The VampIR Book

The VampIR Team

Contents

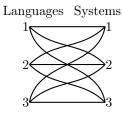
1	What is VampIR and How is it Used?			
	1.1	Why VampIR?	2	
	1.2	Using VampIR	2	
	1.3	Proof Validity and Interaction	4	
2	Programming in VampIR			
	2.1	Basic Arithmetic	5	
	2.2	Functions	6	
	2.3	Tuples	8	
	2.4	Higher-Order Functions	10	
3	Witness Generation Proof Backends		13 14	
4				
5	App	pendix A: Cookbook	15	

1 What is VampIR and How is it Used?

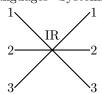
1.1 Why VampIR?

VampIR is, at its core, a language for defining arithmetic circuits over finite fields. It's intended to be a universal intermediate language supporting many proof systems based on arithmetic circuits. Any higher-level language which intends to compile to arithmetic circuits may target VampIR as an intermediate language.

The need for an intermediate language is obvious. Without an adequate intermediate, systems must support every desired proof system individually. This creates an ecosystem like that depicted in the right figure, where the total work required for interconnectedness is quadratic. This also creates more points of failure and more opportunities for support inconsistencies.



Languages Systems



With an adequate intermediate representation, the total amount of work necessary to fill out the ecosystem is only linear. This can be seen in the ecosystem depicted in the left figure. Languages can, instead of targeting a specific proof system, target the intermediate representation. That language would then have support for every proof system targeted by that intermediate representation.

VampIR's goal is to fulfill this need by providing a minimal and expressive interface for describing the core data structure used by most modern zero-knowledge-proof systems. To that end, VampIR aims to be flexible and easy to use but doesn't provide any cryptographic features of its own. It does not presup-

pose any particular implementation or design for a proof system. VampIR files are sufficiently generic that they may even be used for applications that use arithmetic circuits but are not cryptographic in nature.

1.2 Using VampIR

A very basic example of a VampIR program is the following;

```
// This is a comment!
def x = 10;
x = 10;
```

Here, one can see the two main top-level commands available in VampIR. That first line defines a constant, x, to which the value 10 is assigned. The second line is an equation that is expected to be true. Every arithmetic circuit is interpreted as a proposition. Specifically, it is the proposition corresponding to the truth of all the equations appearing in the file which generated it. In this case, the compiled circuit will merely check that 10 = 10. Notice that every line must end in a semicolon. VampIR does not generally care about white space or newlines (beyond spaces separating individual tokens); this example would be interpreted the same if all newlines were removed and everything was put on a single line.

To compile this circuit, of course, VampIR needs to be up and running. First clone VampIR's directory.

```
$ git clone git@github.com:anoma/vamp-ir
```

VampIR is implemented in Rust and can easily be compiled from source using Rust Cargo.

```
$ cd vamp-ir
$ cargo build
```

This will create the VampIR binary at '/target/debug/vamp-ir' within the main VampIR directory. VampIR does not possess any cryptographic capabilities on its own. This means that some specific parameters, such as the field size, cannot be determined by VampIR, but are instead decided at compile time. To compile this circuit, a target must first be chosen. For this example, I will choose PLONK [What is the name of the program? How does one access other targets?].

For the sake of this example, I will assume that the lines in our example program are saved into a file called 'ex1.pir', stored within a new folder called 'examples' within the main VampIR directory.

```
$ mkdir examples
$ printf "def x = 10;\n\nx = 10;">examples/ex1.pir
```

Notice the file ends with '.pir', the standard extension for VampIR files. To compile this into a PLONK circuit, one must first set up public parameters.

```
$ target/debug/vamp-ir setup -o examples/params.pp
> * Setting up public parameters...
> * Public parameter setup success!
```

This will create the file 'params.pp' within our 'examples' directory. The -o argument indicates an output file and is equivalent to --output. One can now create the circuit associated with our file.

This will create our compiled circuit in the file 'circuit.plonk' within the 'examples' directory. The -u argument indicates a universal parameter file and is equivalent to --universal-params. The -s argument indicates a source file and is equivalent to --source.

