let <function fname (param1, ..., paramn) -> ret1, ..., retm block>
    be the function of name $fname visible at the point of the call.
    Let L' be a new local state such that
    L'[$parami] = vi and L'[$reti] = 0 for all i.
    Let G'', L'', mode = E(Gn, L', block)
    G'', Ln, L''[$ret1], ..., L''[$retm]
```

(continues on next page)

(continued from previous page)

```
E(G, L, l: HexLiteral) = G, L, hexString(l),
    where hexString decodes l from hex and left-aligns it into 32 bytes
E(G, L, l: StringLiteral) = G, L, utf8EncodeLeftAligned(l),
    where utf8EncodeLeftAligned performs a utf8 encoding of l
    and aligns it left into 32 bytes
E(G, L, n: HexNumber) = G, L, hex(n)
    where hex is the hexadecimal decoding function
E(G, L, n: DecimalNumber) = G, L, dec(n),
    where dec is the decimal decoding function
```

Type Conversion Functions

Yul has no support for implicit type conversion and therefore functions exist to provide explicit conversion. When converting a larger type to a shorter type a runtime exception can occur in case of an overflow.

Truncating conversions are supported between the following types:

- `bool`
- `u32`
- `u64`
- `u256`
- `s256`

For each of these a type conversion function exists having the prototype in the form of `<input_type>to<output_type>(x:<input_type>) -> y:<output_type>`, such as `u32tobool(x:u32) -> y:bool`, `u256tou32(x:u256) -> y:u32` or `s256tou256(x:s256) -> y:u256`.

Note: `u32tobool(x:u32) -> y:bool` can be implemented as `y := not(iszerou256(x))` and `booltou32(x:bool) -> y:u32` can be implemented as `switch x case true:bool { y := 1:u32 } case false:bool { y := 0:u32 }`

Low-level Functions

The following functions must be available:

Logic	
not(x:bool) -> z:bool	logical not
and(x:bool, y:bool) -> z:bool	logical and
or(x:bool, y:bool) -> z:bool	logical or
xor(x:bool, y:bool) -> z:bool	xor
Arithmetic	
addu256(x:u256, y:u256) -> z:u256	x + y
subu256(x:u256, y:u256) -> z:u256	x - y
mulu256(x:u256, y:u256) -> z:u256	x * y
divu256(x:u256, y:u256) -> z:u256	x / y
divs256(x:s256, y:s256) -> z:s256	x / y, for signed numbers in two's comple
modu256(x:u256, y:u256) -> z:u256	x % y

Table 2 – continued

mods256(x:s256, y:s256) -> z:s256	x % y, for signed numbers in two's compl
signextendu256(i:u256, x:u256) -> z:u256	sign extend from (i*8+7)th bit counting fr
expu256(x:u256, y:u256) -> z:u256	x to the power of y
addmodu256(x:u256, y:u256, m:u256) -> z:u256	(x + y) % m with arbitrary precision arithr
mulmodu256(x:u256, y:u256, m:u256) -> z:u256	(x * y) % m with arbitrary precision arithr
ltu256(x:u256, y:u256) -> z:bool	true if x < y, false otherwise
gtu256(x:u256, y:u256) -> z:bool	true if x > y, false otherwise
sltu256(x:s256, y:s256) -> z:bool	true if x < y, false otherwise (for signed nu
sgtu256(x:s256, y:s256) -> z:bool	true if x > y, false otherwise (for signed nu
equ256(x:u256, y:u256) -> z:bool	true if x == y, false otherwise
iszerou256(x:u256) -> z:bool	true if x == 0, false otherwise
notu256(x:u256) -> z:u256	~x, every bit of x is negated
andu256(x:u256, y:u256) -> z:u256	bitwise and of x and y
oru256(x:u256, y:u256) -> z:u256	bitwise or of x and y
xoru256(x:u256, y:u256) -> z:u256	bitwise xor of x and y
shlu256(x:u256, y:u256) -> z:u256	logical left shift of x by y
shru256(x:u256, y:u256) -> z:u256	logical right shift of x by y
saru256(x:u256, y:u256) -> z:u256	arithmetic right shift of x by y
byte(n:u256, x:u256) -> v:u256	nth byte of x, where the most significant b
Memory and storage	
mload(p:u256) -> v:u256	mem[p..(p+32))
mstore(p:u256, v:u256)	mem[p..(p+32)) := v
mstore8(p:u256, v:u256)	mem[p] := v & 0xff - only modifies a sing
sload(p:u256) -> v:u256	storage[p]
sstore(p:u256, v:u256)	storage[p] := v
msize() -> size:u256	size of memory, i.e. largest accessed mem
Execution control	
create(v:u256, p:u256, s:u256)	create new contract with code mem[p..(p+
call(g:u256, a:u256, v:u256, in:u256, insize:u256, out:u256, outsize:u256) -> r:u256	call contract at address a with input mem[
callcode(g:u256, a:u256, v:u256, in:u256, insize:u256, out:u256, outsize:u256) -> r:u256	identical to `call` but only use the code fr
delegatecall(g:u256, a:u256, in:u256, insize:u256, out:u256, outsize:u256) -> r:u256	identical to `callcode`, but also keep `ca`
abort()	abort (equals to invalid instruction on EVI
return(p:u256, s:u256)	end execution, return data mem[p..(p+s))
revert(p:u256, s:u256)	end execution, revert state changes, return
selfdestruct(a:u256)	end execution, destroy current contract an
log0(p:u256, s:u256)	log without topics and data mem[p..(p+s))
log1(p:u256, s:u256, t1:u256)	log with topic t1 and data mem[p..(p+s))
log2(p:u256, s:u256, t1:u256, t2:u256)	log with topics t1, t2 and data mem[p..(p+
log3(p:u256, s:u256, t1:u256, t2:u256, t3:u256)	log with topics t, t2, t3 and data mem[p..(
log4(p:u256, s:u256, t1:u256, t2:u256, t3:u256, t4:u256)	log with topics t1, t2, t3, t4 and data mem
State queries	
blockcoinbase() -> address:u256	current mining beneficiary
blockdifficulty() -> difficulty:u256	difficulty of the current block
blockgaslimit() -> limit:u256	block gas limit of the current block
blockhash(b:u256) -> hash:u256	hash of block nr b - only for last 256 bloc
blocknumber() -> block:u256	current block number
blocktimestamp() -> timestamp:u256	timestamp of the current block in seconds
txorigin() -> address:u256	transaction sender
txgasprice() -> price:u256	gas price of the transaction
gasleft() -> gas:u256	gas still available to execution

Table 2 – continued

balance(a:u256) -> v:u256	wei balance at address a
this() -> address:u256	address of the current contract / execution
caller() -> address:u256	call sender (excluding delegatecall)
callvalue() -> v:u256	wei sent together with the current call
calldataload(p:u256) -> v:u256	call data starting from position p (32 bytes
calldatasize() -> v:u256	size of call data in bytes
calldatacopy(t:u256, f:u256, s:u256)	copy s bytes from calldata at position f to
codesize() -> size:u256	size of the code of the current contract / e
codecopy(t:u256, f:u256, s:u256)	copy s bytes from code at position f to me
extcodesize(a:u256) -> size:u256	size of the code at address a
extcodecopy(a:u256, t:u256, f:u256, s:u256)	like codecopy(t, f, s) but take code at addr
Others	
discard(unused:bool)	discard value
discardu256(unused:u256)	discard value
splitu256tou64(x:u256) -> (x1:u64, x2:u64, x3:u64, x4:u64)	split u256 to four u64's
combineu64tou256(x1:u64, x2:u64, x3:u64, x4:u64) -> (x:u256)	combine four u64's into a single u256
keccak256(p:u256, s:u256) -> v:u256	keccak(mem[p...(p+s)))

Backends

Backends or targets are the translators from Yul to a specific bytecode. Each of the backends can expose functions prefixed with the name of the backend. We reserve `evm_` and `ewasm_` prefixes for the two proposed backends.

Backend: EVM

The EVM target will have all the underlying EVM opcodes exposed with the *evm_* prefix.

Backend: "EVM 1.5"

TBD

Backend: eWASM

TBD

7.9.2 Specification of Yul Object

Grammar:

```
TopLevelObject = 'object' '{' Code? ( Object | Data )* '}'
Object = 'object' StringLiteral '{' Code? ( Object | Data )* '}'
Code = 'code' Block
Data = 'data' StringLiteral HexLiteral
HexLiteral = 'hex' ('"' ([0-9a-fA-F]{2})* '"' | '\'' ([0-9a-fA-F]{2})* '\'')
StringLiteral = '"' ([^"\r\n\\] | '\\' .)* '"'
```

Above, `Block` refers to `Block` in the Yul code grammar explained in the previous chapter.

An example Yul Object is shown below:

