

# FABRIC: A Functional Language with Algebraic Data Types

CS320 Programming Languages Project #3 (Due: June 28, 2020)

## 1 INTRODUCTION

FABRIC is a toy language for the third project of the CS320 course. FABRIC stands for a **f**unctional language with **algebraic** data types. As the name implies, it features first-class functions and algebraic data types.

We consider FABRIC as a variant of FIBER. They share similar concrete syntax. Also, the operational semantics is the same for the shared features. On the other hand, unlike FIBER, which is a dynamically typed language, FABRIC is a statically typed language. FABRIC allows programmers to define custom types, which is another big difference from FIBER. Due to the presence of custom types, FABRIC lacks some features in FIBER: tuples and lists. The reason is that they can be defined by programmers in FABRIC, and thus including them as language primitives would create redundant work.

While X-FIBER provides control diverters via first-class continuations and exceptions, FABRIC lacks them. The main reason to exclude such features is to reduce your burden to complete the project. However, thinking how to add control diverters in FABRIC would be interesting to you. Also, we hope your previous experience with X-FIBER in Project 2 helps you finish this project, which is quite large, although some concepts, like the continuation passing style, from X-FIBER cannot be directly used in FABRIC. We would like to say  $\text{FIBER} \times \text{FIBER} = \text{FABRIC}$ .

More precisely, FABRIC supports the following features (the bold parts are not from FIBER):

- integers, booleans, **strings, and the unit value**
- basic arithmetic operators, including negation, addition, subtraction, multiplication, division, and modulo
- basic relational operators, including equal-to, not-equal-to, less-than, less-than-or-equal-to, greater-than, and greater-than-or-equal-to.
- basic boolean operators, including negation, conjunction, and disjunction
- **expression sequencing**
- conditional expressions (if and if-else expressions)
- immutable local variables
- **mutually recursive immutable lazy local variables**
- **mutable local variables and assignments**
- first-class functions and function application
- anonymous functions
- mutually recursive functions
- **polymorphic first-order functions**
- **polymorphic algebraic data types and pattern matching**
- **static type checking**

This document defines FABRIC and provides a guide to the project. First, it gives the syntax of FABRIC: Section 2 describes the concrete syntax; Section 3 formalizes the desugaring rules; Section 4 shows the abstract syntax. Second, it describes type checking of FABRIC in Section 5, type erasure of FABRIC in Section 6, and the dynamic semantics of FABRIC in Section 7. Third, it briefly introduces the standard library of FABRIC in Section 8. Finally, Section 9 explains the directory structure and rules for implementation. In addition, Appendix A shows the typing rules of FABRIC, and Appendix B shows the big-step operational semantics of FABRIC.

## 2 CONCRETE SYNTAX

The concrete syntax of FABRIC is written in the [extended Backus–Naur form](#). To improve the readability, we use different colors for different kinds of objects. Syntactic elements of the extended Backus–Naur form, rather than FABRIC, are written in [purple](#). For example, we use `=`, `|`, and `;`. Note that `{ }` denotes a repetition of zero or more times, and `[ ]` denotes an optional existence. Nonterminals are written in [blue](#). For example, `expr` is a nonterminal denoting expressions. Any other objects written in black are terminals. For instance, `"true"` and `"false"` are terminals representing boolean literals.

The parsing phase belongs to the given portion of the interpreter. Therefore, you do not need either to implement a parser or to deal with the concrete syntax directly. However, the concrete syntax helps you write your own test cases, and we recommend you to understand the concrete syntax briefly even though you are free to skip tedious details.

The following is the concrete syntax of FABRIC:

```
ltr  = "A" | "B" | "C" | "D" | "E" | "F" | "G" | "H" | "I" | "J" | "K" | "L"
      | "M" | "N" | "O" | "P" | "Q" | "R" | "S" | "T" | "U" | "V" | "W" | "X"
      | "Y" | "Z" | "a" | "b" | "c" | "d" | "e" | "f" | "g" | "h" | "i" | "j"
      | "k" | "l" | "m" | "n" | "o" | "p" | "q" | "r" | "s" | "t" | "u" | "v"
      | "w" | "x" | "y" | "z" ;
pdgt = "1" | "2" | "3" | "4" | "5" | "6" | "7" | "8" | "9" ;
dgt  = "0" | pdgt ;

sch  = ltr | "_" ;
ch   = sch | dgt ;
id   = sch {ch} ;
tnm  = id ;
tvar = "" id ;

num  = ["-"] dgt {dgt} ;
str  = "" [ch] "" "<" expr "," expr ">" ;

expr = id ["[" type {" type} "]" ]
      | num | "true" | "false" | "(" " )" | str | "-" expr | "!" expr
      | expr "+" expr | expr "-" expr | expr "*" expr | expr "/" expr
      | expr "%" expr | expr "==" expr | expr "!=" expr | expr "<" expr
      | expr "<=" expr | expr ">" expr | expr ">=" expr | expr "&&" expr
      | expr "||" expr | "{" expr {" ;" expr } "}"
      | "if" "(" expr ")" expr ["else" expr]
      | "val" id [":" type] "=" expr ";" expr
      | "var" id [":" type] "=" expr ";" expr
      | rdef {rdef} expr
      | "(" " )" "=>" expr | "(" id ":" type {" id ":" type } ")" "=>" expr
      | id "=" expr
      | expr "(" " )" | expr "(" expr {" , " expr } ")"
      | expr "match" "{" case {case} "}"
      | "(" expr ")" ;

vrnt = "case" id | "case" id "(" type {" type } ")"
```

		==		
*		!=		
/	+	<	&&	
%	-	<=		
		>		
		>=		
← higher		lower		

Fig. 1. Operator Precedence

```

case = "case" id "=>" expr | "case" id "(" id {"", " id"} ")" "=>" expr
rdef = "lazy" "val" id ":" type "=" expr ";"
      | "def" id "[" [" tvar {"", " tvar"} "]" "(" " " ")" ":" type "=" expr ";"
      | "def" id "[" [" tvar {"", " tvar"} "]" "(" id ":" type {"", " id ":" type"} ")"
        ":" type "=" expr ";"
      | "type" tnm "[" [" tvar {"", " tvar"} "]" "{" vrnt {vrnt} "}" ;

type = tnm "[" [" type {"", " type"} "]" | tvar | "Int" | "Boolean" | "Unit"
      | "(" " " " " ">=" type | type ">=" type | "(" type {"", " type"} " " ">=" type
      | "(" type " " " " ;

```

Note that whitespaces, such as ' ', '\t', and '\n', are omitted from the above specification. You can insert any kinds of whitespaces between any two terminals to make a valid program. For example, since we have `expr = num | "-" expr`, if one parses `-1` and `- 1`, then both will succeed, and the results will be the same. On the other hand, because you cannot insert whitespaces at the middle of terminals, `true` cannot be parsed while `true` can be parsed correctly.

Figure 1 shows operator precedence. One appearing earlier in the table precedes one appearing later. In FABRIC, all the binary operators for expressions are left-associative. The only binary operator for types, `=>`, is right-associative.

We provide some example programs following the concrete syntax to help you understand the concrete syntax. At the same time, the examples give you a quick overview of FABRIC.

Expressions separated by semicolons inside a pair of curly braces are sequenced. Note that expression sequencing thus always needs curly braces, and semicolons are not an operator. It acts like an operator only inside curly braces. Sequencing is left-associative, like all the other operators. The following is a typical usage of sequencing and results in 2 if both `f` and `g` terminate normally.

```

{
  f();
  g();
  2
}

```

Strings are not literals in FABRIC. Instead, they are considered as finite foldable collections of integers. Thus, a string requires a value corresponding to the empty string and a function to folds the sequence. The following program results in 294, which is the sum of the ascii codes of 'a', 'b', and 'c'. It may seem weird at the first sight, but if it is used with user-defined lists, then it creates a perfect string, which is a list of integers. Since strings are just syntactic sugar in FABRIC, you can finish the project while ignoring them. If you want to know more detail, refer to Section 3.

"abc"<(x: Int, y: Int) => x + y, 0>

FABRIC allows programmers to omit else branches in conditional expressions. When one else branch is absent, then it is equal to write (), which denotes the unit value, in the else branch. For example, the following programs are equivalent for some function f whose return type is Unit:

<pre> if (1 &gt; 2)   f() else   () </pre>	<pre> if (1 &gt; 2)   f() </pre>
--	----------------------------------

If an else branch is omitted in a nested conditional expressions, the existing else branch is combined with the innermost then branch. For instance, the following programs are equivalent for some functions f and g whose return types are Unit:

<pre> if (1 &lt; 2)   if (1 &gt; 2)     f() else   g() </pre>	<pre> if (1 &lt; 2) {   if (1 &gt; 2)     f()   else     g() } </pre>
---	---

Both of the following programs pass type checking and have the same run-time behavior. The only difference is whether the type annotation of x exists or not. Note that type annotations for variables defined by val are optional.

<pre> val x: Int = 1; x </pre>	<pre> val x = 1; x </pre>
--------------------------------	---------------------------

Variables defined by var are similar to those defined by val except that they are mutable. Their values can change by assignments. Type annotations for them also are optional.

<pre> var x: Int = 1; x = 2 </pre>	<pre> var x = 1; x = 2 </pre>
------------------------------------	-------------------------------

Unlike val and var, variables defined by lazy val always require type annotations. In addition, they can be mutually recursive if they are defined consecutively. Therefore, the following program passes type checking and results in 120 at run time:

```

lazy val fac: Int => Int = (x: Int) => if (x <= 1) 1 else x * fac(x - 1);
fac(5)

```

Just like lazy val, functions defined by def always require parameter type annotations and return type annotations. They can be mutually recursive as well. Thus, you may notice that lazy val and def are quite the same except that lazy val leads to caching, while def does not. You will see their semantics in Section 7. The following program defines the factorial function like before but uses def:

```

def fac(x: Int): Int = if (x <= 1) 1 else x * fac(x - 1);
fac(5)

```

Since FABRIC does not support type inference, every anonymous function requires type annotations for all the parameters of itself. Regardless of the number of parameters, parameters always must appear in a pair of parentheses. The following are example programs using anonymous functions and function applications:

<pre>((x: Int) =&gt; x + x)(2)</pre>	<pre>((x: Int, y: Int) =&gt; x * y)(3, 4)</pre>
--------------------------------------	---

The keyword `type` allows defining custom types. Every custom type in FABRIC is an algebraic data type. Types defined together can be mutually recursive, like `lazy val` and `def`. The following is a typical program defining lists of integers:

```
type IntList {
  case IntNil
  case IntCons(Int, IntList)
}
IntCons(0, IntCons(1, IntCons(2, IntNil)))
```

Pattern matching handles values of custom types. The following defines a function computing the length of a given list.

```
def len(l: IntList): Int = l match {
  case IntNil => 0
  case IntCons(h, t) => 1 + len(t)
};
len(IntCons(0, IntCons(1, IntCons(2, IntNil))))
```

Different kinds of recursive definitions can be mutually recursive. In FABRIC, three kinds of recursive definitions exist: lazy variables, functions, and custom types. For example, consecutively defined functions and lazy variables can be mutually recursive. The following defines functions checking the parity of a given natural number. It works even though `odd` is a lazy variable and `even` is a function.

```
lazy val odd: Int => Boolean = (x: Int) => if (x == 0) false else even(x - 1);
def even(x: Int): Boolean = if (x == 0) true else odd(x - 1);
even(5)
```

Parametric polymorphism is the most interesting and powerful feature in FABRIC. Even though FABRIC supports only a restrictive form of parametric polymorphism, it is still quite powerful. In FABRIC, values cannot be polymorphic, while type definitions and non-anonymous function definitions can be polymorphic. For the ease of implementation, type variables are syntactically distinguished from the names of custom types: type variables always start with `'` (single quotation). This syntax is influenced by Standard ML. The following shows a polymorphic identity function:

```
def f['T](x: 'T): 'T = x
f[Int](1)
```

The following defines polymorphic lists in a typical way:

```
type List['T] {
  case Nil
  case Cons('T, List['T])
}
Cons[Int](0, Cons[Int](1, Cons[Int](2, Nil[Int])))
```

### 3 DESUGARING

To simplify the implementation of the interpreting phase, the parsing phase of the interpreter desugars a given expression. Desugaring rewrites some subexpressions with other expressions. Due to desugaring, the abstract syntax of FABRIC consists of less sorts of expressions than the concrete syntax.

