## Notes on Formal Compiler Construction with the

Christiano Braga

Instituto de Computação, Universidade Federal Fluminense, Niterói, Brazil

July 24, 2020

http://github.com/ChristianoBraga/PiFramework



## 1. Introduction Example

2. II IR expressions
Grammar
Automaton

3. II commands
Grammar
Automaton

4 ∏ IR declarations

Grammar Automaton

5. Π IR abstractions Grammar Automaton

6. II IR recursive abstractions
Grammar
Automaton

7.  $\Pi^2$ :  $\Pi$  Framework in Python

Expressions Commands Declarations Abstractions

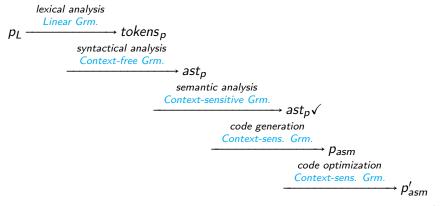
8. IMP language Grammar Examples

## Compiler pipeline

$$\begin{array}{c} p_L \xrightarrow{lexical \ analysis} \ tokens_p \\ & \xrightarrow{syntactical \ analysis} \ ast_p \\ & \xrightarrow{semantic \ analysis} \ ast_p \checkmark \\ & \xrightarrow{code \ generation} \ p_{asm} \\ & \xrightarrow{code \ optimization} \ p'_{asm} \end{array}$$

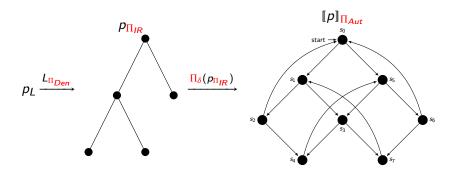


## Compiler pipeline and formal languages





## Compiler pipeline with the $\Pi$ Framework I





## Compiler pipeline with the $\Pi$ Framework $\Pi$

#### Automata



- II IR defines a set of constructions common to many programming languages.
- $\Pi$  IR constructions have a formal automata-based semantics in  $\Pi$  automata.
- One may execute (or validate) a program in a given language by running its associated IT IR program.



## Compiler pipeline with the ∏ Framework III

- Π Framework:
  - http://github.com/ChristianoBraga/PiFramework
- Notes on Formal Compiler Construction with the II Framework: https://github.com/ChristianoBraga/PiFramework/blob/master/notes/notes.pdf.



#### A calculator

We wish to compute simple arithmetic expressions such as 5\*(3+2).



#### A calculator: Lexer

```
\langle digit \rangle ::= [0..9]

\langle digits \rangle ::= \langle digit \rangle^+

\langle boolean \rangle ::= 'true' | 'false'
```



## A calculator: concrete syntax

```
::= \langle aexp \rangle \mid \langle bexp \rangle
\langle exp \rangle
                             ::= \langle aexp \rangle '+' \langle term \rangle | \langle aexp \rangle '-' \langle term \rangle | \langle term \rangle
\langle aexp \rangle
                             ::= \langle term \rangle '*' \langle factor \rangle \langle term \rangle '/' \langle factor \rangle \langle factor \rangle
⟨term⟩
⟨factor⟩
                            ::= '(' \( aexp\) ')' | \( digits\)
                            ::= \langle boolean \rangle \mid ``` \langle bexp \rangle \mid \langle bexp \rangle \langle boolop \rangle \langle bexp \rangle
\langle bexp \rangle
                                     ⟨aexp⟩ ⟨iop⟩ ⟨aexp⟩
⟨boolop⟩
                            ::= '=' | '/\' | '\/'
                            ::= '<' | '>' | '<=' | '>='
\langle iop \rangle
```



## A calculator: abstract syntax

$$\langle exp \rangle \qquad ::= \langle digits \rangle \mid \langle boolean \rangle \mid \langle exp \rangle \langle bop \rangle \langle exp \rangle$$



#### A calculator: II denotations I

Let D in  $\langle digits \rangle$ , B in  $\langle boolean \rangle$  and  $E_1, E_2$  in  $\langle exp \rangle$ ,

$$[E_1 - E_2]_{\Pi} = Sub([E_1]_{\Pi}, [E_2]_{\Pi})$$
(4)

$$\llbracket E_1 * E_2 \rrbracket_{\Pi} = Mul(\llbracket E_1 \rrbracket_{\Pi}, \llbracket E_2 \rrbracket_{\Pi}) \tag{5}$$

$$[E_1/E_2]_{\Pi} = Div([E_1]_{\Pi}, [E_2]_{\Pi})$$
(6)

$$||E_1| < E_2||_{\Pi} = Lt(||E_1||_{\Pi}, ||E_2||_{\Pi})$$
(7)

$$[E_1 <= E_2]_{\Pi} = Le([E_1]_{\Pi}, [E_2]_{\Pi})$$
 (8)

$$[E_1 > E_2]_{\Pi} = Gt([E_1]_{\Pi}, [E_2]_{\Pi})$$
 (9)

$$||E_1\rangle = ||E_2||_{\Pi} = G(||E_1||_{\Pi}, ||E_2||_{\Pi})$$

$$||E_1\rangle = ||E_2||_{\Pi} = Ge(||E_1||_{\Pi}, ||E_2||_{\Pi})$$
(10)

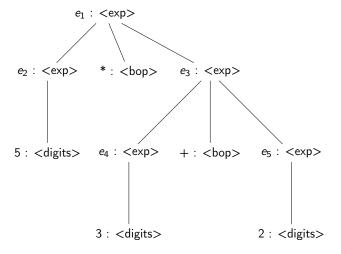


#### A calculator: II denotations II

- $\Pi$  denotations are functions  $\llbracket \cdot \rrbracket_{\Pi} : AST \to \Pi$  IR, where AST denotes the datatype for the abstract syntax tree and  $\Pi$  IR denotes the datatype for  $\Pi$  IR programs.
- Note that  $\llbracket \cdot \rrbracket_{\Pi}$  has *trees* as parameters, instances of *AST*. The example expression 5\*(3+2) becomes