Notice that types for defined expressions are inferred during compilation. VampIR has a simple type system that is mostly implicit. This will be explained in more detail later on. In this simple example, \mathbf{x} is inferred to be an 'int', that is, an integer that will be interpreted as a field element during compilation.

A zero-knowledge proof of circuit correctness can now be synthesized.

This will create our compiled proof in the file 'proof.plonk' within the 'examples' directory. The -c argument indicates a circuit file and is equivalent to --circuit.

The last thing one may want to do is verify the circuit.

This will verify the proof just created. In this case, we've just verified a zero-knowledge proof that 10 = 10. The -p argument indicates a proof file and is equivalent to --proof.

Other than help, every available VampIR command has now been used; setup, compile, prove, and verify. These commands define all current methods for interacting with VampIR. It is a very simple system.

1.3 Proof Validity and Interaction

Values for variables do not need to be given upfront. If our file was instead

```
x = 10;
```

without declaring the value of x, VampIR would interpret this unbound variable as an input needing to be specified during proof generation. If this is saved in the file 'ex2.pir' and compiled to a circuit, one can see that VampIR will ask for an input when it's needed.

```
$ printf "x = 10;">examples/ex2.pir
$ target/debug/vamp-ir compile -u examples/params.pp \
                               -s examples/ex2.pir \
                               -o examples/circuit2.plonk
> * Compiling constraints...
> * Constraint compilation success!
$ target/debug/vamp-ir prove -u examples/params.pp \
                             -c examples/circuit2.plonk \
                             -o examples/proof2.plonk
> * Reading arithmetic circuit...
> * Soliciting circuit witnesses...
> ** x[2] (private):
$ 9
> * Reading public parameters...
> [...]
> * Proof generation success!
```

It asked for the private value for x, to which I input '9'. This should create an invalid proof this time as 9 does not equal 10. If one tries verifying the proof, they will observe that it's invalid.

As you can see, a verification error is given.

2 Programming in VampIR

2.1 Basic Arithmetic

VampIR's arithmetic denotations are fairly straightforward, being very similar to what's present in many mainstream languages.

VampIR provides four different methods to denote numbers. One can use

- standard decimal notation
- hexadecimal notation, so long as the number is prefixed by 0x
- octal notation, so long as the number is prefixed by 00
- binary notation, so long as the number is prefixed by 0b

The following example program demonstrates all the number notations.

Notice that VampIR allows for many top-level equations in the same program.

One can add, subtract, and multiply numbers, using, respectively, the infix operators +, -, and *.

```
20 * 3 - 20 = 20 + 20;
```

VampIR uses a standard order of operations where multiplication takes precedence over addition and subtraction, and everything is left-associated otherwise. Further clarification can be made using parentheses.

```
1 + 2 + 3 - 4 * 5 - 6 + 7 =
((((1 + 2) + 3) - (4 * 5)) - 6) + 7;
```

VampIR also allows for division in arithmetic circuits, denoted with the infix operator /.

```
15/3 = 5;
```

Division has the same precedence as multiplication, and associates to the left, otherwise.

```
2/3*4/5 = ((2/3)*4)/5;
```

This operation will give an output using the division of the underlying field, meaning constraints may differ in truth value depending on the compilation target. If one uses a prime field of size

52435875175126190479447740508185965837690552500527637822603658699938581184513, used, for example, in BLS12-381, then the following is valid.

```
2/3 = 0x26a48d1bb889d46d66689d580335f2ac713f36abaaaa1eaa55555555000000001;
```

If one tries dividing by zero, they will get an error at some point in the process depending on where the division occurs. 1/0 = 1; will produce an error if 'x' is set to 0 during witness solicitation.

VampIR also possesses a single unary operation in the form of negation, denoted with -. Negation must be surrounded by parentheses unless it's iterated with additional negations.

```
(-10) = 0 - 10;

0 - 10 = (--10);
```

Exponentiation is also available via the infix operator ^.

```
2 ^2 = 4;
```

Exponentiation has greater precedence than any other arithmetic operator.

```
5 * 6 ^ 7 = 5 * (6 ^ 7);
```

Only constant exponents are allowed, as they can be translated into iterated multiplication at circuit compilation time. $2 \sim x = 4$; will create an error during type inference.