```
// Code consists of a single object. A single "code" node is the code of the object.
// Every (other) named object or data section is serialized and
// made accessible to the special built-in functions datacopy / dataoffset / datasize
object {
    code {
        let size = datasize("runtime")
        let offset = allocate(size)
        // This will turn into a memory->memory copy for eWASM and
        // a codecopy for EVM
        datacopy(dataoffset("runtime"), offset, size)
        // this is a constructor and the runtime code is returned
        return(offset, size)
    }

    data "Table2" hex"4123"

    object "runtime" {
        code {
            // runtime code

            let size = datasize("Contract2")
            let offset = allocate(size)
            // This will turn into a memory->memory copy for eWASM and
            // a codecopy for EVM
            datacopy(dataoffset("Contract2"), offset, size)
            // constructor parameter is a single number 0x1234
            mstore(add(offset, size), 0x1234)
            create(offset, add(size, 32))
        }

        // Embedded object. Use case is that the outside is a factory contract,
        // and Contract2 is the code to be created by the factory
        object "Contract2" {
            code {
                // code here ...
            }

            object "runtime" {
                code {
                    // code here ...
                }
             }

             data "Table1" hex"4123"
        }
    }
}
```

7.10 Style Guide

7.10.1 Introduction

This guide is intended to provide coding conventions for writing solidity code. This guide should be thought of as an evolving document that will change over time as useful conventions are found and old conventions are rendered obsolete.

Many projects will implement their own style guides. In the event of conflicts, project specific style guides take precedence.

The structure and many of the recommendations within this style guide were taken from python's pep8 style guide.

The goal of this guide is *not* to be the right way or the best way to write solidity code. The goal of this guide is *consistency*. A quote from python's pep8 captures this concept well.

> A style guide is about consistency. Consistency with this style guide is important. Consistency within a project is more important. Consistency within one module or function is most important. But most importantly: know when to be inconsistent – sometimes the style guide just doesn't apply. When in doubt, use your best judgement. Look at other examples and decide what looks best. And don't hesitate to ask!

7.10.2 Code Layout

Indentation

Use 4 spaces per indentation level.

Tabs or Spaces

Spaces are the preferred indentation method.

Mixing tabs and spaces should be avoided.

Blank Lines

Surround top level declarations in solidity source with two blank lines.

Yes:

```
contract A {
    ...
}


contract B {
    ...
}


contract C {
    ...
}
```

No:

```
contract A {
    ...
}
contract B {
    ...
}
```

(continues on next page)

(continued from previous page)

```
contract C {
    ...
}
```

Within a contract surround function declarations with a single blank line.

Blank lines may be omitted between groups of related one-liners (such as stub functions for an abstract contract)

Yes:

```
contract A {
    function spam() public;
    function ham() public;
}


contract B is A {
    function spam() public {
        ...
    }

    function ham() public {
        ...
    }
}
```

No:

```
contract A {
    function spam() public {
        ...
    }
    function ham() public {
        ...
    }
}
```

Maximum Line Length

Keeping lines under the PEP 8 recommendation to a maximum of 79 (or 99) characters helps readers easily parse the code.

Wrapped lines should conform to the following guidelines.

1. The first argument should not be attached to the opening parenthesis.
2. One, and only one, indent should be used.
3. Each argument should fall on its own line.
4. The terminating element, `);`, should be placed on the final line by itself.

Function Calls

Yes:

```
thisFunctionCallIsReallyLong(
    longArgument1,
```

(continues on next page)

(continued from previous page)

```
    longArgument2,
    longArgument3
);
```

No:

```
thisFunctionCallIsReallyLong(longArgument1,
                             longArgument2,
                             longArgument3
);

thisFunctionCallIsReallyLong(longArgument1,
    longArgument2,
    longArgument3
);

thisFunctionCallIsReallyLong(
    longArgument1, longArgument2,
    longArgument3
);

thisFunctionCallIsReallyLong(
longArgument1,
longArgument2,
longArgument3
);

thisFunctionCallIsReallyLong(
    longArgument1,
    longArgument2,
    longArgument3);
```

Assignment Statements

Yes:

```
thisIsALongNestedMapping[being][set][to_some_value] = someFunction(
    argument1,
    argument2,
    argument3,
    argument4
);
```

No:

```
thisIsALongNestedMapping[being][set][to_some_value] = someFunction(argument1,
                                                                   argument2,
                                                                   argument3,
                                                                   argument4);
```

Event Definitions and Event Emitters

Yes:

```
event LongAndLotsOfArgs(
    address sender,
    address recipient,
```

(continues on next page)

(continued from previous page)

```
    uint256 publicKey,
    uint256 amount,
    bytes32[] options
);

LongAndLotsOfArgs(
    sender,
    recipient,
    publicKey,
    amount,
    options
);
```

No:

```
event LongAndLotsOfArgs(address sender,
                        address recipient,
                        uint256 publicKey,
                        uint256 amount,
                        bytes32[] options);

LongAndLotsOfArgs(sender,
                  recipient,
                  publicKey,
                  amount,
                  options);
```

Source File Encoding

UTF-8 or ASCII encoding is preferred.

Imports

Import statements should always be placed at the top of the file.

Yes:

```
import "owned";


contract A {
    ...
}


contract B is owned {
    ...
}
```

No:

```
contract A {
    ...
}
```

(continues on next page)

(continued from previous page)

```
import "owned";


contract B is owned {
    ...
}
```

Order of Functions

Ordering helps readers identify which functions they can call and to find the constructor and fallback definitions easier.

Functions should be grouped according to their visibility and ordered:

- constructor
- fallback function (if exists)
- external
- public
- internal
- private

Within a grouping, place the `view` and `pure` functions last.

Yes:

```
contract A {
    constructor() public {
        ...
    }

    function() external {
        ...
    }

    // External functions
    // ...

    // External functions that are view
    // ...

    // External functions that are pure
    // ...

    // Public functions
    // ...

    // Internal functions
    // ...

    // Private functions
    // ...
}
```

No:

```
contract A {

    // External functions
    // ...

    function() external {
        ...
    }

    // Private functions
    // ...

    // Public functions
    // ...

    constructor() public {
        ...
    }

    // Internal functions
    // ...
}
```

Whitespace in Expressions

Avoid extraneous whitespace in the following situations:

Immediately inside parenthesis, brackets or braces, with the exception of single line function declarations.

Yes:

```
spam(ham[1], Coin({name: "ham"}));
```

No:

```
spam( ham[ 1 ], Coin( { name: "ham" } ) );
```

Exception:

```
function singleLine() public { spam(); }
```

Immediately before a comma, semicolon:

Yes:

```
function spam(uint i, Coin coin) public;
```

No:

```
function spam(uint i , Coin coin) public ;
```

More than one space around an assignment or other operator to align with another:

Yes:

```
x = 1;
y = 2;
long_variable = 3;
```

No:

```
x             = 1;
y             = 2;
long_variable = 3;
```

Don't include a whitespace in the fallback function:

Yes:

```
function() external {
    ...
}
```

No:

```
function () external {
    ...
}
```

Control Structures

The braces denoting the body of a contract, library, functions and structs should:

- open on the same line as the declaration
- close on their own line at the same indentation level as the beginning of the declaration.
- The opening brace should be proceeded by a single space.

Yes:

```
contract Coin {
    struct Bank {
        address owner;
        uint balance;
    }
}
```

No:

```
contract Coin
{
    struct Bank {
        address owner;
        uint balance;
    }
}
```

The same recommendations apply to the control structures `if`, `else`, `while`, and `for`.

Additionally there should be a single space between the control structures `if`, `while`, and `for` and the parenthetic block representing the conditional, as well as a single space between the conditional parenthetic block and the opening brace.

Yes:

```
if (...) {
    ...
}

for (...) {
    ...
}
```

No:

```
if (...)
{
    ...
}

while(...){
}

for (...) {
    ...;}
```

For control structures whose body contains a single statement, omitting the braces is ok *if* the statement is contained on a single line.

Yes:

```
if (x < 10)
    x += 1;
```

No:

```
if (x < 10)
    someArray.push(Coin({
        name: 'spam',
        value: 42
    }));
```

For `if` blocks which have an `else` or `else if` clause, the `else` should be placed on the same line as the `if`'s closing brace. This is an exception compared to the rules of other block-like structures.

Yes:

```
if (x < 3) {
    x += 1;
} else if (x > 7) {
    x -= 1;
} else {
    x = 5;
}


if (x < 3)
    x += 1;
else
    x -= 1;
```

No:

```
if (x < 3) {
    x += 1;
}
else {
    x -= 1;
}
```

Function Declaration

For short function declarations, it is recommended for the opening brace of the function body to be kept on the same line as the function declaration.

The closing brace should be at the same indentation level as the function declaration.

The opening brace should be preceded by a single space.

Yes:

```
function increment(uint x) public pure returns (uint) {
    return x + 1;
}

function increment(uint x) public pure onlyowner returns (uint) {
    return x + 1;
}
```

No:

```
function increment(uint x) public pure returns (uint)
{
    return x + 1;
}

function increment(uint x) public pure returns (uint){
    return x + 1;
}

function increment(uint x) public pure returns (uint) {
    return x + 1;
    }

function increment(uint x) public pure returns (uint) {
    return x + 1;}
```

You should explicitly label the visibility of all functions, including constructors.

Yes:

```
function explicitlyPublic(uint val) public {
    doSomething();
}
```

No:

```
function implicitlyPublic(uint val) {
    doSomething();
}
```

The visibility modifier for a function should come before any custom modifiers.

Yes:

```
function kill() public onlyowner {
    selfdestruct(owner);
}
```

No:

```
function kill() onlyowner public {
    selfdestruct(owner);
}
```

For long function declarations, it is recommended to drop each argument onto it's own line at the same indentation level as the function body. The closing parenthesis and opening bracket should be placed on their own line as well at the same indentation level as the function declaration.

Yes:

```
function thisFunctionHasLotsOfArguments(
    address a,
    address b,
    address c,
    address d,
    address e,
    address f
)
    public
{
    doSomething();
}
```

No:

```
function thisFunctionHasLotsOfArguments(address a, address b, address c,
    address d, address e, address f) public {
    doSomething();
}

function thisFunctionHasLotsOfArguments(address a,
                                        address b,
                                        address c,
                                        address d,
                                        address e,
                                        address f) public {
    doSomething();
}

function thisFunctionHasLotsOfArguments(
    address a,
    address b,
    address c,
    address d,
    address e,
    address f) public {
    doSomething();
}
```

If a long function declaration has modifiers, then each modifier should be dropped to its own line.

Yes:

```
function thisFunctionNameIsReallyLong(address x, address y, address z)
    public
    onlyowner
    priced
    returns (address)
{
    doSomething();
}

function thisFunctionNameIsReallyLong(
    address x,
    address y,
    address z,
)
    public
    onlyowner
    priced
    returns (address)
{
    doSomething();
}
```

No:

```
function thisFunctionNameIsReallyLong(address x, address y, address z)
                                      public
                                      onlyowner
                                      priced
                                      returns (address) {
    doSomething();
}

function thisFunctionNameIsReallyLong(address x, address y, address z)
    public onlyowner priced returns (address)
{
    doSomething();
}

function thisFunctionNameIsReallyLong(address x, address y, address z)
    public
    onlyowner
    priced
    returns (address) {
    doSomething();
}
```

Multiline output parameters and return statements should follow the same style recommended for wrapping long lines found in the *Maximum Line Length* section.

Yes:

```
function thisFunctionNameIsReallyLong(
    address a,
    address b,
    address c
```

(continues on next page)

(continued from previous page)

```
)
    public
    returns (
        address someAddressName,
        uint256 LongArgument,
        uint256 Argument
    )
{
    doSomething()

    return (
        veryLongReturnArg1,
        veryLongReturnArg2,
        veryLongReturnArg3
    );
}
```

No:

```
function thisFunctionNameIsReallyLong(
    address a,
    address b,
    address c
)
    public
    returns (address someAddressName,
             uint256 LongArgument,
             uint256 Argument)
{
    doSomething()

    return (veryLongReturnArg1,
            veryLongReturnArg1,
            veryLongReturnArg1);
}
```

For constructor functions on inherited contracts whose bases require arguments, it is recommended to drop the base constructors onto new lines in the same manner as modifiers if the function declaration is long or hard to read.

Yes:

```
contract A is B, C, D {
    constructor(uint param1, uint param2, uint param3, uint param4, uint param5)
        B(param1)
        C(param2, param3)
        D(param4)
        public
    {
        // do something with param5
    }
}
```

No:

```
contract A is B, C, D {
    constructor(uint param1, uint param2, uint param3, uint param4, uint param5)
    B(param1)
```

(continues on next page)

(continued from previous page)

```
    C(param2, param3)
    D(param4)
    public
    {
        // do something with param5
    }
}

contract A is B, C, D {
    constructor(uint param1, uint param2, uint param3, uint param4, uint param5)
        B(param1)
        C(param2, param3)
        D(param4)
        public {
        // do something with param5
    }
}
```

When declaring short functions with a single statement, it is permissible to do it on a single line.

Permissible:

```
function shortFunction() public { doSomething(); }
```

These guidelines for function declarations are intended to improve readability. Authors should use their best judgement as this guide does not try to cover all possible permutations for function declarations.

Mappings

TODO

Variable Declarations

Declarations of array variables should not have a space between the type and the brackets.

Yes:

```
uint[] x;
```

No:

```
uint [] x;
```

Other Recommendations

- Strings should be quoted with double-quotes instead of single-quotes.

Yes:

```
str = "foo";
str = "Hamlet says, 'To be or not to be...'";
```

No:

```
str = 'bar';
str = '"Be yourself; everyone else is already taken." -Oscar Wilde';
```

- Surround operators with a single space on either side.

Yes:

```
x = 3;
x = 100 / 10;
x += 3 + 4;
x |= y && z;
```

No:

```
x=3;
x = 100/10;
x += 3+4;
x |= y&&z;
```

- Operators with a higher priority than others can exclude surrounding whitespace in order to denote precedence. This is meant to allow for improved readability for complex statement. You should always use the same amount of whitespace on either side of an operator:

Yes:

```
x = 2**3 + 5;
x = 2*y + 3*z;
x = (a+b) * (a-b);
```

No:

```
x = 2** 3 + 5;
x = y+z;
x +=1;
```

7.10.3 Naming Conventions

Naming conventions are powerful when adopted and used broadly. The use of different conventions can convey significant *meta* information that would otherwise not be immediately available.

The naming recommendations given here are intended to improve the readability, and thus they are not rules, but rather guidelines to try and help convey the most information through the names of things.

Lastly, consistency within a codebase should always supersede any conventions outlined in this document.

Naming Styles

To avoid confusion, the following names will be used to refer to different naming styles.

- `b` (single lowercase letter)
- `B` (single uppercase letter)
- `lowercase`
- `lower_case_with_underscores`
- `UPPERCASE`

- `UPPER_CASE_WITH_UNDERSCORES`
- `CapitalizedWords` (or CapWords)
- `mixedCase` (differs from CapitalizedWords by initial lowercase character!)
- `Capitalized_Words_With_Underscores`

Note: When using initialisms in CapWords, capitalize all the letters of the initialisms. Thus HTTPServerError is better than HttpServerError. When using initialisms is mixedCase, capitalize all the letters of the initialisms, except keep the first one lower case if it is the beginning of the name. Thus xmlHTTPRequest is better than XMLHTTPRequest.

Names to Avoid

- `l` - Lowercase letter el
- `O` - Uppercase letter oh
- `I` - Uppercase letter eye

Never use any of these for single letter variable names. They are often indistinguishable from the numerals one and zero.

Contract and Library Names

Contracts and libraries should be named using the CapWords style. Examples: `SimpleToken`, `SmartBank`, `CertificateHashRepository`, `Player`.

Struct Names

Structs should be named using the CapWords style. Examples: `MyCoin`, `Position`, `PositionXY`.

Event Names

Events should be named using the CapWords style. Examples: `Deposit`, `Transfer`, `Approval`, `BeforeTransfer`, `AfterTransfer`.

Function Names

Functions other than constructors should use mixedCase. Examples: `getBalance`, `transfer`, `verifyOwner`, `addMember`, `changeOwner`.

Function Argument Names

Function arguments should use mixedCase. Examples: `initialSupply`, `account`, `recipientAddress`, `senderAddress`, `newOwner`.

When writing library functions that operate on a custom struct, the struct should be the first argument and should always be named `self`.

Local and State Variable Names

Use mixedCase. Examples: `totalSupply`, `remainingSupply`, `balancesOf`, `creatorAddress`, `isPreSale`, `tokenExchangeRate`.

Constants

Constants should be named with all capital letters with underscores separating words. Examples: `MAX_BLOCKS`, `TOKEN_NAME`, `TOKEN_TICKER`, `CONTRACT_VERSION`.

Modifier Names

Use mixedCase. Examples: `onlyBy`, `onlyAfter`, `onlyDuringThePreSale`.

Enums

Enums, in the style of simple type declarations, should be named using the CapWords style. Examples: `TokenGroup`, `Frame`, `HashStyle`, `CharacterLocation`.

Avoiding Naming Collisions

- `single_trailing_underscore_`

This convention is suggested when the desired name collides with that of a built-in or otherwise reserved name.

General Recommendations

TODO

7.11 Common Patterns

7.11.1 Withdrawal from Contracts

The recommended method of sending funds after an effect is using the withdrawal pattern. Although the most intuitive method of sending Ether, as a result of an effect, is a direct `send` call, this is not recommended as it introduces a potential security risk. You may read more about this on the *Security Considerations* page.

This is an example of the withdrawal pattern in practice in a contract where the goal is to send the most money to the contract in order to become the "richest", inspired by King of the Ether.

In the following contract, if you are usurped as the richest, you will receive the funds of the person who has gone on to become the new richest.

