

Mini-Wasm Tutorial

1 Introduction

In this tutorial, we will practice how to write a spec in Wasm-DSL. Here, instead of a full Wasm, a very simplified version of Wasm (which we call Mini-Wasm) is used as our goal.

1.1 Working Directory

The directory `<root>/spectec/tutorial` is where we will work on.

1.2 How To Run

Running `make` in `<root>/spectec` will yield an executable file `watsup` in same directory. You can use this to generate various types of spec. For example, in directory `<root>/spectec`:

```
./watsup ./tutorial/*.watsup --prose
```

will generate prose, using every `./watsup` file in directory `./tutorial` as input. There are various options including `--interpreter`, `--latex`, or `--print-il`. You can see all possible options with command `./watsup -help`.

2 Building Blocks of Wasm-DSL

First, we'll study building blocks of Wasm-DSL, and how to use them.

(For now, You don't have to understand this part completely. You may start from Section 3, and refer to this section when you want to.)

2.1 Syntax Definitions

Syntax definitions describe the grammar of the input language or auxiliary constructs. These are essentially type definitions for Wasm-DSL.

```
syntax <name_of_syntax> = <case> | ... | <case>
```

Defining syntax in Wasm-DSL is basically done as above. Use keyword `syntax`, write the name of syntax in lowercase, and simply list the possible cases with the separator `|`. Here, each of the case can be a nonterminal node (usually written in lowercases) which refers to another syntax, or a terminal node (usually written in uppercases).

2.2 Variable Declarations

Variable declarations ascribe the syntactic class (i.e., type) that meta variables used in rules range over.

```
var <name_of_var> : <type>
```

We can explicitly declare a metavariable as above. Use keyword `var`, and give its name and type. Here, `<type>` may be a complex form, which contains iteration, parametric syntax, etc.

Also, every syntax name is implicitly usable as a variable of the respective type. For example, instead of below code:

```
var t : valtype

rule Step/example:
  z; (CONST t c_1) (CONST t c_2) ~> z; (CONST t c_1)
```

we can write like this:

```
rule Step/example:
  z; (CONST valtype c_1) (CONST valtype c_2) ~> z; (CONST valtype c_1)
```

Also, we can use a metavariable without declaration in suitable situations. In this case, its type will be inferred. Now, we may rewrite the above rule as:

```
rule Step/example:
  z; (CONST anyone c_1) (CONST anyone c_2) ~> z; (CONST anyone c_1)
```

2.3 Relation Declarations

Relation declarations, defining the shape of judgement forms, such as typing or reduction relations. These are essentially type declarations for the meta language. For example:

```
relation Instr_ok: context |- instr : functype
relation Step: config ~> config
```

(TODO: more explanation with examples)

2.4 Rule Definitions

```
rule <name_of_relation>/<name_of_rule>:
  <content_of_rule>
```

We can define the individual rules for each relation as above. `<content_of_rule>` has different form, depending on its relation. Every rule is named, so that it can be referenced. Each premise is introduced by a dash and includes the name of the relation it is referencing, easing checking and processing.

(TODO: more explanation with examples)

2.5 Auxiliary Functions

```
def $<func_name>(<type_of_arg1>, <type_of_arg2>, ... , <type_of_argn>) : <result_type>
def $<func_name>(<arg1>, <arg2>, ... , <argn>) = <result>
```

We can declare auxiliary functions as above. The first line indicates the type of arguments and result of the function. The second line is function body, which describes how the actual result comes out. Function body can be multiple cases, which defines a function by pattern matching. For example, we can define a function `size`, which returns the size of a `valtype` as:

```
def $size(valtype) : nat
def $size(I32) = 32
def $size(I64) = 64
```

Now, the function `size` will return 32 for input `I32`, and 64 for input `I64`.

We can also add condition to a function body.

For example, we can define a function `min`, which returns a smaller integer between two:

```
def $min(nat, nat) : nat
def $min(i, j) = i
  -- if i < j
def $min(i, j) = j
```

Now, the function `min` will return `i` if `i < j`, else `j`.

