# MASR — Meta Abstract Semantics Representation

# Brian Beckman

# 10 Apr 2023

# **Contents**

1	Abstract	3
2	Cheat Sheet 2.1 Multi-Specs	
3	Issues with ASDL 3.1 ASDL is Moribund 3.2 ASDL is Incomplete 3.3 ASDL's ASR is Volatile 3.4 ASDL is Ambiguous	7 7 7 7 7
4	Clojure Solves ASDL Issues	8
5	MASR Definitions	9
6	MASR Tenets	11
7	Base Specs7.1 Atoms: int, float, bool, nat	12 12 13
8	Term-Like Nodes           8.1 dimensions [sic]            8.2 identifier [sic]            8.3 identifiers [sic]	14 14 16 18
9	9.3 <b>TODO</b> symbol	21 22 28 29 29 30

# (c) Brian Beckman, 2023; Creative Commons Attribution-ShareAlike CC-BY-SA

	9.6	<b>TODO</b> deftype	0
	9.7	<b>TODO</b> presence	0
	9.8	abi	1
	9.9	<b>TODO</b> stmt	3
	9.10	<b>TODO</b> expr	3
		ttype	4
		<b>TODO</b> restriction_arg	5
		<b>TODO</b> binop	5
		TODO logicalbinop	5
		<b>TODO</b> cmpop	5
	9.16	<b>TODO</b> integerboz	5
	9.17	<b>TODO</b> arraybound	5
	9.18	<b>TODO</b> arraystorage	5
		<b>TODO</b> cast_kind	5
	9.20	dimension	6
	9.21	<b>TODO</b> alloc_arg	8
		<b>TODO</b> attribute	8
	9.23	<b>TODO</b> attribute_arg	8
	9.24	<b>TODO</b> call_arg	8
		<b>TODO</b> tbind	8
	9.26	<b>TODO</b> array_index	8
		<b>TODO</b> do_loop_head	8
		<b>TODO</b> case_stmt	8
		<b>TODO</b> type_stmt	8
	9.30	<b>TODO</b> enumtype	8
10	Imp	licit Terms 3	8
	10.1	<b>TODO</b> symtab_id	8
	10.2	<b>TODO</b> symbol table	8

# 1 Abstract

Abstract Semantics Representation (ASR) is an innovative intermediate representation<sup>1</sup> (IR) for multiple LCompilers.<sup>2</sup> ASR is independent of the particular programming language under compilation. Current compiler front-ends targeting ASR include LFortran<sup>3</sup> and LPython.<sup>4</sup> ASR is also agnostic to the compiler back end. Current back ends targeted *from* ASR include LLVM, x86, C, and WASM<sup>5</sup>

Being agnostic means that it is easy to write new compilers, both at the front end and the back end. For example, LFortran predates LPython. When the need for a Python compiler arose, only a Python front end was necessary. Within a few days, a new end-to-end compiler, LPython, was created.

LCompiler back ends are completely reusable because ASR eliminates all original language syntax from the IR, in sharp contrast to typical practice, which treat semantics as decorations on syntax trees.

In addition to being more flexible, LCompilers are faster than average because optimizers are not hampered by useless syntactic structure.

Current specifications for ASR are written in ASDL,<sup>6</sup> a metalanguage similar in spirit to yacc but less rich, by design.<sup>7</sup> To build an LCompiler like LFortran or LPython, the ASDL grammar<sup>8</sup> for ASR is parsed, and a library, libasr,<sup>9</sup> in C++, is generated. Compiler front ends call functions in this library to transform and emit ASR trees.

ASDL has several deficiencies, and MASR, <sup>10</sup> described in this document, alleviates them. We aim to replace ASDL with MASR.

This document is pedagogical, both explaining MASR and teaching how to extend and maintain its Clojure code.

This document may lag the Clojure code. It may also lag the current state of libasr, at least until MASR replaces ASDL. The document mirrors an ASDL snapshot.<sup>8</sup>

<sup>1</sup>https://en.wikipedia.org/wiki/Intermediate\_representation

<sup>2</sup>https://github.com/lcompilers/libasr

<sup>3</sup>https://lfortran.org/

<sup>4</sup>https://lpython.org/

 $<sup>^{5}</sup>$ https://webassembly.org/

<sup>6</sup>https://en.wikipedia.org/wiki/Abstract-Type\_and\_Scheme-Definition\_ Language

<sup>&</sup>lt;sup>7</sup>https://en.wikipedia.org/wiki/Yacc

 $<sup>^{8} \</sup>verb|https://github.com/rebcabin/masr/blob/main/ASR\_2023\_APR\_06\_snapshot.asdl|$ 

 $<sup>^9\</sup>mathrm{https://github.com/lfortran/lfortran/tree/c648a8d824242b676512a038bf2257f3b28dad3b/src/libasr$ 

<sup>&</sup>lt;sup>10</sup>pronounced "maser;" it is a Physics pun

# 2 Cheat Sheet

The following table summarizes this document with simple conforming examples via syntactic sugar. There are other ways to construct conforming examples, but sugar is the recommended way.

Equally important are non-conforming examples. See the body of this document, the tests in specs.clj, and the deftest's in core\_tests.clj for non-conforming examples.

All keywords are in namespace masr.specs. For example, ::nat is short for :masr.specs/nat.

Table 1: Non-nested Specs and Multi-Specs

Spec or Multi-Spec	term	Sugared Conforming Example
::nat	NA	(nat 42))
::identifier	NA	(identifier 'boofar)
::identifier-set	NA	(identifier-set ['a 'a])
::identifier-list	NA	(identifier-list ['a 'a])
::identifier-suit	NA	(identifier-suit ['a 'b])
::asr-term	::dimension	(dimension [6 60])
::dimensions	NA	(dimensions [[6 60] [42]])
::asr-term	::intent	(intent 'Local)
::asr-term	::storage-type	(storage-type 'Default)
::asr-term	::abi	(s/conform ::asr-term
		<pre>( abi 'Source :external false))</pre>
::asr-term	::access	(access 'Public)
::asr-term	::presence	(presence 'Required)

Table 2: Nested ASR-Term Multi-Specs

term	head	Sugared Conforming Example	
::ttype	::integer	(Integer 4 [[6 60] [42]])	
::symbol	::variable		

# 2.1 Multi-Specs

Hash-maps that conform to multi-spec ::asr-term have a ::term attribute that matches a defmethod. For instance, the following example is a valid ::asr-term; its ::term attribute is ::intent:

where

```
(s/def ::intent-enum
  #{'Local 'In 'Out 'InOut 'ReturnVar 'Unspecified})
```

Here is another ::asr-term, this time, matching ::abi:

where

```
(s/def ::abi-external ::bool)
```

The other ::asr-term's follow the obvious pattern. The ::term attributes, ::intent, abi, etc., match one of the term defmethod's:

Finally, the multi-spec itself is named ::asr-term:

```
;; name of the mult-spec defmulti tag fn
;; -----
(s/def ::asr-term (s/multi-spec term ::term))
```

# 2.2 Nested Multi-Specs

Contents of multi-specs can, themselves, be multi-specs. Such cases obtain when a term has multiple function-like heads. Examples include ttype and symbol.