#### A calculator: II denotations III





#### A calculator: ☐ denotations IV



## A calculator: executing $\Pi$ IR with $\Pi$ automata

A  $\Pi$  automaton is a 5-tuple  $\mathcal{A} = (G, Q, \delta, q_0, F)$ , where G is a context-free grammar, Q is the set of states,  $q_0$  is the initial state,  $F \subseteq Q$  is the set of final states and

$$\delta: L(G)^* \times L(G)^* \times Store \rightarrow Q$$
,

where L(G) is the language generated by G and Store represents the memory. (Elements in a set  $S^*$  are represented by terms  $[s_1, s_2, ..., s_n]$ .)

```
\delta([Nul(Num(5), Sum(Num(3), Num(2)], \phi, \phi) = \delta([Num(5), Sum(Num(3), Num(2)), \#MUL], \phi, \phi)
 \delta([Num(5), Sum(Num(3), Num(2)), \#MUL], \emptyset, \emptyset) = \delta([Sum(Num(3), Num(2)), \#MUL], [Num(5)], \emptyset)
  \delta([Sum(Num(3), Num(2)), \#MUL], [Num(5)], \phi) = \delta([Num(3), Num(2), \#SUM, \#MUL], [Num(5)], \phi)
\delta([Num(3), Num(2), \#SUM, \#MUL], [Num(5)], \emptyset) = \delta([Num(2), \#SUM, \#MUL], [Num(3), Num(5)], \emptyset)
\delta([Num(2), \#SUM, \#MUL], [Num(3), Num(5)], \emptyset) = \delta([\#SUM, \#MUL], [Num(2), Num(3), Num(5)], \emptyset)
\delta([\#SUM, \#MUL], [Num(2), Num(3), Num(5)], \emptyset) = \delta([\#MUL], [Num(5), Num(5)], \emptyset)
                     \delta([\#MUL], [Num(5), Num(5)], \emptyset) = \delta(\emptyset, [Num(25)], \emptyset)
                                       \delta(\emptyset, [Num(25)], \emptyset) = Num(25)
```



## Excerpt of $\Pi$ IR expressions

```
\langle Statement \rangle ::= \langle Exp \rangle
\langle Exp \rangle
                        ::= <ArithExp> | <BoolExp> <Exp>
\langle ArithExp \rangle ::= 'Num'(\langle digits \rangle) | 'Sum'(\langle Exp \rangle, \langle Exp \rangle)
                                Sub'(\langle Exp \rangle, \langle Exp \rangle) \mid Mul'(\langle Exp \rangle, \langle Exp \rangle)
\langle BoolExp \rangle ::= 'Eq'(\langle Exp \rangle, \langle Exp \rangle) | 'Not'(\langle Exp \rangle)
```



### $\Pi$ automaton semantics for $\Pi$ IR expressions I

• Recall that  $\delta: L(G)^* \times L(G)^* \times Store \rightarrow Q$ , and let  $N, N_i \in \mathbb{N}$ ,  $C, V \in L(G)^*$ ,  $S \in Store$ ,

$$\delta(Num(N) :: C, V, S) = \delta(C, Num(N) :: V, S)$$
(11)

$$\delta(Sum(E_1, E_2) :: C, V, S) = \delta(E_1 :: E_2 :: \#SUM :: C, V, S)$$
 (12)

$$\delta(\#SUM :: C, Num(N_1) :: Num(N_2) :: V, S) = \delta(C, (N_1 + N_2) :: V, S)$$
(13)

• •

$$\delta(Not(E) :: C, V, S) = \delta(E :: \#NOT :: C, V, S)$$
(14)

$$\delta(\#NOT :: C, Boo(true) :: V, S) = \delta(C, Boo(false) :: V, S)$$
(15)

$$\delta(\#NOT :: C, Boo(false) :: V, S) = \delta(C, Boo(true) :: V, S)$$
(16)

- Notation *h*:: *Is* denotes the concatenation of element *h* with the list *Is*.
- *C* represents the *control* stack. *V* represents the *value* stack. *S* denotes the memory store.
- $\delta(\emptyset, V, S)$  denotes an accepting state.



### $\Pi$ automaton semantics for $\Pi$ IR expressions II

• On a particular implementation of the  $\Pi$  Framework, as in Python, <digits> denote built-in numbers in implementation language, such that all arithmetic operations, such as +, are defined. That's why  $N_i$  are in  $\mathbb{N}$ .



#### $\Pi$ IR commands

 Commands are language constructions that require a memory store to be evaluated.

```
⟨Statement⟩ ::= <Cmd>
⟨Exp⟩ := 'Id'(<String>)
⟨Cmd⟩ ::= 'Assign'(<Id>, <Exp>)
| 'Loop'(<BoolExp>, <Cmd>)
| 'CSeq'(<Cmd>, <Cmd>)
```



#### $\Pi$ automaton semantics for $\Pi$ IR commands I

- A location I ∈ Loc denotes a memory cell.
- Storable and Bindable sets denote the data that may be mapped to by identifiers and locations on the memory and environment respectively.
- Store = Loc → Storable, Env = Id → Bindable, Loc ⊆ Storable,
   N ⊆ Loc, Bindable.
- Now the transition function is  $\delta: L(G)^* \times L(G)^* \times Env \times Store \rightarrow Q$ , and let  $W \in String$ ,  $C, V \in L(G)^*$ ,  $S \in Store$ ,  $E \in Env$ ,  $B \in Bindable$ ,  $I \in Loc$ ,  $T \in Storable$ ,  $X \in Exp>$ ,  $M, M_1, M_2 \in Cmd>$ , and