Functions can be used when defining other functions or reduction rules. When using function, write as below form:

```
$<func_name>(<arg1>, <arg2>, ... , <argn>)
```

For example, you can use the function `size` to define the syntax `num_` as below:

```
syntax num_(valtype) = iN($size(valtype))
```

3 Basic Syntax

Now, we will start from writing the basic syntax of Mini-Wasm. Basic syntax of Mini-Wasm is as follows:

$$\begin{aligned}
 N &::= \mathbb{N} \\
 n &::= \mathbb{N} \\
 (\text{integer}) \quad iN(N) &::= 0 \mid \dots \mid 2^N - 1 \\
 (\text{character}) \quad char &::= \text{U+00} \mid \dots \mid \text{U+D7FF} \mid \text{U+E000} \mid \dots \mid \text{U+10FFFF} \\
 (\text{name}) \quad name &::= char^* \\
 (\text{index}) \quad idx &::= iN(32) \\
 (\text{type index}) \quad typeidx &::= idx \\
 (\text{function index}) \quad funcidx &::= idx \\
 (\text{label index}) \quad labelidx &::= idx \\
 (\text{local index}) \quad localidx &::= idx \\
 (\text{number type}) \quad valtype &::= i32 \mid i64 \\
 (\text{function type}) \quad functype &::= valtype^* \rightarrow valtype^* \\
 (\text{external type}) \quad externtype &::= \text{func } functype
 \end{aligned}$$

$$\begin{aligned}\text{size}(\text{i32}) &= 32 \\ \text{size}(\text{i64}) &= 64\end{aligned}$$

$$\text{num}_{\text{valtype}} ::= iN(\text{size}(\text{valtype}))$$

$$\text{binop} ::= \text{add} \mid \text{sub} \mid \text{mul} \mid \text{div}$$

$$\begin{aligned}(\text{instruction}) \quad \text{instr} &::= \text{nop} \\ &\mid \text{drop} \\ &\mid \text{select} \\ &\mid \text{block } \text{functype } \text{instr}^* \\ &\mid \text{loop } \text{functype } \text{instr}^* \\ &\mid \text{if } \text{functype } \text{instr}^* \text{ else } \text{instr}^* \\ &\mid \text{br } \text{labelidx} \\ &\mid \text{br_if } \text{labelidx} \\ &\mid \text{call } \text{funcidx} \\ &\mid \text{return} \\ &\mid \text{const } \text{valtype } \text{num}_{\text{valtype}} \\ &\mid \text{binop } \text{valtype } \text{binop} \\ &\mid \text{local.get } \text{localidx} \\ &\mid \text{local.set } \text{localidx}\end{aligned}$$

$$(\text{expression}) \quad \text{expr} ::= \text{instr}^*$$

$$\begin{aligned}(\text{type}) \quad \text{type} &::= \text{type } \text{functype} \\ (\text{local}) \quad \text{local} &::= \text{local } \text{valtype} \\ (\text{function}) \quad \text{func} &::= \text{func } \text{typeid} \text{ local}^* \text{ expr} \\ (\text{external index}) \quad \text{externidx} &::= \text{func } \text{funcidx} \\ (\text{export}) \quad \text{export} &::= \text{export name } \text{externidx} \\ (\text{module}) \quad \text{module} &::= \text{module } \text{type}^* \text{ func}^* \text{ export}^*\end{aligned}$$

Make a new file `1-syntax.wastup` and write on it.
We'll declare each of the syntax one by one.

3.1 N, n

$$\begin{aligned}N &::= \mathbb{N} \\ n &::= \mathbb{N}\end{aligned}$$

This syntax is simply written in Wasm-DSL like this:

```

syntax N = nat
syntax n = nat

```

Here, `nat` is a pre-defined syntax, which indicates any natural number. This means the syntax `N` and `n` is a natural number.

3.2 iN

(integer) $iN(N) ::= 0 \mid \dots \mid 2^N - 1$

```
syntax iN(N) = 0 | ... | 2^N-1
```

Now, the syntax `iN` is declared in regard with parameter `N`. Also, you can use `...` to indicate a range.

