The book of Revelation

Kajetan Rzepecki 2012-11-01



Contents

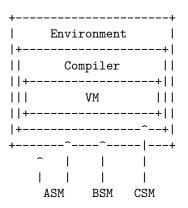
1		ASVM 3						
	1.1	Archit	ecture					
		1.1.1	Environment					
		1.1.2	Compiler					
		1.1.3	VM					
	1.2	Memo	ry model					
		1.2.1	Layout					
		1.2.2	Allocator					
		1.2.3	Garbage collection					
		1.2.4	OpCode encoding					
	1.3		ding					
	1.0	1.3.1	Actor model					
		1.3.2	Processes					
		1.3.3	Message passing					
	1.4		inators					
		1.4.1	Vau calculus					
		1.4.2	Argument evaluation					
	1.5	Forma	l operational semantics					
		1.5.1	Environments					
		1.5.2	Continuations					
		1.5.3	Error handling					
	1.6	Interfa	acing with D					
		1.6.1	Native calls					
		1.6.2	Native types					
		1.6.3	Dynamic FFI					
		1.6.4	Loading any ASM version					
		1.0.4	Loading any ASM version					
2	ΔSI	M proc	gramming language 10					
_		2.1 Phases of evaluation						
	2.1	2.1.1	Lexical analysis					
			· ·					
		2.1.2	Static analysis					
		2.1.3	Code generation					
		2.1.4	Optimisation					
		2.1.5	Evaluation					
	2.2	Lexica	ll					
		2.2.1	Comments					
		2.2.2	Numbers					
		2.2.3	Symbols					
		2.2.4	Identifiers					
		2.2.5	Tuples					
		2.2.6	Vectors					
		2.2.7	Strings					
		2.2.8	Reserved keywords & special tokens					
	2.3		· -					
	۷.ئ							
		2.3.1	Immutability					
		2.3.2	Atoms					
		2.3.3	Combinators					
		2.3.4	Ranges					
		2.3.5	Variables and constants					
		2.3.6	Flow control					

3	Code Cube					
	3.1	cs		14		
		3.1.1	core	14		
		3.1.2	memory	15		
		3.1.3	thread	15		
		3.1.4	error	15		
		3.1.5	reader	15		
		3.1.6	writer	16		
	3.2	cc .		16		
		3.2.1	docs	16		
		3.2.2	test	16		
		3.2.3	dbc	16		
		3.2.4	ranges	17		
		3.2.5	io	19		
		3.2.6	math	20		
		3.2.7	random	20		
		3.2.8	object	20		
		3.2.9	babel	21		
4	Refe	erence	s	23		

1 ASVM

1.1 Architecture

1.1.1 Environment

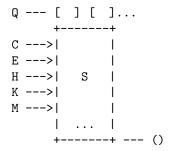


Source types:

- .asm high level ASM.
- .bsm human-readable barebones ASM.
- .csm raw, compiled bytecode

1.1.2 Compiler

1.1.3 VM



Registers & values:

- Q QValue register
- C code pointer
- $\bullet~$ H handler stack pointer
- $\bullet~$ K continuation stack pointer
- M metacontinuation stack pointer
- S store pointer (memory start)
- () fnord value (memory end)

1.2 Memory model

1.2.1 Layout

• Tagged memory

• TValue

Layout:

Variants:

Cons:

- No big numbers without any special treatment.
- Interfacing arrays and native calls to D will be complicated.
- Tag might need additional cons-packing bits.
- Tag appears twice in a pair.

Pros:

- Doesn't need as much padding.
- No dangling null pointers.
- Cons packing is trival.
- Allows for many different kinds of objects to be implemented.
- Allows other kinds of languages to be implemented.
- Maps to Lisp well (was used in Lisp machines).
- TBlobs

Consist of several consecutive QValues.