Exponentiation by 0 will always produce 1, even if the base is 0.

```
2 ^0 0 = 1;

0 ^0 0 = 1;

0 ^0 2 = 0;
```

[Note: this section may describe a bug instead of intended behaviour]

Exponentiation by negatives is a bit weird. They are offset by 1 from what one might expect.

```
2 \quad (-5) = 1 / 2 \quad 4;

6 \quad (-2) = 1 / 6;

5 \quad (-3) = 1 / 5 \quad 2;
```

Exponentiating by -1 will always produce 1, with the exception of 0.

```
2 \quad (-1) = 1;

6 \quad (-1) = 1;

5 \quad (-1) = 1;

0 \quad (-1) = 0;
```

2.2 Functions

The def construct allows one to define, not just constants, but functions as well.

```
def square x = x ^ 2;
square 4 = 16;
```

The name of the function is denoted by the first token after def. Function application is denoted by juxtaposition, a la some functional languages such as Haskell.

During compilation, the type checker will identify square's type as a function from integers to integers.

```
> square[3]: (int -> int)
```

Functions can have as many arguments as you want.

```
def f x y z = x * z + y;
f 4 5 6 = 29;
```

The type of this function indicates multiple arguments.

```
> f[5]: (int -> (int -> int)))
```

This tells us that each argument will take in an integer and produce a function, until the third argument, after which it will produce an integer. This indicates that application is left-associative and functions are curried by default (more on this in section 2.4).

```
((f 4) 5) 6 = f 4 5 6;
```

Definitions can contain definitions inside of them.

```
def g x = {
  def k = 20;
  x * k
};
g 3 = 60;
```

Notice the usage of curly braces to define blocks that scope the function. Also, notice that the semicolon for the definition comes after the closing curly brace, and there is no semicolon for the last line before that brace.

Internal def statements should be used similarly to let statements common in many functional languages. Internal def statements can be functions as well.

```
def cube x = {
  def square x = x * x;
  x * square x
};
cube 3 = 27;
```

Types for internal def statements are not given during compilation, but they are still checked. There is no limit to the number of lines or nestings within a definition.

```
def power x = {
  def hypercube x = {
    def square x = x * x;
    square (square x)
  };
  def cube x = {
    def square x = x * x;
    x * square x
  };
  cube (hypercube x)
  };

power 2 = 4096;
```

Overlapping names are allowed. References to names will refer to the most recent binding.

```
def x = 4;
def x = 8;
x = 8;
```

During type checking, both variables are listed.

```
> x[2]: int
> x[3]: int
```

The numbers next to the names are, in some sense, the "true" names of the variables. Each variable has a unique number identifying it, allowing the system to avoid confusing variables with the same name.

Equations can also appear in definitions. Definitions may or may not return a value and can act as gates for other values.

```
def g1 x = {x = 4; x};
g1 4;

def g2 x = {x = 10};
g2 10;
```

Looking at the types, one sees something interesting.

```
> g2[5]: (int -> ())
```

g2 does, in fact, return a value; something of type (). This will be explained in section 2.3.

This illustrates that a VampIR file does not need any top-level equations, but will still represent a proposition through the equations enforced by function calls. g1 5; will produce an invalid proof, in this case.

Equations aren't enforced if their parent function is never actually called.

```
def j x = {0 = 1; x};
1 = 1;
```

will produce a valid proof because 'j' is never fully instantiated. Equations become part of the circuit when their parent function is fully instantiated. $def k = \{0 = 1; 4\}$; will produce an invalid proof since 'k' is already fully instantiated.

