Typechecking Simple Modules CS496

Modules

For large systems one needs:

- A good way to separate the system into relatively self-contained parts, and to document the dependencies between those parts.
- 2. A better way to control the scope and binding of names.
 - Lexical scoping insufficient when programs may be large or split up over multiple sources
- 3. A way to enforce abstraction boundaries (interface/implem.)
- 4. A way to combine these parts flexibly, so that a single part may be reused in different contexts.

SIMPLE-MODULES

- Adds a simple module system to REC
- A program consists of a sequence of module definitions followed by an expression to be evaluated.

```
module m1
     interface
      [a:int
       b: int
       c : intl
     body
      fa = 33
       x = a-1 (* = 32 *)
       b = a-x (* = 1 *)
       c = x-b] (* =31 *)
10
11
   in ((from m1 take a)-
12
          (from m1 take b))-a
13
```

SIMPLE-MODULES

Each module establishes an abstraction boundary between the module body and the rest of the program.

► The expressions in the module body are inside the abstraction boundary,

10 11

12

13

and everything else is outside the abstraction boundary.

```
module m1
  interface
   [a:int
    b: int
    c : intl
  body
   \Gamma a = 33
    x = a-1
              (* = 32 *)
    b = a - x \quad (* = 1 *)
    c = x-bl (* = 31 *)
in ((from m1 take a)-
        (from m1 take b))-a
```

SIMPLE-MODULES

- ► Each module definition binds a name to a module
- A created module is a set of bindings, much like an environment
- Next we'll see more examples of modules

```
module m1
     interface
       Γa : int
       b: int
        c : intl
    body
       [a = 33]
       x = a-1 (* = 32 *)
       b = a - x \quad (* = 1 *)
        c = x-b] (* =31 *)
10
   let a = 10
11
   in ((from m1 take a)-
12
           (from m1 take b))-a
13
```

```
module m1
     interface
     [a:int
     b : int
      c : int]
   body
    Γa = 33
    x = a-1 (* = 32 *)
    b = a-x (* = 1 *)
       c = x-b] (* =31 *)
10
   let a = 10
11
12
   in ((from m1 take a)-
          (from m1 take b))-a
13
```

- ▶ Has type int and value ((33 1) 10) = 22.
- ▶ from m1 take a and from m1 take b are called qualified variables

```
module m1
interface
    [u : bool]
body
    [u = 33]
```

- ► Type error!
- ► The body of the module must associate each name in the interface with a value of the appropriate type, even if those values are not used elsewhere in the program.

```
module m1
interface
    [u : int
    v : int]
body
    [u = 33]
```

- ► Type error!
- ► The module body must supply bindings for all the declarations in the interface.

```
module m1
interface
    [u : int
    v : int]
body
    [v = 33
    u = 44]
from m1 take u
```

- ► Type error!
- ➤ To keep the implementation simple, our language requires that the module body produce the values in the same order as the interface.

```
module m1
     interface
        [u : int]
3
    body
        [u = 44]
5
6
   module m2
     interface
        [v : int]
10
    body
        [v = (from m1 take u) -11]
11
   (from m1 take u)-(from m2 take v)
12
```

- ► Has type int
- ▶ Modules have let* scoping

```
module m2
     interface
2
        [v : int]
3
     body
5
        [v = (from m1 take u) -11]
6
   module m1
7
     interface
8
        [u : int]
9
     body
10
        [u = 44]
11
   (from m1 take u)-(from m2 take v)
12
```

- ► Type error!
- ▶ from m1 take u is not in scope where it is used in the body of m2.