```
pragma solidity >0.4.24;

contract WithdrawalContract {
    address public richest;
    uint public mostSent;

    mapping (address => uint) pendingWithdrawals;
```

(continues on next page)

(continued from previous page)

```
    constructor() public payable {
        richest = msg.sender;
        mostSent = msg.value;
    }

    function becomeRichest() public payable returns (bool) {
        if (msg.value > mostSent) {
            pendingWithdrawals[richest] += msg.value;
            richest = msg.sender;
            mostSent = msg.value;
            return true;
        } else {
            return false;
        }
    }

    function withdraw() public {
        uint amount = pendingWithdrawals[msg.sender];
        // Remember to zero the pending refund before
        // sending to prevent re-entrancy attacks
        pendingWithdrawals[msg.sender] = 0;
        msg.sender.transfer(amount);
    }
}
```

This is as opposed to the more intuitive sending pattern:

```
pragma solidity >0.4.24;

contract SendContract {
    address public richest;
    uint public mostSent;

    constructor() public payable {
        richest = msg.sender;
        mostSent = msg.value;
    }

    function becomeRichest() public payable returns (bool) {
        if (msg.value > mostSent) {
            // This line can cause problems (explained below).
            richest.transfer(msg.value);
            richest = msg.sender;
            mostSent = msg.value;
            return true;
        } else {
            return false;
        }
    }
}
```

Notice that, in this example, an attacker could trap the contract into an unusable state by causing `richest` to be the address of a contract that has a fallback function which fails (e.g. by using `revert()` or by just consuming more than the 2300 gas stipend). That way, whenever `transfer` is called to deliver funds to the "poisoned" contract, it will fail and thus also `becomeRichest` will fail, with the contract being stuck forever.

In contrast, if you use the "withdraw" pattern from the first example, the attacker can only cause his or her own withdraw to fail and not the rest of the contract's workings.

7.11.2 Restricting Access

Restricting access is a common pattern for contracts. Note that you can never restrict any human or computer from reading the content of your transactions or your contract's state. You can make it a bit harder by using encryption, but if your contract is supposed to read the data, so will everyone else.

You can restrict read access to your contract's state by **other contracts**. That is actually the default unless you declare make your state variables `public`.

Furthermore, you can restrict who can make modifications to your contract's state or call your contract's functions and this is what this section is about.

The use of **function modifiers** makes these restrictions highly readable.

```
pragma solidity ^0.4.22;

contract AccessRestriction {
    // These will be assigned at the construction
    // phase, where `msg.sender` is the account
    // creating this contract.
    address public owner = msg.sender;
    uint public creationTime = now;

    // Modifiers can be used to change
    // the body of a function.
    // If this modifier is used, it will
    // prepend a check that only passes
    // if the function is called from
    // a certain address.
    modifier onlyBy(address _account)
    {
        require(
            msg.sender == _account,
            "Sender not authorized."
        );
        // Do not forget the "_;"! It will
        // be replaced by the actual function
        // body when the modifier is used.
        _;
    }

    /// Make `_newOwner` the new owner of this
    /// contract.
    function changeOwner(address _newOwner)
        public
        onlyBy(owner)
    {
        owner = _newOwner;
    }

    modifier onlyAfter(uint _time) {
        require(
            now >= _time,
            "Function called too early."
```

(continues on next page)

(continued from previous page)

```
        );
        _;
    }

    /// Erase ownership information.
    /// May only be called 6 weeks after
    /// the contract has been created.
    function disown()
        public
        onlyBy(owner)
        onlyAfter(creationTime + 6 weeks)
    {
        delete owner;
    }

    // This modifier requires a certain
    // fee being associated with a function call.
    // If the caller sent too much, he or she is
    // refunded, but only after the function body.
    // This was dangerous before Solidity version 0.4.0,
    // where it was possible to skip the part after `_;`.
    modifier costs(uint _amount) {
        require(
            msg.value >= _amount,
            "Not enough Ether provided."
        );
        _;
        if (msg.value > _amount)
            msg.sender.send(msg.value - _amount);
    }

    function forceOwnerChange(address _newOwner)
        public
        payable
        costs(200 ether)
    {
        owner = _newOwner;
        // just some example condition
        if (uint(owner) & 0 == 1)
            // This did not refund for Solidity
            // before version 0.4.0.
            return;
        // refund overpaid fees
    }
}
```

A more specialised way in which access to function calls can be restricted will be discussed in the next example.

7.11.3 State Machine

Contracts often act as a state machine, which means that they have certain **stages** in which they behave differently or in which different functions can be called. A function call often ends a stage and transitions the contract into the next stage (especially if the contract models **interaction**). It is also common that some stages are automatically reached at a certain point in **time**.

An example for this is a blind auction contract which starts in the stage “accepting blinded bids”, then transitions to

"revealing bids" which is ended by "determine auction outcome".

Function modifiers can be used in this situation to model the states and guard against incorrect usage of the contract.

Example

In the following example, the modifier `atStage` ensures that the function can only be called at a certain stage.

Automatic timed transitions are handled by the modifier `timeTransitions`, which should be used for all functions.

Note: Modifier Order Matters. If atStage is combined with timedTransitions, make sure that you mention it after the latter, so that the new stage is taken into account.

Finally, the modifier `transitionNext` can be used to automatically go to the next stage when the function finishes.

Note: Modifier May be Skipped. This only applies to Solidity before version 0.4.0: Since modifiers are applied by simply replacing code and not by using a function call, the code in the transitionNext modifier can be skipped if the function itself uses return. If you want to do that, make sure to call nextStage manually from those functions. Starting with version 0.4.0, modifier code will run even if the function explicitly returns.

```
pragma solidity ^0.4.22;

contract StateMachine {
    enum Stages {
        AcceptingBlindedBids,
        RevealBids,
        AnotherStage,
        AreWeDoneYet,
        Finished
    }

    // This is the current stage.
    Stages public stage = Stages.AcceptingBlindedBids;

    uint public creationTime = now;

    modifier atStage(Stages _stage) {
        require(
            stage == _stage,
            "Function cannot be called at this time."
        );
        _;
    }

    function nextStage() internal {
        stage = Stages(uint(stage) + 1);
    }

    // Perform timed transitions. Be sure to mention
    // this modifier first, otherwise the guards
    // will not take the new stage into account.
    modifier timedTransitions() {
        if (stage == Stages.AcceptingBlindedBids &&
                    now >= creationTime + 10 days)
```

(continues on next page)

(continued from previous page)

```
            nextStage();
        if (stage == Stages.RevealBids &&
                now >= creationTime + 12 days)
            nextStage();
        // The other stages transition by transaction
        _;
    }

    // Order of the modifiers matters here!
    function bid()
        public
        payable
        timedTransitions
        atStage(Stages.AcceptingBlindedBids)
    {
        // We will not implement that here
    }

    function reveal()
        public
        timedTransitions
        atStage(Stages.RevealBids)
    {
    }

    // This modifier goes to the next stage
    // after the function is done.
    modifier transitionNext()
    {
        _;
        nextStage();
    }

    function g()
        public
        timedTransitions
        atStage(Stages.AnotherStage)
        transitionNext
    {
    }

    function h()
        public
        timedTransitions
        atStage(Stages.AreWeDoneYet)
        transitionNext
    {
    }

    function i()
        public
        timedTransitions
        atStage(Stages.Finished)
    {
    }
}
```

7.12 List of Known Bugs

Below, you can find a JSON-formatted list of some of the known security-relevant bugs in the Solidity compiler. The file itself is hosted in the Github repository. The list stretches back as far as version 0.3.0, bugs known to be present only in versions preceding that are not listed.

There is another file called bugs_by_version.json, which can be used to check which bugs affect a specific version of the compiler.

Contract source verification tools and also other tools interacting with contracts should consult this list according to the following criteria:

- It is mildly suspicious if a contract was compiled with a nightly compiler version instead of a released version. This list does not keep track of unreleased or nightly versions.
- It is also mildly suspicious if a contract was compiled with a version that was not the most recent at the time the contract was created. For contracts created from other contracts, you have to follow the creation chain back to a transaction and use the date of that transaction as creation date.
- It is highly suspicious if a contract was compiled with a compiler that contains a known bug and the contract was created at a time where a newer compiler version containing a fix was already released.

The JSON file of known bugs below is an array of objects, one for each bug, with the following keys:

name Unique name given to the bug

summary Short description of the bug

description Detailed description of the bug

link URL of a website with more detailed information, optional

introduced The first published compiler version that contained the bug, optional

fixed The first published compiler version that did not contain the bug anymore

publish The date at which the bug became known publicly, optional

severity Severity of the bug: very low, low, medium, high. Takes into account discoverability in contract tests, likelihood of occurrence and potential damage by exploits.

conditions Conditions that have to be met to trigger the bug. Currently, this is an object that can contain a boolean value `optimizer`, which means that the optimizer has to be switched on to enable the bug. If no conditions are given, assume that the bug is present.