Using polynomial constraints, one can create sophisticated, reusable checks. The following, for example, checks that a field element is a boolean; either 0 or 1.

```
def isBool x = {
   (x - 1) * x = 0;
   x
};
isBool 0;
isBool 1;
```

2.3 Tuples

The only first-order data structure provided in VampIR is the tuple. One can use them in place of numbers to check multiple things in parallel.

```
def xs = (1, 2);
xs = (1, 2);
During type checking, xs's type is given as
> xs[2]: (int, int)
This indicates that xs is a pair of integers. One can also iterate tuples as much as they want.
def xs = (1, 2, 3, 4, 5, 6, 7, 8, 9);
xs = (1, 2, 3, 4, 5, 6, 7, 8, 9);
During type checking, one can observe the type
> xs[2]: (int, (int, (int, (int, (int, (int, (int, int))))))))
```

This indicates that tuples with more than two entries are, internally, just nested pairs associated to the right.

```
(1, 2, 3) = (1, (2, 3));
```

(1, 2, 3) = ((1, 2), 3); will produce a type error. One cannot compare tuples to numbers, so 1 = (1, 1); will produce an error. One also cannot treat tuples as numbers, so, for example, (1, 2) + (3, 4); will produce an error.

VampIR provides the ability to define functions over tuples using pattern matching.

```
def add (a, b) (x, y) = (a + x, b + y);
add (1, 2) (3, 4) = (4, 6);
```

One can define many standard tuple construction and manipulation functions.

```
def fst (x, y) = x;
def snd (x, y) = y;
def third (x, y, z) = z;
def dup x = (x, x);
def swap (x, y) = (y, x);
def assoc (x, (y, z)) = ((x, y), z);
```

As you can see by assoc, one can pattern match a tuple within another tuple. Additionally, the fact that n-tuples are nested pairs means that fst acts uniformly on tuples; using that definition we'd have fst (1, 2, 3) = 1.

The types here tell us something interesting. One sees, for example, that;

```
> fst[4]: (([24], [25]) -> [24])
> dup[13]: ([43] -> ([43], [43]))
> swap[16]: (([50], [51]) -> ([51], [50]))
```

The numbers in square braces denote variable types. These functions do not have a fixed type, but their types can be instantiated to be different types, depending on the context. Different calls within the same program can also have different types. For example, one can have both of these equations in the same file:

```
fst (1, 2) = 1;
fst ((1, 2), 3) = (1, 2);
```

Even though the same function is called on both lines, different types are returned each time. On the first line, it returns an integer while on the second it returns a pair.

VampIR also provides a 0-tuple denoted (), analogous to the return value of a unit type in many functional languages or a void type in many imperative languages. It's the closest thing to getting a function to return nothing.

```
def tt = ();
def f x = {
    x = 2;
    ()
};

f 2 = tt;
During type inference, the following is shown;
> tt[2]: ()
> f[4]: (int -> ())
```

indicating that tt is an element of the unit type and that f will return an element of the unit type on an integer input. This shows that functions that have only equations do the same thing as functions that explicitly return a (), as mentioned in section 2.2.

Using units, one can design some of our functions to act like list manipulation programs. That is if one thinks of (x, y) as Cons x y and () as Ni1, they can encode a list like '[1, 2, 3, 4]' as (1, 2, 3, 4, ()). The utility of this lies in the fact that each entry will appear in the head of a pair. This allows us to define more flexible programs. Modifying a few of our previous examples, one can have

```
def swap (x, y, r) = (y, x, r);
swap (1, 2, ()) = (2, 1, ());
swap (1, 2, 3, ()) = (2, 1, 3, ());
swap (1, 2, 3, 4, ()) = (2, 1, 3, 4, ());
def tail (x, y) = y;
tail (1, 2, ()) = (2, ());
tail (1, 2, 3, ()) = (2, 3, ());
tail (1, 2, 3, 4, ()) = (2, 3, 4, ());
def total (x, y, r) = x+y;
total (1, 2, ()) = 3;
total (1, 2, 3, ()) = 3;
total (1, 2, 3, 4, ()) = 3;
def add (a, b, q) (x, y, r) = (a + x, b + y, ());
add (1, 2, ()) (3, 4, ()) = (4, 6, ());
add (1, 2, 3, 4, 5, ()) (6, 7, 8, ()) = (7, 9, ());
def zip (a, b, q) (x, y, r) = ((a, x), (b, y), ());
zip (1, 2, 3, 4, 5, ()) (6, 7, 8, ()) = ((1, 6), (2, 7), ());
```

VampIR ultimately generates a fixed, finite data structure, meaning that there can't be a proper variable-length list type. As such, many list manipulation programs that one might take for granted can't be defined in full generality in VampIR. However, this technique can recover some of that functionality in practice.