#### $\Pi$ automaton semantics for $\Pi$ IR commands II

expression  $S' = S/[I \mapsto N]$  means that S' equals to S in all indices but I that is bound to N,

$$\delta(Id(W) :: C, V, E, S) = \delta(C, B :: V, E, S),$$

$$\text{where } E[W] = I \text{ and } S[I] = B,$$

$$(17)$$

$$\delta(Assign(W,X)::C,V,E,S) = \delta(X::\#ASSIGN::C,W::V,E,S'), \tag{18}$$

$$\delta(\#ASSIGN :: C, T :: W :: V, E, S) = \delta(C, V, E, S'),$$
 (19)

where 
$$E[W] = I$$
 and  $S' = S/[I \mapsto T]$ ,

$$\delta(Loop(X,M)::C,V,E,S) = \delta(X::\#LOOP::C,Loop(X,M)::V,E,S), \tag{20}$$

$$\delta(\#LOOP :: C, Boo(true) :: Loop(X, M) :: V, E, S) = \delta(M :: Loop(X, M) :: C, V, E, S), \tag{21}$$

$$\delta(\#LOOP :: C, Boo(false) :: Loop(X, M) :: V, E, S) = \delta(C, V, E, S), \tag{22}$$

$$\delta(CSeq(M_1, M_2) :: C, V, E, S) = \delta(M_1 :: M_2 :: C, V, E, S).$$
(23)



#### $\Pi$ IR declarations

- Declarations are statements that create an environment, binding identifiers to (bindable) values.
- In II IR, a bindable value is either a Boolean value, an integer or a location.
- From a syntactic standpoint, all classes are monotonically extended.

```
\langle Statement \rangle ::= \langle Dec \rangle
```

$$\langle Exp \rangle$$
 ::= 'Ref'( $\langle Exp \rangle$ )> | 'DeRef'( $\langle Id \rangle$ ) | 'ValRef'( $\langle Id \rangle$ )

$$\langle Dec \rangle$$
 ::= 'Bind'( $\langle Id \rangle$ ,  $\langle Exp \rangle$ ) | 'DSeq'( $\langle Dec \rangle$ ,  $\langle Dec \rangle$ )

$$\langle Cmd \rangle$$
 ::= 'Blk'( $\langle Dec \rangle$ ,  $\langle Cmd \rangle$ )



#### $\Pi$ automaton semantics for $\Pi$ IR declarations I

Let  $BlockLocs = \mathcal{P}(Loc)$ , now the transition function is  $\delta : L(G)^* \times L(G)^* \times Env \times Store \times BlockLocs \rightarrow Q$ , and let  $L, L' \in BlockLocs$ ,  $Loc \subseteq Storable$ , and  $S \setminus L$  means the store S without the locations in L,



#### $\Pi$ automaton semantics for $\Pi$ IR declarations II

$$\delta(Ref(X) :: C, V, E, S, L) = \delta(X :: \#REF :: C, V, E, S, L),$$
 (24)

$$\delta(\#REF :: C, T :: V, E, S, L) = \delta(C, I :: V, E, S', L'), \text{ where } S' = S \cup [I \mapsto T], I \notin S, L' = L \cup \{I\},$$

$$\delta(DeRef(Id(W)) :: C, V, E, S, L) = \delta(C, I :: V, E, S, L), \text{ where } I = E[W],$$
(26)

$$\delta(ValRef(Id(W)) :: C, V, E, S, L) = \delta(C, T :: V, E, S, L), \text{ where } T = S[S[E[W]]], \tag{27}$$

$$\delta(Bind(Id(W), X) :: C, V, E, S, L) = \delta(X :: \#BIND :: C, W :: V, E, S, L),$$

$$(\#BIND :: C, B :: W :: H :: V, E, S, L) = \delta(C, [W \mapsto B] :: V, E, S, L),$$
(29)

$$\delta(\#BIND :: C,B :: W :: H :: V,E,S,L) = \delta(C,[W \mapsto B] :: V,E,S,L),$$

$$\delta(DSeq(D_1, D_2), X) :: C, V, E, S, L) = \delta(D_1 :: D_2 :: \#DSEQ :: C, V, E, S, L), \tag{30}$$

$$\delta(\#DSEQ :: C,E :: E' :: H :: V,E,S,L) = \delta(C,(E \cup E') :: H :: V,E,S,L), \text{where } E' \in Env,H \notin Env,$$

$$\delta(Blk(D,M)::C,V,E,S,L) = \delta(D::\#BLKDEC::C,M::L::V,E,S,\phi), \tag{32}$$

$$\delta(\#BLKDEC :: C, E' :: V, E, S, L) = \delta(M :: \#BLKCMD :: C, E :: V, E \setminus E', S, L), \tag{33}$$

$$\delta(\#BLKCMD :: C,E :: L :: V,E',S,L') = \delta(C,V,E,S',L), \text{ where } S' = S \setminus L.$$



(31)

(25)

#### ∏ IR abstractions

- Abstractions extend Bindables by allowing a name to be bound to a list of formal parameters, a list of identifiers, and a block in the environment.
- Such names can be called and applied to actual parameters, a list of expressions.

```
::= 'Bind'(<Id>, <Abs>)
\langle Dec \rangle
\langle Abs \rangle
                     := 'Abs'(<Formals>, <Blk>)
\langle Formals \rangle ::= \langle Id \rangle^*
\langle Cmd \rangle ::= 'Call'(\langle Id \rangle, \langle Actuals \rangle)
\langle Actuals \rangle ::= \langle Exp \rangle^*
```



### $\Pi$ automaton semantics for $\Pi$ IR abstractions — Closures I

We chose a *static binding* semantics for abstractions. Therefore, we interpret abstractions as *closures* formed by an abstraction together with its declaration environment which defines the context in which the abstraction will be evaluated.