3.3 $char, name$

(character) $char ::= U+00 \mid \dots \mid U+D7FF \mid U+E000 \mid \dots \mid U+10FFFF$
(name) $name ::= char^*$

We need a declaration of *name*, which indicates for a general string. We can declare a syntax for a character like this:

```
syntax char = U+0000 | ... | U+D7FF | U+E000 | ... | U+10FFFF
```

This syntax is built-in, and denotes Unicode code points. Now, *name* is simply an iteration of *chars*. Use `*` to represent a sequence.

```
syntax name = char*
```

3.4 $idx, typeidx, funcidx, labelidx, localidx$

(index) $idx ::= iN(32)$
(type index) $typeidx ::= idx$
(function index) $funcidx ::= idx$
(label index) $labelidx ::= idx$
(local index) $localidx ::= idx$

```
syntax idx = I32
```

Since we have four types of index (which are semantically same, but syntactically different), write like this:

```
syntax typeidx = idx
syntax funcidx = idx
syntax labelidx = idx
syntax localidx = idx
```

3.5 $valtype$

(number type) $valtype ::= i32 \mid i64$

```
syntax valtype = I32 | I64
```

3.6 $functype$

(function type) $functype ::= valtype^* \rightarrow valtype^*$

We can use `->` to indicate a function:

```
syntax functype = valtype* -> valtype*
```

3.7 *externtype*

(external type) *externtype* ::= func *functype*

```
syntax externtype = FUNC functype
```

3.8 *size*, *num_valtype*

size(i32) = 32
size(i64) = 64

num_valtype ::= iN(size(*valtype*))

```
def $size(valtype) : nat
def $size(I32) = 32
def $size(I64) = 64

syntax num_(valtype) = iN($size(valtype))
```

The first line declares the type of function `size`, and the next two lines give the actual definition of it. Then, you can use `size` to declare `num_`.

3.9 *binop*

binop ::= add | sub | mul | div

```
syntax binop = ADD | SUB | MUL | DIV
```

3.10 *instr*

(instruction) *instr* ::= nop
| drop
| select
| block *functype instr**
| loop *functype instr**
| if *functype instr** else *instr**
| br *labelidx*
| br_if *labelidx*
| call *funcidx*
| return
| *valtype.const num_valtype*
| *valtype.binop*
| local.get *localidx*
| local.set *localidx*

```

syntax instr =
| NOP
| DROP
| SELECT
| BLOCK functype instr*
| LOOP functype instr*
| IF functype instr* ELSE instr*
| BR labelidx
| BR_IF labelidx
| CALL funcidx
| RETURN
| CONST valtype num_(valtype)
| BINOP valtype binop
| LOCAL.GET localidx
| LOCAL.SET localidx

```

3.11 *expr*

(expression) $expr ::= instr^*$

```

syntax expr = instr*

```

3.12 *module*

(type)	$type ::= type\ functype$
(local)	$local ::= local\ valtype$
(function)	$func ::= func\ typeidx\ local^*\ expr$
(external index)	$externidx ::= func\ funcidx$
(export)	$export ::= export\ name\ externidx$
(module)	$module ::= module\ type^*\ func^*\ export^*$

Declare *module* and its subcomponents as follows:

```

syntax type = TYPE functype
syntax local = LOCAL valtype
syntax func = FUNC typeidx local* expr
syntax externidx = FUNC funcidx
syntax export = EXPORT name externidx
syntax module = MODULE type* func* export*

```

4 Runtime-Related Syntax

Runtime-related syntax of Mini-Wasm is as follows:

(address) $addr ::= \mathbb{N}$
 (function address) $funcaddr ::= addr$

(value) $val ::= const\ valtype\ num_{valtype}$

(external value) $externval ::= func\ funcaddr$

(function instance)	$funcinst$	$::=$	$\{\text{type } func\text{type},$ $\text{module } moduleinst,$ $\text{code } func\}$
(export instance)	$exportinst$	$::=$	$\{\text{name } name,$ $\text{value } externval\}$
(module instance)	$moduleinst$	$::=$	$\{\text{types } func\text{type}^*,$ $\text{funcs } funcaddr^*,$ $\text{exports } exportinst^*\}$
(store)	$store$	$::=$	$\{\text{funcs } funcinst^*\}$
(frame)	$frame$	$::=$	$\{\text{locals } val^*,$ $\text{module } moduleinst\}$
(state)	$state$	$::=$	$store; frame$
(configuration)	$config$	$::=$	$state; admininstr^*$
(administrative instruction)	$admininstr$	$::=$	$instr$ $call_addr\ funcaddr$ $label_n\ \{instr^*\}\ admininstr^*$ $frame_n\ \{frame\}\ admininstr^*$ $trap$

$$funcaddr((s; f)) = f.module.funcs$$

$$local((s; f), x) = f.locals[x]$$

$$with_{local}((s; f), x, v) = s; f[.locals[x] = v]$$

$$funcinst((s; f)) = s.funcs$$

$$default_{i32} = (\text{const } i32\ 0)$$

$$default_{i64} = (\text{const } i64\ 0)$$

Make a new file `2-runtime.wastup` and write on it.