The names of all multi-specs in MASR, nested or not, begin with ::asr- and end with either term or <some-term>-head. Examples: ::asr-term and ::asr-ttype-head. There is only one level of nesting: terms above heads.

Here is the ::asr-term spec for ::ttype; the *nested* multi-spec for contents is named ::asr-type-head.

```
(defmethod term ::ttype [_]
  (s/keys :req [::term ::asr-ttype-head]))
```

where

Here is a conforming example:

Likewise, here is the ::asr-term spec for ::symbol:

Here is a conforming example for :: Variable, abbreviated:

```
(s/valid?
::asr-term {::term ::symbol,
    ::asr-symbol-head
    {::symbol-head ::Variable
    ::symtab-id (nat 2)
    ::varnym (identifier 'x)
    ::intent (intent 'Local)
    ::ttype (ttype (Integer 4 [[0 42])) ,,, }})
```

# 3 Issues with ASDL

#### 3.1 ASDL is Moribund

ASDL has not progressed since originally published in 1987. We know of no other projects adopting ASDL. We should replace ASDL with a modern metalanguage that has a robust, lively ecosystem.

# 3.2 ASDL is Incomplete

Much of the semantics of ASR is currently expressed only in hand-written C++ code. The reason is that ASDL is not sufficiently expressive to cover the needed cases. As usual with such a design, it's more time-consuming and error-prone than necessary to prototype, verify, validate, visualize, modify, and debug. Something more expressive than ASDL is needed to take some responsibility off of hand-written C++ code.

#### 3.3 ASDL's ASR is Volatile

The ASDL for ASR changes frequently, for good reasons. However, stand-aside tools like asr-tester<sup>11</sup> must chase the specification. Just keeping up with ASR-in-ASDL consumes almost all development time for asr-tester. We should unify the language that expresses ASR with the tools that verify and test ASR.

# 3.4 ASDL is Ambiguous

There are many syntactic and semantic ambiguities in the ASDL grammar.<sup>8</sup>. For example, the type notation <code>integer\*</code> might mean, in one place in the grammar, a list of <code>integer</code> with duplicate entries allowed, and, in another place in the grammar, a set of <code>integer</code> with duplicate entries not allowed..

ASDL is not sufficient to express such distinctions. In practice, the hand-written C++ implementations implicitly make these distinctions, hiding them from view and making them difficult to revise. It is bad practice to hide fine distinctions that have observable effects in the implementations. Instead, we should express those distinctions directly in the specifications. Because ASDL cannot express such distinctions, we must adopt something more expressive than ASDL.

<sup>11</sup>https://github.com/rebcabin/asr-tester

# 4 Clojure Solves ASDL Issues

ASR expressions, being trees, have a natural representation in S-Expressions. <sup>12</sup> Clojure, being a modern Lisp, natively handles S-Expressions.

Clojure is up-to-date, lively, and production-ready.

Clojure.spec, <sup>13</sup> is a *force majeure* for precision, completeness, verification, and validation. The collection of MASR specs amounts to a meta-type system for ASR.

Clojure specs are arbitrary predicate functions. Clojure specs can easily express the difference between *list* and *set*, solving the ambiguity issue outlined in Section 3.4. Clojure specs, moreover, can flexibly express type-system features beyond the logics of typical, hard-coded type systems. That flexibility affords new long-term opportunities, say for experiments in dependent types and concurrency types.<sup>14</sup>. In the short run, clojure.spec will make type constraints for ASDL explicit and manifest, and will relieve the burden on C++ programmers to manage implicit constraints.

 $<sup>^{12} \</sup>verb|https://en.wikipedia.org/wiki/S-expression|$ 

<sup>13</sup>https://clojuredocs.org/clojure.spec.alpha

<sup>14</sup>https://rholang.io/

# 5 MASR Definitions

**Definition 1.** A *spec* is a predicate function that tests an expression for conformance. *Spec* is a synonym for *type* in this document.

**Definition 2.** *Terms* are the "objects" or "productions" of ASR, like symbol or dimension.

Names of terms appear to the left of equals signs in the ASDL grammar.<sup>8</sup> Names of terms are generally in lower-case.

Table 3 exhibits terms, ambiguous types, and term-like types. Ambiguous types and term-like types are used but not defined in the ASDL grammar, but are explicitly defined in MASR.

The ambiguous types, symbol\_table and symtab\_id, are called out. The ASDL grammar conflates these two, having only symbol\_table to mean either a full hash-map entity or an integer ID, depending on criteria hidden in hand-written C++ code. A primary objective of MASR is to remove this kind of ambiguity. This kind of ambiguity is not a deficiency of ASDL like that explained in Section 3.4. Unlike the difference between a list and a set, ASDL can express the difference between a hash-map and an integer ID. The failure to do so is a design flaw in the current ASDL grammar.

The contents of Table 3 have been greatly abbreviated and edited for presentation.

**Definition 3.** *Heads* are expressions like Local and CaseStmt, generally in PascalCase, that appear on the right-hand sides of equals signs in Table 3.

See the blog post in the footnote<sup>15</sup> for an informal description of *PascalCase*.

There are of two kinds of heads:

function-like heads — have parentheses and typed parameters,
 e.g., CaseStmt (expr\*, stmt\*)

enum-like heads — no parentheses, e.g., Local

MASR has a Clojure spec and syntactic sugar for each head. There are about 250 heads by a recent count.

**Definition 4.** An *ASR entity* is a compound type like CaseStmt (expr\*, stmt\*), with a function-like head and zero-or more arguments, possibly with names, that require recursive conformance.

 $<sup>^{15} \</sup>texttt{https://alok-verma6597.medium.com/case-styles-in-development-camel-pascal-snake-and-kebab-case-edded} \\$ 