Closure : Formals  $\times$  Blk  $\times$  Env  $\rightarrow$  Bindable



## $\Pi$ automaton semantics for $\Pi$ IR abstractions — Example I

```
1 ImΠ source code:
2 # In this example we encapsulate the iterative calculation
3 # of the factorial within a function call.
4 let var z = 1
5 in
6 let fn f(x) =
7 let var y = x
8 in
9 while not (y == 0)
10 do
11 z := z * y
12 y := y - 1
13 in f(10)
```



## $\Pi$ automaton semantics for $\Pi$ IR abstractions — Example II

Automaton

```
1 Π IR AST:
  Blk(Bind(Id(z), Ref(Num(1))),
    Blk(Bind(Id(f), Abs(Id(x),
      Blk(Bind(Id(y), Ref(Id(x))),
        Loop(Not(Eq(Id(y), Num(0))),
          CSeq(Assign(Id(z), Mul(Id(z), Id(y))),
              Assign(Id(y), Sub(Id(y), Num(1)))))))
      Call(Id(f), Num(10)))
```



#### $\Pi$ automaton semantics for $\Pi$ IR abstractions I

Let  $F \in Formals$ ,  $B \in Blk$ ,  $I \in Id$ ,  $A \in Actuals$ ,  $V_i \in Value$ ,  $1 \le i \le n$ ,  $n \in \mathbb{N}$ ,

$$\delta(Abs(F,B)::C,V,E,S,L) = \delta(C,Closure(F,B,E)::V,E,S,L)$$
 (35)

$$\delta(Call(I,[X_1,X_2,...,X_n])) :: C, V, E, S, L) =$$

$$\delta(X_n :: X_{n-1} :: ... :: X_1 :: \#CALL(I,n) :: C, V, E, S, L)$$
(36)

$$\delta(\#CALL(I,n) :: C, (V_1 :: V_2 :: ... V_n :: V), E, S, L) =$$

$$\delta(B :: \#BLKCMD :: C, E :: V, E', S, L)$$

$$\mathbf{where} \ E = \{I \mapsto Closure(F, B, E_1)\} \cup E_2,$$

$$(37)$$

 $E' = E/E_1/match(F, [V_1, V_2, ..., V_n])$ 



#### $\Pi$ automaton semantics for $\Pi$ IR abstractions II

```
match: Id^* \times Values^* \rightarrow Env
match(fl, al) = if|fl| \neq |al| \text{ than } \{\} \text{ else } \_match(fl, al, \{\})\}
\_match: Id^* \times Values^* \times Env \rightarrow Env
\_match([], [], E) = E
\_match(f, a, E) = \{f \mapsto a\} \cup E
\_match(f :: fl, a :: al, E) = \_match(fl, al, \{f \mapsto a\} \cup E)
```



#### 

 Abstractions can be recursive to allow for the declaration of recursive functions.

$$\langle Dec \rangle$$
 ::= 'Rbnd'( $\langle Id \rangle$ ,  $\langle Abs \rangle$ )



## $\Pi$ automaton semantics for $\Pi$ IR recursive abstractions — Recursive closures I

In the context of *static binding* semantics for abstractions, in a call to a recursive function, the evaluation of identifiers needs to be reminded about the binding of the function name to a closure.

Rec : Formals  $\times$  Blk  $\times$  Env  $\times$  Env  $\rightarrow$  Bindable

 $unfold : Env \rightarrow Env$ 

 $reclose_E : Env \rightarrow Env$ 



# $\Pi$ automaton semantics for $\Pi$ IR recursive abstractions — Recursive closures II

$$unfold(E) = reclose_E(E)$$
 (38)

$$reclose_E(I \mapsto Closure(F, B, E')) = (I \mapsto Rec(F, B, E', E))$$
 (39)

$$reclose_{E}(I \mapsto Rec(F, B, E', E'')) = (I \mapsto Rec(F, B, E', E)) \tag{40}$$

$$reclose_E(I \mapsto v) = (I \mapsto v) \text{ if } v \neq Closure(F, B, E)$$
 (41)

$$reclose_E(E_1 \cup E_2) = reclose_E(E_1) \cup reclose_E(E_2)$$
 (42)

$$reclose_{E}(\emptyset) = \emptyset$$
 (43)



# $\Pi$ automaton semantics for $\Pi$ IR recursive abstractions — Recursive closures I

$$\delta(Rbnd(I,Abs(F,B)) :: C,V,E,S,L) =$$

$$\delta(C,unfold(I \mapsto Closure(F,B,E)) :: V,E,S,L) \qquad (44)$$

$$\delta(\#CALL(I,n) :: C,V_1 :: V_2 :: ... :: V_n :: V,E,S,L) =$$

$$\delta(B :: \#BLKCMD :: C,E :: V,E',S,L) \qquad (45)$$

$$\mathbf{where} \ E = \{I \mapsto Rec(F,B,E_1,E_2)\} \cup E_3,$$

$$E' = E/E_1/unfold(E_2)/match(F,[V_1,V_2,...,V_n])$$



## 

```
class Statement:
    def __init__(self, *args):
        self.opr =args

def __str__(self):
    ret =str(self.__class__.__name__)+"("
    for o in self.opr:
        ret +=str(o)
    ret +=")"
    return ret
class Exp(Statement): pass
class ArithExp(Exp): pass
```



```
class Num(ArithExp):
    def __init__(self, f):
        assert(isinstance(f, int))
        ArithExp.__init__(self,f)

class Sum(ArithExp):
    def __init__(self, e1, e2):
        assert(isinstance(e1, Exp) and isinstance(e2, Exp))
        ArithExp.__init__(self, e1, e2)

...
```



### Π IR expressions in Python III

```
class BoolExp(Exp): pass
class Eq(BoolExp):
def __init__(self, e1, e2):
    assert(isinstance(e1, Exp) and isinstance(e2, Exp))
    BoolExp.__init__(self, e1, e2)
...
```



```
exp =Sum(Num(1), Mul(Num(2), Num(4)))
print(exp)

Sum(Num(1)Mul(Num(2)Num(4)))
```



```
_{1} \exp 2 = Mul(2, 1)
3 AssertionError Traceback (most recent call last)
4 <ipython-input-7-00fd40a79a54> in <module>()
5 \longrightarrow 1 \exp 2 = Mul(2, 1)
7 <ipython-input-5-42a82e58862f> in __init__(self, e1, e2)
       28 class Mul(ArithExp):
       29 def __init__(self, e1, e2):
10 --->30 assert(isinstance(e1, Exp) and isinstance(e2, Exp))
       31 ArithExp.__init__(self, e1, e2)
       32 class BoolExp(Exp): pass
12
4 AssertionError:
```



