Syntax and Semantics of the cat Language

Jade Alglave, Patrick Cousot, and Luc Maranget
© 2018 HSA Foundation. All rights reserved.

The contents of this document are provided in connection with the HSA Foundation specifications. This specification is protected by copyright laws and contains material proprietary to the HSA Foundation. It or any components may not be reproduced, republished, distributed, transmitted, displayed, broadcast or otherwise exploited in any manner without the express prior written permission of HSA Foundation. You may use this specification for implementing the functionality therein, without altering or removing any trademark, copyright or other notice from the specification, but the receipt or possession of this specification does not convey any rights to reproduce, disclose, or distribute its contents, or to manufacture, use, or sell anything that it may describe, in whole or in part.

HSA Foundation grants express permission to any current Founder, Promoter, Supporter Contributor, Academic or Associate member of HSA Foundation to copy and redistribute UNMODIFIED versions of this specification in any fashion, provided that NO CHARGE is made for the specification and the latest available update of the specification for any version of the API is used whenever possible. Such distributed specification may be re-formatted AS LONG AS the contents of the specification are not changed in any way. The specification may be incorporated into a product that is sold as long as such product includes significant independent work developed by the seller. A link to the current version of this specification on the HSA Foundation web-site should be included whenever possible with specification distributions.

HSA Foundation makes no, and expressly disclaims any, representations or warranties, express or implied, regarding this specification, including, without limitation, any implied warranties of merchantability or fitness for a particular purpose or non-infringement of any intellectual property. HSA Foundation makes no, and expressly disclaims any, warranties, express or implied, regarding the correctness, accuracy, completeness, timeliness, and reliability of the specification. Under no circumstances will the HSA Foundation, or any of its Founders, Promoters, Supporters, Academic, Contributors, and Associates members or their respective partners, officers, directors, employees, agents or representatives be liable for any damages, whether direct, indirect, special or consequential damages for lost revenues, lost profits, or otherwise, arising from or in connection with these materials.
Acknowledgments

This document is the result of the contributions of many people. Here is a partial list of the contributors, including the companies that they represented at the time of their contribution.

INRIA

- Luc Maranget

Microsoft Research Cambridge and University College London

- Jade Alglave

New York University and Ecole Normale Superieure

- Patrick Cousot
## Contents

Acknowledgments .............................................................................................................. 3

About this document: Syntax and Semantics of the cat Language ......................... 7
  Audience ....................................................................................................................... 7
  HSA Information Sources ............................................................................................ 7

1. Specifications and statements .................................................................................. 8
  1.1 Evaluating a cat specification ................................................................................. 8
  1.2 Specifications .......................................................................................................... 8
  1.3 Statements ............................................................................................................... 9

2. Typed values and semantic domains ........................................................................ 10
  2.1 Typed values .......................................................................................................... 10
  2.2 Auxiliaries ............................................................................................................. 12

3. Expressions ............................................................................................................... 13
  3.1 Identifiers ............................................................................................................... 14
    3.1.1 Predefined identifiers denoting sets of events ................................................. 15
    3.1.2 Predefined identifiers denoting relations on events ..................................... 16
    3.1.3 Primitives to manipulate sets and relations over events .............................. 17
  3.2 Tags ........................................................................................................................ 18
    3.2.1 Tags .................................................................................................................. 18
    3.2.2 Declarations ................................................................................................... 19
    3.2.3 Scopes .............................................................................................................. 20
    3.2.4 Matching over tags ......................................................................................... 20
  3.3 Functions and procedures ..................................................................................... 21
    3.3.1 Functions ......................................................................................................... 21
    3.3.2 Function calls .................................................................................................. 22
    3.3.3 Procedures ..................................................................................................... 22
    3.3.4 Procedure calls .............................................................................................. 23
  3.4 Sets and tuples ....................................................................................................... 24
    3.4.1 Sets .................................................................................................................. 24
    3.4.2 Matching over sets ......................................................................................... 25
    3.4.3 Tuples .............................................................................................................. 26
    3.4.4 Grouping ......................................................................................................... 26
  3.5 Bindings ................................................................................................................... 26
    3.5.1 Binding definitions ......................................................................................... 27
    3.5.2 Binding expressions ....................................................................................... 28
  3.6 Operators on sets and relations .......................................................................... 29
    3.6.1 Unary operators ............................................................................................. 29
    3.6.2 Binary operators ............................................................................................ 30

4. Constraints ................................................................................................................. 33
  4.1 Checks .................................................................................................................... 33
  4.2 Iteration over sets ................................................................................................. 34
5. Requirements ................................................................. 36
  5.1 Statements .................................................................. 36
  5.2 Quantification via with binding .................................. 36