Witness solicitation interacts in a somewhat interesting way with pairs. If one compiles the program

```
x = (1, 2);
```

it seems like they will need to input a pair. However, during proof creation, VampIR will ask;

```
> * Soliciting circuit witnesses...
> ** x.0[9] (private):
> ** x.1[10] (private):
```

It splits the variable into two sub-variables, named x.0 (the first element of x) and x.1 (the second element of x). So the compiled circuit does not, in fact, have a hole in the shape of a pair; rather it has two holes corresponding to the elements of the pair.

Without a clear first-order datatype for each uninstantiated variable, circuit generation is not possible. If one tries compiling

```
x = y;
```

they will get an error indicating such.

2.4 Higher-Order Functions

VampIR has rather robust support for higher-order functions. One can manipulate functions more or less like any other value. One can save a partially instantiated function.

```
def f x y = x * y;
def g = f 2;
g 3 = 6;
```

Notice that g is defined in terms of f with only one of its two arguments applied. The types are given as

```
> f[4]: (int -> (int -> int))
> g[5]: (int -> int)
```

one can see that g's type is exactly the return type of f. This is no different than if f returned an integer instead of a function.

One can also write functions that take a function as input.

```
def app2 f x = f (f x);
def times2 x = 2 * x;
app2 times2 3 = 12;
Looking at the types;
> f[4]: app2[4]: (([10] -> [10]) -> ([10] -> [10]))
```

one can see that app2 takes a function from the (variable) type 10 onto itself, and creates a new function, also from 10 onto itself.

As mentioned in section 2.2, the constraints of a function are not called unless the function is fully substantiated. The following program produces a valid proof.

```
def f x y = {0 = 1; x};
f 2;
```

While f is called, it's not fully instantiated, and so 0 = 1 does not become part of the generated circuit.

There are many higher-order combinators that may be useful for some applications.

```
def comp f g x = f (g x);
def const x y = x;
def flip f x y = f y x;
def delta f x = f x x;
```

VampIR provides a notation for anonymous functions using the fun keyword.

```
(fun x {x}) 15 = 15;

(fun x y {x + y}) 15 20 = 35;

flip (fun x y {x - y}) 15 3 = (-12);

(fun (x, y) {x + y}) 15 3 = 18;
```

Notice that one can bind multiple arguments and pattern match on tuples using anonymous functions. These can also be used as return values in definitions.

```
def id = fun x {x};
id 2 = 2;
one can also place constraints inside anonymous functions.
(fun x {x = 2; x}) 2;
(fun x {x = 5}) 5 = ();
```

as you can see, anonymous functions don't need to return data. Like defs, funs that don't return a value still return a ().

We cannot use functions in comparisons; only first-order data can be compared.

VampIR does not give unrestricted access to higher-order functions; any appearing in a program must be simply typed. For example, $fun x \{ x x \}$; will produce a typing error. This means that, in particular, one cannot define a fixed-point combinator, which would be necessary to implement general recursion. Still, there are many powerful techniques made available by what is given. In particular, one can define and utilize lambda-encoded data.