Like the concrete syntax, you do not need to care about desugaring in detail. However, understanding desugaring helps you find how test cases are transformed.

$\llbracket -e \rrbracket = \llbracket e \rrbracket * -1$	$\llbracket e_1 > e_2 \rrbracket = \llbracket \neg (\llbracket e_1 \rrbracket \leq \llbracket e_2 \rrbracket) \rrbracket$
$\llbracket !e \rrbracket = \text{if } (\llbracket e \rrbracket) \text{ false else true}$	$\llbracket e_1 \geq e_2 \rrbracket = \llbracket \neg (\llbracket e_1 \rrbracket < \llbracket e_2 \rrbracket) \rrbracket$
$\llbracket e_1 - e_2 \rrbracket = \llbracket e_1 \rrbracket + \llbracket -e_2 \rrbracket$	$\llbracket e_1 \ \&\& \ e_2 \rrbracket = \text{if } (\llbracket e_1 \rrbracket) \llbracket e_2 \rrbracket \text{ else false}$
$\llbracket e_1 != e_2 \rrbracket = \llbracket \neg (\llbracket e_1 \rrbracket == \llbracket e_2 \rrbracket) \rrbracket$	$\llbracket e_1 \    \ e_2 \rrbracket = \text{if } (\llbracket e_1 \rrbracket) \text{ true else } \llbracket e_2 \rrbracket$
$\llbracket e_1 \leq e_2 \rrbracket = \text{val } \underline{x_1} = \llbracket e_1 \rrbracket;$ $\text{val } \underline{x_2} = \llbracket e_2 \rrbracket;$ $\llbracket x_1 == x_2 \    \ x_1 < x_2 \rrbracket$	$\llbracket \text{if } (e_1) \ e_2 \rrbracket = \text{if } (\llbracket e_1 \rrbracket) \llbracket e_2 \rrbracket \text{ else } ()$ $\llbracket "n_1 \dots n_i" < e_1, \ e_2 > \rrbracket = \text{val } \underline{x} = \llbracket e_1 \rrbracket;$ $x(n_1 \dots, \ x(n_i, \ \llbracket e_2 \rrbracket)) \dots$

Any other cases recursively desugar their subexpressions.

Fig. 2. Desugaring Rules

Identifier	$x \in Id$	Expression	$e ::= x[\bar{\tau}]$	(type application)
Number	$n \in \mathbb{Z}$		$  \ n$	(integer)
Boolean	$b ::= \text{true}$ $  \ \text{false}$		$  \ b$	(boolean)
			$  \ ()$	(unit)
			$  \ e + e$	(addition)
Type Name	$t \in TN$		$  \ e \times e$	(multiplication)
Type Variable	$\alpha \in TV$		$  \ e \div e$	(division)
Type	$\tau ::= t[\bar{\tau}]$ $  \ \alpha$ $  \ \text{Int}$ $  \ \text{Boolean}$ $  \ \text{Unit}$ $  \ \bar{\tau} \rightarrow \tau$		$  \ e \bmod e$	(modulo)
			$  \ e = e$	(equal-to)
			$  \ e < e$	(less-than)
			$  \ e ; e$	(sequence)
			$  \ \text{if } e \ e \ e$	(conditional)
			$  \ \mu \ x[\bar{\tau}] = e \text{ in } e$	(local variable)
Mutability	$\mu ::= \text{var} \   \ \text{val}$		$  \ \bar{d} \text{ in } e$	(recursive definitions)
Variant	$w ::= x(\bar{\tau})$		$  \ \lambda \bar{x} : \bar{\tau}. e$	(anonymous function)
Case	$c ::= x(\bar{x}) \rightarrow e$		$  \ x := e$	(assignment)
Recursive Definition	$d ::= \text{lazy } x : \tau = e$ $  \ \text{def } x[\bar{\alpha}](\bar{x} : \bar{\tau}) : \tau = e$ $  \ \text{type } t[\bar{\alpha}] \ \bar{w}$		$  \ e(\bar{e})$	(function application)
			$  \ e \text{ match } \bar{c}$	(pattern matching)

Fig. 3. Abstract Syntax

Figure 2 defines desugaring of FABRIC expressions. Let  $e$  and  $x$  respectively denote an expression and an identifier. An expression  $e$  is desugared to  $\llbracket e \rrbracket$ .

#### 4 ABSTRACT SYNTAX

We use an overline above mathematical material to indicate an appropriately punctuated sequence of zero or more repetitions of the material, with a subscript added to each metavariable in the material. For example,  $t[\bar{\tau}]$  equals  $t[\tau_1, \dots, \tau_n]$  for some natural number  $n$ . Note that  $n$  can be zero, which allows  $t[\bar{\tau}]$  to be the same as  $t[]$ .

We use a box around mathematical material to indicate an optional existence. For instance,  $\mu \ x[\bar{\tau}] = e \text{ in } e$  can be either  $\mu \ x = e \text{ in } e$  or  $\mu \ x : \tau = e \text{ in } e$  since the existence of  $\tau$  is optional.

Figure 3 describes the abstract syntax of FABRIC. Metavariable  $x$  ranges over identifiers;  $n$  ranges over integers;  $b$  ranges over boolean literals, which are either true or false;  $t$  ranges over the names

of custom types;  $\alpha$  ranges over type variables, which appear when polymorphic functions and methods are defined;  $\tau$  ranges over types;  $\mu$  ranges over mutability tags for strict local variables, which are either `var` or `val`;  $w$  ranges over variant definitions;  $c$  ranges over cases in pattern matching;  $d$  ranges over recursive definitions, which are either lazy variables, functions, or custom types;  $e$  ranges over expressions.

The following briefly describes types:

- $t[\bar{\tau}]$  is a first-order type application.  $t$  is the name of a custom type;  $\tau$ 's are type arguments. If type arguments do not exist, then  $t$  is a monomorphic type. For example, `Fruit[]` can be a type where `Fruit` is the name of a monomorphic type. If there are one or more type arguments, then  $t$  is a polymorphic type. For instance, `List[Int]` can be a type where `List` is the name of a polymorphic type with one type parameter. `Int` is the only type argument in this case.
- $\alpha$  is a type variable.
- `Int` is an integer type.
- `Boolean` is a boolean type.
- `Unit` is a unit type.
- $\bar{\tau}_p \rightarrow \tau_r$  is a function type.  $\tau_p$ 's are the types of the parameters;  $\tau_r$  is the return type. For example, `Int  $\rightarrow$  Boolean` is the type of a function from an integer to a boolean value.

The following briefly describes expressions:

- $x[\bar{\tau}]$  is an identifier instantiated with type arguments.  $x$  is the name of a variable, a parameter, a function, or a variant;  $\tau$ 's are type arguments. If type arguments do not exist, then  $x$  is a monomorphic entity. For example, `x[]` can denote a variable whose name is `x`. If there are one or more type arguments, then  $x$  is a polymorphic entity. For instance, when there is a polymorphic function whose name is `f`, `f[Int]` denotes the function instantiated with `Int`. Note that all variables and parameters are monomorphic, while functions and variants can be either monomorphic or polymorphic.
- $n$  is an integral literal;  $b$  is a boolean literal; `()` is the unit value.
- $e_1 + e_2$ ,  $e_1 \times e_2$ ,  $e_1 \div e_2$ ,  $e_1 \bmod e_2$ ,  $e_1 = e_2$ , and  $e_1 < e_2$  are binary operations on integers.
- $e_1 ; e_2$  is a sequence of expressions.  $e_1$  is evaluated before  $e_2$  is evaluated.
- `if  $e_1$   $e_2$   $e_3$`  is a conditional expression.  $e_1$  is the condition;  $e_2$  is the true branch;  $e_3$  is the false branch.
- $\mu x = e_1$  in  $e_2$  is a local variable definition without a type annotation. It defines a local variable whose name is  $x$  and scope is  $e_2$ .  $e_1$  determines the value of  $x$ .  $\mu$  is either `val` or `var`. If  $\mu$  is `val`, then  $x$  is an immutable variable. Otherwise, it is a mutable variable.  $\mu x : \tau = e$  in  $e$  is a local variable definition with a type annotation for  $x$ .  $\tau$  is the type of  $x$ .
- $\bar{d}$  in  $e$  defines recursive definitions. Each  $d$  is either a lazy variable definition, a function definition, or a type definition. Variables, functions, and types defined by  $d$ 's can be used in  $e$ . Identifiers introduced by  $d$ 's must have distinct names; type identifiers introduced by  $d$ 's must have distinct names.
  - lazy  $x : \tau = e$  is a lazy variable definition.  $x$  is the name of the variable;  $\tau$  is the type of the variable;  $e$  determines the value of the variable.
  - `def  $x[\bar{\alpha}](\bar{x}_p : \bar{\tau}_p) : \tau_r = e$`  is a function definition.  $x$  is the name of the function;  $\alpha$ 's are the type parameters of the function;  $x_p$ 's are the parameters of the function;  $\tau_p$ 's are type annotations for the parameters;  $\tau_r$  is the return type of the function;  $e$  is the body of the function.  $\alpha$ 's must be distinct type variables;  $x_p$ 's must have distinct names.
  - type `t[ $\bar{\alpha}$ ]  $\bar{w}$`  is a type definition.  $t$  is the name of the type;  $\alpha$ 's are the type parameters of the type;  $w$ 's are the variants of the type.  $\alpha$ 's must be distinct type variables; all the variants

must have distinct names. There must be one or more variants. Each variant definition  $w$  has the syntax of  $x(\bar{\tau})$ .  $x$  is the name of the variant;  $\tau$ 's are the types of the parameters of the variant.

- $\lambda\bar{x}:\bar{\tau}.e$  is an anonymous function.  $x$ 's are the parameters of the function;  $\tau$ 's are type annotations for the parameters;  $e$  is the body of the function.  $x$ 's must have distinct names.
- $x := e$  updates the value of a variable denoted by  $x$ .  $e$  determines the new value of  $x$ .
- $e(\bar{e}_a)$  is a function application.  $e$  is a function to be applied to arguments;  $e_a$ 's are arguments for the function application.
- $e \text{ match } \bar{c}$  is a pattern matching expression on  $e$ .  $c$ 's are the cases of the pattern matching. All the cases must have distinct variant names. There must be one or more cases. Each case  $c$  has the syntax of  $x(\bar{x}_p) \rightarrow_{e_c}$ .  $x$  is the name of a variant; values carried by a variant value can be bound to  $x_p$ 's.  $e_c$  determines the result of the pattern matching. The scope of each  $x_p$  is  $e_c$ .