4.1 $addr, funcaddr$

(address)	$addr$	$::=$	\mathbb{N}
(function address)	$funcaddr$	$::=$	$addr$

```
syntax addr = nat
syntax funcaddr = addr
```

4.2 val

(value)	val	$::=$	$\text{const } val\text{type } num_{val\text{type}}$
---------	-------	-------	--

```
syntax val = CONST valtype num_(valtype)
```


4.3 *externval*

(external value) *externval* ::= *func funcaddr*

```
syntax externval = FUNC funcaddr
```

4.4 *funcinst, exportinst, moduleinst*

(function instance) *funcinst* ::= {type *functype*,
module *moduleinst*,
code *func*}

(export instance) *exportinst* ::= {name *name*,
value *externval*}

(module instance) *moduleinst* ::= {types *functype**,
funcs *funcaddr**,
exports *exportinst**}

```
syntax funcinst =  
  { TYPE functype,  
    MODULE moduleinst,  
    CODE func }  
syntax exportinst =  
  { NAME name,  
    VALUE externval }  
syntax moduleinst =  
  { TYPES functype*,  
    FUNCS funcaddr*,  
    EXPORTS exportinst* }
```

4.5 *store, frame, state, config*

(store) *store* ::= {funcs *funcinst**}

(frame) *frame* ::= {locals *val**,
module *moduleinst*}

(state) *state* ::= *store; frame*

(configuration) *config* ::= *state; admininstr**

```
syntax store = { FUNCS funcinst* }  
  
syntax frame =  
  { LOCALS val*,  
    MODULE moduleinst }  
  
syntax state = store; frame  
syntax config = state; admininstr*
```

4.6 *admininstr*

(administrative instruction) *admininstr* ::= *instr*
| *call_addr funcaddr*
| *label_n {instr*} admininstr**
| *frame_n {frame} admininstr**
| *trap*

```

syntax admininstr =
| instr
| CALL_ADDR funcaddr
| LABEL_ n '{instr*} admininstr*'
| FRAME_ n '{frame} admininstr*'
| TRAP

```

4.7 funcaddr

$$\text{funcaddr}((s; f)) = f.\text{module.funcs}$$

```

def $funcaddr(state) : funcaddr*
def $funcaddr((s; f)) = f.MODULE.FUNCS

```

4.8 local

$$\text{local}((s; f), x) = f.\text{locals}[x]$$

```

def $local(state, localidx) : val
def $local((s; f), x) = f.LOCALS[x]

```

4.9 with_{local}

$$\text{with}_{\text{local}}((s; f), x, v) = s; f[\text{locals}[x] = v]$$

```

def $with_local(state, localidx, val) : state
def $with_local((s; f), x, v) = s; f[.LOCALS[x] = v]

```

4.10 funcinst

$$\text{funcinst}((s; f)) = s.\text{funcs}$$

```

def $funcinst(state) : funcinst*
def $funcinst((s; f)) = s.FUNCS

```

4.11 default_{valtype}

$$\begin{aligned} \text{default}_{i32} &= (\text{const } i32 \ 0) \\ \text{default}_{i64} &= (\text{const } i64 \ 0) \end{aligned}$$

```

def $default_(valtype) : val
def $default_(I32) = (CONST I32 0)
def $default_(I64) = (CONST I64 0)

```

5 Validation Rules

We're done with writing syntax of Mini-Wasm. Now, we will declare validation rules of Mini-Wasm. Make a new file `3-typing.watup` and write on it. First, declare some variables to use in validation rules as below:

```

var i : nat
var x : idx
var l : labelidx
var t : valtype
var ft : functype
var in : instr
var e : expr
var a : addr
var fa : funcaddr
var v : val
var xv : externval
var mm : moduleinst
var fi : funcinst
var xi : exportinst
var s : store
var f : frame
var z : state
var ty : type
var loc : local
var ex : export
var xx : externidx