term partial expansion  1 unit TranslationUnit(symbol_table, node*)  2 symbolmany heads  3 storage_type Default Save Parameter Allocatable  4 access Public Private  5 intent Local In Out InOut   6 deftype Implementation Interface  7 presence Required Optional  8 abi Source LFortranModule  Intrinsic  9 stmtmany heads  10 exprmany heads  11 ttype Integer(int, dimension*)   12 restriction_arg RestrictionArg(ident, symbol)  13 binop Add Sub  BitRShift  14 logicalbinop And Or Xor NEqv Eqv  15 cmpop Eq NotEq Lt LtE Gt GtE  16 integerboz Binary Hex Octal  17 arraybound LBound UBound  18 arraystorage RowMajor ColMajor  19 cast_kind RealToInteger IntegerToReal   20 dimension (expr? start, expr? length)  21 alloc_arg (expr a, dimension* dims)  22 attribute arg (ident arg)  23 attribute_arg (expr? value)  25 tbind Bind(string lang, string name)  26 array_index (expr? value)  27 do_loop_head (expr? value)  28 case_stmt CaseStmt (expr? right, expr? step)  29 type_stmt TypeStmtName(symbol, stmt*)   29 type_stmt TypeStmtName(symbol, stmt*)   29 type_stmt TypeStmtName(symbol, stmt*)   30 enumtype IntegerConsecutiveFromZero   ambiguous  31 symbol_table Clojure maps  32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists  1 int float bool nat bignat  2 identifier by regex  1 identifiers identifier*, via Clojure sets		Table 3: Nodes in the ASDL Grammar			
2 symbol			partial expansion		
Storage_type   Defauit   Save   Parameter   Allocatable   4 access   Public   Private   5 intent   Local   In   Out   InOut					
4 access		symbol	•		
Source   Implementation   Interface		storage_type			
Implementation   Interface   Required   Optional   Source   LFortranModule     Intrinsic     Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic     Intrinsic     Source   LFortranModule     Intrinsic     Intrinsic     Source   LFortranModule     Intrinsic       Source   LFortranModule     Intrinsic     Intrinsic     Source   LFortranModule     Intrinsic     Intrinsic     Source   LFortranModule     Intrinsic     Source   LFortranModule     Intrinsic     Intrinsic     Source   Left   Substitute     Source   Left   Left   Substitute     Source   Left   Left   Substitute     Source   Left   Left   Substitute     Source   Left   Left   Substitute   Left   Left   Substitute   Left   Left   Substitute   Left		access	Public Private		
Required   Optional   Source   LFortranModule     Intrinsic     Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic   Source   LFortranModule     Intrinsic     Integer (ident , symbol)     Source   LFortranModule     Integer (Intrinsic   Source   LFortranModule     Integer (Intrinsic   Source   LEVE   LEVE   Source   LEVE		intent	Local   In   Out   InOut		
Source   LFortranModule     Intrinsic  stmt		deftype	Implementation   Interface		
<pre>9 stmt</pre>		presence	Required   Optional		
10 expr		abi	Source LFortranModule  Intrinsic		
Integer(int, dimension*)   restriction_arg  binop  Add Sub  BitRShift  logicalbinop  Add Sub  BitRShift  cmpop  Eq NotEq Lt LtE Gt GtE  integerboz  Binary Hex Octal  arraybound  RestrictionArg(ident, symbol)  Add Sub  BitRShift  And Or Xor NEqv Eqv  Eq NotEq Lt LtE Gt GtE  integerboz  Binary Hex Octal  Binary Hex Octal  ColMajor  cast_kind  RealToInteger IntegerToReal   dimension  (expr? start, expr? length)  attribute_arg  (expr a, dimension* dims)  attribute_arg  (ident arg)  attribute_arg  (expr? value)  bind(string lang, string name)  (expr? vefic, expr? right, expr? step)  case_stmt  (expr? vefic, expr? right, expr? step)  case_stmt  CaseStmt(expr*, stmt*)   ytype_stmt  TypeStmtName(symbol, stmt*)   ambiguous  symtab_id  clojure maps  int (new in MASR; not in ASDL)  *term-like  dimensions  dimension*, via Clojure vectors or lists  int float bool nat bignat  by regex	9	stmt			
12 restriction_arg RestrictionArg(ident, symbol) 13 binop Add Sub  BitRShift 14 logicalbinop And Or Xor NEqv Eqv 15 cmpop Eq NotEq Lt LtE Gt GtE 16 integerboz Binary Hex Octal 17 arraybound LBound UBound 18 arraystorage RowMajor ColMajor 19 cast_kind RealToInteger IntegerToReal  20 dimension (expr? start, expr? length) 21 alloc_arg (expr a, dimension* dims) 22 attribute Attribute(ident name, attr-arg* args) 23 attribute_arg (ident arg) 24 call_arg (expr? value) 25 tbind Bind(string lang, string name) 26 array_index (expr? value) 27 do_loop_head (expr? v, expr? start expr? end, expr? step) 28 case_stmt CaseStmt(expr*, stmt*)  29 type_stmt TypeStmtName(symbol, stmt*)  30 enumtype IntegerConsecutiveFromZero  ambiguous 31 symbol_table Clojure maps 32 symtab_id int (new in MASR; not in ASDL) *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int  float bool nat bignat 0 identifier by regex		expr	many heads		
Add Sub  BitRShift		ttype	<pre>Integer(int, dimension*)  </pre>		
14 logicalbinop		restriction_arg	RestrictionArg(ident, symbol)		
15 cmpop	13	binop	Add Sub  BitRShift		
Binary   Hex   Octal   17   arraybound   LBound   UBound   18   arraystorage   RowMajor   ColMajor   19   cast_kind   RealToInteger   IntegerToReal     20   dimension   (expr? start, expr? length)   21   alloc_arg   (expr a, dimension* dims)   22   attribute   Attribute (ident name, attr-arg* args)   (ident arg)   (expr? value)   25   tbind   Bind(string lang, string name)   (expr? value)   26   array_index   (expr? left, expr? right, expr? step)   (expr? v, expr? start expr? end, expr? step)   28   case_stmt   CaseStmt (expr*, stmt*)     29   type_stmt   TypeStmtName(symbol, stmt*)     30   enumtype   IntegerConsecutiveFromZero     ambiguous   31   symbol_table   Symtab_id   Clojure maps   int (new in MASR; not in ASDL)   *term-like   0   dimensions   dimension*, via Clojure vectors or lists   oidentifier   by regex   others   others	14	logicalbinop	And   Or   Xor   NEqv   Eqv		
17 arraybound 18 arraystorage 19 cast_kind 20 dimension 21 alloc_arg 22 attribute 23 attribute_arg 24 call_arg 25 tbind 26 array_index 27 do_loop_head 28 case_stmt 29 type_stmt 30 enumtype 31 symbol_table 32 attmable 33 symtab_id 34 call_arg 35 case_stm 36 dimensions 37 do_loop_table 38 case_stmt 39 do_loop_table 30 dimensions 30 dimensions 31 symbol_table 32 dimensions 33 dimension*, via Clojure vectors or lists 34 dimensions 35 dimensions 36 dimensions 37 dimensions 38 dimension*, via Clojure vectors or lists 39 identifier 40 dimensions 40 dimensions 40 dimensions 40 dimensions 40 dimension*, via Clojure vectors or lists 40 identifier 41 dimensions 42 colMajor 42 colMajor 42 colMajor 42 colMajor 42 colMajor 42 (expr? start, expr? length) 42 (expr? value) 43 lidentifier 44 call_arg 44 (expr? value) 44 call_arg 45 (ident arg) 46 (expr? value) 47 lident arg 48 case_stmt pand(string lang, string name) 48 (expr? right, expr? step) 49 case_string lang, string name) 40 (expr? value) 40 case_string lang, string name) 40 (expr? value) 51 case_stmt (expr? right, expr? step) 60 case_stmt (expr?, stmt*)   60 case_stmt (expr*, stmt*)   61 case_stmt (expr*, stmt*)   62 case_stmt (expr*, stmt*)   63 case_stmt (expr*, stmt*)   64 case_stmt (expr*, stmt*)   65 case_stmt (expr*, stmt*)   67 case_stmt* 67 case_stmt 68 case_stmt 68 case_stmt 68 case_stmt 69 case_stmt 69 case_stmt	15	cmpop	Eq NotEq Lt LtE Gt GtE		