## 5 TYPE CHECKING

This section explains type checking of FABRIC in a natural language. See Appendix A to find the formal static semantics of FABRIC.

First, well-formedness of types has to be defined. We use the following metavariable and terminologies:

- Metavariable  $\Gamma$  ranges over type environments. A type environment contains two kinds of mappings:
  - A mapping from an identifier to a pair of a type scheme and a mutability tag. The following is the definition of a type scheme:

$$\text{Type Scheme } T ::= \forall \bar{\alpha}.\tau$$

A type scheme is a type followed by a list of zero or more type variables. The type of a type scheme is parametrized by the type variables of the type scheme. We use type schemes to represent both monomorphic and polymorphic types. A type scheme representing a monomorphic type  $\tau$  is  $\forall.\tau$ . A type scheme representing a polymorphic type parametrized by a single type variable is  $\forall\alpha.\tau$ . Note that type schemes exist only inside type environments. When a polymorphic type is used in an expression, the type has to be instantiated with a proper number of type arguments. Therefore, the type of every expression is a type, which is always monomorphic, not a type scheme.

- A mapping from a type name to a type definition, whose syntax is type  $t[\bar{\alpha}] \bar{w}$ . A type definition type  $t[\bar{\alpha}] \bar{w}$  creates a mapping from  $t$  to type  $t[\bar{\alpha}] \bar{w}$ .

In addition, a type environment contains type variables.

- Substituting  $\alpha$  with  $\tau'$  in  $\tau$  produces a type that is the same as  $\tau$  except that every occurrence of  $\alpha$  has been replaced with  $\tau'$ . For example, substituting  $\alpha$  with  $\text{Boolean}$  in  $\alpha, \alpha \rightarrow \text{Int}$  results in  $\text{Boolean}, \text{Boolean} \rightarrow \text{Int}$ . If multiple substitutions happen simultaneously, each substitution does not affect the others. For instance, consider  $\alpha \rightarrow \beta$ . If one substitutes  $\alpha$  with  $\beta$  and  $\beta$  with  $\gamma$  simultaneously, the result is  $\beta \rightarrow \gamma$ , not  $\gamma \rightarrow \gamma$ .
- Every sentence explaining type checking assumes use of the current type environment, which is given to a case including the sentence at the beginning unless a particular type environment is specified.
- If we say “it is well-formed” while explaining well-formedness of  $\tau$ , then  $\tau$  is well-formed.
- We use the word “must” to represent requirements. If a requirement is violated while well-formedness of  $\tau$  is explained, then  $\tau$  is not well-formed.

The following explains how well-formedness of each type is checked.



**Case  $t[\tau_1, \dots, \tau_n]$ :**

- (1) Let  $\Gamma$  be the current type environment.
- (2) For every  $1 \leq i \leq n$ ,
  - (a)  $\tau_i$  must be well-formed.
- (3)  $t$  must be in the domain of  $\Gamma$ .
- (4) Let  $\Gamma(t)$  be type  $t[\alpha_1, \dots, \alpha_m] \bar{w}$ .
- (5)  $n$  must be the same as  $m$ . (The number of the type arguments must be the same as the number of the type parameters.)
- (6) It is well-formed.

**Case  $\alpha$ :**

- (1) Let  $\Gamma$  be the current type environment.
- (2)  $\alpha$  must be in  $\Gamma$ .

(3) It is well-formed.

**Case Int:**

- (1) It is well-formed.

**Case Boolean:**

- (1) It is well-formed.

**Case Unit:**

- (1) It is well-formed.

**Case  $\tau_1, \dots, \tau_n \rightarrow \tau$ :**

- (1) For every  $1 \leq i \leq n$ ,
  - (a)  $\tau_i$  must be well-formed.
- (2)  $\tau$  must be well-formed.
- (3) It is well-formed.

Now, type checking of expressions can be defined. We use the following terminologies:

- If we say “the type is  $\tau$ ” while explaining type checking of  $e$ , then the type of  $e$  is  $\tau$ , and  $e$  is well-typed.
- We use the word “must” to represent requirements. If a requirement is violated while type checking of  $e$  is explained, then  $e$  is not well-typed.
- If we say “the type of  $e$  must be  $\tau$ ,” then it implicitly adds the requirement that “ $e$  must be well-typed.”
- If we say “let  $\tau$  be the type of  $e$ ,” then it implicitly adds the requirement that “ $e$  must be well-typed.”

The following explains how each expression is type-checked.

**Case  $x[\tau_1, \dots, \tau_n]$ :**

- (1) Let  $\Gamma$  be the current type environment.
- (2) For every  $1 \leq i \leq n$ ,
  - (a)  $\tau_i$  must be well-formed.
- (3)  $x$  must be in the domain of  $\Gamma$ .
- (4) Let  $\Gamma(x)$  be  $(\forall \alpha_1, \dots, \alpha_m. \tau, \mu)$ .
- (5)  $n$  must be the same as  $m$ . (The number of the type arguments must be the same as the number of the type parameters.)
- (6) For every  $1 \leq i \leq n$ ,
  - (a) Substitute every occurrence of  $\alpha_i$  in  $\tau$  with  $\tau_i$ .  
(Every substitution happens simultaneously.)
- (7) Let  $\tau'$  be the result of substitution.
- (8) The type is  $\tau'$ .

**Case  $n$ :**

- (1) The type is Int.

**Case  $b$ :**

- (1) The type is Boolean.

**Case  $()$ :**

- (1) The type is Unit.

**Case  $e_1 \oplus e_2$ :**

- (\*) Suppose that  $\oplus \in \{+, \times, \div, \text{mod}\}$ .
- (1) The type of  $e_1$  must be Int.
- (2) The type of  $e_2$  must be Int.
- (3) The type is Int.

**Case  $e_1 \otimes e_2$ :**

- (\*) Suppose that  $\otimes \in \{=, <\}$ .
- (1) The type of  $e_1$  must be Int.
- (2) The type of  $e_2$  must be Int.
- (3) The type is Boolean.

**Case  $e_1 ; e_2$ :**

- (1)  $e_1$  must be well-typed.
- (2) Let  $\tau$  be the type of  $e_2$ .
- (3) The type is  $\tau$ .

**Case if  $e_1 \ e_2 \ e_3$ :**

- (1) The type of  $e_1$  must be Boolean.
- (2) Let  $\tau$  be the type of  $e_2$ .
- (3) The type of  $e_3$  must be  $\tau$ .
- (4) The type is  $\tau$ .

**Case**  $\mu x:\tau'_1=e_1$  in  $e_2$ :

- (1) Let  $\Gamma$  be the current type environment.
- (2) If  $\tau'_1$  exists, then
  - (a)  $\tau'_1$  must be well-formed.
- (3) Let  $\tau_1$  be the type of  $e_1$ .
- (4) If  $\tau'_1$  exists, then
  - (a)  $\tau'_1$  must be equal to  $\tau_1$ .
- (5) Add a mapping from  $x$  to  $(\forall.\tau_1, \mu)$  to  $\Gamma$ .
- (6) Let  $\Gamma'$  be the new type environment.
- (7) Let  $\tau$  be the type of  $e_2$  under  $\Gamma'$ .
- (8) The type is  $\tau$ .

**Case**  $d_1, \dots, d_n$  in  $e$ :

- (1) Let  $\Gamma_0$  be the current type environment.
- (2) For every  $1 \leq i \leq n$ ,
  - (a) If  $d_i$  is a type definition, then
    - (i) Let  $d_i$  be type  $t[\bar{\alpha}] \bar{w}$ .
    - (ii)  $t$  must not be in the domain of  $\Gamma_0$ .
  - (b) Let  $d_i$  creates  $\Gamma'_i$ .
  - (c) Add every mapping in  $\Gamma'_i$  to  $\Gamma_{i-1}$ .
  - (d) Let  $\Gamma_i$  be the new type environment.
- (3) For every  $1 \leq i \leq n$ ,
  - (a)  $d_i$  must be well-formed under  $\Gamma_n$ .
- (4) Let  $\tau$  be the type of  $e$  under  $\Gamma_n$ .
- (5)  $\tau$  must be well-formed under  $\Gamma$ .
- (6) The type is  $\tau$ .

**Case**  $\lambda x_1:\tau_1, \dots, x_n:\tau_n.e$ :

- (1) Let  $\Gamma_0$  be the current type environment.
- (2) For every  $1 \leq i \leq n$ ,
  - (a)  $\tau_i$  must be well-formed under  $\Gamma_0$ .
- (3) For every  $1 \leq i \leq n$ ,
  - (a) Add a mapping from  $x_i$  to  $(\forall.\tau_i, \text{val})$  to  $\Gamma_{i-1}$ .
  - (b) Let  $\Gamma_i$  be the new type environment.
- (4) Let  $\tau$  be the type of  $e$  under  $\Gamma_n$ .
- (5) The type is  $\tau_1, \dots, \tau_n \rightarrow \tau$ .

**Case**  $x := e$ :

- (1) Let  $\Gamma$  be the current type environment.
- (2)  $x$  must be in the domain of  $\Gamma$ .
- (3) Let  $\Gamma(x)$  be  $(\forall \bar{\alpha}.\tau, \mu)$ .
- (4) The length of  $\bar{\alpha}$  must be zero. (There must be no type parameter.)
- (5)  $\mu$  must be var.
- (6)  $\tau$  must be the type of  $e$ .
- (7) The type is Unit.

**Case**  $e(e_1, \dots, e_n)$ :

- (1) Let  $\tau_f$  be the type of  $e$ .
- (2)  $\tau_f$  must be a function type.
- (3) Let  $\tau_f$  be  $\tau_1 \dots \tau_m \rightarrow \tau$ .
- (4)  $n$  must be the same as  $m$ . (The number of the arguments must be the same as the number of the parameters.)
- (5) For every  $1 \leq i \leq n$ ,
  - (a)  $\tau_i$  must be the type of  $e_i$ .
- (6) The type is  $\tau$ .

**Case**  $e$  match  $c_1, \dots, c_n$ :

- (1) Let  $\Gamma$  be the current type environment.
- (2) Let  $\tau_m$  be the type of  $e$ .
- (3)  $\tau_m$  must be a type application.
- (4) Let  $\tau_m$  be  $t[\tau_1, \dots, \tau_m]$ .
- (5)  $t$  must be in the domain of  $\Gamma$ .
- (6) Let  $\Gamma(t)$  be type  $t[\alpha_1, \dots, \alpha_l] w_1, \dots, w_k$ .
- (7)  $m$  must be the same as  $l$ . (The number of the type arguments must be the same as the number of the type parameters.)
- (8)  $n$  must be the same as  $k$ . (The number of the cases must be the same as the number of the variants.)
- (9) For every  $1 \leq i \leq n$ ,
  - (a) Let  $\tau'_i$  be the type of  $c_i$  when  $w_1, \dots, w_k, \alpha_1, \dots, \alpha_l$ , and  $\tau_1, \dots, \tau_m$  are given.
- (10) For every  $1 \leq i \leq n$ ,
  - (a)  $\tau'_i$  must be equal to  $\tau'_1$ .
- (11) The type is  $\tau'_1$ .

In addition to the types of expressions, we need the types of cases. We use the following terminologies:

- If we say “the type is  $\tau$ ” while explaining type checking of  $c$ , then the type of  $c$  is  $\tau$ , and  $c$  is well-typed.
- We use the word “must” to represent requirements. If a requirement is violated while type checking of  $c$  is explained, then  $c$  is not well-typed.
- If we say “let  $\tau$  be the type of  $e$ ,” then it implicitly adds the requirement that “ $e$  must be well-typed.”