```

(TODO: add validation rules)

6 Reduction Rules

Now, let's write the reduction rules for the instructions. First, make a new file `4-reduction.wastup` and write as follows:

```
relation Step: config ~> config
```

It means that rules of `relation Step` receives `config` as its input and then yields new `config` as its output. Here, `config` is defined in file `4-runtime.watup`:

```

syntax config = state; admininstr*
syntax admininstr =
  | instr
  | CALL_ADDR funcaddr
  | LABEL_ n '{instr*} admininstr*
  | FRAME_ n '{frame} admininstr*
  | TRAP

```

As you can see, `config` is a pair of `state` and sequence of `admininstr`. `admininstr` is a superset of `instr`, which contains some additional administrative instructions.

Declaring reduction rule in Wasm-DSL is basically done like this:

```

rule Step/<rule name>:
  <input state>; <input instructions> ~> <output state>; <output instructions>
  -- if <condition>
  ...
  -- if <condition>

```

This means that if every **condition** is satisfied, then the state and instructions from stack is reduced to the right hand side. There may be no conditions.
Now we'll declare each of the reduction rules one by one.

6.1 NOP

Use **eps** to indicate an empty instruction sequence.
Use metavariable **z** for a **state**.

Answer:

```
rule Step/nop:
  z; NOP ~> z; eps
```

6.2 DROP

Use following syntax **val**, which is defined in `4-runtime.watsup`:

```
syntax val = CONST valtype val_(valtype)
```

as a variable.

Answer:

```
rule Step/drop:
  z; val DROP ~> z; eps
```

6.3 SELECT

When we get **SELECT** from stack, we have two cases: condition is true or false.
Define each of the case as separate rule as follows:

```
rule Step/select-true:
  ...

rule Step/select-false:
  ...
```

Write like **val_1**, **val_2** (and so on) to distinguish multiple **vals**.
Use **!=** and **=** for integer comparison.

Answer:

```
rule Step/select-true:
  z; val_1 val_2 (CONST I32 c) SELECT ~> z; val_1
  -- if c != 0

rule Step/select-false:
  z; val_1 val_2 (CONST I32 c) SELECT ~> z; val_2
  -- if c = 0
```

As you can see here, you can distinguish multiple metavariables with same type by adding **_1**, **_2** (and so on) after the metavariable name. Also, the names of two rules should be the same before hyphen (-), so that they can be posed as the same reduction rule.

6.4 BINOP

When we get BINOP from stack, there are three cases for `binop_valtype`: ADD, SUB, MUL. Define each of the case as separate rule, and use built-in functions `iadd`, `isub`, `imul` as follows:

```
def $iadd(N, iN(N), iN(N)) : iN(N)
def $isub(N, iN(N), iN(N)) : iN(N)
def $imul(N, iN(N), iN(N)) : iN(N)
def $iadd(N, c_1, c_2) = $((c_1 + c_2) \ 2^N)
def $isub(N, c_1, c_2) = $((c_1 - c_2) \ 2^N)
def $imul(N, c_1, c_2) = $((c_1 * c_2) \ 2^N)

rule Step/binop-add:
  ...

rule Step/binop-sub:
  ...

rule Step/binop-mul:
  ...
```

Use function `size`, which is defined from subsection 4.1.

Answer:

```
def $iadd(N, iN(N), iN(N)) : iN(N)
def $isub(N, iN(N), iN(N)) : iN(N)
def $imul(N, iN(N), iN(N)) : iN(N)

rule Step/binop-add:
  z; (CONST t c_1) (CONST t c_2) (BINOP t ADD) ~> z; (CONST t c)
  -- if $iadd($size(t), c_1, c_2) = c

rule Step/binop-sub:
  z; (CONST t c_1) (CONST t c_2) (BINOP t SUB) ~> z; (CONST t c)
  -- if $isub($size(t), c_1, c_2) = c

rule Step/binop-mul:
  z; (CONST t c_1) (CONST t c_2) (BINOP t MUL) ~> z; (CONST t c)
  -- if $imul($size(t), c_1, c_2) = c
```