SIMPLE-MODULES: Concrete Syntax

SIMPLE-MODULES: Concrete Syntax

```
\langle ModuleDefn \rangle ::= module \langle Identifier \rangle interface \langle Iface \rangle body \langle ModuleBody \rangle
```

Syntax for the interface of a module:

```
\langle \textit{Iface} \rangle ::= [\{\langle \textit{Decl} \rangle\}^*] 
\langle \textit{Decl} \rangle ::= \langle \textit{Identifier} \rangle : \langle \textit{Type} \rangle
```

Syntax for the body of a module

```
\langle ModuleBody \rangle ::= [\{\langle Defn \rangle\}^*] 
\langle Defn \rangle ::= \langle Identifier \rangle = \langle Expression \rangle
```

${\bf SIMPLE\text{-}MODULES:} \ \textbf{Concrete Syntax}$

```
\langle \textit{Expression} \rangle ::= ...as before... \langle \textit{Expression} \rangle ::= from \langle \textit{Identifier} \rangle take \langle \textit{Identifier} \rangle
```

SIMPLE-MODULES: Abstract Syntax

```
type prog = AProg of (mdecl list)*expr \langle \textit{Program} \rangle \ ::= \ \{\langle \textit{ModuleDefn} \rangle\}^* \ \langle \textit{Expression} \rangle
```

SIMPLE-MODULES: Abstract Syntax

```
type mdecl = AMod of string*interface*module_body \langle \textit{ModuleDefn} \rangle \ ::= \ \begin{array}{c} \text{module} \ \langle \textit{Identifier} \rangle \\ \text{interface} \ \langle \textit{Iface} \rangle \\ \text{body} \ \langle \textit{ModuleBody} \rangle \end{array}
```

SIMPLE-MODULES: Abstract Syntax

Concrete vs Abstract Syntax

```
module m1
    interface
    ſu : int
3
    v : int]
   body
6
   [v = 33]
     u = 44
  from m1 take u
  AProg
   ([AMod ("m1",
     ASimpleInterface [AValDecl ("u", IntType); AValDecl ("v",
3
         → IntType)],
     AModBody [AValDef ("v", Int 33); AValDef ("u", Int 44)])],
   QualVar ("m1", "u"))
```

Simple Modules

The Interpreter

The Type-Checker

```
Two part process
                                 module m1
                                    interface
    1. Module body evaluation
                                     ſa : int
    2. Body evaluation
                                      b: int
Module body evaluation
                                      c : int]
                                    body
    Will produce an
                                     fa = 33
       environment consisting of
                                      x = a-1 (* = 32 *)
       all the bindings exported
                                      b = a - x \quad (* = 1 *)
       by the module.
                                      c = x-b] (* =31 *)
                              10
                                  let a = 10
                              11
Body Evaluation
                                  in ((from m1 take a)-
                              12
    Will produce an expressed 13
                                          (from m1 take b))-a
      value.
```

- Two part process
 - 1. Module body evaluation
 - 2. Body evaluation
- Module body evaluation
 - Will produce an environment consisting of all the bindings exported by the module.

10

11

12

- Body Evaluation
 - Will produce an expressed 13 value.

```
module m1
  interface
   [a:int
    b: int
    c : int]
  body
   fa = 33
    x = a-1 (* = 32 *)
    b = a - x \quad (* = 1 *)
    c = x-b] (* =31 *)
let a = 10
in ((from m1 take a)-
       (from m1 take b))-a
```

- Two part process
 - 1. Module body evaluation
 - 2. Body evaluation
- Module body evaluation
 - Will produce an environment consisting of all the bindings exported by the module.

10

11

12

- Body Evaluation
 - Will produce an expressed 13 value.

```
module m1
  interface
   [a:int
    b: int
    c : int]
  body
   fa = 33
    x = a-1 (* = 32 *)
    b = a - x \quad (* = 1 *)
    c = x-b] (* =31 *)
let a = 10
in ((from m1 take a)-
       (from m1 take b))-a
```