- Type checking of  $c$  is possible only when a list of variant definitions, a list of type variables, and a list of types are given.

The following explains type checking of cases.

**Case**  $x(x_1, \dots, x_n) \rightarrow e$ :

- (1) Let  $w_1, \dots, w_k$ ,  $\alpha_1, \dots, \alpha_l$ , and  $\tau_1, \dots, \tau_m$  be given.
- (2)  $l$  must be the same as  $m$ .
- (3) Let  $\Gamma_0$  be the current type environment.
- (4) There must be  $w$  among  $w_1, \dots, w_k$  such that the name of  $w$  is  $x$ .
- (5) Let  $w$  be  $x(\tau'_1, \dots, \tau'_h)$ .
- (6)  $n$  must be the same as  $h$ .
- (7) For every  $1 \leq i \leq h$ ,

- (a) For every  $1 \leq j \leq l$ ,
  - (i) Substitute every occurrence of  $\alpha_j$  in  $\tau'_i$  with  $\tau_j$ .  
(Every substitution happens simultaneously.)
- (b) Let  $\tau''_i$  be the result of substitution.
- (c) Add a mapping from  $x_i$  to  $(\forall. \tau''_i, \text{val})$  to  $\Gamma_{i-1}$ .
- (d) Let  $\Gamma_i$  be the new type environment.
- (8) Let  $\tau$  be the type of  $e$  under  $\Gamma_h$ .
- (9) The type is  $\tau$ .

We need to define well-formedness of recursive definitions. We use the following terminologies:

- If we say “it is well-formed” while explaining well-formedness of  $d$ , then  $d$  is well-formed.
- We use the word “must” to represent requirements. If a requirement is violated while well-formedness of  $d$  is explained, then  $d$  is not well-formed.

The following explains well-formedness of recursive definitions.

**Case** lazy  $x:\tau=e$ :

- (1)  $\tau$  must be well-formed.
- (2)  $\tau$  must be the type of  $e$ .
- (3) It is well-formed.

- (a) Add a mapping from  $x_i$  to  $(\forall. \tau_i, \text{val})$  to  $\Gamma_{i-1}$ .
- (b) Let  $\Gamma_i$  be the new type environment.
- (8)  $\tau$  must be the type of  $e$  under  $\Gamma_n$ .
- (9) It is well-formed.

**Case** def  $x[\alpha_1, \dots, \alpha_n](x_1:\tau_1, \dots, x_m:\tau_m):\tau=e$ :

- (1) Let  $\Gamma$  be the current type environment.
- (2) For every  $1 \leq i \leq n$ ,
  - (a)  $\alpha_i$  must not be in  $\Gamma$ .
- (3) Add  $\alpha_1, \dots, \alpha_n$  to  $\Gamma$ .
- (4) Let  $\Gamma_0$  be the new type environment.
- (5) For every  $1 \leq i \leq m$ ,
  - (a)  $\tau_i$  must be well-formed under  $\Gamma_0$ .
- (6)  $\tau$  must be well-formed under  $\Gamma_0$ .
- (7) For every  $1 \leq i \leq m$ ,

**Case** type  $t[\alpha_1, \dots, \alpha_n] w_1, \dots, w_m$ :

- (1) Let  $\Gamma$  be the current type environment.
- (2) For every  $1 \leq i \leq n$ ,
  - (a)  $\alpha_i$  must not be in  $\Gamma$ .
- (3) Add  $\alpha_1, \dots, \alpha_n$  to  $\Gamma$ .
- (4) Let  $\Gamma'$  be the new type environment.
- (5) For every  $1 \leq i \leq m$ ,
  - (a) Let  $w_i$  be  $x(\tau_1, \dots, \tau_l)$ .
  - (b) For every  $1 \leq j \leq l$ ,
    - (i)  $\tau_j$  must be well-formed under  $\Gamma'$ .
- (6) It is well-formed.

We need to define type environments created by recursive definitions as well. We use the following terminology:

- If we say “the result is  $\Gamma$ ” while explaining  $d$ , then  $d$  creates  $\Gamma$ .

The following explains type environments created by recursive definitions.

**Case** lazy  $x:\tau=e$ :

- (1) Add a mapping from  $x$  to  $(\forall. \tau, \text{val})$  to the empty type environment.

- (2) Let  $\Gamma$  be the new type environment.
- (3) The result is  $\Gamma$ .

**Case** def  $x[\bar{\alpha}](\bar{x}:\bar{\tau}):\tau=e$ :

Variant	$W ::= x()$	$  E \div E$	(division)
	$  x(\_)$	$  E \bmod E$	(modulo)
Case	$C ::= x(\bar{x}) \rightarrow E$	$  E = E$	(equal-to)
	Recursive Definition $D ::= \text{lazy } x=E$	$  E < E$	(less-than)
		$  E ; E$	(sequence)
	$  \text{type } \bar{W}$	$  \text{if } E E E$	(conditional)
Expression	$E ::= x$ (variable)	$  \text{let } x=E \text{ in } E$	(local variable)
	$  n$ (integer)	$  \bar{D} \text{ in } E$	(recursive definition)
	$  b$ (boolean)	$  \lambda \bar{x}. E$	(anonymous function)
	$  ()$ (unit)	$  x := E$	(assignment)
	$  E + E$ (addition)	$  E(\bar{E})$	(function application)
	$  E \times E$ (multiplication)	$  E \text{ match } \bar{C}$	(pattern matching)

Fig. 4. Abstract Syntax after Type Erasure

- (1) Add a mapping from  $x$  to  $(\forall \bar{\alpha}. \bar{\tau} \rightarrow \tau, \text{val})$  to the empty type environment.
  - (2) Let  $\Gamma$  be the new type environment.
  - (3) The result is  $\Gamma$ .
- Case type  $t[\bar{\alpha}] w_1, \dots, w_n$ :**
- (1) Add a mapping from  $t$  to type  $t[\bar{\alpha}] w_1, \dots, w_n$  to the empty type environment.
  - (2) Let  $\Gamma_0$  be the new type environment.
  - (3) For every  $1 \leq i \leq n$ ,
    - (a) If  $w_i$  does not take any parameter, then
      - (i) Let  $w_i$  be  $x()$ .
      - (ii) Add a mapping from  $x$  to  $(\forall \bar{\alpha}. t[\bar{\alpha}], \text{val})$  to  $\Gamma_{i-1}$ .
      - (iii) Let  $\Gamma_i$  be the new type environment.
    - (b) Else if  $w_i$  takes one or more parameters, then
      - (i) Let  $w_i$  be  $x(\bar{\tau})$ .
      - (ii) Add a mapping from  $x$  to  $(\forall \bar{\alpha}. \bar{\tau} \rightarrow t[\bar{\alpha}], \text{val})$  to  $\Gamma_{i-1}$ .
      - (iii) Let  $\Gamma_i$  be the new type environment.
  - (4) The result is  $\Gamma_n$ .

## 6 TYPE ERASURE

Type erasure removes type annotations from programs. FABRIC uses type erasure since type annotations take no role at run time. They are used for only type checking. In this section, we define type erasure for FABRIC. Since type erasure is implemented in the provided portion of the code, you do not need to care about type erasure in detail. However, understanding type erasure helps you find how test cases are transformed.

First, we define an untyped version, which lacks type annotations, of FABRIC. Figure 4 describes the abstract syntax of the untyped version. Metavariable  $W$  ranges over untyped variant definitions;  $C$  ranges over untyped cases in pattern matching;  $D$  ranges over untyped recursive definitions;  $E$  ranges over untyped expressions.

Most of the untyped version is similar to the typed version except that they lack type annotations. The following briefly describes notable changes:

- A variant definition  $W$  is either  $x()$  or  $x(\_)$ . The former,  $x()$ , represents a variant without any parameters; the latter,  $x(\_)$ , represents a variant with one or more parameters. It lacks not only type annotations but also the exact number of parameters. However, whether parameters exist or not is important at run time. If there is no parameter, the name of the variant itself denotes the only value of the variant. On the other hand, if there are one or more parameters, the name of the variant has to be applied to a proper number of arguments to produce a value of the variant. For example, when type `List[] Nil(), Cons(Int, List)` is given, both `Nil` and

$\begin{aligned} \text{erase}(e) &= E \\ \text{erase}(d) &= D \\ \text{erase}(w) &= W \\ \text{erase}(c) &= C \end{aligned}$	
$\text{erase}(x[\bar{\tau}]) = x$	$\text{erase}(e_1 ; e_2) = \text{erase}(e_1) ; \text{erase}(e_2)$
$\text{erase}(n) = n$	$\text{erase}(\text{if } e_1 \ e_2 \ e_3) = \text{if } \text{erase}(e_1) \ \text{erase}(e_2) \ \text{erase}(e_3)$
$\text{erase}(b) = b$	$\text{erase}(\mu \ x[\bar{\tau}] = e_1 \text{ in } e_2) = \text{let } x = \text{erase}(e_1) \text{ in } \text{erase}(e_2)$
$\text{erase}(( )) = ( )$	$\text{erase}(\bar{d} \text{ in } e) = \text{erase}(\bar{d}) \text{ in } \text{erase}(e)$
$\text{erase}(e_1 + e_2) = \text{erase}(e_1) + \text{erase}(e_2)$	$\text{erase}(\lambda \bar{x} : \bar{\tau}. e) = \lambda \bar{x}. \text{erase}(e)$
$\text{erase}(e_1 \times e_2) = \text{erase}(e_1) \times \text{erase}(e_2)$	$\text{erase}(x := e) = x := \text{erase}(e)$
$\text{erase}(e_1 \div e_2) = \text{erase}(e_1) \div \text{erase}(e_2)$	$\text{erase}(e(\bar{e})) = \text{erase}(e)(\text{erase}(\bar{e}))$
$\text{erase}(e_1 \bmod e_2) = \text{erase}(e_1) \bmod \text{erase}(e_2)$	$\text{erase}(e \text{ match } \bar{c}) = \text{erase}(e) \text{ match } \text{erase}(\bar{c})$
$\text{erase}(e_1 = e_2) = \text{erase}(e_1) = \text{erase}(e_2)$	$\text{erase}(\text{lazy } x : \tau = e) = \text{lazy } x = e$
$\text{erase}(e_1 < e_2) = \text{erase}(e_1) < \text{erase}(e_2)$	$\text{erase}(\text{def } x[\bar{\alpha}](\bar{x} : \bar{\tau}) : \tau = e) = \text{def } x(\bar{x}) = e$
	$\text{erase}(\text{type } t[\bar{\alpha}] \ \bar{w}) = \text{type } \text{erase}(\bar{w})$
	$\text{erase}(x()) = x()$
	$\text{erase}(x(\bar{\tau})) = x(\_) \quad (\text{if } \#(\bar{\tau}) > 0)$
	$\text{erase}(x(\bar{x}) \rightarrow e) = x(\bar{x}) \rightarrow \text{erase}(e)$

Fig. 5. Type Erasure

Value	$v \in \mathbb{V}$	Address	$a \in \mathcal{A}$
$v ::= n \mid b \mid ( ) \mid \langle \lambda \bar{x}. E, \sigma \rangle$		Environment	$\sigma \in Id \hookrightarrow \mathcal{A}$
$\mid \langle \langle E, \sigma \rangle \rangle \mid \langle x \rangle \mid x(\bar{v})$		Store	$M \in \mathcal{A} \hookrightarrow \mathbb{V}$

Fig. 6. Definitions for Dynamic Semantics

$\text{Cons}(1)$  are well-typed expressions that result in values of the type `List`, while `Nil()` is not a well-typed expression.