If VampIR didn't already have build-in pairs, one could implement them as

```
def pair x y = fun p { p x y };
```

By issuing different arguments to pair x y, one can get the underlying components.

```
def fst p = p (fun x y {x});
fst (pair 1 2) = 1;

def snd p = p (fun x y {y});
snd (pair 1 2) = 2;
```

While one cannot compare these pairs using = directly, one can define a bespoke comparison function.

```
def eq p q = {
  fst p = fst q;
  snd p = snd q
};
eq (pair 1 2) (pair 1 2);
```

Further structural manipulation functions can easily be defined.

```
def swap p = p (fun x y {pair y x});
eq (swap (pair 1 2)) (pair 2 1);

def assoc p =
  fun f {p (fun x yz {yz (fun y z {
     f (pair x y) z
  })});
eq (fst (assoc (pair 1 (pair 2 3)))) (pair 1 2);
snd (assoc (pair 1 (pair 2 3))) = 3;
```

Functions converting between the two formats can easily be defined;

```
def lpair (x, y) = pair x y;
def upair p = p (fun x y {(x, y)});
eq (lpair (1, 2)) (pair 1 2);
(1, 2) = upair (pair 1 2);
```

This is all interesting but doesn't give us anything that wasn't already provided. A variant of this idea can encode arbitrary length lists.

```
def nil = fun n c {n};
def cons x xs = fun n c {c x (xs n c)};

def ex_list = cons 1 (cons 2 (cons 3 nil));
```

The utility of this construction is that the length of the list is not encoded in the structure of the type, so it can be arbitrarily long. One can also issue arguments into the list, which represents folding a function over the list.

```
ex_list 0 (fun x y {x + y}) = 6;
One can define many well-founded recursive functions.
def append 11 12 = fun n c { 11 (12 n c) c };
def map f l = fun n c { 1 n (fun x xs { c (f x) xs }) };
```

One cannot define generic translations between lists and built-in tuples, so a bespoke method must be made for each length. But only the translations need to be length-specific. The operations over lists can be generic.

```
def tup2list_2 (x1, x2) = (cons x1 (cons x2 nil));
def tup2list_3 (x1, x2, x3) = (cons x1 (cons x2 (cons x3 nil)));
def tup2list_4 (x1, x2, x3, x4) = (cons x1 (cons x2 (cons x3 (cons x4 nil))));
def replaceFst a (x, y) = (a, y);
def rotate_2 (x1, x2) = (x2, x1);
def rotate_3 (x1, x2, x3) = (x3, x1, x2);
def rotate_4 (x1, x2, x3, x4) = (x4, x1, x2, x3);
def list2tup_2 1 = 1 (0, 0)
                     (fun x xs {replaceFst x (rotate_2 xs)});
def list2tup_3 1 = 1 (0, 0, 0)
                     (fun x xs {replaceFst x (rotate_3 xs)});
def list2tup_4 l = 1 (0, 0, 0, 0)
                     (fun x xs {replaceFst x (rotate_4 xs)});
list2tup_2 (tup2list_2 (1, 2)) = (1, 2);
list2tup_3 (tup2list_3 (1, 2, 3)) = (1, 2, 3);
list2tup_4 (tup2list_4 (1, 2, 3, 4)) = (1, 2, 3, 4);
def sum 1 = 1 0 (fun x y \{x + y\});
sum (tup2list_2 (1, 2)) = 3;
sum (tup2list_3 (1, 2, 3)) = 6;
sum (tup2list_4 (1, 2, 3, 4)) = 10;
```

Other variants of this construction could be made. Any kind of well-founded tree can be encoded and folded over via similar methods, for example.

It should be noted that there are some fundamental limitations to this method. Mainly, it can't interact non-trivially with the underlying field. One cannot, for example, translate a field element into a Church-encoded number or a boolean. While one can certainly implement those in VampIR (in fact, app2 from earlier is a Church numeral, and const is a Church boolean), they cannot be used to actually reason about field elements. The methods described in this section are exclusively useful for data structure manipulation, but not data manipulation in general. This is because data structures are fixed at circuit generation time, but field values are not completely fixed until proof generation, which occurs after circuit generation. This implies that the circuit's structure cannot depend on field values.

3 Witness Generation

How to make your circuits more fresh... [TODO]

4 Proof Backends

Addition, Multiplication, Division, Oh my! [TODO]

5 Appendix A: Cookbook

Get a load of these delicious recipes... [TODO]

References