6. References ...................................................................... 38

Figures

Figure 1–1 Semantics of specifications .............................................. 8
Figure 1–2 Semantics of statements .................................................. 9
Figure 2–1 Typed values ............................................................... 10
Figure 2–2 Semantic domains ....................................................... 11
Figure 2–3 Auxiliaries used to define operators’ semantics ............... 12
Figure 3–1 Simple expressions ..................................................... 13
Figure 3–2 Match clauses ............................................................ 13
Figure 3–3 Expressions ............................................................... 14
Figure 3–4 Definitions ............................................................... 14
Figure 3–5 Reserved identifiers ................................................... 14
Figure 3–6 Semantics of identifiers .............................................. 15
Figure 3–7 Predefined sets of events ........................................... 15
Figure 3–8 Semantics of predefined sets .................................... 16
Figure 3–9 Predefined relations over events ................................ 16
Figure 3–10 Semantics of predefined relations .............................. 17
Figure 3–11 Semantics of primitives .......................................... 17
Figure 3–12 Tags .................................................................... 18
Figure 3–13 Auxiliary over tags `id ............................................... 18
Figure 3–14 Value of tag `id ...................................................... 18
Figure 3–15 Gathering all events bearing the same tag `id .............. 18
Figure 3–16 Declarations ........................................................... 19
Figure 3–17 Semantics of enum .................................................. 19
Figure 3–18 Semantics of instructions ........................................ 19
Figure 3–19 Building the scope instance of level tag ...................... 20
Figure 3–20 Tag matching ........................................................ 21
Figure 3–21 Anonymous functions ............................................. 21
Figure 3–22 Patterns ................................................................ 21
Figure 3–23 Values of functions .................................................. 22
Figure 3–24 Semantics of function calls ...................................... 22
Figure 3–25 Procedures ............................................................ 23
Figure 3–26 Declarations of procedures ...................................... 23
Figure 3–27 Semantics of procedure calls .................................... 24
Figure 3–28 Sets .................................................................... 24
Figure 3–29 Semantics of sets .................................................... 25
Figure 3–30 Set matching ........................................................... 25
Figure 3–31 Tuples .................................................................. 26
Figure 3–32 Semantics of tuples ............................................... 26
Figure 3–33 Semantics of grouping ............................................ 26
Figure 3–34 Bindings ................................................................ 27
Figure 3–35 Value of a binding .................................................... 27
Figure 3–36 Binding definitions, including recursive function binding and recursive set/relation binding ... 28
Figure 3–37 Binding expressions ................................................ 29
Figure 3–38 Unary and binary operators ....................................... 29
Figure 3–39 Unary operators on relations ...................................... 30
Figure 3–40 Complement of a set or relation ............................... 30
Figure 3–41 Sequence of relations ............................................. 31
Figure 3-42 Adding an element to a set .................................................................................. 31
Figure 3-43 Cartesian product of two sets of events ................................................................. 31
Figure 3-44 Binary operators relative to both sets and relations .............................................. 32
Figure 4-1 Constraints ........................................................................................................... 33
Figure 4-2 Value of checks .................................................................................................... 33
Figure 4-3 Checks ................................................................................................................ 33
Figure 4-4 Check condition on relation .................................................................................. 34
Figure 4-5 cat check relation ................................................................................................. 34
Figure 4-6 Iteration over sets ............................................................................................... 35
Figure 4-7 cat iterate ............................................................................................................. 35
Figure 5-1 Requirements ...................................................................................................... 36
Figure 5-2 Semantics of requirements ................................................................................... 36
Figure 5-3 Verdict at top level ............................................................................................... 37
About this document: Syntax and Semantics of the cat Language

This document describes the syntax and semantics of the cat language.

The cat language is inspired by OCaml¹, featuring immutable bindings, first-class functions, and pattern matching. However, cat is a domain-specific language with important differences from OCaml:

- Base values are specialized into sets of events, relations over events, first-class-functions and tags (akin to C enumerations or OCaml constant constructors). There are also two structured values: sets of values and tuples of values.
- The cat language distinguishes among expressions that evaluate to some value, statements that introduce new definitions or constraints, and requirements that introduce quantifiers over statements.

The following notations are used:

- Square brackets [ .. ] to denote optional components
- Parentheses ( .. ) to denote grouping
- ( .. ) * (resp. ( .. ) + ) to denote zero, one, or several (resp. one or several) repetitions of the enclosed components

Audience

This document is written for system and component architects who are interested in supporting the HSA infrastructure (hardware and software) within platform designs.

HSA Information Sources

- HSA Platform System Architecture Specification Version 1.2 describes the HSA system architecture.

¹Xavier Leroy, Damien Doligez, Alain Frisch, Jacques Garrigue, Didier Rémy, and Jérôme Vouillon. The OCaml system, release 4.02, Documentation and user’s manual. caml.inria.fr, 24 September 2014.
1. Specifications and statements

Specifications (or cat files) filter candidate executions. In other words, the semantics of a cat specification are defined with respect to a candidate execution.

1.1 Evaluating a cat specification

Evaluating a cat specification means allowing or forbidding that candidate execution. More precisely, evaluating a cat file makes a result object \((j, f, \rho)\) evolve, where:

- \(j\) is a judgment, allowed or forbidden
- \(f\) is a set of flags that is recorded to signal certain executions (e.g., executions with data races)
- \(\rho\) is an environment that is augmented with new definitions as evaluation progresses

Initially, the judgment is allowed, the set of flags is empty, and predefined identifiers are bound to event sets and relations over events as described in 3.1 Identifiers (on page 14).

1.2 Specifications

Specifications are lists of requirements preceded by an identifier. The syntax and semantics of specifications are shown in Figure 1–1 (below).

Figure 1–1 Semantics of specifications

```
Specifications —

\( \text{cat} \in \text{Cat} \)
\( \text{cat} ::= \text{identifier} \)
\( \quad | \text{identifier requirements} \)

\( \bowtie[\text{identifier}] X \triangleq (\text{allowed}, \emptyset) \)
\( \bowtie[\text{identifier requirements}] X \triangleq \bowtie[\text{requirements}] X (\emptyset, \emptyset) \)
\( \bowtie[\text{cat}] X \triangleq \exists f \in \mathcal{F}. \bowtie[\text{cat}] X = (\text{allowed}, f) \)
\( \bowtie[\text{cat}] \triangleq \{X \in \text{Candidate} | \bowtie[\text{cat}] X \} \)
```

Requirements are evaluated in sequence until one requirement raises error or forbidden or until the end of the requirement list. In that latter case, the specification accepts the candidate execution and raises allowed. The semantics of a cat specification \(\text{cat}\) is the set of candidate executions \(X\) such that evaluating the cat specification \(\text{cat}\) over \(X\) raises the judgment allowed:

\( \bowtie[\text{cat}] \triangleq \{X \in \text{Candidate} | \exists f \in \mathcal{F}. \bowtie[\text{cat}] X = (\text{allowed}, f)\} \)
Requirements are lists of two object types: statements or with x from S (a special quantifier). The special quantifier is described in 5.2 Quantification via with binding (on page 36).