Pairs, triples, etc:

```
[pair|--PTR--] [fnord|--PTR--]
[triple|--PTR--] [fnord|--PTR--]
```

Cons packing:

- Tuples and Lists Primitives:
 - %car first pointer of a pair
 - %cdr second pointer of a pair
- Arrays and Vectors

Primitives:

- %ref array/vector pointer and offset pair
- %slice array/vector pointer, start and end pointer triple
- Compound types

Used to implement sealer/unsealer pattern.

```
# Could be a hash.
(var *compound-type* 0)
# Could use a separate Q type (Type?) and make use of unique references and is? predicate.
(function make-type ()
  (do (var t *compound-type*)
      (set! *compound-type* (+ 1 *compound-type*))
      (tuple t
             (lambda (o)
                (cons t o))
             (lambda (o)
                (if (and (tuple? o)
                         (equal? (car o) t))
                    (cdr o)
                    (error "Type mismatch.")))))
(function typeof (o)
  (when (pair? o)
    (car o)))
(var (T sealT unsealT) (make-type))
(var foo (sealT (tuple 1 2 3)))
# Might facilitate predicate-based type pattern matching.
(function baz (v)
  (case (typeof v)
    (T (unsealT v))
   (X (unsealX v))
    ...))
```

1.2.2 Allocator

1.2.3 Garbage collection

• GC bits

```
[00|--VAL--] ---> unmanaged (pinned)
[01|--VAL--] ---> undecided
[10|--VAL--] ---> undecided
[11|--VAL--] ---> undecided
```

1.2.4 OpCode encoding

Always pairs —> type part of the tag can be used as the operator type. Example:

1.3 Threading

1.3.1 Actor model

- Threading
 - %tid returns current threads ID.
 - %spawn spawns a thread evaluating given bytecode.
 - %send sends a bunch of immutable data to a thread.
 - %receive receives a bunch of data.

1.3.2 Processes

1.3.3 Message passing

1.4 Combinators

1.4.1 Vau calculus

Basics:

env)))

Primitives

- %vau creates a lexically scoped operative combinator taking dynamic environment.
- $\bullet~$ %wrap induces arg evaluation allowing for applicative combinators.

1.4.2 Argument evaluation

1.5 Formal operational semantics

1.5.1 Environments

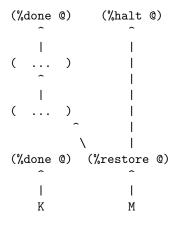
1.5.2 Continuations

• @ register

Instead of value stores the return address where the value should be stored.

• Metacontinuations

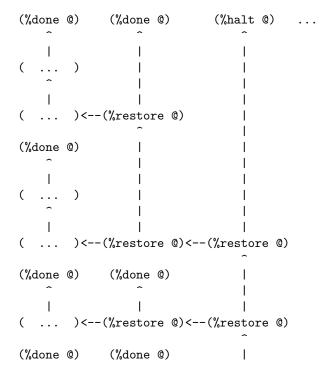
Additional M stack containing continuation segments.

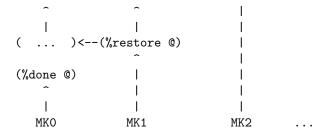


Primitives:

- % restore - sets K to the stored continuation stack segment and applies it to the continuation hole.
- %done ends the current continuation segment and invokes the M register.
- %halt ends the flow of the program.
- Generalized metacontinuations

 $\label{eq:multiple metacontinuation stacks} \ \mbox{with multiple segments each}.$





Primitives:

- %restore pushes a stored continuation stack segment onto the MK register.
- %done pops the MK register leaving the rest of the meta-stack.
- %halt ends the flow of the program.

Possible primitives:

- %done-if premature MK register poping (if @ != ()).
- %select depending on @ pushes one of its children onto the MK stack.

Notes:

- Might be really cool. Especially because it doesn't require constant consing of the continuation stack.
- All the code can be pre-transformed into dataflow format and then executed with no further transformations.
- Used to implement delimited continuations.