```
[
    {
        "name": "OneOfTwoConstructorsSkipped",
        "summary": "If a contract has both a new-style constructor (using the␣
↪constructor keyword) and an old-style constructor (a function with the same name as␣
↪the contract) at the same time, one of them will be ignored.",
        "description": "If a contract has both a new-style constructor (using the␣
↪constructor keyword) and an old-style constructor (a function with the same name as␣
↪the contract) at the same time, one of them will be ignored. There will be a␣
↪compiler warning about the old-style constructor, so contracts only using new-style␣
↪constructors are fine.",
        "introduced": "0.4.22",
        "fixed": "0.4.23",
        "severity": "very low"
    },
    {
        "name": "ZeroFunctionSelector",
```

(continues on next page)

(continued from previous page)

```
        "summary": "It is possible to craft the name of a function such that it is␣
↪executed instead of the fallback function in very specific circumstances.",
        "description": "If a function has a selector consisting only of zeros, is␣
↪payable and part of a contract that does not have a fallback function and at most␣
↪five external functions in total, this function is called instead of the fallback␣
↪function if Ether is sent to the contract without data.",
        "fixed": "0.4.18",
        "severity": "very low"
    },
    {
        "name": "DelegateCallReturnValue",
        "summary": "The low-level .delegatecall() does not return the execution␣
↪outcome, but converts the value returned by the functioned called to a boolean␣
↪instead.",
        "description": "The return value of the low-level .delegatecall() function is␣
↪taken from a position in memory, where the call data or the return data resides.␣
↪This value is interpreted as a boolean and put onto the stack. This means if the␣
↪called function returns at least 32 zero bytes, .delegatecall() returns false even␣
↪if the call was successuful.",
        "introduced": "0.3.0",
        "fixed": "0.4.15",
        "severity": "low"
    },
    {
        "name": "ECRecoverMalformedInput",
        "summary": "The ecrecover() builtin can return garbage for malformed input.",
        "description": "The ecrecover precompile does not properly signal failure for␣
↪malformed input (especially in the 'v' argument) and thus the Solidity function can␣
↪return data that was previously present in the return area in memory.",
        "fixed": "0.4.14",
        "severity": "medium"
    },
    {
        "name": "SkipEmptyStringLiteral",
        "summary": "If \"\" is used in a function call, the following function␣
↪arguments will not be correctly passed to the function.",
        "description": "If the empty string literal \"\" is used as an argument in a␣
↪function call, it is skipped by the encoder. This has the effect that the encoding␣
↪of all arguments following this is shifted left by 32 bytes and thus the function␣
↪call data is corrupted.",
        "fixed": "0.4.12",
        "severity": "low"
    },
    {
        "name": "ConstantOptimizerSubtraction",
        "summary": "In some situations, the optimizer replaces certain numbers in the␣
↪code with routines that compute different numbers.",
        "description": "The optimizer tries to represent any number in the bytecode␣
↪by routines that compute them with less gas. For some special numbers, an incorrect␣
↪routine is generated. This could allow an attacker to e.g. trick victims about a␣
↪specific amount of ether, or function calls to call different functions (or none at␣
↪all).",
        "link": "https://blog.ethereum.org/2017/05/03/solidity-optimizer-bug/",
        "fixed": "0.4.11",
        "severity": "low",
        "conditions": {
            "optimizer": true
```

(continues on next page)

(continued from previous page)

```
        }
    },
    {
        "name": "IdentityPrecompileReturnIgnored",
        "summary": "Failure of the identity precompile was ignored.",
        "description": "Calls to the identity contract, which is used for copying␣
↪memory, ignored its return value. On the public chain, calls to the identity␣
↪precompile can be made in a way that they never fail, but this might be different␣
↪on private chains.",
        "severity": "low",
        "fixed": "0.4.7"
    },
    {
        "name": "OptimizerStateKnowledgeNotResetForJumpdest",
        "summary": "The optimizer did not properly reset its internal state at jump␣
↪destinations, which could lead to data corruption.",
        "description": "The optimizer performs symbolic execution at certain stages.␣
↪At jump destinations, multiple code paths join and thus it has to compute a common␣
↪state from the incoming edges. Computing this common state was simplified to just␣
↪use the empty state, but this implementation was not done properly. This bug can␣
↪cause data corruption.",
        "severity": "medium",
        "introduced": "0.4.5",
        "fixed": "0.4.6",
        "conditions": {
            "optimizer": true
        }
    },
    {
        "name": "HighOrderByteCleanStorage",
        "summary": "For short types, the high order bytes were not cleaned properly␣
↪and could overwrite existing data.",
        "description": "Types shorter than 32 bytes are packed together into the same␣
↪32 byte storage slot, but storage writes always write 32 bytes. For some types, the␣
↪higher order bytes were not cleaned properly, which made it sometimes possible to␣
↪overwrite a variable in storage when writing to another one.",
        "link": "https://blog.ethereum.org/2016/11/01/security-alert-solidity-
↪variables-can-overwritten-storage/",
        "severity": "high",
        "introduced": "0.1.6",
        "fixed": "0.4.4"
    },
    {
        "name": "OptimizerStaleKnowledgeAboutSHA3",
        "summary": "The optimizer did not properly reset its knowledge about SHA3␣
↪operations resulting in some hashes (also used for storage variable positions) not␣
↪being calculated correctly.",
        "description": "The optimizer performs symbolic execution in order to save re-
↪evaluating expressions whose value is already known. This knowledge was not␣
↪properly reset across control flow paths and thus the optimizer sometimes thought␣
↪that the result of a SHA3 operation is already present on the stack. This could␣
↪result in data corruption by accessing the wrong storage slot.",
        "severity": "medium",
        "fixed": "0.4.3",
        "conditions": {
            "optimizer": true
        }
```

(continues on next page)

(continued from previous page)

```
    },
    {
        "name": "LibrariesNotCallableFromPayableFunctions",
        "summary": "Library functions threw an exception when called from a call that␣
↪received Ether.",
        "description": "Library functions are protected against sending them Ether␣
↪through a call. Since the DELEGATECALL opcode forwards the information about how␣
↪much Ether was sent with a call, the library function incorrectly assumed that␣
↪Ether was sent to the library and threw an exception.",
        "severity": "low",
        "introduced": "0.4.0",
        "fixed": "0.4.2"
    },
    {
        "name": "SendFailsForZeroEther",
        "summary": "The send function did not provide enough gas to the recipient if␣
↪no Ether was sent with it.",
        "description": "The recipient of an Ether transfer automatically receives a␣
↪certain amount of gas from the EVM to handle the transfer. In the case of a zero-
↪transfer, this gas is not provided which causes the recipient to throw an exception.
↪",
        "severity": "low",
        "fixed": "0.4.0"
    },
    {
        "name": "DynamicAllocationInfiniteLoop",
        "summary": "Dynamic allocation of an empty memory array caused an infinite␣
↪loop and thus an exception.",
        "description": "Memory arrays can be created provided a length. If this␣
↪length is zero, code was generated that did not terminate and thus consumed all gas.
↪",
        "severity": "low",
        "fixed": "0.3.6"
    },
    {
        "name": "OptimizerClearStateOnCodePathJoin",
        "summary": "The optimizer did not properly reset its internal state at jump␣
↪destinations, which could lead to data corruption.",
        "description": "The optimizer performs symbolic execution at certain stages.␣
↪At jump destinations, multiple code paths join and thus it has to compute a common␣
↪state from the incoming edges. Computing this common state was not done correctly.␣
↪This bug can cause data corruption, but it is probably quite hard to use for␣
↪targeted attacks.",
        "severity": "low",
        "fixed": "0.3.6",
        "conditions": {
            "optimizer": true
        }
    },
    {
        "name": "CleanBytesHigherOrderBits",
        "summary": "The higher order bits of short bytesNN types were not cleaned␣
↪before comparison.",
        "description": "Two variables of type bytesNN were considered different if␣
↪their higher order bits, which are not part of the actual value, were different. An␣
↪attacker might use this to reach seemingly unreachable code paths by providing␣
↪incorrectly formatted input data.",
```

(continues on next page)

(continued from previous page)