1.3 Statements

Statements are evaluated for their effects: adding new definitions or checking constraints. Figure 1–2 (below) lists semantics of statements. Note that once an error has been raised, that state is retained.

```latex
\begin{align*}
\text{Statements} & \quad \in \quad \text{Statement} \\
\text{statement} & \quad ::= \quad \text{definition} \\
& \quad | \quad \text{constraint}
\end{align*}
```

\[\mathcal{\text{Sem}}[\text{statement} \ \{\text{statement}\}^+] \ X \ <j, f, \rho> \triangleq \]
\[\text{let } <j', f', \rho'> = \mathcal{\text{Sem}}[\text{statement}] \ X \ <j, f, \rho> \text{ in} \]
\[\text{if } j' \text{ = allowed then} \]
\[\mathcal{\text{Sem}}[\{\text{statement}\}^+] \ X \ <j', f', \rho'> \]
\[\text{else } <j', f', \rho'>\]

Error —

\[\mathcal{\text{Sem}}[\text{statement}] \ X \ \text{error} \triangleq \ \text{error}\]
2. Typed values and semantic domains

2.1 Typed values

Typed values (Figure 2–1 (below)) gathered in the set $\mathcal{V}$ (Figure 2–2 (on the next page)). Events of type $\text{evt}$ belong to the set $\text{Evt}$. There are no operations on events, so the type $\text{evt}$ can be used only to type elements of relations or sets.

Typed values include:

- the $\text{error}$ symbol
- tags (of type $\text{tag}$), which belong to $\text{Tag}$
- relations over events (of type $\text{rel}$) which belong to $\mathcal{P}(\text{Evt} \times \text{Evt})$
- sets (of type $\text{set}$) of values, which belong to $\mathcal{P}(\mathcal{V})$; sets must be homogeneous and cannot be sets of functions or procedures as reflected by the predicate $\text{well-formed}$
- tuples (of type $\text{tuple}$) of values, which belong to $\bigcup_{n \in \mathbb{N}} \prod_{i=1}^{n} \mathcal{V}$
- enumerations of tags (of type $\text{enum}$), which belong to $\mathcal{P}(\text{Tag})$
- functions (of type $\text{fun}$)
- procedures (of type $\text{proc}$)

![Typed values diagram]

The value of a function or procedure is a closure memorizing its:
2. Typed values and semantic domains  2.1 Typed values

- Parameter, which belongs to \( Pat \)
- Body, which, for functions, belongs to \( Expr \), and for procedures, can be a list of elements of \( Statement \)
- Calling environment, which belongs to \( \mathcal{E} \)

Environments associate identifiers, which belong to the set \( Identifier \), to typed values. More precisely, environments are partial functions from identifiers to these values: \( \mathcal{E} \triangleq Identifier \rightarrow \mathcal{V} \).

**Judgments** can be of two types:
- allowed when a candidate execution passes all the checks imposed by the \texttt{cat} specification
- forbidden when a candidate execution fails one of the checks of the \texttt{cat} specification

In addition:
- **Flagged checks** collect identifiers of checks that have been flagged.
- **Results** either collect judgments, flagged checks, and environments or raise \texttt{error} if needed.

![Figure 2–2 Semantic domains](image)

---

11 | Syntax and Semantics of the \texttt{cat} Language, Version 1.2
2.2 Auxiliaries

Figure 2–3 (below)\(^1\) shows the auxiliaries used to define the semantics of operators over sets and relations.

\[
\begin{align*}
\mathbb{I}_X & \triangleq \{ \langle e, e \rangle \mid e \in X \} & \text{identity relation on set } X \\
r \circ r' & \triangleq \{ \langle e, e' \rangle \mid \exists e''. \langle e, e'' \rangle \in r \wedge \langle e'', e' \rangle \in r' \} & \text{sequence of relations} \\
\text{dom}(r) & \triangleq \{ x \mid \exists y. \langle x, y \rangle \in r \} & \text{domain of relation } r \\
\text{range}(r) & \triangleq \{ y \mid \exists x. \langle x, y \rangle \in r \} & \text{range of relation } r \\
\text{fld}(r) & \triangleq \text{dom}(r) \cup \text{range}(r) & \text{field of relation } r \\
\text{lfp} \subseteq F = \bigcap \{ X \in \mathcal{P}(S) \mid F(X) \subseteq X \} & \text{the least fixpoint of the } \subseteq\text{-increasing} \\
& \text{operator } F \text{ on the powerset } \mathcal{P}(S)
\end{align*}
\]

3. Expressions

Expressions let the user build new sets or relations over tags and events. Figure 3–1 (below) through Figure 3–4 (on the facing page) list the syntax of expressions.

Several constructs are non-deterministic:

- set matching (see 3.4.2 Matching over sets (on page 25))
- iteration over sets (see 4.2 Iteration over sets (on page 34))
- quantification via with binding (see 5.2 Quantification via with binding (on page 36))

The semantics of an expression is error when the semantics of any of its subexpressions is error. To leave this check implicit, it is assumed that the mathematical construct \( \text{let } t_1; \ldots; t_l = \langle \circ \rangle [expr_i] X \rho, \quad i \in [1, l] \) in ... equals error when there exists \( i \) in \([1, l]\) such that \( \langle \circ \rangle [expr_i] = \) error.

Figure 3–1 Simple expressions

| Simple expressions | — |
|:—— | — |
| simple ∈ Simples | — |
| simple ::= | — |
| | id | — |
| | tag | — |
| | function | anonymous functions |
| | procedure | procedures |
| | set | sets |
| | tuple | tuples |

Figure 3–2 Match clauses

| Match clauses | — |
|:—— | — |
| clause ∈ Clauses | — |
| clause ::= | — |
### 3.1 Identifiers

Identifiers are either predefined or defined by the user through `definition` statements.