1.5.3 Error handling

1.6 Interfacing with D

1.6.1 Native calls

Implementation:

1.6.2 Native types

Implementation:

Example usage:

```
struct Test {
    int bar;
    string foo;
}
// ...
    ASM.defineType!Test;
    ASM.define("foobar", (scpe, args) {
        if(args.car.type == Type.UserDefined)
        if(args.car.userType == typeid(Type))
        // Do shit
        return ASM.fnord;
    });
    ASM.doString(q{
        (var baz (scope
                   (var _inner (newTest))
                   (function getFoo ()
                      (getTestFoo _inner))
                   (function setFoo (newVal)
                      (setTestFoo _inner newVal))
                   (function getBar ()
                      (getTestBar _inner))
                   (function setBar (newVal)
                      (setTestBar _inner newVal))))
        ((baz setFoo) "Test")
        (foobar baz)
        (writeln (baz getBar))
    });
```

1.6.3 Dynamic FFI

1.6.4 Loading any ASM version

2 ASM programming language

- 2.1 Phases of evaluation
- 2.1.1 Lexical analysis
- 2.1.2 Static analysis
- 2.1.3 Code generation
- 2.1.4 Optimisation
- 2.1.5 Evaluation
- 2.2 Lexical
- 2.2.1 Comments
 - Metadata
 - Opts
 - Expression comments
 - Shebang parameters problem
 - Multiline comments
- 2.2.2 Numbers
- 2.2.3 Symbols
- 2.2.4 Identifiers
- **2.2.5** Tuples
- 2.2.6 Vectors
- 2.2.7 Strings
- 2.2.8 Reserved keywords & special tokens
- 2.3 Semantics
- 2.3.1 Immutability
- 2.3.2 Atoms
 - Numbers
 - Symbols
 - Booleans
 - Unit Type

2.3.3 Combinators

• Applicative combinators

```
(function (foo bar baz)
  (* bar baz))
```

• Named call parameters syntax

```
(foo (bar . 23) (baz . 5))
      (foo bar=23 baz=5)
      (foo --bar 23 --baz 5)
   • Operative combinators
     According to Vau calculus as defined by John Shutt.
       (var lambda (%vau (args body) env
                      (%wrap (eval `($%vau $args ignored $body)
                                   env))))
       (var wrap (%lambda (combinator)
                   (%lambda args
                      (eval `($combinator $@args)))))
2.3.4 Ranges
As defined by Andrei Alexandrescu in On iteration.
   Input:
           InputRange ---> front, empty?, popFront!
          ForwardRange ---> save (deep copy)
 BidirectionalRange -+-> back, popBack!
           InfiniteRandomAccessRange ---> [] (indexing)
            FiniteRandomAccessRange ---> [] (indexing)
   Output:
 OutputRange ---> put
   Output ranges could be dropped in favour of impure functions -
 (put 23) ---> (foo 23 'bar '(1 2 3)).
   • With macro
      # Could be generalized to all scopes.
      (with someRange
                           (do
                               ((someRange pop!))
        (pop!)
        (put! 'foo))
                               ((someRange put!) 'foo))
  • Example
      (function circular (tpl)
        (scope (var offset tpl)
               (function empty? ()
                 '())
               (functin front ()
                 (car offset))
               (function popFront! ()
                 (set! offset (cdr offset))
                 (when (not offset)
                   (set! offset tpl)))))
      (var foo (circular '(1 2 3)))
```

Together with type tagging/boxing/sealing might prove to be quite nice:

- Tuples
- Vectors

Vectores are heterogenous.

[vec|stored-type][int|length][ptr|data]

Vec is the vector tag. Stored-type contains ASM typeinfo of the types of stored elements or None for an empty vector or Any for a vector containing different typed-values. Length contains the length of the data in the vector. Ptr is the data pointer.