Now, we may combine them by declaring a new function `binop` as follows:

```

def $iadd(N, iN(N), iN(N)) : iN(N)
def $isub(N, iN(N), iN(N)) : iN(N)
def $imul(N, iN(N), iN(N)) : iN(N)

def $binop(valtype, binop_(valtype), val_(valtype), val_(valtype)) : val_(valtype)*
def $binop(t, ADD, c_1, c_2) = $iadd($size(t), c_1, c_2)
def $binop(t, SUB, c_1, c_2) = $isub($size(t), c_1, c_2)
def $binop(t, MUL, c_1, c_2) = $imul($size(t), c_1, c_2)

rule Step/binop:
  z; (CONST t c_1) (CONST t c_2) (BINOP t binop) ~> z; (CONST t c)
  -- if $binop(t, binop, c_1, c_2) = c

```

6.5 BLOCK

To indicate a label, refer to following syntax:

```

LABEL_ n '{instr*} admininstr*

```

Answer:

```

rule Step/block-eps:
  z; (BLOCK eps instr*) ~> z; (LABEL_ 0 '{eps} instr*)
rule Step/block-val:
  z; (BLOCK t instr*) ~> z; (LABEL_ 1 '{eps} instr*)

```

Here, you may combine them by using /\ for 'and', \/ for 'or', and ? for an optional argument:

```

rule Step/block:
  z; (BLOCK t? instr*) ~> z; (LABEL_ n '{eps} instr*)
  -- if t? = eps /\ n = 0 \/ t? != eps /\ n = 1

```

6.6 LOOP

Answer:

```

rule Step/loop:
  z; (LOOP t? instr*) ~> z; (LABEL_ 0 '{LOOP t? instr*} instr*)

```

6.7 IF

Answer:

```

rule Step/if-true:
  z; (CONST I32 c) (IF t? instr_1* ELSE instr_2*) ~> z; (BLOCK t? instr_1*)
  -- if c != 0

rule Step/if-false:
  z; (CONST I32 c) (IF t? instr_1* ELSE instr_2*) ~> z; (BLOCK t? instr_2*)
  -- if c = 0

```

6.8 BR

For arithmetic expressions, you should write like $\$(1+1)$, instead of $1+1$.
Use \wedge for a sequence with given length. e.g. `instr \wedge n`.

Answer:

```
rule Step/br-zero:
z; (LABEL_ n '{instr}* val'* val $\wedge$ n (BR 0) instr*)  $\sim$ > z; val $\wedge$ n instr'*

rule Step/br-succ:
z; (LABEL_ n '{instr}* val* (BR  $\$(1+1)$ ) instr*)  $\sim$ > z; val* (BR 1)
```

6.9 BR_IF

Answer:

```
rule Step/br_if-true:
  z; (CONST I32 c) (BR_IF 1)  $\sim$ > z; (BR 1)
  -- if c  $\neq$  0

rule Step/br_if-false:
  z; (CONST I32 c) (BR_IF 1)  $\sim$ > z; eps
  -- if c = 0
```

6.10 CALL

To get a sequence of `funcaddr` from state `z`, write `$funcaddr(z)`.
To get `n`th element of sequence `seq`, write `seq[n]`.

Answer:

```
rule Step/call:
  z; (CALL x)  $\sim$ > z; (CALL_ADDR $funcaddr(z)[x])
```

6.11 FRAME

To indicate a frame, refer to following syntax:

```
FRAME_ n '{frame} admininstr*
```

Answer:

```
rule Step/frame-vals:
  z; (FRAME_ n '{f} val $\wedge$ n)  $\sim$ > z; val $\wedge$ n
```

6.12 RETURN

Answer:

```
rule Step/return-frame:
  z; (FRAME_ n '{f} val'* val^n RETURN instr*) ~> z; val^n

rule Step/return-label:
  z; (LABEL_ n '{instr'}* val* RETURN instr*) ~> z; val* RETURN
```

6.13 LOCAL.GET

To get a local value from **state** *z* and **idx** *x*, write `$local(z, x)`.

Answer:

```
rule Step/local.get:
  z; (LOCAL.GET x) ~> z; $local(z, x)
```

6.14 LOCAL.SET

To get a new state which is exactly same with **state** *z* except that its local value with **idx** *x* is **val** *v*, write `$with_local(z, x, v)`.

Answer:

```
rule Step/local.set:
  z; val (LOCAL.SET x) ~> $with_local(z, x, val); eps
```