Module Body Evaluation

Evaluation of a module body will produce an environment consisting of all the bindings exported by the module.

```
env =
| EmptyEnv |
| ExtendEnv of string*exp_val*env |
| ExtendEnvRec of string*string*Ast.expr*env |
| ExtendEnvMod of string*env*env
```

```
module m1
     interface
      [a:int
     b : int
     c : int]
5
    body
7
     [a = 33]
8
      b = 44
     c = 551
   module m2
10
     interface
11
     [a : int
12
      b : int]
13
   body
14
15
     [a = 66]
       b = 771
16
17
   let z = 99
   in z-((from m1 take a)-(from m2 take a))
18
```

Module Body Evaluation

Environment and Store extant at line 17

```
1 Environment:
2 [Module m2 [(b,NumVal 77)(a,NumVal 66)];
3 Module m1 [(c,NumVal 55)(b,NumVal 44)(a,NumVal 33)];
4 (i,NumVal 1);
5 (v,NumVal 5);
6 (x,NumVal 10)]
7 Store:
```

Body Evaluation

 Same as before except we now have to deal with qualified variables

from m take var

- We use apply_env_qual
- ► This first looks up the module m in the current environment, and then looks up var in the resulting environment.

```
eval_prog (Ast.AProg(ms,e)) =
    eval_expr (add_module_definitions init_env ms) e
```

```
eval_prog (Ast.AProg(ms,e)) =
eval_expr (add_module_definitions init_env ms) e
```

► Prepares the environment with the result of evaluating all the modules

```
add_module_definition (en:env) (AModBody vdefs):env =
     snd @@ List.fold_left (fun (env,renv) (AValDef(var,decl))
2
         → ->
3
          let ev=eval_expr env decl
          in (ExtendEnv(var,ev,env),ExtendEnv(var,ev,renv))) (en
4
              → .EmptvEnv)
5
       vdefs
6
   and
     add_module_definitions (en:env) (ms:mdecl list):env =
7
         List.fold_left (fun curr_en (AMod(mname, minterface,
8
             \hookrightarrow mbodv)) \rightarrow
          ExtendEnvMod (mname, add_module_definition curr_en mbody
9
              → ,curr_en) ) en
10
       ms
```

add_module_definition evaluates each definition in the body of the module Simple Modules

The Interpreter

The Type-Checker

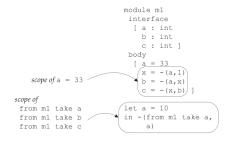
The Type-Checker

Make sure that

- each module body satisfies its interface,
- each variable is used consistently with its type.

The scoping rules:

- Modules follow let* scoping, putting into scope qualified variables for each of the bindings exported by the module.
- ▶ Declarations and definitions both follow let* scoping as well



Type-Checking

- 1. Obtain type of interface of module body
- 2. Obtain interface of module
- 3. Compare them

Let us revisit some examples from above

```
module m1
interface
    [u : bool]
body
    [u = 33]
```

- ► Type error!
- The body of the module must associate each name in the interface with a value of the appropriate type, even if those values are not used elsewhere in the program.

```
module m1
interface
    [u : int
    v : int]
body
    [u = 33]
4
```

- Type error!
- ► The module body must supply bindings for all the declarations in the interface.

```
module m1
interface
    [u : int]
body
[v = 2
    u = 33-v]
4
```

- ► Ok!
- ▶ v is private to module body

- ► Type error!
- ▶ v is private to module body

Example 4

```
module m1
interface
    [u : int
    v : int]
body
    [v = 33
    u = 44]
from m1 take u
```

- ► Type error!
- ➤ To keep the implementation simple, our language requires that the module body produce the values in the same order as the interface.

Example 6

```
module m2
     interface
2
        [v : int]
3
     body
5
        [v = (from m1 take u) -11]
6
   module m1
7
     interface
8
        [u : int]
9
     body
10
        [u = 44]
11
    (from m1 take u)-(from m2 take v)
12
```

- ▶ Type error!
- ▶ from m1 take u is not in scope where it is used in the body of m2.