RowMajor   ColMajor  19 cast_kind RealToInteger   IntegerToReal    20 dimension (expr? start, expr? length)  21 alloc_arg (expr a, dimension* dims)  22 attribute Attribute (ident name, attr-arg* args)  23 attribute_arg (expr? value)  25 tbind Bind(string lang, string name)  26 array_index (expr? left, expr? right, expr? step)  27 do_loop_head (expr? v, expr? start expr? end, expr? step)  28 case_stmt CaseStmt(expr*, stmt*)    29 type_stmt TypeStmtName(symbol, stmt*)    30 enumtype IntegerConsecutiveFromZero    ambiguous  31 symbol_table Clojure maps  32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists  0 atoms int   float   bool   nat   bignat  0 identifier by regex		integerboz			
19 cast_kind RealToInteger IntegerToReal  20 dimension (expr? start, expr? length) 21 alloc_arg (expr a, dimension* dims) 22 attribute Attribute(ident name, attr-arg* args) 23 attribute_arg (ident arg) 24 call_arg (expr? value) 25 tbind Bind(string lang, string name) 26 array_index (expr? left, expr? right, expr? step) 27 do_loop_head (expr? v, expr? start expr? end, expr? step) 28 case_stmt CaseStmt(expr*, stmt*)  29 type_stmt TypeStmtName(symbol, stmt*)  30 enumtype IntegerConsecutiveFromZero   ambiguous  31 symbol_table Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int  float bool nat bignat 0 identifier by regex	17	arraybound	LBound   UBound		
dimension (expr? start, expr? length)  21 alloc_arg (expr a, dimension* dims)  22 attribute Attribute(ident name, attr-arg* args)  23 attribute_arg (ident arg)  24 call_arg (expr? value)  25 tbind Bind(string lang, string name)  26 array_index (expr? left, expr? right, expr? step)  27 do_loop_head (expr? v, expr? start expr? end, expr? step)  28 case_stmt (expr*, stmt*)    29 type_stmt TypeStmtName(symbol, stmt*)    30 enumtype IntegerConsecutiveFromZero    ambiguous  31 symbol_table Clojure maps  32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists  0 atoms int   float   bool   nat   bignat  0 identifier by regex	18	arraystorage	RowMajor ColMajor		
21 alloc_arg (expr a, dimension* dims) 22 attribute 23 attribute_arg (ident arg) 24 call_arg (expr? value) 25 tbind Bind(string lang, string name) 26 array_index (expr? left, expr? right, expr? step) 27 do_loop_head (expr? v, expr? start expr? end, expr? step) 28 case_stmt CaseStmt(expr*, stmt*)   29 type_stmt TypeStmtName(symbol, stmt*)   30 enumtype IntegerConsecutiveFromZero    ambiguous 31 symbol_table Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex	19	cast_kind	RealToInteger IntegerToReal		
22 attribute Attribute (ident name, attr-arg* args) 23 attribute_arg (ident arg) 24 call_arg (expr? value) 25 tbind Bind(string lang, string name) 26 array_index (expr? left, expr? right, expr? step) 27 do_loop_head (expr? v, expr? start expr? end, expr? step) 28 case_stmt CaseStmt(expr*, stmt*)   29 type_stmt TypeStmtName(symbol, stmt*)   30 enumtype IntegerConsecutiveFromZero    ambiguous 31 symbol_table Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex	20	dimension	(expr? start, expr? length)		
23 attribute_arg (ident arg) 24 call_arg (expr? value) 25 tbind Bind(string lang, string name) 26 array_index (expr? left, expr? right, expr? step) 27 do_loop_head (expr? v, expr? start expr? end, expr? step) 28 case_stmt CaseStmt(expr*, stmt*)   29 type_stmt TypeStmtName(symbol, stmt*)   30 enumtype IntegerConsecutiveFromZero    ambiguous  31 symbol_table Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex		alloc_arg	(expr a, dimension* dims)		
24 call_arg (expr? value) 25 tbind Bind(string lang, string name) 26 array_index (expr? left, expr? right, expr? step) 27 do_loop_head (expr? v, expr? start expr? end, expr? step) 28 case_stmt CaseStmt(expr*, stmt*)   29 type_stmt TypeStmtName(symbol, stmt*)   30 enumtype IntegerConsecutiveFromZero    ambiguous  31 symbol_table Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex	22	attribute	Attribute(ident name, attr-arg* args)		
25 tbind Bind(string lang, string name) 26 array_index (expr? left, expr? right, expr? step) 27 do_loop_head (expr? v, expr? start expr? end, expr? step) 28 case_stmt CaseStmt(expr*, stmt*)   29 type_stmt TypeStmtName(symbol, stmt*)   30 enumtype IntegerConsecutiveFromZero    ambiguous 31 symbol_table Symtab_id Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex		attribute_arg	( ident arg)		
26 array_index (expr? left, expr? right, expr? step) 27 do_loop_head (expr? v, expr? start expr? end, expr? step) 28 case_stmt CaseStmt(expr*, stmt*)   29 type_stmt TypeStmtName(symbol, stmt*)   30 enumtype IntegerConsecutiveFromZero    ambiguous 31 symbol_table Symtab_id Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex		call_arg	(expr? value)		
27 do_loop_head (expr? v, expr? start expr? end, expr? step) 28 case_stmt CaseStmt(expr*, stmt*)   29 type_stmt TypeStmtName(symbol, stmt*)   30 enumtype IntegerConsecutiveFromZero    ambiguous  31 symbol_table Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex		tbind			
28 case_stmt CaseStmt(expr*, stmt*)  29 type_stmt TypeStmtName(symbol, stmt*)  30 enumtype IntegerConsecutiveFromZero   ambiguous  31 symbol_table Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex	26	array_index	(expr? left, expr? right, expr? step)		
29 type_stmt TypeStmtName(symbol, stmt*)  30 enumtype IntegerConsecutiveFromZero   ambiguous  31 symbol_table Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex		do_loop_head	( expr? v, expr? start expr? end, expr? step)		
30 enumtype IntegerConsecutiveFromZero   ambiguous  31 symbol_table Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int float bool nat bignat 0 identifier by regex		case_stmt	CaseStmt(expr*, stmt*)		
ambiguous  31 symbol_table Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex		type_stmt	<pre>TypeStmtName(symbol, stmt*)  </pre>		
31 symbol_table Clojure maps 32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex	30	enumtype	IntegerConsecutiveFromZero		
32 symtab_id int (new in MASR; not in ASDL)  *term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex					
*term-like  0 dimensions dimension*, via Clojure vectors or lists 0 atoms int   float   bool   nat   bignat 0 identifier by regex		symbol_table			
0 dimensions dimension*, via Clojure vectors or lists 0 atoms int float bool nat bignat 0 identifier by regex	32	symtab_id	int (new in MASR; not in ASDL)		
<pre>0 atoms</pre>		*term-like			
0 identifier by regex	0	dimensions	dimension*, via Clojure vectors or lists		
	0	atoms	int float bool nat bignat		
0 identifiers identifier*, via Clojure sets	0	identifier			
	0	identifiers	identifier*, via Clojure sets		