Reserved identifiers are shown in Figure 3–5 (below). User-defined identifiers cannot be reserved identifiers and are bound in the environment $\rho$ (Figure 3–6 (on the next page)).
3. Expressions

3.1 Identifiers

3.1.1 Predefined identifiers denoting sets of events

Predefined identifiers denoting sets of events are shown in Figure 3–7 (below). These include:

- the universal sets
- the set of all write, read, memory, branch, and fence events
- the sets of initial and final writes

The semantics of these identifiers (Figure 3–8 on the facing page) are straightforward and denote the eponymous sets of events.

```
Identifiers —

id ∈ Identifier  (Identifier ∩ Reserved = Ø)

\left[ id \right]_X \rho = \begin{cases} 
\text{if } id ∈ \text{dom}(\rho) \text{ then} \\
\quad \text{let } \text{type} = \rho(id) \text{ in} \\
\quad \quad \text{if } \text{type} = \text{enum} \text{ then set } v \text{ else type } v \\
\quad \quad \text{else error}
\end{cases}

Figure 3–6 Semantics of identifiers
```
3.1.2 Predefined identifiers denoting relations on events

Predefined identifiers denoting relations on events are shown in Figure 3–9 (below). These include:

- the empty and identity relations
- the relation over events that access the same memory location
- the relation over events with different pids
- the program order
- the read-from relation
- the initial and final writes

The semantics of these predefined identifiers (Figure 3–10 (on the next page)) is relatively straightforward:

- 0 is the empty relation
- id is the identity relation
- loc is the relation between events accessing the same variable
- ext is the relation between events from different threads
Expressions

3.1 Identifiers

- $\text{po}$ is the program order relation
- $\text{rf}$ the read-from relation

Figure 3-10 Semantics of predefined relations

Semantics of predefined relations —

\[
\begin{align*}
\text{po}[0] X \rho & \triangleq \text{rel: } \emptyset \\
\text{po}[\text{id}] X \rho & \triangleq \text{rel: } \text{evts-of}(X) \triangleq \text{rel: } \{(e, e) \mid e \in \text{evts-of}(X)\} \\
\text{po}[\text{loc}] X \rho & \triangleq \text{rel: } \{(e, e') \in \text{evts-of}(X) \times \text{evts-of}(X) \mid \text{loc-of}(e) = \text{loc-of}(e')\} \\
\text{po}[\text{ext}] X \rho & \triangleq \text{rel: } \{(e, e') \in \text{evts-of}(X) \times \text{evts-of}(X) \mid \text{pid-of}(e) \neq \text{pid-of}(e')\} \\
\text{po}[\text{po}] X \rho & \triangleq \text{rel: po-of}(X) \\
\text{po}[\text{rf}] X \rho & \triangleq \text{rel: rf-of}(X)
\end{align*}
\]

3.1.3 Primitives to manipulate sets and relations over events

Primitives to manipulate sets and relations over events are shown in Figure 3-11 (below). There are four primitives \(\triangleq\), \{\text{classes, linearisations, tag2events, tag2scope}\}.

Primitives for \text{tag2events} and \text{tag2scopes} are described in 3.2 Tags (on the facing page). For the other two primitives:

- \text{classes} takes as argument an expression $\text{expr}$, which should evaluate as a relation $r$. If $r$ is an equivalence relation, the equivalence classes of $r$ are returned; otherwise, an error is raised.

- \text{linearisations} takes as argument a pair of two expressions: $\text{expr}_1$, which should evaluate to a set $S$, and $\text{expr}_2$, which should evaluate to a relation $r$. If this relation is acyclic, all possible linearisations (topological sorts) of $r$ over $S$ are returned; otherwise, the empty set is returned.

Figure 3-11 Semantics of primitives

Semantics of primitive functions —

\[
\begin{align*}
\text{classes} \ \text{expr} \ X \rho & \triangleq \text{let type: } r = \text{classes} \ \text{expr} \ X \rho \text{ in} \\
& \quad \text{if } \text{type } = \text{rel} \land (\text{fld}(r) \subseteq r \land (r)^{-1} \subseteq r \land r \not\subseteq r) \text{ then} \\
& \quad \quad \text{set: } \{\text{set: } \{\text{evt: } e \in \text{fld}(r) \mid (e, e') \in r \} \mid e' \in \text{fld}(r)\} \\
& \quad \text{else error} \\
\text{linearisations} \ \text{expr} \ X \rho & \triangleq \text{let type: } v = \text{linearisations} \ \text{expr} \ X \rho \text{ in} \\
& \quad \text{if } \text{type: } v = \text{tuple: } (\text{set: } s, \text{rel: } r) \land (\forall \text{type}_e: v \in s . \text{type}_e = \text{evts}) \text{ then} \\
& \quad \quad \text{if } r^+ \cap \text{fld} = \emptyset \text{ then} \\
& \quad \quad \quad \text{set: } \{\text{rel: } r' \in \text{evts}(s \times s) \mid r \cap (s \times s) \subseteq r' \land r' \cap r' \subseteq r' \land \\
& \quad \quad \quad \quad \quad \forall e \neq e' \in s: (e, e') \in r' \lor (e', e) \in r'\} \\
& \quad \quad \text{else error}
\end{align*}
\]
3.2 Tags

3.2.1 Tags

Tags are identifiers preceded by a quote ’; which are gathered in sets (Figure 3–12 (below)). We first define an auxiliary over a tag tag: is-tag-declared checks that tag has been defined in an environment ρ, i.e., belongs to an enumeration tag-set-name in ρ (Figure 3–13 (below)).

Figure 3–12 Tags

Tags —

\[
\begin{align*}
tag & \in \text{Tag} \\
\text{tag} & ::= \ 'id \\
s\text{-}\text{tag} & ::= \text{tag}
\end{align*}
\]

scope tags

Figure 3–13 Auxiliary over tags ’id

Auxiliary over tags ’id —

\[
\text{is-tag-declared}['id] \rho \triangleq \exists \text{tag-set} \in \text{dom}(\rho) . \rho(\text{tag-set}) = \text{enum}: T \land \text{tag}: 'id \in T
\]

Next, the value of a tag ’id is the corresponding typed value if the tag has been declared in the environment ρ, or an error if not (Figure 3–14 (below)).