- The syntax of a type definition is  $\text{type } \bar{W}$ , which lacks the name of the type. Since the names of types are unnecessary at run time, each type definition contains only a list of variants.
- The syntax of a variable definition is  $\text{let } x = E \text{ in } E$ . It lacks not only a type annotation for the variable but also the mutability tag. Whether a variable is mutable or immutable matters only during type checking. If a program passes type checking, it is guaranteed that the value of every immutable variable never changes. Therefore, mutability tags are not required at run time.

Figure 5 defines type erasure for FABRIC.  $\text{erase}(e)$  denotes an untyped expression created by erasing types in  $e$ . Similarly,  $\text{erase}(d)$ ,  $\text{erase}(w)$ , and  $\text{erase}(c)$  respectively denotes applying type erasure to a recursive definition, a variant definition, and a case in pattern matching. Type erasure is defined in a straightforward recursive manner.

## 7 DYNAMIC SEMANTICS

This section explains the dynamic semantics of FABRIC in a natural language. See Appendix B to find the formal big-step operational semantics of FABRIC.

Figure 6 defines values, addresses, environments, and stores. Metavariable  $v$  ranges over values. A value is one of the following:

- an integer  $n$
- a boolean value  $b$
- the unit value  $( )$

- a closure  $\langle \lambda \bar{x}.E, \sigma \rangle$
- a delayed computation  $\langle \langle E, \sigma \rangle \rangle \in ExprV$
- a constructor  $\langle x \rangle$
- a variant value  $x(\bar{v})$

$\mathbb{V}$  is the set of every value. Metavariable  $a$  ranges over addresses, which are elements of an infinite set  $\mathcal{A}$ . Metavariable  $\sigma$  ranges over environments, which are finite maps from identifiers to addresses. Metavariable  $M$  ranges over stores, which are finite maps from addresses to values.

We use the following terminologies:

- Every sentence explaining the semantics assumes use of the current environment, which is given to a case including the sentence at the beginning unless a particular environment is specified.
- Every sentence assumes use of the most recently created store.
- If we say “the result is  $v$ ” while explaining evaluation of  $e$ , then  $v$  is the result of  $e$ .
- We use the word “must” to represent requirements. If a requirement is violated, then a run-time error occurs. Any occurrence of a run-time error immediately terminates the execution.

The following explains evaluation of expressions.

**Case  $x$ :**

- (1) Let  $\sigma$  be the current environment.
- (2) Let  $M$  be the current store.
- (3)  $x$  must be in the domain of  $\sigma$ .
- (4) Let  $a$  be  $\sigma(x)$ .
- (5)  $a$  must be in the domain of  $M$ .
- (6) Let  $v$  be  $M(a)$ .
- (7) If  $v$  is not a delayed computation, then
  - (a) The result is  $v$ .
- (8) Else if  $v$  is a delayed computation, then
  - (a) Let  $v$  be  $\langle \langle E, \sigma' \rangle \rangle$ .
  - (b) Evaluate  $E$  under  $\sigma'$ .
  - (c) Let  $v_1$  be the result of  $E$ .
  - (d) Let  $M_1$  be the current store.
  - (e) Add a mapping from  $a$  to  $v_1$  to  $M_1$ .
  - (f) The result is  $v_1$ .

**Case  $n$ :**

- (1) The result is  $n$ .

**Case  $b$ :**

- (1) The result is  $b$ .

**Case  $()$ :**

- (1) The result is  $()$ .

**Case  $E_1 \otimes E_2$ :**

- (\*) Suppose that  $\otimes \in \{+, \times, \div, \text{mod}, =, <\}$ .
- (1) Evaluate  $E_1$ .
- (2) Let  $v_1$  be the result of  $E_1$ .
- (3) Evaluate  $E_2$ .
- (4) Let  $v_2$  be the result of  $E_2$ .

- (5)  $(v_1, v_2)$  must be in the domain of  $\otimes$ . Note that  $+, \times \in (\mathbb{Z}, \mathbb{Z}) \rightarrow \mathbb{Z}$ ,  $\div, \text{mod} \in (\mathbb{Z}, \mathbb{Z} \setminus \{0\}) \rightarrow \mathbb{Z}$ , and  $=, < \in (\mathbb{Z}, \mathbb{Z}) \rightarrow \{\text{true}, \text{false}\}$ .
- (6) The result is  $v_1 \otimes v_2$ .

**Case  $E_1 ; E_2$ :**

- (1) Evaluate  $E_1$ .
- (2) Evaluate  $E_2$ .
- (3) Let  $v_2$  be the result of  $E_2$ .
- (4) The result is  $v_2$ .

**Case if  $E_1 E_2 E_3$ :**

- (1) Evaluate  $E_1$ .
- (2) Let  $v_1$  be the result of  $E_1$ .
- (3)  $v_1$  must be a boolean.
- (4) If  $v_1$  is true, then
  - (a) Evaluate  $E_2$ .
  - (b) Let  $v_2$  be the result of  $E_2$ .
  - (c) The result is  $v_2$ .
- (5) Else if  $v_1$  is false, then
  - (a) Evaluate  $E_3$ .
  - (b) Let  $v_3$  be the result of  $E_3$ .
  - (c) The result is  $v_3$ .

**Case let  $x=E_1$  in  $E_2$ :**

- (1) Let  $\sigma$  be the current environment.
- (2) Evaluate  $E_1$ .
- (3) Let  $v_1$  be the result of  $E_1$ .
- (4) Let  $M_1$  be the current store.
- (5) Let  $a$  not be in the domain of  $M_1$ .
- (6) Add a mapping from  $x$  to  $a$  to  $\sigma$ .

- (7) Let  $\sigma'$  be the new environment.
- (8) Add a mapping from  $a$  to  $v_1$  to  $M_1$ .
- (9) Evaluate  $E_2$  under  $\sigma'$ .
- (10) Let  $v_2$  be the result of  $E_2$ .
- (11) The result is  $v_2$ .

**Case**  $D_1, \dots, D_n$  in  $E$ :

- (1) Let  $\sigma$  be the current environment.
- (2) Let  $\sigma_0$  be the empty environment.
- (3) Let  $M$  be the current store.
- (4) For every  $1 \leq i \leq n$ ,
  - (a) Let  $D_i$  creates  $\sigma'_i$ . (Every address in the range of  $\sigma'_i$  is not in the domain of  $M$  and the range of  $\sigma_{i-1}$ . In addition, all the addresses in the range of  $\sigma'_i$  are distinct from each other.)
  - (b) Add every mapping in  $\sigma'_i$  to  $\sigma_{i-1}$ .
  - (c) Let  $\sigma_i$  be the new environment.
- (5) Add every mapping in  $\sigma_n$  to  $\sigma$ .
- (6) Let  $\sigma'$  be the new environment.
- (7) For every  $1 \leq i \leq n$ ,
  - (a) Let  $D_i$  creates  $M'_i$  when  $\sigma'$  is given.
  - (b) Let  $M_i$  be the current store.
  - (c) Add every mapping in  $M'_i$  to  $M_i$ .
- (8) Evaluate  $E$  under  $\sigma_n$ .
- (9) Let  $v$  be the result of  $E$ .
- (10) The result is  $v$ .

**Case**  $\lambda \bar{x}.E$ :

- (1) Let  $\sigma$  be the current environment.
- (2) The result is  $\langle \lambda \bar{x}.E, \sigma \rangle$ .

**Case**  $x := E$ :

- (1) Let  $\sigma$  be the current environment.
- (2)  $x$  must be in the domain of  $\sigma$ .
- (3) Evaluate  $E$ .
- (4) Let  $v$  be the result of  $E$ .
- (5) Let  $M_1$  be the current store.
- (6) Add a mapping from  $\sigma(x)$  to  $v$  to  $M_1$ .
- (7) The result is  $()$ .

**Case**  $E(E_1, \dots, E_n)$ :

- (1) Evaluate  $E$ .
- (2) Let  $v$  be the result of  $E$ .
- (3)  $v$  must be either a closure or a constructor.
- (4) For every  $1 \leq i \leq n$ ,
  - (a) Evaluate  $E_i$ . (If  $i > 1$ , it happens after the evaluation of  $E_{i-1}$ .)
  - (b) Let  $v_i$  be the result of  $E_i$ .

(5) If  $v$  is a closure, then

- (a) Let  $v$  be  $\langle \lambda x_1, \dots, x_m.E_c, \sigma_0 \rangle$ .
  - (b)  $n$  must be the same as  $m$ . (The number of the arguments must be the same as the number of the parameters.)
  - (c) Let  $M_n$  be the current store.
  - (d) For every  $1 \leq i \leq n$ ,
    - (i) Let  $a_i$  be an address that is not in the domain of  $M_n$  and different from every  $a_j$  where  $i \neq j$ .
  - (e) For every  $1 \leq i \leq n$ ,
    - (i) Add a mapping from  $x_i$  to  $a_i$  to  $\sigma_{i-1}$ .
    - (ii) Let  $\sigma_i$  be the new environment.
  - (f) For every  $1 \leq i \leq n$ ,
    - (i) Let  $M'_i$  be the current store.
    - (ii) Add a mapping from  $a_i$  to  $v_i$  to  $M'_i$ .
  - (g) Evaluate  $E_c$  under  $\sigma_n$ .
  - (h) Let  $v_c$  be the result of  $E_c$ .
  - (i) The result is  $v_c$ .
- (6) If  $v$  is a constructor, then
- (a) Let  $v$  be  $\langle x \rangle$ .
  - (b) The result is  $x(v_1, \dots, v_n)$ .

**Case**  $E$  match  $C_1, \dots, C_n$ :

- (1) Let  $\sigma_0$  be the current environment.
- (2) Evaluate  $E$ .
- (3) Let  $v$  be the result of  $E$ .
- (4)  $v$  must be a variant value.
- (5) Let  $v$  be  $x(v_1, \dots, v_m)$ .
- (6) There must be  $C$  among  $C_1, \dots, C_n$  such that the name of  $C$  is  $x$ .
- (7) Let  $C$  be  $x(x_1, \dots, x_k) \rightarrow E_c$ .
- (8)  $m$  must be the same as  $k$ .
- (9) Let  $M_1$  be the current store.
- (10) For every  $1 \leq i \leq m$ ,
  - (a) Let  $a_i$  be an address that is not in the domain of  $M_1$  and different from every  $a_j$  where  $i \neq j$ .
- (11) For every  $1 \leq i \leq m$ ,
  - (a) Add a mapping from  $x_i$  to  $a_i$  to  $\sigma_{i-1}$ .
  - (b) Let  $\sigma_i$  be the new environment.
- (12) For every  $1 \leq i \leq m$ ,
  - (a) Let  $M'_i$  be the current store.
  - (b) Add a mapping from  $a_i$  to  $v_i$  to  $M'_i$ .
- (13) Evaluate  $E_c$  under  $\sigma_m$ .
- (14) Let  $v_c$  be the result of  $E_c$ .
- (15) The result is  $v_c$ .

We need to define environments created by recursive definitions. We use the following terminology:

- If we say “the result is  $\sigma$ ” while explaining  $D$ , then  $D$  creates  $\sigma$ .

The following explains environments created by recursive definitions.

**Case lazy  $x=E$ :**

- (1) Add a mapping from  $x$  to a new address  $a$  to the empty environment.
- (2) Let  $\sigma$  be the new environment.
- (3) The result is  $\sigma$ .