```
        "severity": "medium/high",
        "fixed": "0.3.3"
    },
    {
        "name": "ArrayAccessCleanHigherOrderBits",
        "summary": "Access to array elements for arrays of types with less than 32␣
↪bytes did not correctly clean the higher order bits, causing corruption in other␣
↪array elements.",
        "description": "Multiple elements of an array of values that are shorter than␣
↪17 bytes are packed into the same storage slot. Writing to a single element of such␣
↪an array did not properly clean the higher order bytes and thus could lead to data␣
↪corruption.",
        "severity": "medium/high",
        "fixed": "0.3.1"
    },
    {
        "name": "AncientCompiler",
        "summary": "This compiler version is ancient and might contain several␣
↪undocumented or undiscovered bugs.",
        "description": "The list of bugs is only kept for compiler versions starting␣
↪from 0.3.0, so older versions might contain undocumented bugs.",
        "severity": "high",
        "fixed": "0.3.0"
    }
]
```

7.13 Contributing

Help is always appreciated!

To get started, you can try *Building from Source* in order to familiarize yourself with the components of Solidity and the build process. Also, it may be useful to become well-versed at writing smart-contracts in Solidity.

In particular, we need help in the following areas:

- Improving the documentation
- Responding to questions from other users on StackExchange and the Solidity Gitter
- Fixing and responding to Solidity's GitHub issues, especially those tagged as up-for-grabs which are meant as introductory issues for external contributors.

7.13.1 How to Report Issues

To report an issue, please use the GitHub issues tracker. When reporting issues, please mention the following details:

- Which version of Solidity you are using
- What was the source code (if applicable)
- Which platform are you running on
- How to reproduce the issue
- What was the result of the issue
- What the expected behaviour is

Reducing the source code that caused the issue to a bare minimum is always very helpful and sometimes even clarifies a misunderstanding.

7.13.2 Workflow for Pull Requests

In order to contribute, please fork off of the `develop` branch and make your changes there. Your commit messages should detail *why* you made your change in addition to *what* you did (unless it is a tiny change).

If you need to pull in any changes from `develop` after making your fork (for example, to resolve potential merge conflicts), please avoid using `git merge` and instead, `git rebase` your branch.

Additionally, if you are writing a new feature, please ensure you write appropriate Boost test cases and place them under `test/`.

However, if you are making a larger change, please consult with the Solidity Development Gitter channel (different from the one mentioned above, this on is focused on compiler and language development instead of language use) first.

New features and bugfixes should be added to the `Changelog.md` file: please follow the style of previous entries, when applicable.

Finally, please make sure you respect the coding style for this project. Also, even though we do CI testing, please test your code and ensure that it builds locally before submitting a pull request.

Thank you for your help!

7.13.3 Running the compiler tests

Solidity includes different types of tests. They are included in the application called `soltest`. Some of them require the `cpp-ethereum` client in testing mode, some others require `libz3` to be installed.

`soltest` reads test contracts that are annotated with expected results stored in `./test/libsolidity/syntaxTests`. In order for soltest to find these tests the root test directory has to be specified using the `--testpath` command line option, e.g. `./build/test/soltest -- --testpath ./test`.

To disable the z3 tests, use `./build/test/soltest -- --no-smt --testpath ./test` and to run a subset of the tests that do not require `cpp-ethereum`, use `./build/test/soltest -- --no-ipc --testpath ./test`.

For all other tests, you need to install cpp-ethereum and run it in testing mode: `eth --test -d /tmp/testeth`.

Then you run the actual tests: `./build/test/soltest -- --ipcpath /tmp/testeth/geth.ipc --testpath ./test`.

To run a subset of tests, filters can be used: `soltest -t TestSuite/TestName -- --ipcpath /tmp/testeth/geth.ipc --testpath ./test`, where `TestName` can be a wildcard `*`.

Alternatively, there is a testing script at `scripts/test.sh` which executes all tests and runs `cpp-ethereum` automatically if it is in the path (but does not download it).

Travis CI even runs some additional tests (including `solc-js` and testing third party Solidity frameworks) that require compiling the Emscripten target.

Writing and running syntax tests

As mentioned above, syntax tests are stored in individual contracts. These files must contain annotations, stating the expected result(s) of the respective test. The test suite will compile and check them against the given expectations.

Example: `./test/libsolidity/syntaxTests/double_stateVariable_declaration.sol`

```
contract test {
    uint256 variable;
    uint128 variable;
}
// ----
// DeclarationError: Identifier already declared.
```

A syntax test must contain at least the contract under test itself, followed by the separator `----`. The additional comments above are used to describe the expected compiler errors or warnings. This section can be empty in case that the contract should compile without any errors or warnings.

In the above example, the state variable `variable` was declared twice, which is not allowed. This will result in a `DeclarationError` stating that the identifier was already declared.

The tool that is being used for those tests is called `isoltest` and can be found under `./test/tools/`. It is an interactive tool which allows editing of failing contracts using your preferred text editor. Let's try to break this test by removing the second declaration of `variable`:

```
contract test {
    uint256 variable;
}
// ----
// DeclarationError: Identifier already declared.
```

Running `./test/isoltest` again will result in a test failure:

```
syntaxTests/double_stateVariable_declaration.sol: FAIL
    Contract:
        contract test {
            uint256 variable;
        }

    Expected result:
        DeclarationError: Identifier already declared.
    Obtained result:
        Success
```

which prints the expected result next to the obtained result, but also provides a way to change edit / update / skip the current contract or to even quit. `isoltest` offers several options for failing tests:

- edit: `isoltest` will try to open the editor that was specified before using `isoltest --editor /path/to/editor`. If no path was set, this will result in a runtime error. In case an editor was specified, this will open it such that the contract can be adjusted.
- update: Updates the contract under test. This will either remove the annotation which contains the exception not met or will add missing expectations. The test will then be run again.
- skip: Skips the execution of this particular test.
- quit: Quits `isoltest`.

Automatically updating the test above will change it to

```
contract test {
    uint256 variable;
}
// ----
```

and re-run the test. It will now pass again:

```
Re-running test case...
syntaxTests/double_stateVariable_declaration.sol: OK
```

Note: Please choose a name for the contract file, that is self-explainatory in the sense of what is been tested, e.g. `double_variable_declaration.sol`. Do not put more than one contract into a single file. `isoltest` is currently not able to recognize them individually.

7.13.4 Running the Fuzzer via AFL

Fuzzing is a technique that runs programs on more or less random inputs to find exceptional execution states (segmentation faults, exceptions, etc). Modern fuzzers are clever and do a directed search inside the input. We have a specialized binary called `solfuzzer` which takes source code as input and fails whenever it encounters an internal compiler error, segmentation fault or similar, but does not fail if e.g. the code contains an error. This way, internal problems in the compiler can be found by fuzzing tools.

We mainly use AFL for fuzzing. You need to download and install AFL packages from your repos (afl, afl-clang) or build them manually. Next, build Solidity (or just the `solfuzzer` binary) with AFL as your compiler:

```
cd build
# if needed
make clean
cmake .. -DCMAKE_C_COMPILER=path/to/afl-gcc -DCMAKE_CXX_COMPILER=path/to/afl-g++
make solfuzzer
```

At this stage you should be able to see a message similar to the following:

```
Scanning dependencies of target solfuzzer
[ 98%] Building CXX object test/tools/CMakeFiles/solfuzzer.dir/fuzzer.cpp.o
afl-cc 2.52b by <lcamtuf@google.com>
afl-as 2.52b by <lcamtuf@google.com>
[+] Instrumented 1949 locations (64-bit, non-hardened mode, ratio 100%).
[100%] Linking CXX executable solfuzzer
```

If the instrumentation messages did not appear, try switching the cmake flags pointing to AFL's clang binaries:

```
# if previously failed
make clean
cmake .. -DCMAKE_C_COMPILER=path/to/afl-clang -DCMAKE_CXX_COMPILER=path/to/afl-clang++
make solfuzzer
```

Othwerise, upon execution the fuzzer will halt with an error saying binary is not instrumented:

```
afl-fuzz 2.52b by <lcamtuf@google.com>
... (truncated messages)
[*] Validating target binary...