Figure 3–14 Value of tag ’id

Value of a tag ’id —

\[
\begin{align*}
\text{val}[‘id] X \rho & \triangleq \\
\text{if } \text{is-tag-declared}['id] \rho \text{ then } \text{tag}: ‘id \text{ else error}
\end{align*}
\]

Finally, the primitive tag2events gathers all events bearing the tag tag if the tag tag is declared in the environment ρ (Figure 3–15 (below)).

Figure 3–15 Gathering all events bearing the same tag ’id

Gathering all events bearing the same tag ’id —

\[
\begin{align*}
\text{gather}[\text{tag2events}(‘id)] X \rho & \triangleq \\
\text{if } \text{is-tag-declared}['id] \rho \text{ then } \{ e \in \text{evts-of}(X) \mid ‘id \in \text{annot-of}(e) \} \text{ else error}
\end{align*}
\]
3.2.2 Declarations

Enumerations of tags named by an identifier are declared with the construct `enum`. These tags can be used to annotate LISA instructions using the eponymous `instructions` construct (Figure 3–16 (below)).

![Figure 3–16 Declarations](image)

Declarations augment the environment. An `enum` declaration (Figure 3–17 (below)) extends the environment with the corresponding set of tags. In other words, the semantics of `enum id = \[\ldots\] tag_1 \ldots \| tag_n` augments the environment \(\rho\) with the set of typed tags `tag_1,\ldots,tag_n` under the name `id`.

![Figure 3–17 Semantics of enum](image)

The semantics of an `instruction` declaration (Figure 3–18 (below)) is as follows:

- If the declaration contains a tag that is not in the environment, an error is raised.
- If there is an event whose \(i\)th tag is not in the \(i\)th tag set, an error is raised.

![Figure 3–18 Semantics of instructions](image)
3.2.3 Scopes

The scope tags s-tag is distinguished semantically from other tags (Figure 3–19 (below)). Thus for enum declarations, the identifier scopes is reserved for declaring scopes. If an enum scopes declaration is provided, the two functions narrower and wider must be declared on scope tags to define the scope hierarchy.

In addition, the primitive tag2scope builds the relation between events coming from instructions that belong to the same scope (viz., the scope instances of that scope) relative to a scope tree appearing in the original program.

Scope semantics are shown in Figure 3–19 (below).

![Figure 3–19 Building the scope instance of level tag](image)

```
Building the scope instance of level tag —

\( \tau([\text{tag2scope}(s\text{-tag})] X \rho) = \)
\[
\begin{align*}
\text{if } \text{is-tag-declared}[s\text{-tag}] \rho & \text{ then } \text{sr.of}(X) s\text{-tag} \\
\text{else } & \text{error}
\end{align*}
\]
```

3.2.4 Matching over tags

Matching over tags is as follows:

```latex
\begin{align*}
\text{match } expr \text{ with } \\
& | \tag_1 \rightarrow \text{expr}_1 \\
& | \ldots \\
& | \tag_n \rightarrow \text{expr}_n \\
& | \_ \rightarrow \text{expr}_d \\
\end{align*}
```

The value of the match expression is computed as follows:

- Evaluate expr to some value \( \nu \), which must be a tag \( t \).
- Compare \( \nu \) to the tags \( \tag_1, \ldots, \tag_n \) in that order.
- If some tag pattern \( \tag_i \) equals \( t \), the value of the match is the value of the corresponding expression \( \text{expr}_i \). Otherwise, the value of the match is the value of the default expression \( \text{expr}_d \). Because the default clause \( \_ \rightarrow \text{expr}_d \) is optional, the match construct may fail in error.

The semantics of matching over tags are shown in Figure 3–20 (on the next page).
3.3 Functions and procedures

3.3.1 Functions

Functions are first-class values as reflected by the anonymous function construct \texttt{fun pat -> expr} (Figure 3-21 (below)). The expression \texttt{expr} is the \texttt{body} of the function. A function takes only the argument \texttt{pat} (Figure 3-22 (below)). When this argument is a tuple, it may be destructured by a tuple pattern \((id_1, \ldots, id_n)\).

The value of a function (Figure 3-23 (on the facing page)) is its \texttt{closure}. \((\lambda id \cdot body, \rho')\) is written for a closure with parameter \texttt{id}, body \texttt{body}, and declaration environment \(\rho'\). This closure can also be understood as a triple \((id, body, \rho')\).
3.3.2 Function calls

Function calls are written \( expr_1 \ expr_2 \). Functions are of arity one (i.e., they take only one argument) and the application operator is left implicit.

Notice that function application binds tighter than all binary operators (see \[3.6.2\] Binary operators (on page 30)) and looser than postfix operators (see \[3.6.1\] Unary operators (on page 29)). Furthermore, the implicit application operator is left-associative.

The \texttt{cat} language has call-by-value semantics. The effective parameter \( expr_2 \) is evaluated before being bound to the function formal parameter (Figure 3-24 (below)).

3.3.3 Procedures

Procedures are used to define constraints that are checked later (Figure 3-25 (on the next page)).

Procedures have no result; the body of a procedure is a list of statements, and the procedure will be invoked to apply the constraints within its body.
3. Expressions  3.3 Functions and procedures

Figure 3–25 Procedures

\[
\begin{align*}
\text{Procedures} & \quad - \\
\text{procedure} & \quad \in \quad \text{Procedure} \\
\text{procedure} & \quad ::= \quad \text{procedure id pat} = \{\text{statement}\}^+ \quad \text{end}
\end{align*}
\]

Figure 3–26 (below) shows the semantics of procedures. As with functions, procedure declarations simply augment the environment \(\rho\) with their closure.