**Case def  $x(\bar{x})=E$ :**

- (1) Add a mapping from  $x$  to a new address  $a$  to the empty environment.
- (2) Let  $\sigma$  be the new environment.
- (3) The result is  $\sigma$ .

**Case type  $W_1, \dots, W_n$ :**

- (1) Let  $\sigma_0$  be the empty environment.

- (a) If  $W$  is a variant definition without a parameter, then
    - (i) Let  $W$  be  $x()$ .
    - (ii) Add a mapping from  $x$  to a new address  $a$  to  $\sigma_{i-1}$ .
    - (iii) Let  $\sigma_i$  be the new environment.
  - (b) Else if  $W$  is a variant definition with one or more parameters, then
    - (i) Let  $W$  be  $x(\_)$ .
    - (ii) Add a mapping from  $x$  to a new address  $a$  to  $\sigma_{i-1}$ .
    - (iii) Let  $\sigma_i$  be the new environment.
- (2) The result is  $\sigma_n$ .

We need to define also stores created by recursive definitions. We use the following terminology:

- If we say “the result is  $M$ ” while explaining  $D$ , then  $D$  creates  $M$ .
- $D$  can create a store only when an environment is given.

The following explains stores created by recursive definitions.

**Case lazy  $x=E$ :**

- (1) Let  $\sigma$  be a given environment.
- (2) Add a mapping from  $\sigma(x)$  to  $\langle\langle E, \sigma \rangle\rangle$  to the empty store.
- (3) Let  $M$  be the new store.
- (4) The result is  $M$ .

**Case def  $x(\bar{x})=E$ :**

- (1) Let  $\sigma$  be a given environment.
- (2) Add a mapping from  $\sigma(x)$  to  $\langle\lambda\bar{x}.E, \sigma\rangle$  to the empty store.
- (3) Let  $M$  be the new store.
- (4) The result is  $M$ .

**Case type  $W_1, \dots, W_n$ :**

- (1) Let  $\sigma$  be a given environment.

- (2) Let  $M_0$  be the empty store.
  - (a) If  $W$  is a variant definition without a parameter, then
    - (i) Let  $W$  be  $x()$ .
    - (ii) Add a mapping from  $\sigma(x)$  to  $x()$  to  $M_{i-1}$ .
    - (iii) Let  $M_i$  be the new environment.
  - (b) Else if  $W$  is a variant definition with one or more parameters, then
    - (i) Let  $W$  be  $x(\_)$ .
    - (ii) Add a mapping from  $\sigma(x)$  to  $\langle x \rangle$  to  $M_{i-1}$ .
    - (iii) Let  $M_i$  be the new environment.
- (3) The result is  $M_n$ .

## 8 STANDARD LIBRARY

We provide the standard library of FABRIC. The library consists of eight parts: utility functions for primitive types, which are integers, boolean values, and the unit value, pair types and utility functions, option types and utility functions, box types and utility functions, list types and utility functions, map types and utility functions, utility functions for strings, and utility functions to build parser combinators. This sections gives a brief overview of the FABRIC standard library.



### 8.1 Primitive Types

```
def intEquals(l: Int, r: Int): Boolean
def intMax(l: Int, r: Int): Int
def intMin(l: Int, r: Int): Int
def booleanEquals(l: Boolean, r: Boolean): Boolean
def unitEquals(l: Unit, r: Unit): Boolean
```

### 8.2 Pairs

```
type Pair['T, 'S] {
  case Pair('T, 'S)
}
def pairFst['T, 'S](p: Pair['T, 'S]): 'T
def pairSnd['T, 'S](p: Pair['T, 'S]): 'S
def pairEquals['T, 'S](f: ('T, 'T) => Boolean, g: ('S, 'S) => Boolean):
  (Pair['T, 'S], Pair['T, 'S]) => Boolean
```

### 8.3 Options

```
type Option['T] {
  case None
  case Some('T)
}
def optionEquals['T](f: ('T, 'T) => Boolean): (Option['T], Option['T]) => Boolean
def optionFilter['T](o: Option['T], f: 'T => Boolean): Option['T]
def optionFilterNot['T](o: Option['T], f: 'T => Boolean): Option['T]
def optionFlatMap['T, 'S](o: Option['T], f: 'T => Option['S]): Option['S]
def optionFlatten['T](o: Option[Option['T]]): Option['T]
def optionForeach['T](o: Option['T], f: 'T => Unit): Unit
def optionGetOrElse['T](o: Option['T], t: 'T): 'T
def optionIsEmpty['T](o: Option['T]): Boolean
def optionMap['T, 'S](o: Option['T], f: 'T => 'S): Option['S]
def optionNonEmpty['T](o: Option['T]): Boolean
def optionToList['T](o: Option['T]): List['T]
```

### 8.4 Boxes

```
type Box['T] {
  case Box_(Option['T] => 'T)
}
def Box['T](t: 'T): Box['T]
def boxGet['T](b: Box['T]): 'T
def boxSet['T](b: Box['T], t: 'T): 'T
```

### 8.5 Lists

```
type List['T] {
  case Nil
  case Cons('T, List['T])
}
def List0['T](): List['T]
```

```

def List1['T](t1: 'T): List['T]
def List2['T](t1: 'T, t2: 'T): List['T]
def List3['T](t1: 'T, t2: 'T, t3: 'T): List['T]
def List4['T](t1: 'T, t2: 'T, t3: 'T, t4: 'T): List['T]
def List5['T](t1: 'T, t2: 'T, t3: 'T, t4: 'T, t5: 'T): List['T]
def listAppended['T](l: List['T], t: 'T): List['T]
def listConcat['T](l: List['T], r: List['T']): List['T]
def listCount['T](l: List['T], f: 'T => Boolean): Int
def listDrop['T](l: List['T], i: Int): List['T]
def listEquals['T](f: ('T, 'T) => Boolean): (List['T], List['T']) => Boolean
def listExists['T](l: List['T], f: 'T => Boolean): Boolean
def listFilter['T](l: List['T], f: 'T => Boolean): List['T]
def listFilterNot['T](l: List['T], f: 'T => Boolean): List['T]
def listFind['T](l: List['T], f: 'T => Boolean): Option['T]
def listFlatMap['T, 'S](l: List['T], f: 'T => List['S']): List['S]
def listFlatten['T](l: List[List['T]]): List['T]
def listFoldLeft['T, 'S](a: 'S, l: List['T], f: ('S, 'T) => 'S): 'S
def listFoldRight['T, 'S](l: List['T], a: 'S, f: ('T, 'S) => 'S): 'S
def listForall['T](l: List['T], f: 'T => Boolean): Boolean
def listForeach['T](l: List['T], f: 'T => Unit): Unit
def listGet['T](l: List['T], i: Int): Option['T]
def listIsEmpty['T](l: List['T]): Boolean
def listLength['T](l: List['T']): Int
def listMap['T, 'S](l: List['T], f: 'T => 'S): List['S]
def listNonEmpty['T](l: List['T']): Boolean
def listPrepended['T](l: List['T], t: 'T): List['T]
def listReverse['T](l: List['T']): List['T]
def listTake['T](l: List['T], i: Int): List['T]
def listUnzip['T, 'S](l: List[Pair['T, 'S]]): Pair[List['T], List['S]]
def listZip['T, 'S](l: List['T], r: List['S']): List[Pair['T, 'S]]
def listZipWithIndex['T](l: List['T']): List[Pair['T, Int]]

```

## 8.6 Maps

```

type Map['T, 'S] {
  case Map(List[Pair['T, 'S]], ('T, 'T) => Boolean)
}
def Map0['T, 'S](f: ('T, 'T) => Boolean): Map['T, 'S]
def Map1['T, 'S](f: ('T, 'T) => Boolean, t1: 'T, s1: 'S): Map['T, 'S]
def Map2['T, 'S](f: ('T, 'T) => Boolean, t1: 'T, s1: 'S, t2: 'T, s2: 'S): Map['T, 'S]
def Map3['T, 'S](
  f: ('T, 'T) => Boolean, t1: 'T, s1: 'S, t2: 'T, s2: 'S, t3: 'T, s3: 'S
): Map['T, 'S]
def Map4['T, 'S](
  f: ('T, 'T) => Boolean, t1: 'T, s1: 'S, t2: 'T, s2: 'S, t3: 'T, s3: 'S, t4: 'T, s4: 'S
): Map['T, 'S]
def Map5['T, 'S](
  f: ('T, 'T) => Boolean,

```

```

t1: 'T, s1: 'S, t2: 'T, s2: 'S, t3: 'T, s3: 'S, t4: 'T, s4: 'S, t5: 'T, s5: 'S
): Map['T, 'S]
def mapFold['T, 'S, 'R](r: 'R, m: Map['T, 'S], f: ('R, 'T, 'S) => 'R): 'R
def mapGet['T, 'S](m: Map['T, 'S], t: 'T): Option['S]
def mapRemoved['T, 'S](m: Map['T, 'S], t: 'T): Map['T, 'S]
def mapToList['T, 'S](m: Map['T, 'S]): List[Pair['T, 'S]]
def mapUpdated['T, 'S](m: Map['T, 'S], t: 'T, s: 'S): Map['T, 'S]

```

## 8.7 Strings

```

val STRP = Cons[Int]
val EOS = Nil[Int]
val stringEquals = listEquals[Int](intEquals)
def substring(s: List[Int], i: Int, j: Int): List[Int]

```

## 8.8 Parser Combinators

```

type Parser['T] {
  case Parser(List[Int] => ParseResult['T])
}
type ParseResult['T] {
  case Success('T, List[Int])
  case Failure
}
def parse['T](parser: Parser['T], str: List[Int]): ParseResult['T]
def parseAll['T](parser: Parser['T], str: List[Int]): Option['T]
def parserConst(c: Int): Parser[Int]
def parserCond(p: Int => Boolean): Parser[Int]
def parserThen['T, 'S](l: () => Parser['T], r: () => Parser['S]): Parser[Pair['T, 'S]]
def parserOr['T](l: () => Parser['T], r: () => Parser['T]): Parser['T]
def parserMap['T, 'S](l: () => Parser['T], f: 'T => 'S): Parser['S]

```

# 9 GUIDE TO THE PROJECT

## 9.1 Directory and File Structures

To start the project, you have to place the `proj03` directory under the `src/main/scala/cs320` directory of an SBT project for Exercise #1. The `proj03` directory contains four files: `Project03.scala`, `package.scala`, `StdLib.scala`, and `proj03.pdf`.

The `Project03.scala` file defines both typed and untyped versions of the abstract syntax of FABRIC. In addition, it implements the parser, the desugarer, and the type erasure of FABRIC. In the file, you can see the following code:

```

def run(str: String): String = {
  val e = Expr(str)
  typeCheck(e)
  interp(erase(e)).toString
}
def runWithStdLib(str: String): String = run(StdLib.code + str)

object Typed { ... }
object Untyped { ... }

```

The `run` function parses a given string to get a typed expression. It type-checks the expression with the `typeCheck` function. If the type checking succeeds, it erases types from the expression with the `erase` function to obtain an untyped expression. It evaluates the untyped expression with the `interp` function and returns the result after converting it to a string.

The `runWithStdLib` function does a similar thing with `run`, but it prepends the source code of the standard library in front of a given string before running it.

Definitions related to the typed version can be found in the `Typed` object. Definitions related to the untyped version can be found in the `Untyped` object.