# 6 MASR Tenets

- **Entity Hash-Maps** ASR entities<sup>16</sup> shall be hash-maps with fully-qualified keywords as keys (see Section 9.1 for motivating example).
- **Multi-Specs** ASR entity hash-maps shall be recursively checked and generated via Clojure multi-specs. <sup>16</sup>
- **Explicit** ASR entity hash-maps shall contain all necessary information, even at the cost of verbosity. Defaults are not permitted.
- **Syntax Sugar** Extra constructor functions for ASR entities may allow default values for positional and keyword arguments. See Section 9.11 for an example and see Issue 3 on MASR's GitHub repo.<sup>17</sup>

<sup>16</sup>https://clojure.org/guides/spec

<sup>17</sup>https://github.com/rebcabin/masr/issues/3

# 7 Base Specs

The specs in this section are the atoms in the term-like grouping in Table 3

# 7.1 Atoms: int, float, bool, nat

The specs for int, float, and bool are straightforward:

```
(s/def ::int int?) ;; java.lang.Long
(s/def ::float float?)
(s/def ::bool boolean?)
```

### 7.1.1 Sugar

We restrict the spec, nat, for natural numbers, to int, for practical reasons:

```
(tests
(s/valid? ::nat (nat 42))
                                             := true
(s/valid? :: nat (nat -42))
                                             := false
(s/valid? ::nat (nat 0))
                                             := true
(s/valid? ::nat (nat 0xffffffffffffffff))
                                            := false
(s/valid? ::nat (nat -0xFFFFFFFFFFFFFF)) := false
(s/valid?
 ::nat
 (nat (unchecked-long 0xFFFFFFFFFFFFFF))) := false
(s/valid?
 ::nat
 (nat (unchecked-long -0xfffffffffffffffff))) := true
 (s/valid? ::nat (nat 0x7FFFFFFFFFFFFFF)) := true)
```

#### 7.2 Notes

A Clojure *int* is a Java *Long*, with some peculiar behavior for hex literals.<sup>18</sup> The gist is that hex literals for negative numbers in Clojure must have explicit minus signs, lest they become clojure.lang.BigInt, which we disallow for dimension (9.20) and dimensions (8.1) in MASR. To get negative java.lang.Long, one employs Clojure's unchecked-long.

 $<sup>^{18} \</sup>texttt{https://clojurians.slack.com/archives/C03S1KBA2/p1681690965585429}$ 

# 8 Term-Like Nodes

This section of the document exhibits specs for the *term-like nodes* in Table 3: namely dimensions (plural), identifier, and identifiers. These are not terms, but share some similarities with terms. Note carefully the singulars and plurals in the names of the specs. dimension (singular) is a term and covered in Section 9.20.

### 8.1 dimensions [sic]

A MASR dimensions [sic], dimension\* in ASDL, is a homogeneous ordered collection (list or vector) of zero or more dimension instances (9.20). Because ::dimensions [sic] is not a term, we do not need nested multi-specs. However, because ::dimension [sic] is a term, the spec for ::dimensions must ensure that the elements of its collection conform to ::dimension, which is an asr-term multi-spec. We ensure so with a function that selects terms that match a given spec, dimension in this case. We may reuse that function in other specs that represent collections that are not, themselves, terms.

```
(defn term-selector-spec [kwd]
  (s/and ::asr-term
    #(= kwd (::term %))))
```

(The notation # (...%...) is Clojure shorthand for an anonymous function (lambda) with a positional argument denoted by %, and positional arguments %1, %2, ... when there are two or more arguments. Applying a keyword like ::term as a function picks that keyword out of its hash-map argument.)

Here is the spec, ::dimensions, for dimensions. We limit the number of dimensions to 9 for practical reasons. The meaning of a ::dimensions instance with 0 dimensions is an open question (Issue  $6^{19}$ ).

<sup>19</sup>https://github.com/rebcabin/masr/issues/6

#### 8.1.1 FullForm

The following tests show a couple of ways of writing out a ::dimensions instance in full-form. The first is necessary in files other than specs.clj, say in core\_tests.clj. The second can be used in specs.clj:

### 8.1.2 Sugar

The following tests illustrate the sugar for ::dimensions:

### 8.2 identifier [sic]

An ASR identifier is a C or Fortran identifier, which begins with an alphabetic glyph or an underscore, and has alpha-numeric characters or underscores following. The only complication in the spec is the need to generate instances via s/with-gen. The spec solves the generation problem for identifiers, plus shows a pattern for other specs that need custom generators.

```
(let [alpha-re #"[a-zA-Z_]" ;; "let over lambda."
     alphameric-re #"[a-zA-Z0-9_] *"]
 (def alpha?
   #(re-matches alpha-re %))
 (def alphameric?
    #(re-matches alphameric-re %))
  (defn identifier? [sy]
   ;; exclude strings, numbers, quoted numbers
    (and (symbol? sy)
         (let [s (str sy)]
           (and (alpha? (subs s 0 1))
                (alphameric? (subs s 1))))))
 (def identifier-generator
    (tgen/let [c (gen/char-alpha)
               s (gen/string-alphanumeric)]
      (symbol (str c s))))
 (s/def ::identifier
    (s/with-gen
     identifier?
     ;; fn wrapping a macro:
      (fn [] identifier-generator))))
```

The following tests illustrate validation and generation:

```
(tests
  (s/valid? :masr.specs/identifier 'foobar) := true
  (s/valid? :masr.specs/identifier '_f__547) := true
  (s/valid? :masr.specs/identifier '1234) := false)
#_
  (gen/sample (s/gen :masr.specs/identifier))
;; => (e c Q G Z2qP fXzg1 sRx2J6 YIhKlV k6 f7k1Xl4)
;; => (k hM LV QWC qW0X RGk3u W Kg6X Q2YvF0621 ODUt9)
```

# 8.2.1 Sugar

We define a simple function for creating conforming identifiers and illustrate it with a couple of tests:

### 8.3 identifiers [sic]

ASDL identifier\* is ambiguous. There are three kinds of identifier collections in MASR:  $^{20}\,$ 

**identifier-set** — unordered, no duplicates

**identifier-list** — ordered, duplicates allowed (we use vector)

identifier-suit — ordered, duplicates not allowed

For all three kinds, we limit the number of identifiers to 99 for practical purposes:

```
(def MIN-NUMBER-OF-IDENTIFIERS 0)
(def MAX-NUMBER-OF-IDENTIFIERS 99)
```

#### 8.3.1 identifier-set

The spec for a set of identifiers is straightforward because of Clojure's literal syntax, #{\ldots}, for sets, including the empty set:

See the code for uninteresting details of the sugar-function, identifier-set. The following tests show it at work:

 $<sup>^{20} \</sup>verb|https://github.com/rebcabin/masr/issues/1|$ 

#### 8.3.2 identifier-list

The spec for a list of identifiers is almost the same as the spec for a set of identifiers. It differs only in the :into clause — into a vector rather than into a set:

The implementation of the sugar-function for identifier-list is uninteresting. The following tests show it at work:

```
(tests
(let [x (identifier-list ['a 'a])]
   (s/valid? ::identifier-list x) := true
   (vector? x)
                                  := true
  (count
           x)
                                  := 2)
(let [x (identifier-list [])]
  (s/valid? ::identifier-list x) := true
   (vector? x)
  (count
           x)
                                  := 0)
 (let [x (identifier-list ['a '1])]
  (s/valid? :: identifier-list x) := false
  x := ::invalid-identifier-list))
```

# 8.3.3 identifier-suit

The spec for an identifier-suit is almost the same as for identifier-list, only checking that there are no duplicate elements

Here are the tests for the (uninteresting) sugar-function:

```
(tests
(let [x (identifier-suit ['a 'a])]
  (s/valid? ::identifier-suit x) := false
   (vector? x)
(let [x (identifier-suit ['a 'b])]
  (s/valid? ::identifier-suit x) := true
   (vector? x)
                                  := true
   (count
           x)
                                  := 2)
(let [x (identifier-suit [])]
   (s/valid? ::identifier-suit x) := true
   (vector? x)
                                   := true
                                  := 0)
  (count
           x)
 (let [x (identifier-suit ['a '1])]
  (s/valid? ::identifier-suit x) := false
  x := ::invalid-identifier-suit))
```

# 9 Specs

The following sections

- summarize the Clojure specs for all ASR terms and heads
- pedagogically explain the architecture and approach taken in the Clojure code so that anyone may extend and maintain it.

The architecture is the remainder from several experiments. For example, defrecord and defprotocol for polymorphism were tried and discarded in favor of multi-specs. $^{16}$ 

The tests in <code>core\_test.clj</code> exhibit many examples that pass and, more importantly, fail the specs. We also keep lightweight, load-time tests inline to the source file for the specs, <code>specs.clj</code>. The balance between inline tests and separate tests is fluid.

The best way to learn the code is to study the tests and to run them in the Clojure REPL or in the CIDER debugger in Emacs.<sup>21</sup>

We present the terms somewhat out of the order of Table 3. First is *intent*, as it is the archetype for several enum-like terms and heads.

 $<sup>^{21} \</sup>verb|https://docs.cider.mx/cider/debugging/debugger.html|$ 

#### 9.1 intent

#### 9.1.1 Sets for Contents

An ASR intent is one of the symbols

```
Local, In, Out, InOut, ReturnVar, Unspecified.
```

The spec for the *contents* of an intent is simply this set of enum-like heads. Any Clojure set (e.g., in  $\#\{\dots\}$  brackets) doubles as a predicate function for set membership. In the following two examples, the set appears in the function position of the usual Clojure function-call syntax (*function args*\*):

If a candidate member is in a set, the result of calling the set like a function is the candidate member.

```
(#{'Local 'In 'Out 'InOut 'ReturnVar 'Unspecified} 'Local)
```

Local

When the candidate element, say fubar, is not in the set, the result is nil, which does not print:

```
(#{'Local 'In 'Out 'InOut 'ReturnVar 'Unspecified} 'fubar)
```

Any predicate function can be registered as a Clojure spec. <sup>13</sup> Therefore the spec for *intent contents* is just the set of valid members.

#### 9.1.2 Specs have Fully Qualified Keyword Names

The name of the spec is ::intent-enum. The double colon in ::intent-enum is shorthand. In the file specs.clj, double colon implicitly signifies that a keyword like intent-enum is in the namespace masr.specs. In other files, like core\_test.clj, the same keyword is spelled :masr.specs/intent-enum.

The names of all Clojure specs must be fully qualified in namespaces.

```
(s/def ::intent-enum
  #{'Local 'In 'Out 'InOut 'ReturnVar 'Unspecified})
```

#### 9.1.3 How to Use Specs

To check an expression like 'Local against the ::intent-enum spec, write

```
(s/valid? ::intent-enum 'Local)
;; => true
(s/valid? ::intent-enum 'fubar)
;; => false
```

To produce conforming or non-conforming (invalid) entities in other code, write

```
(s/conform ::intent-enum 'Local)
;; => Local
(s/conform ::intent-enum 'fubar)
;; => :clojure.spec.alpha/invalid
```

To generate a few conforming samples, write

```
(gen/sample (s/gen ::intent-enum) 5)
;; => (Unspecified Unspecified Out Unspecified Local)
```

or, with conformance explanation (trivial in this case):

```
(s/exercise ::intent-enum 5)
;; => ([Out Out]
;;        [ReturnVar ReturnVar]
;;        [In In]
;;        [Local Local]
;;        [ReturnVar ReturnVar])
```

Strip out the conformance information as follows:

```
(map second (s/exercise ::intent-enum 5))
;; => (In ReturnVar Out In ReturnVar)
```

s/valid?, s/conform, gen/sample, and s/exercise pertain to any Clojure specs, no matter how complex or rich.

#### 9.1.4 The Spec that Contains the Contents

::intent-enum is just the spec for the *contents* of an intent, not for the intent itself. The spec for the intent itself is an implementation of a polymorphic Clojure *multi-spec*, <sup>16</sup>::asr-term.

#### 9.1.5 Multi-Specs

A multi-spec is like a tagged union in C. The multi-spec, ::asr-term, pertains to all Clojure hash-maps<sup>22</sup> that have a tag named ::term with a value like ::intent or ::storage-type, etc. The values, if themselves fully qualified keywords, are recursively checked.

A multi-spec has three components:

defmulti<sup>23</sup> — a polymorphic interface that declares the tag-fetcher function,
 ::term in this case. The tag-fetcher function fetches a tag's value from
 any candidate hash-map. The defmulti dispatches to a defmethod that
 matches the fetched tag value, ::intent in this case. ::term is a fully qual ified keyword of course, but all keywords double as tag-fetchers for hash maps.<sup>24</sup>

**defmethod**<sup>25</sup> — individual specs, each implementing the interface; in this case, if the ::term of a hash-map matches ::intent, then the corresponding defmethod is invoked (see Section 9.1.7 below).