Figure 3–26 Declarations of procedures

\[
\begin{align*}
\text{Declarations of procedures} & \quad - \\
\text{\(\in\)procedure id}_1 \ id_2 = \{\text{statement}\}^+ \quad \text{end} & \quad X \ (j, f, \rho) \triangleq \\
\langle j, f, \rho[id_1 := \text{proc}: \langle \lambda \ id_2 \cdot \{\text{statement}\}^+, \rho]\rangle & \\
\text{\(\in\)procedure id ( ) = \{\text{statement}\}^+ \quad \text{end} & \quad X \ (j, f, \rho) \triangleq \\
\langle j, f, \rho[id := \text{proc}: \langle \lambda ( ) \cdot \{\text{statement}\}^+, \rho]\rangle & \\
\text{\(\in\)procedure id} (id_1, \ldots, id_\ell) = \{\text{statement}\}^+ \quad \text{end} & \quad X \ (j, f, \rho) \triangleq \\
\langle j, f, \rho[id := \text{proc}: \langle \lambda \langle id_1, \ldots, id_\ell\rangle \cdot \{\text{statement}\}^+, \rho]\rangle
\end{align*}
\]

3.3.4 Procedure calls

Procedure calls are written \text{call} \ id \ expr, where \(id\) is the name of a previously defined procedure (Figure 3–27 (on the facing page)). The bindings performed during a procedure call are discarded when the procedure returns, and all other effects (e.g., checks or flags as described in 4.1 Checks (on page 33)) are performed are retained. Procedures cannot be recursive.
3.4 Sets and tuples

3.4.1 Sets

Sets (Figure 3–28 (below)) are written as \(\{expr_1, expr_2, \ldots, expr_n\}\) with \(n\) greater than 0.

Because events are not values, a set of events cannot be built using explicit set expressions. Sets are homogeneous in that they contain elements of the same type. Set semantics are shown in Figure 3–29 (on the next page). The value of \(\{}\) is the empty set, and the value of \(\{expr_1, \ldots, expr_n\}\) is the set of values \(\{v_1 \ldots v_n\}\) where the \(v_i\) are the values of \(expr_i\).
3.4.2 Matching over sets

Matching over sets is as follows:

```plaintext
match expr with
  | {} → expr₁
  | id₁ ++ id₂ → expr₂
end
```

The value of the `match` is computed as follows:

- Evaluate `expr` to some value `v`, which must be a set. If `v` is the empty set `{}`, the value of the `match` is the value of `expr₁`.
- If `v` is a non-empty set `S`, let `e` be some element in `S` and `S'` be the set `S` minus the element `e`.

The value of the `match` is the value of `expr₂` in a context where `id₁` is bound to `e` and `id₂` is bound to `S'`.

**Figure 3-30 (below)** shows the semantics of matching over sets, where the non-deterministic choice `e ∈ s` is arbitrary and unknown.

```plaintext
match expr₀ with [|| {}] {→ expr₁|| id₁ ++ id₂ → expr₂ end] X ρ ⇒
  let type: s = (expr₀) X ρ in
  if type ≠ set then error
  else if s = {} then (expr₁) X ρ
  else let e ∈ s in
    (expr₂) X ρ[id₂ := set: (s \ {e})][id₁ := e]
```
3.4.3 Tuples

Tuples (Figure 3–31 (below)) include the empty tuple () and constructed tuples \((expr_1, expr_2, \ldots, expr_n)\), with \(n\) greater than 2. There is no tuple of size one.

![Figure 3–31 Tuples](image)

Figure 3–32 (below) shows the semantics of tuples. The value of () is the empty tuple (), and the value of \(<expr_1, \ldots, expr_n>\) is the tuple of values \(<v_1, \ldots, v_n>\) where the \(v_i\) are the values of \(expr_i\). The values \(<v_1, \ldots, v_n>\) are not required to have the same type.

![Figure 3–32 Semantics of tuples](image)

3.4.4 Grouping

Grouping is straightforward, as shown in Figure 3–33 (below). The semantics of a parenthesized expression \((expr)\) are the semantics of \(expr\), idem for begin \(expr\) end.

![Figure 3–33 Semantics of grouping](image)

3.5 Bindings

Bindings (Figure 3–34 (on the next page)) are of the form \(pat = expr\) or \(id pat = expr\), where \(id pat = expr\) is syntactic sugar for \(id = \mathbf{fun} \ pat \to expr\).

Bindings simply update the environment \(\rho\). The bindings for \(pat = expr\) are as follows:

- If \(pat\) is (), then \(expr\) must evaluate to the empty tuple.
- If \(pat\) is \(id\) or \(id\) \(id\), then \(id\) is bound to the value of \(expr\).
3. Expressions  3.5 Bindings

- If \( \text{pat} \) is a proper tuple pattern \((\text{id}_1, \ldots, \text{id}_n)\) with \(n\) greater than 2, then \(\text{expr}\) must evaluate to a tuple value of size \(n\) \((\nu_1, \ldots, \nu_n)\), and the names \(\text{id}_1, \ldots, \text{id}_n\) are bound to the values \(\nu_1, \ldots, \nu_n\).

---

Figure 3–34 Bindings

\[
\text{binding} \in \text{Binding} \quad \text{binding} ::= \text{pat} = \text{expr} \mid \text{id} \text{pat} = \text{expr} \\
\text{where id pat} = \text{expr} \triangleq \text{id} = \text{fun pat} \rightarrow \text{expr}
\]

Figure 3–35 Value of a binding

\[
\begin{align*}
\varnothing\text{[pat = expr]}X\rho & \triangleq \text{match pat with} \\
| () & \rightarrow \text{tuple: ()} \\
| \text{id} & \rightarrow \rho[\text{id} := \varnothing\text{[expr]}X\rho] \\
| (\text{id}_1, \ldots, \text{id}_m) & \rightarrow \\
& \text{match (\varnothing\text{[expr]}X\rho)} \text{ with} \\
& | \text{tuple: (e}_1, \ldots, e_m) & \rightarrow \rho[\text{id}_1 := e_1] \ldots [\text{id}_m := e_m] \\
& | \_ & \rightarrow \text{error} \\
\varnothing\text{[id pat = expr]}X\rho & \triangleq \varnothing\text{[id = fun pat} \rightarrow \text{expr]}X\rho
\end{align*}
\]

3.5.1 Binding definitions

Binding definitions (Figure 3–34 (above), Figure 3–35 (above), and Figure 3–36 (on the facing page)) happen through the \texttt{let} and \texttt{let rec} constructs, which bind value names for the rest of a specification evaluation.