The package `scala` file contains the empty `typeCheck` and `interp` functions. It also provides some test cases. You must complete the `typeCheck` and `interp` functions. In the file, you can see the following code:

```
def typeCheck(e: Typed.Expr): Typed.Type = T.typeCheck(e)
def interp(e: Untyped.Expr): Untyped.Value = U.interp(e)

object T {
  import Typed._
  def typeCheck(expr: Expr): Type = ???
}
object U {
  import Untyped._
  def interp(expr: Expr): Value = ???
}
```

We recommend you to implement everything related to type checking inside the object `T` and to implement everything related to interpretation inside the object `U`.

Since passing all the given test cases does not guarantee that your type checker and interpreter are perfect, we highly recommend you to write your own test cases. As we provide [reference type checker and interpreter](#) of FABRIC, you can find the correct results of your new test cases.

The `StdLib.scala` contains the source code of the FABRIC standard library.

## 9.2 Rules for Implementation

- If the type of  $e$  is  $\tau$ , then `typeCheck( $e$ )` must be equal to  $\tau$ .
- If  $e$  is not well-typed, then `typeCheck( $e$ )` must terminate by calling the error function. Error messages can be any strings.
- If  $E$  results in  $v$ , then `interp( $E$ )` must be equal to  $v$ .
- If execution of  $E$  terminates with a run-time error, then `interp( $E$ )` must terminate by calling the error function. Error messages can be any strings.
- If execution of  $E$  does not terminate, then `interp( $E$ )` must not terminate or terminate due to stack overflow.
- You can define helper functions.
- Do not import anything.
- Do not use while loops, for loops, and break statements.
- Do not use mutable variables and mutable collections.
- Do not modify `Project03.scala` and `StdLib.scala`.

## A STATIC SEMANTICS

Figure 7 defines sets  $\mathbb{M}$ ,  $\mathbb{T}$ , and  $\mathbb{D}$ , and metavariables  $\delta$  and  $\Delta$ . It formally defines type substitutions and type environments.

Mutability	$\mu$	$\in$	$\mathbb{M}$
Type Definition	$\Delta$	$\in$	$\mathbb{D}$
	$\Delta$	$::=$	$\text{type } t[\overline{\alpha}] \ \overline{w}$
Type Substitution	$\delta$	$::=$	$[\overline{\alpha} \leftarrow \overline{\tau}]$
Type Environment	$\Gamma$	$\in$	$(Id \hookrightarrow (\mathbb{T} \times \mathbb{M})) + (TN \hookrightarrow \mathbb{D}) + (TV \hookrightarrow \{\cdot\})$

Fig. 7. Definitions

Figure 8 defines some notations used in the appendices. In the figure,  $S$  denotes an arbitrary set.  $i, j$ , and  $k$  denote arbitrary mathematical objects.

$S_1 + S_2$	$\triangleq$	$\{i_1 \cup i_2 : i_1 \in S_1 \wedge i_2 \in S_2\}$	$\Gamma[\alpha]$	$\triangleq$	$\Gamma \cup \{(\alpha, \cdot)\}$
$S - i$	$\triangleq$	$\{(j, k) : (j, k) \in S \wedge j \neq i\}$	$\Gamma[x:\tau]$	$\triangleq$	$\Gamma[x:(\forall.\tau, \text{val})]$
$S[i:j]$	$\triangleq$	$(S - i) \cup \{(i, j)\}$	$\Gamma[x:(\tau, \mu)]$	$\triangleq$	$\Gamma[x:(\forall.\tau, \mu)]$
$S[i=j]$	$\triangleq$	$(S - i) \cup \{(i, j)\}$	$\Gamma[x:T]$	$\triangleq$	$\Gamma[x:(T, \mu)]$ for some $\mu$
$S[i \mapsto j]$	$\triangleq$	$(S - i) \cup \{(i, j)\}$	$\Gamma(x) = T$	$\triangleq$	$\Gamma(x) = (T, \mu)$ for some $\mu$
$\#(i_1, \dots, i_n)$	$\triangleq$	$n$	$\Gamma(x) = (\tau, \mu)$	$\triangleq$	$\Gamma(x) = (\forall.\tau, \mu)$

Fig. 8. Notations

From now on, we use underlines in addition to overlines. An underline below submaterial within overlined material prevents an overline from attaching subscripts to the metavariables in the submaterial; the effect is to require the underlined material to be the same in each copy. For example,  $\tau \underline{\delta}$  represents  $\tau_1 \delta, \dots, \tau_n \delta$  for any  $n \geq 0$ .

Figure 9 formally defines the results of substituting types.

$\tau \delta = \tau$	$t[\overline{\tau}] \delta = t[\overline{\tau \underline{\delta}}]$	$\text{Int} \delta = \text{Int}$
$\alpha[\dots \alpha \leftarrow \tau \dots] = \tau$	$\text{Boolean} \delta = \text{Boolean}$	$\text{Unit} \delta = \text{Unit}$
$\alpha[\overline{\alpha \leftarrow \tau}] = \alpha$ (if $\wedge \overline{\alpha \neq \underline{\alpha}}$ )	$(\overline{\tau \rightarrow \tau}) \delta = \overline{\tau \underline{\delta} \rightarrow \tau \delta}$	

Fig. 9. Type Substitutions

We use overlines and underlines for inference rules as well. If a rule contains an overline, it is, in fact, not a rule; it is a rule schema, which generates a set of infinitely many rules. For instance, the rule schema

$$\frac{\overline{\Gamma \vdash \tau} \quad \Gamma \vdash \tau}{\Gamma \vdash \overline{\tau \rightarrow \tau}}$$

generates the following rules:

$$\frac{\Gamma \vdash \tau}{\Gamma \vdash \tau \rightarrow \tau} \quad \frac{\Gamma \vdash \tau_1 \quad \Gamma \vdash \tau}{\Gamma \vdash \tau_1 \rightarrow \tau} \quad \dots \quad \frac{\Gamma \vdash \tau_1 \quad \dots \quad \Gamma \vdash \tau_{10} \quad \Gamma \vdash \tau}{\Gamma \vdash \tau_1, \dots, \tau_{10} \rightarrow \tau} \quad \dots$$

Figure 10 formally defines the well-formedness of types; Figure 11 formally defines the types of expressions and cases.

$$\begin{array}{c}
\boxed{\Gamma \vdash \tau} \qquad \frac{\overline{\Gamma \vdash \tau} \quad t \in \text{Domain}(\Gamma) \quad \Gamma(t) = \text{type } t[\overline{\alpha}] \quad \overline{w} \quad \#(\overline{\alpha}) = \#(\overline{\tau})}{\Gamma \vdash t[\overline{\tau}]}
\\[10pt]
\frac{\alpha \in \text{Domain}(\Gamma)}{\Gamma \vdash \alpha} \qquad \Gamma \vdash \text{Int} \qquad \Gamma \vdash \text{Boolean} \qquad \Gamma \vdash \text{Unit} \qquad \frac{\overline{\Gamma \vdash \tau} \quad \Gamma \vdash \tau}{\Gamma \vdash \overline{\tau} \rightarrow \tau}
\end{array}$$

Fig. 10. Well-formedness of Types

$$\begin{array}{c}
\boxed{\Gamma \vdash e : \tau} \qquad \frac{\overline{\Gamma \vdash \tau} \quad x \in \text{Domain}(\Gamma) \quad \Gamma(x) = \forall \overline{\alpha}. \tau}{\Gamma \vdash x[\overline{\tau}] : \tau[\overline{\alpha} \leftarrow \overline{\tau}]}
\\[10pt]
\Gamma \vdash n : \text{Int} \qquad \Gamma \vdash b : \text{Boolean} \qquad \Gamma \vdash () : \text{Unit}
\\[10pt]
\frac{\Gamma \vdash e_1 : \text{Int} \quad \Gamma \vdash e_2 : \text{Int}}{\Gamma \vdash e_1 \oplus e_2 : \text{Int}} \qquad \frac{\Gamma \vdash e_1 : \text{Int} \quad \Gamma \vdash e_2 : \text{Int}}{\Gamma \vdash e_1 \odot e_2 : \text{Boolean}} \qquad \frac{\Gamma \vdash e_1 : \tau_1 \quad \Gamma \vdash e_2 : \tau_2}{\Gamma \vdash e_1 ; e_2 : \tau_2}
\\[10pt]
(\oplus \in \{+, \times, \div, \text{mod}\}) \qquad (\odot \in \{=, <\})
\\[10pt]
\frac{\Gamma \vdash e_1 : \text{Boolean} \quad \Gamma \vdash e_2 : \tau \quad \Gamma \vdash e_3 : \tau}{\Gamma \vdash \text{if } e_1 \text{ } e_2 \text{ } e_3 : \tau} \qquad \frac{\boxed{\Gamma \vdash \tau_1} \quad \Gamma \vdash e_1 : \tau_1 \quad \Gamma[x : (\tau_1, \mu)] \vdash e_2 : \tau_2}{\Gamma \vdash \mu \ x : \boxed{\tau_1} = e_1 \text{ in } e_2 : \tau_2}
\\[10pt]
\frac{\overline{\Gamma \vdash d \text{ name ok}} \quad \Gamma' = \Gamma; \overline{tbind(d)} \quad \overline{\Gamma' \vdash d} \quad \Gamma' \vdash e : \tau \quad \Gamma \vdash \tau}{\Gamma \vdash \overline{d} \text{ in } e : \tau} \qquad \frac{\overline{\Gamma \vdash \tau} \quad \Gamma[x : \overline{\tau}] \vdash e : \tau}{\Gamma \vdash \lambda \overline{x} : \overline{\tau}. e : \overline{\tau} \rightarrow \tau}
\\[10pt]
\frac{x \in \text{Domain}(\Gamma) \quad \Gamma(x) = (\tau, \text{var}) \quad \Gamma \vdash e : \tau}{\Gamma \vdash x := e : \text{Unit}} \qquad \frac{\Gamma \vdash e : \overline{\tau} \rightarrow \tau \quad \overline{\Gamma \vdash e : \tau}}{\Gamma \vdash e(\overline{e}) : \tau}
\\[10pt]
\frac{\Gamma \vdash e : t[\overline{\tau}] \quad t \in \text{Domain}(\Gamma) \quad \Gamma(t) = \text{type } t[\overline{\alpha}] \quad \overline{w} \quad \#(\overline{w}) = \#(\overline{c}) \quad \delta = [\overline{\alpha} \leftarrow \overline{\tau}] \quad \overline{\Gamma, \overline{w}, \delta \vdash c : \tau}}{\Gamma \vdash e \text{ match } \overline{c} : \tau}
\\[10pt]
\boxed{\Gamma, \overline{w}, \delta \vdash c : \tau} \qquad \frac{w_i = x(\overline{\tau}) \quad \Gamma[x : \overline{\tau} \delta] \vdash e : \tau}{\Gamma, \overline{w}, \delta \vdash x(\overline{x}) \rightarrow e : \tau}
\end{array}$$

Fig. 11. Types of Expressions

Note that metavariables under the same overline have to be consistent in their numbers. For example, the rule schema

$$\frac{\overline{\Gamma \vdash \tau} \quad x \in \text{Domain}(\Gamma) \quad \Gamma(x) = \forall \overline{\alpha}. \tau}{\Gamma \vdash x[\overline{\tau}] : \tau[\overline{\alpha} \leftarrow \overline{\tau}]}$$

generates

$$\frac{\Gamma \vdash \tau_1 \quad x \in \text{Domain}(\Gamma) \quad \Gamma(x) = \forall \alpha_1. \tau_1}{\Gamma \vdash x[\tau_1] : \tau[\alpha_1 \leftarrow \tau_1]}$$

but not

$$\frac{\Gamma \vdash \tau_1 \quad x \in \text{Domain}(\Gamma) \quad \Gamma(x) = \forall \alpha_1, \alpha_2. \tau_1}{\Gamma \vdash x[\tau_1] : \tau[\alpha_1 \leftarrow \tau_1, \alpha_2 \leftarrow \tau_2]}$$