**s/multi-spec** — tying together the defmulti and, redundantly, the tag-fetcher. <sup>26</sup>

#### 9.1.6 Specs for All Terms

Start with a spec for ::term:

```
;; like ::intent, ::symbol, ::expr, ... (s/def ::term qualified-keyword?)
```

The spec says that any fully qualified keyword, like ::intent, is a MASR term. This spec leaves room for growth of MASR by adding more fully qualified keywords for more MASR types-qua-terms.

s/def stands for clojure.spec.alpha/def, the def macro in the clojure.spec.alpha namespace. The namespace is aliased to s.

Next, specify the defmulti polymorphic interface, term, (no colons) for all term specs:

```
(defmulti term ::term)
```

This defmulti dispatches to a defmethod based on the results of applying the keyword-qua-function ::term to a hash-map:

```
(::term {::term ::intent ...})
```

equals::intent.

<sup>22</sup>https://clojuredocs.org/clojure.core/hash-map

 $<sup>^{23} \</sup>texttt{https://clojuredocs.org/clojure.core/defmulti}$ 

 $<sup>^{24} \</sup>texttt{https://stackoverflow.com/questions/6915531}$ 

<sup>&</sup>lt;sup>25</sup>https://clojuredocs.org/clojure.core/defmethod

<sup>&</sup>lt;sup>26</sup>Multi-specs allow re-tagging, but we do not need that level of generality.

The spec is named ::term and the tag-fetcher is named ::term. They don't need to be the same. They could have different names.

### 9.1.7 Spec for intent

If applying ::term to a Clojure hash-map produces ::intent, the following spec, which specifies all intents, will be invoked. It ignores its argument, \_:

```
(defmethod term ::intent [_]
  (s/keys :req [::term ::intent-enum]))
```

This spec states that an *intent* is a Clojure hash-map with a ::term keyword and an ::intent-enum keyword.

### 9.1.8 The Multi-Spec Itself: ::asr-term

s/multi-spec ties defmulti term to the tag-fetcher ::term. The multi-spec
itself is named ::asr-term:

```
;; name of the mult-spec defmulti tag fn
;; -----
(s/def ::asr-term (s/multi-spec term ::term))
```

### 9.1.9 Examples of Intent

The following shows a valid example:

true

Here is an invalid sample:

false

Generate a few valid samples:

### 9.1.10 Another asr-term: a Pattern Emerges

To define another asr-term, specify the contents and write a defmethod. The one multi-spec, ::asr-term, suffices for all.

For example, another asr-term for an enum-like is storage-type:

```
(s/def ::storage-type-enum
  #{'Default, 'Save, 'Parameter, 'Allocatable})

(defmethod term ::storage-type [_]
  (s/keys :req [::term ::storage-type-enum]))
```

All enum-like specs follow this pattern.

### 9.1.11 Syntax Sugar

{::term ::intent, ::intent-enum 'Local}, a valid asr-term entity, is long and ugly. Write a short function, intent, via s/conform, explained in Section 9.1.3:

Entities have shorter expression with the sugar:

All our specs are like that: a long-form hash-map and a short-form sugar function that does a conformance check.

#### 9.1.12 Capture the Enum-Like Pattern in a Macro

All enum-likes have a *contents* spec, a defmethod term, and a syntax-sugar function. The following macro pertains to all such enum-like multi-specs:

```
(defmacro enum-like [term, heads]
 (let [ns "masr.specs"
       tkw (keyword ns (str term))
       tke (keyword ns (str term "-enum"))
       tki (keyword ns (str "invalid-" term))]
   ` (do
       (s/def ~tke ~heads)
                                ;; the set
       (defmethod term ~tkw [_#] ;; the multi-spec
         (s/keys :req [:masr.specs/term ~tke]))
       (defn ~term [it#]
                              ;; the syntax
         (let [st# (s/conform
                    :masr.specs/asr-term
                    {:masr.specs/term ~tkw
                     ~tke it#})]
           (if (s/invalid? st#) ~tki, st#))))))
```

Use the macro like this:

```
(enum-like
  intent
  #{'Local 'In 'Out 'InOut 'ReturnVar 'Unspecified})
(enum-like
  storage-type
  #{'Default, 'Save, 'Parameter, 'Allocatable})
```

# 9.2 TODO unit

# 9.3 TODO symbol

# 9.3.1 **TODO** Variable

# 9.4 storage\_type

The following tests illustrate conformance:

The implementation is covered in Section 9.1.12 on the enum-like macro.

# 9.5 TODO access

```
(enum-like access #{'Public 'Private})
(tests
  (let [public (access 'Public)]
      (s/conform ::asr-term public) := public)
  (access 'foobar) := ::invalid-access)
```

# 9.6 **TODO** deftype

# 9.7 TODO presence

```
(enum-like presence #{'Required 'Optional})
(tests
(let [required (presence 'Required)]
   (s/conform ::asr-term required) := required)
(presence 'fubar) := ::invalid-presence)
```

#### 9.8 abi

*Abi* is a rich case. It is enum-like, similar to *intent* (Section 9.1), but with restrictions. Its heads include several *external-abis*:

and one *internal-abi*, specified as a Clojure set to get the membership-test functionality:

```
(def internal-abis #{'Source})
```

The *abi-enum* spec for the contents of an *abi* term is the unions of these two sets:

```
(s/def ::abi-enum
  (set/union external-abis internal-abis))
```

Specify an additional key in a conforming *abi* hash-map with a ::bool predicate:

```
(s/def ::abi-external ::bool)
```

Add a convenience function for logic:

```
(defn iff [a b]
(or (and a b)
(not (or a b))))
```

Specify the defmethod for the *abi* itself with a hand-written generator (clojure.spec is not quite strong enough to create the generator automatically):

Generate a few conforming samples:

```
(gen/sample (s/gen (s/and
                    ::asr/asr-term
                    #(= ::asr/abi (::asr/term %))))
            5)
;; => (#::asr{:term ::asr/abi,
             :abi-enum Interactive, :abi-external true}
       #::asr{:term ::asr/abi,
;;
              :abi-enum Source, :abi-external false}
;;
       #::asr{:term ::asr/abi,
;;
              :abi-enum Source, :abi-external false}
;;
       #::asr{:term ::asr/abi,
;;
              :abi-enum Source, :abi-external false}
;;
       #::asr{:term ::asr/abi,
;;
;;
              :abi-enum Interactive, :abi-external true})
```

#### 9.8.1 Syntax Sugar

The sugar for *abi* uses Clojure destructuring<sup>27, 28</sup> for keyword arguments.