- First, the construct \texttt{let binding}_1 and \ldots and \texttt{binding}_n (i.e., \texttt{let pat}_1 = \texttt{expr}_1 and \ldots and \texttt{pat}_n = \texttt{expr}_n), evaluates \texttt{expr}_1, \ldots, \texttt{expr}_n and binds the names in the patterns \texttt{pat}_1, \ldots, \texttt{pat}_n to the resulting values.

- Second, for recursive function bindings \texttt{let rec id}_1 \texttt{pat}_1 = \texttt{expr}_1 and \ldots and \texttt{id}_n \texttt{pat}_n = \texttt{expr}_n, we follow Milner and Tofte\(^1\) where the proof of existence and unicity of the infinite closure \(c^\infty\) is based on Aczel\(^2\).

- Third, for recursive set or relation bindings \texttt{let rec id}_1 = \texttt{expr}_1 and \ldots and \texttt{id}_n = \texttt{expr}_n, the


least solution of the equations is computed \( id_1 = expr_1, \ldots, id_n = expr_n \) on sets or relations using inclusion for ordering.

The recursive bindings may be mutually recursive. These recursive definitions should be well-formed, i.e., terminating. The result of ill-formed definitions is undefined.

3.5.2 Binding expressions

Binding expressions (Figure 3–37 (on the next page)) happen through the construct `let [rec] bindings in expr`, which locally binds the names defined by `bindings` to evaluate `expr`. Both non-recursive and recursive bindings are allowed.
3.6 Operators on sets and relations

Operators can be unary or binary. **Figure 3-38 (below)** lists them and describes their semantics.

### 3.6.1 Unary operators

An expression denoting a relation can be built as follows:

- Identity closure with the operator ?
- Reflexive-transitive closure with the operator *
- Transitive closure with +
- Complement with ~
- Inverse with ^-1
3. Expressions

3.6 Operators on sets and relations

These operators are postfix and are defined on relations only (Figure 3–39 below), except for the complement, which can apply to sets of events or tags as well (Figure 3–40 below). As described in 3.1.2 Predefined identifiers denoting relations on events (on page 16), the value of identifier 0 is the empty relation.

### Figure 3–39 Unary operators on relations

Unary operators on relations —

\[
\widehat{\cdot} X \rho \triangleq \\
\text{let type: } r = \widehat{\cdot} X \rho \text{ in} \\
\text{if type } = \text{rel then} \\
\text{rel: cat-operation-of } X \rho \\
\text{else error}
\]

### Figure 3–40 Complement of a set or relation

Complement of a set or relation —

\[
\text{let type: } s = \widehat{\cdot} X \rho \text{ in} \\
\text{if type } = \text{set } \land \forall \text{type: } e \in s . \text{type } = \text{evt} \text{ then} \\
\text{set: } \{(\text{evt}: e | e \in \text{evts-of}(X)) \setminus s\} \\
\text{else if type } = \text{set } \land \exists \text{id } \in \text{dom}(\rho) . \rho(id) = \text{enum: } T \land s \subseteq T \text{ then} \\
\text{set: } (T \setminus s) \\
\text{else if type } = \text{rel } \text{then} \\
\text{rel: } (\text{evts-of}(X) \times \text{evts-of}(X)) \setminus s \\
\text{else error}
\]

3.6.2 Binary operators

The sequence of two expressions can be built (or composed as shown in Figure 3–41 (on the next page)) with the operator ; , which is defined on relations only. In addition:
• An element can be added to a set using the addition operator \( expr_1 + + expr_2 \), which operates on sets. The value of \( expr_2 \) must be a set of values \( S \) and the operator returns the set \( S \) augmented with the value of \( expr_1 \). (See Figure 3–42 (below).)

• A new relation can be built out of the Cartesian product of two sets of events with the infix operator \( \ast \). (See Figure 3–43 (below).)

**Figure 3–41** Sequence of relations

```
Sequence of relations —

\( \circ [expr_1 ; expr_2] \ X \rho \triangleq \)

let type_r = \( \circ [expr_1] \ X \rho \), \( i \in [1, 2] \) in

if type_r = type_r = rel then
  rel: r_1 \( \ast r_2 \)
else error
```

**Figure 3–42** Adding an element to a set

```
Adding an element to a set —

\( \circ [expr_1 + + expr_2] \ X \rho \triangleq \)

let type_v = \( \circ [expr_1] \ X \rho \), \( i \in [1, 2] \) in

let \( S = \text{set}: (\{ \text{type_v: v_1} \} \cup v_2) \) in

if well-formed(S) then \( S \) else error
```

**Figure 3–43** Cartesian product of two sets of events

```
Cartesian product of two sets of events —

\( \circ [expr_1 \ast expr_2] \ X \rho \triangleq \)

let type_s = \( \circ [expr_1] \ X \rho \), \( i \in [1, 2] \) in

if type_s = type_s = set \( \land \forall \text{type_v: v} \in s_1 \cup s_2 \). \text{type = evr} \) then
  rel: s_1 \( \times s_2 \)
else error
```

We can build the union, intersection, and difference of the sets and relations (**Figure 3–44 (on the facing page)**). The semantics of \( expr_1 \mathbin{\circ} expr_2 \) is the operator \( \circ \) applied to the sets (resp. relations) \( s_1 \) and \( s_2 \), viz., the values of \( expr_1 \) and \( expr_2 \).
<table>
<thead>
<tr>
<th>op</th>
<th>cat-operation-of [op]</th>
</tr>
</thead>
<tbody>
<tr>
<td>`</td>
<td><code> </code></td>
</tr>
<tr>
<td><code>&amp;</code></td>
<td><code>∩</code></td>
</tr>
<tr>
<td><code>\</code></td>
<td><code>\</code></td>
</tr>
</tbody>
</table>