Pros:

- merges vectors and arrays
- O(1) best case typeinfo
- the type info will be used by D code facilitating data-packing

Cons:

- implies immutability
- 2-3 word header in addition to the actual data
- Vectors of multi-word values should be un-allowed (using vectors of pointers instead)

Optimisations: Data-pack same-typed-data vectors of basic types not to wrap them with an ASM typeinfo (requiring simple-type packing, unpacking, vector copy and vector slice).

Store vec and stored-type in length making it 32 bit long instead of 48 bits:

[vec|stored-type|length][ptr|data]

Pros:

- 1-2 word header

Cons:

- 32 bits for length

Use additional cons-packing value to indicate a vectorized type (followed by length and data pointer).

```
[int|1] <-- int
[int*|4][ptr|--->] ... \footnote{DEFINITION NOT FOUND: 1 }\textsuperscript{,}\,\footnote{DEFINITION NOT FOUND: 1 }\textsu
```

Pros:

- 1-2 word header
- 48 bit length
- cons-packed tuples will easly vectorize

Cons:

- ?

Empty vector:

[vec|None][int|0]

- Multidimentional vectors

```
 [vec|int] [ptr|shape] [ptr|-->] \dots \\ footnotemark[1] \\ textsuperscript{,}\\,\\footnotemark[2]\\ textsuperscri
```

Stored type:

AnyBasicType - assumes data-packing, requires packing on vectorization and unpacking on extension. Any - assumes no data-packing (disallows multi-word values). Ptr - assumes an array tag less pointers to ASM values. None - reserved for Fnord only?

- Strings
 An optimised vector of 8bit integers.
- Environments / Hashes
- Iteration

2.3.5 Variables and constants

• Common syntax for variables and function definition Scheme-like, Scala-eque:

2.3.6 Flow control

• Pattern matching

- \square Binds escaped symbols from pattern to the actual objects.
- -? Escaped symbols = embeded symbols.
- ? Returns a Scope with the symbols defined in it.
- Error handling

A pair of condition predicate and condition handler. Signalizing condition invokes iteratively each predicate in the handler stack until it one is true and runs its corresponding handler.

- Continuations
- Backtracking
 - ? Triggered by backtrack expression.
 - -? Extended syntax ?.

3 Code Cube

3.1 cs

3.1.1 core

- Type predicates
- Type conversions
- Basic math
- import Imports symbols, loads modules, manages scopes: (import func from ModuleA as AFunc all from ModuleB)
- let Immutable let and mutable var:

```
(let ((foo bar))
  # foo is immutable
)

(var ((foo bar))
  # foo is mutable
)
```

• module/program/class/application etc Wraps a bunch of functions and state into a single, named unit:

Dependancy injection:

• case/switch/match/type-dispatch Switch-like control structure, with fallthrough + case goto, case ranges etc:

```
(switch a
  case b (foo bar baz)
  case c (faz baz baz)
  default foo)
```

3.1.2 memory

- GC
 - $-\Box$ collect! does a collection.
 - \square minimise! minimises memory use.
 - □ disable! stops GC.
 - \Box enable! resumes GCs work.
- Allocator

3.1.3 thread

3.1.4 error

- \Box (handle e handler) handler = (error-object handling-function)
- □ (raise error-object)
- □ warn runtime warning
- \square assert check condition and rise errors

3.1.5 reader

Based on dynamic PEG parser generator, because it doesn't need separate lexing phase. Reader macros will be grammar based.

Implementation:

Name arrow Rules

transform:

Any code, really.

Name:

Rule name - used inside of it for transforms and outside for parsing.

Rules - implicitly wrappend in a sequence:

- (a b c ...) sequence
- (/ a b c . . .) ordered choice
- (* Rules) zero or more repeats of the Rules
- (+ Rules) one or more repeats of the Rules
- (? Rules) optional Rules
- (! Rules) not Rules
- $\bullet~$ (& Rules) and Rules
- $\bullet~$ (: Rules) consumes input and drops captures

• (~ Rules) - concatenates captures

arrow:

- <- basic
- < spacing consuming
- < concatenative

3.1.6 writer

Using pattern matching and string embeds, possibly sewn together with the reader. Migth be of use for the bytecode/crosscode compiler.