### $\Pi$ automaton for $\Pi$ IR expressions I

```
1 ## Expressions
2 class ValueStack(list): pass
3 class ControlStack(list): pass
4 class ExpKW:
5 SUM = "#SUM"
6 SUB = "#SUB"
7 MUL = "#MUL"
8 EQ = "#EQ"
9 NOT = "#NOT"
```



### $\Pi$ automaton for $\Pi$ IR expressions II

```
1 class ExpPiAut(dict):
     def __init__(self):
         self["val"] =ValueStack()
         self["cnt"] =ControlStack()
4
     def __evalSum(self, e):
5
         e1 =e.opr[0]
6
         e2 =e.opr[1]
         self.pushCnt(ExpKW.SUM)
         self.pushCnt(e1)
         self.pushCnt(e2)
LO
     def pushCnt(self, e):
         cnt =self.cnt()
         cnt.append(e)
۱4
```



### $\Pi$ automaton for $\Pi$ IR expressions III

```
1 ea =ExpPiAut()
2 print(exp)
3 ea.pushCnt(exp)
4 while not ea.emptyCnt():
5     ea.eval()
6     print(ea)
```



### $\Pi$ automaton for $\Pi$ IR expressions IV

```
1 Sum(Num(1)Mul(Num(2)Num(4)))
2 {'val': [], 'cnt': ['#SUM', <__main__.Num object at 0x111851470>, <
                                      __main__.Mul object at 0x1118516d8>]
3 {'val': [], 'cnt': ['#SUM', <__main__.Num object at 0x111851470>, '#MUL'
                                      , <__main__.Num object at
                                      0x111851630>, <__main__.Num object
                                      at 0x1118516a0>]}
4 {'val': [4], 'cnt': ['#SUM', <__main__.Num object at 0x111851470>, '#MUL
                                      ', <__main__.Num object at
                                      0x111851630>]}
5 {'val': [4, 2], 'cnt': ['#SUM', <__main__.Num object at 0x111851470>, '#
                                      MUL'1}
6 {'val': [8], 'cnt': ['#SUM', <__main__.Num object at 0x111851470>]}
7 {'val': [8, 1], 'cnt': ['#SUM']}
8 {'val': [9], 'cnt': []}
```



#### ∏ IR commands I

```
1 class Cmd(Statement): pass
2 class Id(Exp):
    def __init__(self, s):
3
         assert(isinstance(s, str))
         Exp.__init__(self, s)
6 class Assign(Cmd):
     def __init__(self, i, e):
7
         assert(isinstance(i, Id) and isinstance(e, Exp))
         Cmd.__init__(self, i, e)
10 class Loop(Cmd):
     def __init__(self, be, c):
11
12
         assert(isinstance(be, BoolExp) and isinstance(c, Cmd))
         Cmd.__init__(self, be, c)
13
14 class CSeq(Cmd):
     def __init__(self, c1, c2):
15
         assert(isinstance(c1, Cmd) and isinstance(c2, Cmd))
16
         Cmd.__init__(self, c1, c2)
17
```



### ∏ IR commands II

```
cmd =Assign(Id("x"), Num(1))
print(type(cmd))
print(cmd)

<class '__main__.Assign'>
5 Assign(Id(x)Num(1))
```



#### $\Pi$ automaton for $\Pi$ IR commands I

Environment, Location, Store and commands opcodes.

```
1 ## Commands
2 class Env(dict): pass
3 class Loc(int): pass
4 class Sto(dict): pass
5 class CmdKW:
    ASSIGN = "#ASSIGN"
    LOOP = "#LOOP"
```



#### $\Pi$ automaton for $\Pi$ IR commands II

 $\Pi$  automaton for commands extends the  $\Pi$  automaton for expressions.

```
1 class CmdPiAut(ExpPiAut):
     def __init__(self):
         self["env"] =Env()
         self["sto"] =Sto()
         ExpPiAut.__init__(self)
     def env(self):
6
         return self["env"]
     def getLoc(self, i):
         en =self.env()
         return en[i]
LO
     def sto(self):
11
         return self["sto"]
     def updateStore(self, 1, v):
         st =self.sto()
         st[1] =v
```

#### $\Pi$ automaton for $\Pi$ IR commands III

 $\Pi$  semantics for assignment.

```
\begin{split} \delta(Assign(W,X) :: C,V,E,S) &= \delta(X :: \#ASSIGN :: C,W :: V,E,S'), \\ \delta(\#ASSIGN :: C,T :: W :: V,E,S) &= \delta(C,V,E,S'), \\ \text{where } E[W] &= I \text{ and } S' = S/[I \mapsto T]. \end{split}
```

```
def __evalAssign(self, c):
    i =c.opr[0]
    e =c.opr[1]
    self.pushVal(i.opr[0])
    self.pushCnt(CmdKW.ASSIGN)
    self.pushCnt(e)

def __evalAssignKW(self):
    v =self.popVal()
    i =self.popVal()
    l =self.getLoc(i)
```

### $\Pi$ automaton for $\Pi$ IR commands IV

self.updateStore(1, v)



#### $\Pi$ automaton for $\Pi$ IR commands V

 $\Pi$  semantics for identifiers.

```
\delta(Id(W)::C,V,E,S) = \delta(C,B::V,E,S), where E[W] = I and S[I] = B.
```

```
def __evalId(self, i):
    s =self.sto()
    l =self.getLoc(i)
    self.pushVal(s[1])
```



#### $\Pi$ automaton for $\Pi$ IR commands VI

 $\Pi$  semantics for loop: recursive step.

```
\delta(Loop(X,M)::C,V,E,S) = \delta(X::\#LOOP::C,Loop(X,M)::V,E,S).
```

```
def __evalLoop(self, c):
    be =c.opr[0]
    bl =c.opr[1]
    self.pushVal(Loop(be, bl))
    self.pushVal(bl)
    self.pushCnt(CmdKW.LOOP)
    self.pushCnt(be)
```