Conforming examples:

```
(abi 'Source :external false)
(abi 'LFortranModule :external true)
(abi 'GFortranModule :external true)
(abi 'BindC :external true)
(abi 'Interactive :external true)
(abi 'Intrinsic :external true)
```

Non-conforming due to incorrect boolean:

```
(abi 'Source :external true)
(abi 'LFortranModule :external false)
(abi 'GFortranModule :external false)
(abi 'BindC :external false)
(abi 'Interactive :external false)
(abi 'Intrinsic :external false)
```

<sup>27</sup>https://clojure.org/guides/destructuring

<sup>&</sup>lt;sup>28</sup>https://gist.github.com/rebcabin/a3c24be3e17135f355348c834ab14141

Non-conforming due to incorrect types or structure:

```
(abi 'Source :external 42)  ;; types are not ::bool
(abi 'Source :external "foo") ;;   |
(abi 'Source :external 'foo) ;; -=-
(abi 'Source false) ;; no :external keyword
(abi 'Source true) ;;   |
(abi 'Source 42) ;;   |
(abi 'foo true) ;;   |
(abi 'foo false) ;; -=-
```

We don't show tests of incorrect arity.

Here is the implementation of the sugar, exhibiting the destructuring technique:

# 9.9 TODO stmt

# **9.10 TODO expr**

# 9.11 ttype

Ttype [*sic*] has a nested multi-spec. Ttype is an archetype for all function-like heads, just as *intent* is an archetype for all enum-like heads.

```
(defmulti ttype-head ::ttype-head)
(defmethod ttype-head ::Integer [_]
  (s/keys :req [::ttype-head ::bytes-kind ::dimensions]))
(s/def ::asr-ttype-head
  (s/multi-spec ttype-head ::ttype-head))
```

```
(defmethod term ::ttype [_]
  (s/keys :req [::term ::asr-ttype-head]))
```

#### 9.11.1 Full Form

One may always write out ttype specs in full:

### 9.11.2 Sugar for Integer, Real, Complex, Logical

Sugar for ttypes comes in two varieties, *light sugar* and *full sugar*. Names for light-sugar specs have trailing hyphens. Light sugar requires specs with keywords, as in:

```
(ttype (Integer- {:dimensions [], :kind 4}))
(ttype (Integer- {:kind 4, :dimensions []}))
```

Names for full-sugar specs do not have trailing hyphens. Full sugar uses positional arguments, as in

```
(ttype (Integer))
(ttype (Integer 4))
(ttype (Integer 2 []))
(ttype (Integer 8 [[6 60] [42]]))
```

See the tests for many examples.

# (c) Brian Beckman, 2023; Creative Commons Attribution-ShareAlike CC-BY-SA

- 9.11.3 TODO Character
- 9.12 **TODO** restriction\_arg
- 9.13 TODO binop
- 9.14 TODO logicalbinop
- 9.15 TODO cmpop
- 9.16 **TODO** integerboz
- 9.17 TODO arraybound
- 9.18 TODO arraystorage
- 9.19 TODO cast\_kind

#### 9.20 dimension

A *dimension* is 0, 1, or 2 nats in a Clojure list or vector:

If there is one nat, it specifies the length of any array dimension that enjoys the instance. For example, in the ttype (Integer 4 [[42]]) (9.11), the one dimension in the dimensions [sic] (8.1) of the ttype is [42]. The ttype specifies a rank-1 array of 42 4-byte integers, with indices starting at 1 and running through 42.

If there are two nats, the first nat specifies the starting index of any array dimension that enjoys the instance, and the second nat specifies the length. For example, in the ttype (Integer 4 [[6 60]]) (9.11), the one dimension in the dimensions [sic] (8.1) of the ttype is [6 60]. The ttype specifies a rank-1 array of 60 4-byte integers with indices starting at 6 and running through 65.

If there are no nats, i.e., the array dimension of any array enjoying the instance is **unspecified**. For an example, consider the ttype (Integer 4 [[]]) (9.11). This unspecified type is an open question, Issue  $6.^{19}$  Also unspecified is a dimension of zero length.<sup>29</sup>

#### 9.20.1 TODO: Issue 6: Empty Dimension

See Issue 6 in MASR's GitHub repo $^{19}$  for discussion of the meaning of (Integer 4 [[]]). Empty dimensions [sic], as in (Integer 4 []), are also unspecified and further discussed in Section 8.1.

### 9.20.2 TODO: Issue 7: Zero Length

The following specs, in context of a ttype (9.11) for convenience, are legal in the ASDL grammar.<sup>8</sup> The meaning is **unspecified**:

```
(Integer 4 [[0]])
(Integer 4 [[6 0]])
```

See Issue 7 in MASR's GitHub repo.<sup>29</sup>

### 9.20.3 TODO: Issue 5: Unspecified Starting Index

If there is only 1 nat, the starting index is unspecified. In most programming languages, the default is 0. In Fortran and Mathematica (Wolframscript<sup>30</sup>), the default

<sup>&</sup>lt;sup>29</sup>https://github.com/rebcabin/masr/issues/7

<sup>30</sup>https://www.wolfram.com/wolframscript/

is 1. We provisionally adopt that convention while the issue is open. See Issue 5 in MASR's GitHub repo.  $^{31}$ 

#### 9.20.4 FullForm

The following tests illustrate the full form for *dimension*:

### 9.20.5 Sugar

The following tests illustrate the syntactic sugar for *dimension*:

```
(s/conform ::asr-term
          {::term ::dimension,
           ::dimension-content '(1 60)}) :=
(dimension '(1 60))
(s/valid? ::asr-term (dimension 60))
                                               := false
(s/valid? ::asr-term (dimension [[]]))
                                               := false
(s/valid? ::asr-term (dimension 'foobar))
                                                := false
(s/valid? ::asr-term (dimension ['foobar])) := false
;; throw arity (s/valid? ::asr-term (dimension)) := false
(s/valid? ::asr-term (dimension []))
                                               := true
(s/valid? ::asr-term (dimension [60]))
                                               := true
(s/valid? ::asr-term (dimension [0]))
                                               := true
(s/valid? ::asr-term (dimension '(1 60)))
                                                := true
(s/valid? ::asr-term (dimension '()))
                                                := true)
```

<sup>31</sup>https://github.com/rebcabin/masr/issues/5

- 9.21 TODO alloc\_arg
- 9.22 **TODO** attribute
- 9.23 **TODO** attribute\_arg
- 9.24 **TODO** call\_arg
- 9.25 TODO tbind
- 9.26 **TODO** array\_index
- 9.27 TODO do\_loop\_head
- 9.28 TODO case\_stmt
- 9.29 TODO type\_stmt
- 9.30 TODO enumtype

# 10 Implicit Terms

Terms used, explicitly or implicitly, but not defined in ASDL. Some items specified in ASDL as *symbol\_table* are actually *symtab\_id*.

- 10.1 TODO symtab\_id
- 10.2 TODO symbol\_table

(s/def ::symbol-table map?