3.2 cc

3.2.1 docs

Used for documenting code, using... code in the comments. Something along these lines (needs more work):

```
#? (ASMdoc
#? This function does some stuff and returns other stuff.
#? --params
#? bar - an integer,
#? --returns - another integer,
#? --example
#? (var baz (foo 23))
#? )
(function foo (bar)
   (doStuff bar))
```

3.2.2 test

Automated unittest runner:

```
(unittest Foo
    assert (equal? bar baz)
    assert (foo bar baz)
  test Bar
    assert (foo bar baz)
    assert (foo bar baz)
    log "herp derp"
  test Baz
    assert (bar foo faz)
  finally (derp herp))
```

3.2.3 dbc

function macro - creates a function with all kinds of cool stuff:

```
(function (foo bar baz)
  in (equal? bar 23)
  in (> baz bar)
  out (< result bar)
  body (bar baz))</pre>
```

• \square erforce - makes sure an operation will succeed.

3.2.4 ranges

- Collection manipulation
 - $\boxminus join$ if the second argument is a collection prepends it the first argument, if it's not a collection joins both arguments into a pair. Creates a new collection. Examples:
 - * (join 1 '[1 2 3]) -> [1 1 2 3]
 - * (join '(a b) '[1 2 3]) -> [(a b) 1 2 3]
 - * (join 'a 'b) -> (a b)
 - \boxminus append if argument types match appends element or a collection to another collection, if types don't match appends the second argument to the collection. Creates a new collection. Examples:
 - * (append ' $[1\ 2\ 3]\ 4) -> [1\ 2\ 3\ 4]$
 - * (append '[1 2 3] '[4 5 6]) -> [1 2 3 4 5 6]
 - * (append '(2 3) '[2 3]) -> (2 3 [2 3])
 - \boxtimes first returns a reference to the first element of a mutable collection, or its value for an immutable collection.
 - \boxtimes rest returns a new collection referencing the rest part of the old one.
 - $-\boxtimes$ second, third, fourth etc.
 - \boxtimes nth returns nth element of a collection.
 - $-\boxtimes$ map maps an operation to a collection collecting results.
 - \boxtimes reduce maps an operation to a collection reducing it to a single value.
 - \square ? slice slices a collection creating subcollection.
 - \boxminus ? push, push-back, pop, pop-back.
 - ? etc

• Collection creation

- \boxtimes list returns a list consisting of the call args.
- \boxtimes tuple returns a tuple consisting of the call args.
- $-\boxtimes$ set returns a set consisting of the call args.
- $-\boxtimes$ scope reuturns a scope with call args defined in it.
- ? etc.

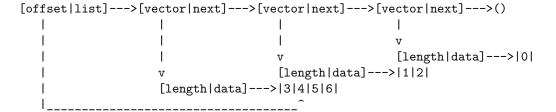
• APL-esque array processing

- + monadic conjugate, dyadic plus
- - monadic negate, dyadic minus
- $\div \text{div m reciprocal, d divide}$
- $\times \text{mul m sign of, d multiply}$
- upstil m ceiling, d maximum
- downstil m floor, d minimum
- * exp m exponential, d power
- -! bang m factorial, d binomial
- | stile m magnitude, d residue
- log m natural logarithm, d logarithm
- circle m times pi, d circular function
- domino m matrix inverse of, d matrix division