 $\Pi$  semantics for loop: basic steps.

```
\delta(\#LOOP :: C, Boo(true) :: Loop(X, M) :: V, E, S) = \delta(M :: Loop(X, M) :: C, V, E, S), \delta(\#LOOP :: C, Boo(false) :: Loop(X, M) :: V, E, S) = \delta(C, V, E, S).
```

```
def __evalLoopKW(self):
    t =self.popVal()
    if t:
        c =self.popVal()
        lo =self.popVal()
        self.pushCnt(lo)
        self.pushCnt(c)
    else:
        self.popVal()
    self.popVal()
    self.popVal()
```

#### Π automaton for Π IR commands VIII

 $\Pi$  semantics for command composition.

```
\delta(CSeq(M_1, M_2) :: C, V, E, S) = \delta(M_1 :: M_2 :: C, V, E, S).
```

```
def __evalCSeq(self, c):
    c1 =c.opr[0]
    c2 =c.opr[1]
    self.pushCnt(c2)
    self.pushCnt(c1)
```



#### $\Pi$ automaton for $\Pi$ IR commands IX

Commands are now on the top of the food chain.

```
def eval(self):
         c =self.popCnt()
         if isinstance(c, Assign):
             self.__evalAssign(c)
         elif c ==CmdKW.ASSIGN:
5
             self.__evalAssignKW()
6
         elif isinstance(c, Id):
             self.__evalId(c.opr[0])
         elif isinstance(c, Loop):
             self.__evalLoop(c)
         elif c == CmdKW.I.OOP:
11
             self.__evalLoopKW()
         elif isinstance(c, CSeq):
             self.__evalCSeq(c)
         else:
             self.pushCnt(c)
             ExpPiAut.eval(self)
١7
```

 $\Pi$  automaton for  $\Pi$  IR commands X



### ∏ IR declarations in Python I

#### DeRef and ValRef not implemented yet.

```
1 ## Declarations
2 class Dec(Statement): pass
3 class Bind(Dec):
    def __init__(self, i, e):
        assert (isinstance(i, Id) and isinstance(e, Exp))
        Dec.__init__(self, i, e)
7 class Ref(Exp):
    def __init__(self, e):
8
        assert (isinstance(e, Exp))
        Exp.__init__(self, e)
LO
11 class Blk(Cmd):
    def __init__(self, d, c):
12
13
        assert (isinstance(d, Dec) and isinstance(c, Cmd))
۱4
        Cmd.__init__(self, d, c)
```



### ∏ IR declarations in Python II

```
class DSeq(Dec):
    def __init__(self, d1, d2):
        assert (isinstance(d1, Dec) and isinstance(d2, Dec))
        Dec.__init__(self, d1, d2)
```



```
1 ## Declarations
2 class DecExpKW(ExpKW):
    REF ="#REF"
3
 class DecCmdKW(CmdKW):
    BLKDEC ="#BLKDEC"
6
    BLKCMD ="#BLKCMD"
7
 class DecKW:
    BTND ="#BTND"
LO
    DSEQ ="#DSEQ"
11
class DecPiAut(CmdPiAut):
14
    def __init__(self):
        self["locs"] =[]
        CmdPiAut.__init__(self)
۱6
    def locs(self):
```

```
return self["locs"]
20
     def pushLoc(self, 1):
21
         ls =self.locs()
         ls.append(1)
24
     def __evalRef(self, e):
         ex = e.opr[0]
26
         self.pushCnt(DecExpKW.REF)
         self.pushCnt(ex)
28
29
     def __newLoc(self):
30
         sto =self.sto()
31
         if sto:
32
             return max(list(sto.keys())) +1
33
         else:
34
             return 0.0
35
36
     def __evalRefKW(self):
```

```
v =self.popVal()
38
         1 =self.__newLoc()
39
         self.updateStore(1, v)
10
         self.pushLoc(1)
11
         self.pushVal(1)
12
13
     def __evalBind(self, d):
14
         i =d.opr[0]
         e =d.opr[1]
16
         self.pushVal(i)
         self.pushCnt(DecKW.BIND)
18
         self.pushCnt(e)
19
50
     def __evalBindKW(self):
51
         1 =self.popVal()
         i =self.popVal()
53
         x = i.opr[0]
54
         self.pushVal({x: 1})
55
```

61 / 85

```
def __evalDSeq(self, ds):
57
        d1 =ds.opr[0]
        d2 =ds.opr[1]
         self.pushCnt(DecKW.DSEQ)
         self.pushCnt(d2)
         self.pushCnt(d1)
53
     def __evalDSeqKW(self):
        d2 =self.popVal()
        d1 =self.popVal()
        d1.update(d2)
         self.pushVal(d1)
59
     def __evalBlk(self, d):
70
        ld =d.opr[0]
        c = d.opr[1]
        1 =self.locs()
        self.pushVal(list(1))
74
         self.pushVal(c)
```

50

51

52

54

56

57

# Π IR declarations in Python VII

76

77 78

79

30

31

33 34

35

36

37

38

39

90 91

92

```
self.pushCnt(DecCmdKW.BLKDEC)
   self.pushCnt(ld)
def __evalBlkDecKW(self):
   d =self.popVal()
   c =self.popVal()
   1 =self.locs()
   self.pushVal(1)
   en =self.env()
   ne =en.copy()
   ne.update(d)
   self.pushVal(en)
   self["env"] =ne
   self.pushCnt(DecCmdKW.BLKCMD)
   self.pushCnt(c)
def __evalBlkCmdKW(self):
   en =self.popVal()
   ls =self.popVal()
```

```
self["env"] =en
   s =self.sto()
   s ={k: v for k, v in s.items() if k not in ls}
   self["sto"] =s
   # del 1s
   ols =self.popVal()
   self["locs"] =ols
def eval(self):
   d =self.popCnt()
   if isinstance(d, Bind):
       self.__evalBind(d)
   elif d ==DecKW.BIND:
       self.__evalBindKW()
   elif isinstance(d, DSeq):
       self.__evalDSeq(d)
   elif d ==DecKW.DSEQ:
       self.__evalDSeqKW()
   elif isinstance(d, Ref):
```