Figure 3.44 Binary operators relative to both sets and relations

Binary operators relative to both sets and relations —

\[
\begin{align*}
\text{let } & \text{type}_1, \text{type}_2 = \left\langle \left\langle [expr_1] X \rho \right\rangle \right\rangle, \quad i \in [1, 2] \text{ in} \\
& \text{if } \text{type}_1 = \text{type}_2 = \text{set} \land \text{well-formed}(\text{set} \cup \text{set}) \text{ then} \\
& \quad \text{set: } \text{set} \left(\text{cat-operation-of [op]}\right) \text{ set} \\
& \text{else if } \text{type}_1 = \text{type}_2 = \text{rel} \text{ then} \\
& \quad \text{rel: } \text{set} \left(\text{cat-operation-of [op]}\right) \text{ rel} \\
& \text{else error}
\end{align*}
\]
4. Constraints

Constraints can be checks, procedure calls, or iteration over sets (Figure 4-1 (below)).

4.1 Checks

Checks (Figure 4-2 (below)) happen through the construct `check expr`, which evaluates `expr` and applies the check `check`. The six checks (Figure 4-3 (below)) include:

- acyclicity (keyword `acyclic`) and its negation
- irreflexivity (keyword `irreflexive`) and its negation
- emptiness (keyword `empty`) and its negation

If the check succeeds (Figure 4-4 (on the facing page)), the candidate execution is allowed so far. Otherwise, the candidate execution is forbidden.
A check can optionally be named \textit{id} using the keyword \textit{as}.

\textbf{Figure 4–4 Check condition on relation}

\begin{verbatim}
check-condition-on-relation[check] \triangleq
  \begin{array}{ll}
    \text{match check with} \\
    \text{acyclic} \rightarrow \text{check-condition-on-relation[irreflexive]} \ R^+ \\
    \text{irreflexive} \rightarrow (\forall e \in \text{fld}(R). \langle e, e \rangle \not\in R) \\
    \text{empty} \rightarrow R = \emptyset \\
    \sim \text{check}’ \rightarrow \neg\text{check-condition-on-relation[check']} \ R
  \end{array}
\end{verbatim}

A check can be flagged by prefixing it with the \textit{flag} keyword (\textbf{Figure 4–5 (below)}). Flagged checks must be named with the \textit{as} construct. Failed flagged checks do not stop evaluation. Instead, failed flagged checks are recorded under their names in the component \textit{f} of the semantics of the \textit{cat} specification (e.g., using the \textit{herd7} tool to handle flagged candidate executions later).

Flagged checks are useful for specifications with statements that impact the semantics of an entire program, e.g., specifications phrased in terms of data races, such as those of C++ or HSA.

\textbf{Figure 4–5 cat check relation}

\begin{verbatim}
cat-check-relation[flagoption check \ R as id] \langle j, f, \rho \rangle \triangleq
  \begin{cases}
    \text{if (check-condition-on-relation[check] \ R) then} \\
    \langle j, f, \rho \rangle \\
    \text{else if flagoption = flag then} \\
    \langle j, f \cup \{id\}, \rho \rangle \\
    \text{else (forbidden, f, \rho)}
  \end{cases}
\end{verbatim}

\textbf{4.2 Iteration over sets}

Checks over sets can be iterated using the \textit{forall} construct:

\begin{verbatim}
forall id in expr do
  statements
end
\end{verbatim}

As shown in \textbf{Figure 4–6 (on the next page)}, the expression \textit{expr} must evaluate to a set \text{S}. The list of statements \textit{statements} is then evaluated for all bindings of the name \textit{id} to some element \textit{e} of \text{S}. 

4. Constraints  4.2 Iteration over sets

In practice, as failed checks forbid the candidate execution, this amounts to checking the conjunction of the checks within statements for all the elements of $S$. Similar to procedure calls, the bindings performed during an iteration are discarded when iteration ends, and all other effects (e.g., checks) are retained.

**Figure 4–7 (below)** shows the semantics of iteration. The iteration is non-deterministic because the choice $e$ in $S$ is arbitrary and unknown.
5. Requirements

Requirements (Figure 5–1 (below)) are the constitutive blocks of a cat specification. Requirements can be statements or with bindings. Figure 5–2 (below) shows the semantics of requirements.

5.1 Statements

Statements and their semantics are presented in the prior sections.

At the level of requirements, lists of statements are evaluated, evaluation is gathered \((j, f, \rho)\), and the environment \(\rho\) is forgotten.

5.2 Quantification via with binding

Quantification via with binding happens through the construct with id from expr. This construct extends the current environment by one binding (Figure 5–2 (above)).
The grammar allows `with` bindings to occur only at the top level. The expression `expr` is evaluated to a set `S`. Then the remainder of the specification is evaluated for each choice of element `e` in `S` in an environment extended by a binding of the name `id` to `e`.

The verdict is given at the top level as follows (Figure 5–3 (below)): If `S` is empty, `forbidden` is returned (with unmodified flags); otherwise, all possible evaluations are gathered `<j_i, f_i, \rho_i>`, with `i` ranging over some set `\Delta` modeling all possibilities of the `with` quantification). Then:

- If for one of the possible `with` choices `e` in `S`, the judgment `j_i` is `allowed`, `<allowed, f>` is returned, where `f` is the set of flags of all possible `allowed` choices in `S`.
- Otherwise, if the (e.g., the judgment `j_i` is `forbidden` for all possible choices), `forbidden` (with unmodified flags) is returned.

![Figure 5-3 Verdict at top level](image)

**Verdict at top-level**

- `verdict \emptyset \triangleq forbidden`
- `verdict S \triangleq error` when `error \in S`
- `verdict \{ (j_i, f_i, \rho_i) \mid i \in \Delta \} \triangleq` otherwise

- `let f = \bigcup \{ f_i \mid i \in \Delta \land j_i = allowed \}` in
- `if \exists i \in \Delta . j_i = allowed` then `<allowed, f>`
- `else` `<forbidden, \emptyset>`
6. References