- decode d decode
- encode, d encode
- ? roll m roll, d deal
- land d lang/lowest common multiple
- lor d logical or/greatest common divisor
- nand d nand
- nor d nor
- < less d less than
- > greater d greater than
- leq d less or equal
- geq d greater or equal
- = eq d equal to
- neq d not qual to
- equnderbar m depth of, d matches
- nequnderbar d not match
- rho m shape of, d reshape
- , comma m ravel, d cotenate/laminate
- commabar m columnize, d catenate along first axis
- circle stile m reverse, d rotate
- circle bar m reverse along first axis, d rotate along first axis
- transpose m transpose, d general transpose
- $-\uparrow$ up m mix, d take
- $-\downarrow$ down m split, d drop
- left shoe m enclose, d partitioned enclose
- epsilon m enlist, d member of
- index m array/default, d index
- right shoe m first, d pick
- / slash d replicate, o reduce
- slashbar d replicate along first axis, o reduce along first axis
- − \ slope d expand, o scan
- slopebar d expand along first axis, o scan along first axis
- ~ tilde m not, d without
- union m unique, d union
- intersection d intersection
- left tack m same, d left
- right take m same, d right
- iota m index generator, d index of
- epsilon underbar d find
- grade up m grade up, d collated grade up
- grade down m gnade down, d collated grade down
- high minus same as minus
- ' quote string delimiter
- $-\leftarrow$ left d assignment
- zilde niladic empty numeric vector (same as (iota 0))
- thorn m format, d format by specification

- diamond statement separator
- comment comment
- del self reference
- alpha left argument of a dyadic function
- omega right argument of a dyadic function
- quad system name prefix
- " dieresis o each
- dieresis tilde o commute
- dierosis star o power
- . dot o inner product ((dot jot) produces outer product)
- jot o compose ((dot jot) produces outer product)
- VLists

O(log n) indexing. If offset is 0, vlist prealocates additional chunk of data.



• vectorize

• loop

Common Lisp like loop macro:

```
(loop for foo in bar
    for baz being each hash-key of goo
    when gaz
    do gar)
```

• for

3.2.5 io

- Input
 - $-\Box$ readln Unformatted (string) reads.
 - \square read Formatted reads.
 - $-\Box$? load/open Loads a file for reading (as a Scope/Stream with read defined acordingly).
 - $-\ \square$? close closes an imput stream.

```
-\Box eof? - returns 'yup/the object if it has reached EOF.
```

- ? etc
- Output
 - $-\boxtimes$ write writes string representation of the args.
 - ? etc

3.2.6 math

- □ sqrt
- □ pow
- □ exp
- □ min/max/clamp
- □ etc

3.2.7 random

- \square Marsane Twister
- Gaussian distribution

3.2.8 object

• opElvis syntax

```
foo ? bar == (if-non-fnord foo bar) == (if foo foo bar)
```

Implementation:

(macro if-non-fnord (foo bar)

Example:

```
(let ((leny (length map))
            (lenx (let ((__GENSYMO (length (car map))))
                     (if __GENSYMO
                        __GENSYMO
                         1))))
        (let ((__GENSYM1 (do-stuff lenx leny)))
          (if __GENSYM1
              __GENSYM1
              23)))
   • Dynamic dispatch
      (defmethod foo (bar baz) body) (foo bar baz) <=> ((get bar 'foo) baz) # Dynamic dispatch
     or
      (with bar (foo baz)) # With macro
     or
      (function foo (baz) body) (connect bar foo) # Slots
3.2.9 babel
   • JSON
   • XML
   • SVG
     Returns a wellformed SVG string:
      (SVG 100 100
           (circle 50 50
                   '(255 255 100)))
   • YAML
   • LATEX
     Returns a wellformed LATEX string:
      (LaTeX
        "The following equation is herp derp derp:"
        (equation "a^2 + b^2 = c^2")
        (equation "\herp = \derp"))
   • dot
       - ASM AST/module dependancy -> graphviz utility.
   iexpr
      (package foo
```

(function (bar arg0 arg1)
 (if (and (atom? arg0))

(atom? arg1))

4 References