Also, if a rule schema contains boxes, then the boxes have to be consistent, i.e. all of them exist simultaneously, or none of them exist. Therefore, the rule schema

$$\frac{\boxed{\Gamma \vdash \tau_1} \quad \Gamma \vdash e_1 : \tau_1 \quad \Gamma[x : (\tau_1, \mu)] \vdash e_2 : \tau_2}{\Gamma \vdash \mu x : \boxed{\tau_1} = e_1 \text{ in } e_2 : \tau_2}$$

never generates

$$\frac{\Gamma \vdash e_1 : \tau_1 \quad \Gamma[x : (\tau_1, \mu)] \vdash e_2 : \tau_2}{\Gamma \vdash \mu x : \tau_1 = e_1 \text{ in } e_2 : \tau_2}$$

Figure 12, 13, and 14 define additional judgements for the static semantics.

$$\boxed{\Gamma \vdash d \text{ name ok}} \quad \Gamma \vdash \text{lazy } x : \tau = e \text{ name ok} \quad \Gamma \vdash \text{def } x[\bar{\alpha}](\bar{x} : \bar{\tau}) : \tau = e \text{ name ok} \quad \frac{t \notin \text{Domain}(\Gamma)}{\Gamma \vdash \text{type } t[\bar{\alpha}] \bar{w} \text{ name ok}}$$

Fig. 12. Well-formedness of Type Names

$$\boxed{\Gamma \vdash d \quad \Gamma \vdash w} \quad \frac{\Gamma \vdash \tau \quad \Gamma \vdash e : \tau}{\Gamma \vdash \text{lazy } x : \tau = e} \quad \frac{\overline{\alpha \notin \text{Domain}(\Gamma)} \quad \Gamma' = \Gamma[\bar{\alpha}] \quad \overline{\Gamma' \vdash \tau} \quad \Gamma' \vdash \tau \quad \Gamma'[\bar{x} : \bar{\tau}] \vdash e : \tau}{\Gamma \vdash \text{def } x[\bar{\alpha}](\bar{x} : \bar{\tau}) : \tau = e}$$

$$\frac{\overline{\Gamma \vdash \tau}}{\Gamma \vdash x(\bar{\tau})} \quad \frac{\overline{\alpha \notin \text{Domain}(\Gamma)} \quad \Gamma' = \Gamma[\bar{\alpha}] \quad \overline{\Gamma' \vdash w}}{\Gamma \vdash \text{type } t[\bar{\alpha}] \bar{w}}$$

Fig. 13. Well-formedness of Recursive Definitions

$$\boxed{tbind(d) = \Gamma \quad tbind(w, t[\bar{\alpha}]) = \Gamma}$$

$$tbind(\text{lazy } x : \tau = e) = [x : \tau]$$

$$tbind(\text{def } x[\bar{\alpha}](\bar{x} : \bar{\tau}) : \tau = e) = [x : \forall \bar{\alpha}. \bar{\tau} \rightarrow \tau]$$

$$tbind(\text{type } t[\bar{\alpha}] \bar{w}) = [t = \text{type } t[\bar{\alpha}] \bar{w}; tbind(w, t[\bar{\alpha}])]$$

$$tbind(x(\cdot), t[\bar{\alpha}]) = [x : \forall \bar{\alpha}. t[\bar{\alpha}]]$$

$$tbind(x(\bar{\tau}), t[\bar{\alpha}]) = [x : \forall \bar{\alpha}. \bar{\tau} \rightarrow t[\bar{\alpha}]] \quad (\text{if } \#(\bar{\tau}) > 0)$$

Fig. 14. Binding of Recursive Definitions

## B BIG-STEP OPERATIONAL SEMANTICS

Figure 15 and 16 define the operational semantics of FABRIC.

$$\begin{array}{c}
\boxed{\sigma, M \vdash e \Rightarrow v, M} \quad \frac{x \in \text{Domain}(\sigma) \quad \sigma(x) \in \text{Domain}(M) \quad M(\sigma(x)) \notin \text{Expr}V}{\sigma, M \vdash x \Rightarrow M(\sigma(x)), M} \\
\\
\frac{x \in \text{Domain}(\sigma) \quad \sigma(x) \in \text{Domain}(M) \quad M(\sigma(x)) = \langle \langle E, \sigma' \rangle \rangle \quad \sigma', M \vdash E \Rightarrow v, M_1}{\sigma, M \vdash x \Rightarrow v, M_1[\sigma(x) \mapsto v]} \\
\\
\sigma, M \vdash n \Rightarrow n, M \quad \sigma, M \vdash b \Rightarrow b, M \quad \sigma, M \vdash () \Rightarrow (), M \\
\\
\frac{\sigma, M \vdash E_1 \Rightarrow n_1, M_1 \quad \sigma, M_1 \vdash E_2 \Rightarrow n_2, M_2}{\sigma, M \vdash E_1 \odot E_2 \Rightarrow n_1 \odot n_2, M_2} \quad (\odot \in \{+, \times, =, <\}) \\
\\
\frac{\sigma, M \vdash E_1 \Rightarrow n_1, M_1 \quad \sigma, M_1 \vdash E_2 \Rightarrow n_2, M_2 \quad n_2 \neq 0}{\sigma, M \vdash E_1 \otimes E_2 \Rightarrow n_1 \otimes n_2, M_2} \quad (\otimes \in \{\div, \text{mod}\}) \\
\\
\frac{\sigma, M \vdash E_1 \Rightarrow v_1, M_1 \quad \sigma, M_1 \vdash E_2 \Rightarrow v_2, M_2}{\sigma, M \vdash E_1 ; E_2 \Rightarrow v_2, M_2} \\
\\
\frac{\sigma, M \vdash E_1 \Rightarrow \text{true}, M_1 \quad \sigma, M_1 \vdash E_2 \Rightarrow v_2, M_2}{\sigma, M \vdash \text{if } E_1 E_2 E_3 \Rightarrow v_2, M_2} \quad \frac{\sigma, M \vdash E_1 \Rightarrow \text{false}, M_1 \quad \sigma, M_1 \vdash E_3 \Rightarrow v_2, M_2}{\sigma, M \vdash \text{if } E_1 E_2 E_3 \Rightarrow v_2, M_2} \\
\\
\frac{\sigma, M \vdash E_1 \Rightarrow v_1, M_1 \quad a \notin \text{Domain}(M_1) \quad \sigma[x \mapsto a], M_1[a \mapsto v_1] \vdash E_2 \Rightarrow v_2, M_2}{\sigma, M \vdash \text{let } x = E_1 \text{ in } E_2 \Rightarrow v_2, M_{M_2}} \\
\\
\frac{\overline{D \text{ creates } \sigma} \quad \sigma'' = \overline{\sigma} \quad \text{Domain}(\sigma'') = \{\overline{x}\} \quad M \cap \{\overline{\sigma''(x)}\} = \emptyset \quad \text{distinct}(\overline{\sigma''(x)})}{\sigma' = \sigma; \sigma'' \quad \overline{D, \sigma'} \text{ create } M' \quad M_1 = M; \overline{M'} \quad \sigma', M_1 \vdash E \Rightarrow v, M_2}}{\sigma, M \vdash \overline{D} \text{ in } E \Rightarrow v, M_2} \\
\\
\sigma, M \vdash \lambda \overline{x}. E \Rightarrow \langle \lambda \overline{x}. E, \sigma \rangle, M \quad \frac{x \in \text{Domain}(\sigma) \quad \sigma, M \vdash E \Rightarrow v, M_1}{\sigma, M \vdash x := E \Rightarrow (), M_1[\sigma(x) \mapsto v]} \\
\\
\frac{\sigma, M \vdash E \Rightarrow \langle \lambda \overline{x}. E', \sigma' \rangle, M_0 \quad \overline{\sigma, M_{i-1} \vdash E_i \Rightarrow v_i, M_i}^{1 \leq i \leq n} \quad \{\overline{a}\} \cap \text{Domain}(M_n) = \emptyset \quad \text{distinct}(\overline{a}) \quad \sigma'[\overline{x} \mapsto \overline{a}], M_n[\overline{a} \mapsto \overline{v}] \vdash E' \Rightarrow v, M'}{\sigma, M \vdash E(\overline{E}) \Rightarrow v, M'}
\end{array}$$

Fig. 15. Evaluation of Expressions (1/2)

The overline notation may explicitly bind index variables by using a range expression. In such cases, the index variables may appear explicitly in the covered material, and subscripts are not automatically attached to metavariables. For example,  $\overline{x}^{1 \leq i \leq n}$  represents  $x_1, x_2, \dots, x_n$  for any  $n \geq 0$ . In addition, the statement  $\text{distinct}(k_1, \dots, k_n)$  is true if and only if  $k_i \neq k_j$  for every  $1 \leq i < j \leq n$ .



$$\begin{array}{c}
\sigma, M \vdash E \Rightarrow \langle x \rangle, M_0 \quad \overline{\sigma, M_{i-1} \vdash E_i \Rightarrow v_i, M_i}^{1 \leq i \leq n} \\
\hline
\sigma, M \vdash E(\bar{E}) \Rightarrow x(\bar{v}), M_n \\
\\
\sigma, M \vdash E \Rightarrow x(\bar{v}), M_1 \quad C_i = x(\bar{x}) \rightarrow E' \\
\{\bar{a}\} \cap \text{Domain}(M_1) = \emptyset \quad \text{distinct}(\bar{a}) \quad \sigma[x \mapsto \bar{a}], M_1[\bar{a} \mapsto \bar{v}] \vdash E' \Rightarrow v, M_2 \\
\hline
\sigma, M \vdash E \text{ match } \bar{C} \Rightarrow v, M_2
\end{array}$$

Fig. 16. Evaluation of Expressions (2/2)

Figure 17 formally defines environments and stores created by recursive definitions.

$D \text{ creates } \sigma \quad W \text{ creates } \sigma \quad D, \sigma \text{ create } M \quad W, \sigma \text{ create } M$	
$\text{lazy } x=E \text{ creates } [x \mapsto a]$	$\text{lazy } x=E, \sigma \text{ create } [\sigma(x) \mapsto \langle \langle E, \sigma \rangle \rangle]$
$\text{def } x(\bar{x})=E \text{ creates } [x \mapsto a]$	$\text{def } x(\bar{x})=E, \sigma \text{ create } [\sigma(x) \mapsto \langle \lambda \bar{x}. E, \sigma \rangle]$
$\frac{\overline{W \text{ creates } \sigma}}{\text{type } \bar{W} \text{ creates } \bar{\sigma}}$	$\frac{\overline{W, \underline{\sigma} \text{ create } M}}{\text{type } \bar{W}, \sigma \text{ create } \bar{M}}$
$x() \text{ creates } [x \mapsto a]$	$x(), \sigma \text{ create } [\sigma(x) \mapsto x()]$
$x(\_) \text{ creates } [x \mapsto a]$	$x(\_), \sigma \text{ create } [\sigma(x) \mapsto \langle x \rangle]$

Fig. 17. Binding of Recursive Definitions