98

99

00

02

)5

06

07

80

9

LO

```
self.__evalRef(d)
elif d ==DecExpKW.REF:
    self.__evalRefKW()
elif isinstance(d, Blk):
    self.__evalBlk(d)
elif d ==DecCmdKW.BLKDEC:
    self.__evalBlkDecKW()
elif d ==DecCmdKW.BLKCMD:
    self.__evalBlkCmdKW()
else:
    self.pushCnt(d)
    CmdPiAut.eval(self)
```



۱6

18

20

22

23

24

## $\Pi$ IR declarations in Python X

```
1 dc =DecPiAut()
2 fac =Loop(Not(Eq(Id("y"), Num(0))),
           CSeq(Assign(Id("x"), Mul(Id("x"), Id("y"))),
                Assign(Id("v"), Sub(Id("v"), Num(1)))))
5 dec =DSeq(Bind(Id("x"), Ref(Num(1))),
           Bind(Id("v"), Ref(Num(200))))
7 fac_blk =Blk(dec, fac)
8 dc.pushCnt(fac_blk)
9 while not dc.emptyCnt():
     aux =dc.copy()
LO
    dc.eval()
11
    if dc.emptyCnt():
        print(aux)
13
```



## $\Pi$ IR abstractions in Python I

```
1
2 class Formals(list):
     def __init__(self, f):
         if isinstance(f, list):
             for a in f:
                 if not isinstance(a, Id):
6
                     raise IllFormed(self, a)
7
             self.append(f)
         else:
             raise IllFormed(self, f)
LO
12 class Abs:
     def __init__(self, f, b):
         if isinstance(f, list):
             if isinstance(b, Blk):
                 self._opr =[f, b]
             else:
                 raise IllFormed(self, b)
```

```
else:
             raise IllFormed(self, f)
20
      def formals(self):
         return self._opr[0]
24
      def blk(self):
         return self._opr[1]
26
      def __str__(self):
28
         ret =str(self.__class__.__name__) +"("
         formals =self.formals()
30
         ret +=str(formals[0]) # First formal argument
31
         for i in range(1, len(formals)):
32
             ret +=". "
33
             ret +=str(formals[i]) # Remaining formal arguments
34
         ret +=". "
35
         ret +=str(self.blk()) # Abstraction block
36
         ret +=")"
```

```
return ret
38
39
o class BindAbs(Bind):
      ,,,
11
      BindAbs is a form of bind but that receives an Abs instead of an
12
      expression.
13
      ,,,
      def __init__(self, i, p):
15
          if isinstance(i, Id):
16
              if isinstance(p, Abs):
                  Dec.__init__(self, i, p)
18
19
              else:
                  raise IllFormed(self, p)
50
          else:
51
              raise IllFormed(self, i)
53
64 class Actuals(list):
     def __init__(self, a):
55
          if isinstance(a, list):
```

```
for e in a:
                 if not isinstance(e, Exp):
                     raise IllFormed(self, e)
             self.append(a)
50
          else:
51
             raise IllFormed(self, a)
52
53
64 class Call(Cmd):
      def __init__(self, f, actuals):
          if isinstance(f, Id):
56
              if isinstance(actuals, list):
                 Cmd.__init__(self, f, actuals)
58
              else:
59
                 raise IllFormed(self, actuals)
70
          else:
             raise IllFormed(self, f)
73
      def caller(self):
74
         return self.operand(0)
```

```
def actuals(self):
77
         return self.operand(1)
78
79
class Closure(dict):
     def __init__(self, f, b, e):
31
         if isinstance(f, list):
             if isinstance(b, Blk):
33
34
                 # I wanted to write assert(isinstance(e, Env)) but it
                                                        fails.
                 if isinstance(e, dict):
                     self["for"] =f # Formal parameters
36
                     self["env"] =e # Current environment
                     self["block"] =b # Procedure block
38
                 else:
39
                     raise IllFormed(self, e)
90
             else:
91
                 raise IllFormed(self, b)
         else:
```

```
raise IllFormed(self, f)
94
     def __str__(self):
96
         ret =str(self.__class__.__name__) +"("
         formals =self.formals()
98
         fst_formal =formals[0] # First formal argument
99
         ret +=str(fst_formal)
         for i in range(1, len(formals)):
             ret +=", "
             formal =formals[i] # Remaining formal arguments
             ret +=str(formal)
         ret +=". "
)5
         ret +=str(self.blk()) # Closure block
06
         ret +=")"
18
         return ret
9
     def formals(self):
LO
         return self['for']
```

```
def env(self):
13
         return self['env']
     def blk(self):
۱6
         return self['block']
١7
class AbsPiAut(DecPiAut):
     def __evalAbs(self, a):
20
         if not isinstance(a, Abs): # p must be an abstraction
             raise EvaluationError(self, "Function __evalAbs called with
                                                  no abstraction but with "
                                                   , a, " instead.")
         else:
             f =a.formals() # Formal parameters
             b =a.blk() # Body
             e =self.env() # Current environment
26
             # Closes the given abs. with the current env
             c =Closure(f, b, e)
             # Closure c is pushed to the value stack such that
```

```
self.pushVal(c)
       # a BIND may create a new binding to a given identifier.
def __match(self, f, a):
    , , ,
   Given a list of formal parameters and a list of actual parameters
   it returns an environment relating the elements of the former
                                          with the latter.
    ,,,
   if isinstance(f, list):
       if isinstance(a, list):
           if len(f) == 0:
               return {}
           if len(f) ==len(a) and len(f) >0:
           # For some reason, f[0] is a tuple, not an Id.
               f0 = f[0]
               a0 = a[0]
               b0 = \{f0.id(): a0.num()\}
```