[-] Looks like the target binary is not instrumented! The fuzzer depends on
    compile-time instrumentation to isolate interesting test cases while
    mutating the input data. For more information, and for tips on how to
    instrument binaries, please see /usr/share/doc/afl-doc/docs/README.

    When source code is not available, you may be able to leverage QEMU
    mode support. Consult the README for tips on how to enable this.
```

(continues on next page)

(continued from previous page)

```
    (It is also possible to use afl-fuzz as a traditional, "dumb" fuzzer.
    For that, you can use the -n option - but expect much worse results.)

[-] PROGRAM ABORT : No instrumentation detected
         Location : check_binary(), afl-fuzz.c:6920
```

Next, you need some example source files. This will make it much easier for the fuzzer to find errors. You can either copy some files from the syntax tests or extract test files from the documentation or the other tests:

```
mkdir /tmp/test_cases
cd /tmp/test_cases
# extract from tests:
path/to/solidity/scripts/isolate_tests.py path/to/solidity/test/libsolidity/
↪SolidityEndToEndTest.cpp
# extract from documentation:
path/to/solidity/scripts/isolate_tests.py path/to/solidity/docs docs
```

The AFL documentation states that the corpus (the initial input files) should not be too large. The files themselves should not be larger than 1 kB and there should be at most one input file per functionality, so better start with a small number of input files. There is also a tool called `afl-cmin` that can trim input files that result in similar behaviour of the binary.

Now run the fuzzer (the `-m` extends the size of memory to 60 MB):

```
afl-fuzz -m 60 -i /tmp/test_cases -o /tmp/fuzzer_reports -- /path/to/solfuzzer
```

The fuzzer will create source files that lead to failures in `/tmp/fuzzer_reports`. Often it finds many similar source files that produce the same error. You can use the tool `scripts/uniqueErrors.sh` to filter out the unique errors.

7.13.5 Whiskers

Whiskers is a templating system similar to Mustache. It is used by the compiler in various places to aid readability, and thus maintainability and verifiability, of the code.

The syntax comes with a substantial difference to Mustache: the template markers `{{` and `}}` are replaced by < and > in order to aid parsing and avoid conflicts with *Inline Assembly* (The symbols < and > are invalid in inline assembly, while `{` and `}` are used to delimit blocks). Another limitation is that lists are only resolved one depth and they will not recurse. This may change in the future.

A rough specification is the following:

Any occurrence of `<name>` is replaced by the string-value of the supplied variable `name` without any escaping and without iterated replacements. An area can be delimited by `<#name>...</name>`. It is replaced by as many concatenations of its contents as there were sets of variables supplied to the template system, each time replacing any `<inner>` items by their respective value. Top-level variables can also be used inside such areas.

7.14 Frequently Asked Questions

This list was originally compiled by fivedogit.

7.14.1 Basic Questions

Is it possible to do something on a specific block number? (e.g. publish a contract or execute a transaction)

Transactions are not guaranteed to happen on the next block or any future specific block, since it is up to the miners to include transactions and not up to the submitter of the transaction. This applies to function calls/transactions and contract creation transactions.

If you want to schedule future calls of your contract, you can use the alarm clock.

What is the transaction "payload"?

This is just the bytecode "data" sent along with the request.

Is there a decompiler available?

There is no exact decompiler to Solidity, but Porosity is close. Because some information like variable names, comments, and source code formatting is lost in the compilation process, it is not possible to completely recover the original source code.

Bytecode can be disassembled to opcodes, a service that is provided by several blockchain explorers.

Contracts on the blockchain should have their original source code published if they are to be used by third parties.

Create a contract that can be killed and return funds

First, a word of warning: Killing contracts sounds like a good idea, because "cleaning up" is always good, but as seen above, it does not really clean up. Furthermore, if Ether is sent to removed contracts, the Ether will be forever lost.

If you want to deactivate your contracts, it is preferable to **disable** them by changing some internal state which causes all functions to throw. This will make it impossible to use the contract and ether sent to the contract will be returned automatically.

Now to answering the question: Inside a constructor, `msg.sender` is the creator. Save it. Then `selfdestruct(creator);` to kill and return funds.

example

Note that if you `import "mortal"` at the top of your contracts and declare `contract SomeContract is mortal { ...` and compile with a compiler that already has it (which includes Remix), then `kill()` is taken care of for you. Once a contract is "mortal", then you can `contractname.kill.sendTransaction({from:eth.coinbase})`, just the same as my examples.

Can you return an array or a `string` from a solidity function call?

Yes. See array_receiver_and_returner.sol.

What is problematic, though, is returning any variably-sized data (e.g. a variably-sized array like `uint[]`) from a function **called from within Solidity**. This is a limitation of the EVM and will be solved with the next protocol update.

Returning variably-sized data as part of an external transaction or call is fine.

Is it possible to in-line initialize an array like so: `string[] myarray = ["a", "b"];`

Yes. However it should be noted that this currently only works with statically sized memory arrays. You can even create an inline memory array in the return statement. Pretty cool, huh?

Example:

```
pragma solidity ^0.4.16;

contract C {
    function f() public pure returns (uint8[5] memory) {
        string[4] memory adaArr = ["This", "is", "an", "array"];
        return ([1, 2, 3, 4, 5]);
    }
}
```

Can a contract function return a `struct`?

Yes, but only in `internal` function calls.

If I return an `enum`, I only get integer values in web3.js. How to get the named values?

Enums are not supported by the ABI, they are just supported by Solidity. You have to do the mapping yourself for now, we might provide some help later.

Can state variables be initialized in-line?

Yes, this is possible for all types (even for structs). However, for arrays it should be noted that you must declare them as static memory arrays.

Examples:

```
pragma solidity ^0.4.0;

contract C {
    struct S {
        uint a;
        uint b;
    }

    S public x = S(1, 2);
    string name = "Ada";
    string[4] adaArr = ["This", "is", "an", "array"];
}

contract D {
    C c = new C();
}
```

How do structs work?

See struct_and_for_loop_tester.sol.

How do for loops work?

Very similar to JavaScript. Such as the following example:

```
for (uint i = 0; i < a.length; i ++) { a[i] = i; }
```

See struct_and_for_loop_tester.sol.

What are some examples of basic string manipulation (`substring`, `indexOf`, `charAt`, etc)?

There are some string utility functions at stringUtils.sol which will be extended in the future. In addition, Arachnid has written solidity-stringutils.

For now, if you want to modify a string (even when you only want to know its length), you should always convert it to a `bytes` first:

```
pragma solidity ^0.4.0;

contract C {
    string s;

    function append(byte c) public {
        bytes(s).push(c);
    }

    function set(uint i, byte c) public {
        bytes(s)[i] = c;
    }
}
```

Can I concatenate two strings?

You have to do it manually for now.

Why is the low-level function `.call()` less favorable than instantiating a contract with a variable (`ContractB b;`) and executing its functions (`b.doSomething();`)?

If you use actual functions, the compiler will tell you if the types or your arguments do not match, if the function does not exist or is not visible and it will do the packing of the arguments for you.

See ping.sol and pong.sol.

Is unused gas automatically refunded?

Yes and it is immediate, i.e. done as part of the transaction.

When returning a value of say `uint` type, is it possible to return an `undefined` or "null"-like value?

This is not possible, because all types use up the full value range.

You have the option to `throw` on error, which will also revert the whole transaction, which might be a good idea if you ran into an unexpected situation.

If you do not want to throw, you can return a pair:

```
pragma solidity >0.4.23 <0.5.0;

contract C {
    uint[] counters;

    function getCounter(uint index)
        public
        view
        returns (uint counter, bool error) {
            if (index >= counters.length)
                return (0, true);
            else
                return (counters[index], false);
    }

    function checkCounter(uint index) public view {
        (uint counter, bool error) = getCounter(index);
        if (error) {
            // ...
        } else {
            // ...
        }
    }
}
```

Are comments included with deployed contracts and do they increase deployment gas?

No, everything that is not needed for execution is removed during compilation. This includes, among others, comments, variable names and type names.

What happens if you send ether along with a function call to a contract?

It gets added to the total balance of the contract, just like when you send ether when creating a contract. You can only send ether along to a function that has the `payable` modifier, otherwise an exception is thrown.

Is it possible to get a tx receipt for a transaction executed contract-to-contract?

No, a function call from one contract to another does not create its own transaction, you have to look in the overall transaction. This is also the reason why several block explorer do not show Ether sent between contracts correctly.

What is the `memory` keyword? What does it do?

The Ethereum Virtual Machine has three areas where it can store items.

The first is "storage", where all the contract state variables reside. Every contract has its own storage and it is persistent between function calls and quite expensive to use.

The second is "memory", this is used to hold temporary values. It is erased between (external) function calls and is cheaper to use.

The third one is the stack, which is used to hold small local variables. It is almost free to use, but can only hold a limited amount of values.

For almost all types, you cannot specify where they should be stored, because they are copied every time they are used.

The types where the so-called storage location is important are structs and arrays. If you e.g. pass such variables in function calls, their data is not copied if it can stay in memory or stay in storage. This means that you can modify their content in the called function and these modifications will still be visible in the caller.

There are defaults for the storage location depending on which type of variable it concerns:

- state variables are always in storage
- function arguments are in memory by default
- local variables of mapping type reference storage by default
- local variables of value type (i.e. neither array, nor struct nor mapping) are stored in the stack

For local variables of struct or array type the storage location has to be stated explicitly.

Example:

```
pragma solidity ^0.4.0;

contract C {
    uint[] data1;
    uint[] data2;

    function appendOne() public {
        append(data1);
    }

    function appendTwo() public {
        append(data2);
    }

    function append(uint[] storage d) internal {
        d.push(1);
    }
}
```

The function `append` can work both on `data1` and `data2` and its modifications will be stored permanently. If you remove the `storage` keyword, the default is to use `memory` for function arguments. This has the effect that at the point where `append(data1)` or `append(data2)` is called, an independent copy of the state variable is created in memory and `append` operates on this copy (which does not support `.push` - but that is another issue). The modifications to this independent copy do not carry back to `data1` or `data2`.

Warning: Prior to version 0.5.0, a common mistake was to declare a local variable and assume that it will be created in memory, although it will be created in storage. Using such a variable without initializing could lead to unexpected behavior. Starting from 0.5.0, however, the storage location for local variables has to be specified explicitly and local storage variables have to be initialized, which should prevent these kinds of mistakes.

7.14.2 Advanced Questions

How do you get a random number in a contract? (Implement a self-returning gambling contract.)

Getting randomness right is often the crucial part in a crypto project and most failures result from bad random number generators.

If you do not want it to be safe, you build something similar to the coin flipper but otherwise, rather use a contract that supplies randomness, like the RANDAO.

Get return value from non-constant function from another contract

The key point is that the calling contract needs to know about the function it intends to call.

See ping.sol and pong.sol.

Get contract to do something when it is first mined

Use the constructor. Anything inside it will be executed when the contract is first mined.

See replicator.sol.

How do you create 2-dimensional arrays?

See 2D_array.sol.

Note that filling a 10x10 square of `uint8` + contract creation took more than `800,000` gas at the time of this writing. 17x17 took `2,000,000` gas. With the limit at 3.14 million... well, there's a pretty low ceiling for what you can create right now.

Note that merely "creating" the array is free, the costs are in filling it.

Note2: Optimizing storage access can pull the gas costs down considerably, because 32 `uint8` values can be stored in a single slot. The problem is that these optimizations currently do not work across loops and also have a problem with bounds checking. You might get much better results in the future, though.

What happens to a `struct`'s mapping when copying over a `struct`?

This is a very interesting question. Suppose that we have a contract field set up like such:

```
struct User {
    mapping(string => string) comments;
}

function somefunction public {
   User user1;
   user1.comments["Hello"] = "World";
   User user2 = user1;
}
```

In this case, the mapping of the struct being copied over into the userList is ignored as there is no "list of mapped keys". Therefore it is not possible to find out which values should be copied over.

How do I initialize a contract with only a specific amount of wei?

Currently the approach is a little ugly, but there is little that can be done to improve it. In the case of a `contract A` calling a new instance of `contract B`, parentheses have to be used around `new B` because `B.value` would refer to a member of `B` called `value`. You will need to make sure that you have both contracts aware of each other's presence and that `contract B` has a `payable` constructor. In this example:

```
pragma solidity >0.4.24;

contract B {
    constructor() public payable {}
```

(continues on next page)

(continued from previous page)

```
}

contract A {
    address child;

    function test() public {
        child = (new B).value(10)(); //construct a new B with 10 wei
    }
}
```

Can a contract function accept a two-dimensional array?

This is not yet implemented for external calls and dynamic arrays - you can only use one level of dynamic arrays.

What is the relationship between `bytes32` and `string`? Why is it that `bytes32 somevar = "stringliteral";` works and what does the saved 32-byte hex value mean?

The type `bytes32` can hold 32 (raw) bytes. In the assignment `bytes32 samevar = "stringliteral";`, the string literal is interpreted in its raw byte form and if you inspect `somevar` and see a 32-byte hex value, this is just `"stringliteral"` in hex.

The type `bytes` is similar, only that it can change its length.

Finally, `string` is basically identical to `bytes` only that it is assumed to hold the UTF-8 encoding of a real string. Since `string` stores the data in UTF-8 encoding it is quite expensive to compute the number of characters in the string (the encoding of some characters takes more than a single byte). Because of that, `string s; s.length` is not yet supported and not even index access `s[2]`. But if you want to access the low-level byte encoding of the string, you can use `bytes(s).length` and `bytes(s)[2]` which will result in the number of bytes in the UTF-8 encoding of the string (not the number of characters) and the second byte (not character) of the UTF-8 encoded string, respectively.

Can a contract pass an array (static size) or string or `bytes` (dynamic size) to another contract?

Sure. Take care that if you cross the memory / storage boundary, independent copies will be created:

```
pragma solidity ^0.4.16;

contract C {
    uint[20] x;

    function f() public {
        g(x);
        h(x);
    }

    function g(uint[20] memory y) internal pure {
        y[2] = 3;
    }

    function h(uint[20] storage y) internal {
        y[3] = 4;
    }
}
```

The call to `g(x)` will not have an effect on `x` because it needs to create an independent copy of the storage value in memory. On the other hand, `h(x)` successfully modifies `x` because only a reference and not a copy is passed.

Sometimes, when I try to change the length of an array with ex: `arrayname.length = 7;` I get a compiler error `Value must be an lvalue`. Why?

You can resize a dynamic array in storage (i.e. an array declared at the contract level) with `arrayname.length = <some new length>;`. If you get the "lvalue" error, you are probably doing one of two things wrong.

1. You might be trying to resize an array in "memory", or
2. You might be trying to resize a non-dynamic array.

```
// This will not compile

pragma solidity ^0.4.18;

contract C {
    int8[] dynamicStorageArray;
    int8[5] fixedStorageArray;

    function f() {
        int8[] memory memArr;        // Case 1
        memArr.length++;             // illegal

        int8[5] storage storageArr = fixedStorageArray;   // Case 2
        storageArr.length++;                              // illegal

        int8[] storage storageArr2 = dynamicStorageArray;
        storageArr2.length++;                     // legal


    }
}
```

Important note: In Solidity, array dimensions are declared backwards from the way you might be used to declaring them in C or Java, but they are access as in C or Java.

For example, `int8[][5] somearray;` are 5 dynamic `int8` arrays.

The reason for this is that `T[5]` is always an array of 5 `T`'s, no matter whether `T` itself is an array or not (this is not the case in C or Java).

Is it possible to return an array of strings (`string[]`) from a Solidity function?

Not yet, as this requires two levels of dynamic arrays (`string` is a dynamic array itself).

If you issue a call for an array, it is possible to retrieve the whole array? Or must you write a helper function for that?

The automatic *getter function* for a public state variable of array type only returns individual elements. If you want to return the complete array, you have to manually write a function to do that.

What could have happened if an account has storage value(s) but no code? Example: http://test.ether.camp/account/5f740b3a43fbb99724ce93a879805f4dc89178b5

The last thing a constructor does is returning the code of the contract. The gas costs for this depend on the length of the code and it might be that the supplied gas is not enough. This situation is the only one where an "out of gas" exception does not revert changes to the state, i.e. in this case the initialisation of the state variables.

https://github.com/ethereum/wiki/wiki/Subtleties

After a successful CREATE operation's sub-execution, if the operation returns x, 5 * len(x) gas is subtracted from the remaining gas before the contract is created. If the remaining gas is less than 5 * len(x), then no gas is subtracted, the code of the created contract becomes the empty string, but this is not treated as an exceptional condition - no reverts happen.

What does the following strange check do in the Custom Token contract?

```
require((balanceOf[_to] + _value) >= balanceOf[_to]);
```

Integers in Solidity (and most other machine-related programming languages) are restricted to a certain range. For `uint256`, this is `0` up to `2**256 - 1`. If the result of some operation on those numbers does not fit inside this range, it is truncated. These truncations can have serious consequences, so code like the one above is necessary to avoid certain attacks.

More Questions?

If you have more questions or your question is not answered here, please talk to us on gitter or file an issue.

Index

R

S

T

U

V

W

Y