31

33

34

35

37 88

39

10

11

13

```
if len(f) ==1:
17
                      return b0
18
                  else:
19
                      # For some reason, f[0] is a tuple, not an Id.
50
                      f1 =f[1]
51
                      a1 = a[1]
52
                      b1 = \{f1.id(): a1.num()\}
                      e =b0.update(b1)
54
                      for i in range(2, len(f)):
                          fi =f[i][0]
56
                          ai =a[i][0]
                          e.update({fi.id(): ai.num()})
58
                      return e
              else:
50
                  raise EvaluationError("Call to '__match' on " +str(self) +
                                                          ": " +"formals and
                                                          actuals differ in
                                                          size.")
          else:
```

## $\Pi$ IR abstractions in Python X

```
raise EvaluationError("Call to '__match' on " +str(self) +":
                                             " +" no formals, but with
                                              ", f, " instead.")
def __evalCall(self, c):
    ,,,
   Essentially, a call is translated into a block.
   If we were progrmming pi in a symbolic language,
   we could simply crete a proper block and push it to the control
                                         stack.
   However, the environment is not symbolic: is a dictionary of
                                         objects.
   To create a block we would need to "pi-IR-fy" it, that is,
                                         recreate the
   pi IR tree from the concrete environmnet and joint it with
                                         matches created
   also at pi IR level. These would be pushed back into the control
                                         stack and
                                                                 1:UIIIDULACAO
```

55

56

57

58

59

70

72

# ∏ IR abstractions in Python XI

```
reobjectifyed. Thus, to avoid pi-IRfication and reevaluatuation
                                     of the
environment we manipulate it at the object level, which is
                                    dangerous but
seems to be correct.
In this implementation, actual parameters are already evaluated.
,,,
if not isinstance(c, Call): # c must be a Call object
   raise EvaluationError("Call to __evalCall with no Call object
                                         but with ", c, " instead
else:
   # Procedure to be called
   caller =c.caller()
   # Retrieves the current environment.
   e =self.env()
   # Retrieves the closure associated with the caller function.
   clos =e[caller.id()]
```

75

76

78

79

30

31

33

34

35

36

90

33

94

00

01

```
# Retrieves the actual parameters from the call.
a =c.actuals()
# Retrieves the formal parameters from the closure.
f =clos.formals()
# Matches formals and actuals, creating an environment.
d =self.__match(f, a)
# Retrives the closure's environment.
ce =clos.env()
# The caller's block must run on the closures environment
# overwritten with the matches.
d.update(ce)
self["env"] =d
self.pushVal(self.locs())
# Saves the current environment in the value stack.
self.pushVal(e)
# Pushes the keyword BLKCMD for block completion.
self.pushCnt(DecCmdKW.BLKCMD)
# Pushes the body of the caller function into the control
                                     stack.
```

```
self.pushCnt(clos.blk())

def eval(self):
    d =self.popCnt()
    if isinstance(d, Abs):
        self.__evalAbs(d)
    elif isinstance(d, Call):
        self.__evalCall(d)
    else:
        self.pushCnt(d)
        DecPiAut.eval(self)
```



# Complete Imp grammar in EBNF notation I

```
1 @@grammar::IMP
2 @@eol_comments ::/#.*?/start = @:cmd ;
6 cmd =nop | let | assign | loop | call;
8 call =i:identifier '(' { a:actual }* ')';
10 actual =e1:expression { ',' e2:expression }* | {} ;
12 nop ='nop';
14 loop =op:'while' ~ e:expression 'do' { c:cmd }+ ;
assign =id:identifier op:':=' ~ e:expression ;
18 let =op:'let' ~ d:dec 'in' { c:cmd }+ ;
```

## Complete Imp grammar in EBNF notation II

```
19
20 dec =var | fn :
21
var =op: 'var' ~ id:identifier '=' e:expression ;
23
fn =op:'fn' ~ id:identifier '(' f:formal ')' '=' c:cmd ;
26 formal =i1:identifier { ',' i2:identifier }* | {} ;
27
expression =0:bool_expression;
9
so bool_expression =negation | equality | conjunction | disjunction
                  lowereq | greatereq | lowerthan | greaterthan
31
                 | add_expression ;
33
g4 equality =left:add_expression op:"==" ~ right:bool_expression ;
35
se conjunction =left:add_expression op:"and" ~ right:bool_expression ;
37
```

## Complete Imp grammar in EBNF notation III

```
38 disjunction =left:add_expression op:"or" ~ right:bool_expression ;
39
lowereq =left:add_expression op:"<=" ~ right:add_expression ;
11
greatereq =left:add_expression op:">=" ~ right:add_expression ;
13
lowerthan =left:add_expression op:"<" ~ right:add_expression ;
15
greaterthan =left:add_expression op:">" ~ right:add_expression ;
parentesisexp = '(' ~ @:bool_expression ')';
negation =op: 'not' ~ b:bool_expression ;
add_expression =addition | subtraction | @:mult_expression ;
53
addition =left:mult_expression op:"+" ~ right:add_expression ;
55
subtraction =left:mult_expression op:"-" ~ right:add_expression ;
```

### Complete Imp grammar in EBNF notation IV

```
mult_expression =multiplication | division
59
                  atom
                parentesisexp;
50
multiplication =left:atom op:"*" ~ right:mult_expression ;
division =left:atom op:"/" ~ right:mult_expression ;
atom =number | truth | identifier ;
number =/d+/;
identifier =/(?!\d)\w+/;
r2 truth = 'True' | 'False';
```



#### Example: iterative factorial

```
# The classic iterative factorial example let var z=1 in let var y=10 in while not (y==0) do z:=z*y y:=y-1
```



### Example: iterative factorial within a function

```
# In this example we encapsulate the iterative calculation
# of the factorial within a function call.

let var z = 1

in

let fn f(x) = 

let var y = x

in

while not (y == 0)

do

z := z * y

11

y := y - 1

in f(10)
```