Typing System for SIMPLE-MODULES

- We are going to introduce typing judgements
- ► Then typing rules
- Finally, we are going to implement a type-checker by using the typing rules as specification

Summary of Typing Judgements

▶ Judgements for typing expressions

Judgements for typing programs

▶ Judgements for typing list of module declarations

Typing Judgements for Expressions

Before

Now

- tenv is the standard type environment from before
- m_tenv is a module type environment and is required for typing the expression from m take x

Typing Programs in Expressions

▶ Module Type

$$m[u_1:t_1,\ldots,u_n:t_n]$$

► Module type environment (m_tenv)

$$m_1[u_{1,1}:t_1,\ldots,u_{1,n_1}:t_{n_1}]\ldots m_k[u_{k,1}:t_1,\ldots,u_{k,n_k}:t_{n_k}]$$

- ▶ If $m[u_1:t_1,\ldots,u_n:t_n] \in \mathtt{m_tenv}$, then
 - ▶ We say $m \in Dom(m_tenv)$
 - ightharpoonup m_tenv $(m,u_i)=t_i$

Typing System for Expressions

$$\frac{\textit{m} \in \textit{Dom}(\texttt{m_tenv}) \quad \texttt{m_tenv}(\textit{m},\textit{x}) = \texttt{t}}{\texttt{m_tenv}; \texttt{tenv} \vdash \texttt{from} \ \texttt{m} \ \texttt{take} \ \texttt{x} :: \texttt{t}} \ \textit{TFromTake}$$

Typing Judgements for Programs

▶ M is the list of declared modules

Typing Judgements for Module Declarations

$$m_{tenv_1} \vdash M :: m_{tenv_2}$$

- m_tenv2 is the type of the list of modules M
- ▶ m_tenv₁ is the type of the list of modules that M can use
- ▶ list1 < list2 means that list1 is a sublist of list2
- ▶ $[x_i]_{i \in I} \triangleleft [y_j]_{j \in J}$ determines an injective, order preserving function $f: I \rightarrow J$

$$(\texttt{m_tenv}; [y_1 := s_1] \dots [y_{j-1} := s_{j-1}] \texttt{tenv} \vdash e_j :: s_j)_{j \in J}$$

$$(t_i = s_{f(i)})_{i \in I}$$

$$\underline{m[x_i : t_i]_{i \in I} \text{ m_tenv} \vdash \texttt{M} :: \texttt{m_tenv}}$$

$$\underline{m[x_i : t_i]_{i \in I} [y_j = e_j]_{j \in J}}$$

$$\underline{m[x_i : t_i]_{i \in I} [y_j = e_j]_{j \in J}}$$

$$\underline{m[x_i : t_i]_{i \in I} [y_j = e_j]_{j \in J}}$$

$$\underline{m[x_i : t_i]_{i \in I} \text{ m_tenv}}$$

Summary of Typing Judgements

▶ Judgements for typing expressions

Judgements for typing programs

▶ Judgements for typing list of module declarations

Implementing the Type-Checker

- ▶ We will use the typing rules as a guideline
- Rather than having to deal with two different type environments (m_tenv and env below)

$$\frac{m \in Dom(\texttt{m_tenv}) \quad \texttt{m_tenv}(m,x) = t}{\texttt{m_tenv;tenv} \vdash \texttt{from m take } x :: t} TFromTake$$

we will have just one environment where we can lookup variables and modules

Type Environments

```
type tenv =

EmptyTEnv

ExtendTEnv of string*texpr*tenv

ExtendTEnvMod of string*tenv*tenv
```

Type-Checking Qualified Variables

from m take var

```
\frac{m \in Dom(\texttt{m\_tenv}) \quad \texttt{m\_tenv}(m, var) = t}{\texttt{m\_tenv}; \texttt{env} \vdash \texttt{from m take var} :: t} TFrom Take
```

Type-Checking Qualified Variables

from m take var

- first lookup m in the type environment,
- ▶ then lookup up the type of var in the resulting interface.

```
\frac{m \in Dom(\texttt{m\_tenv}) \quad \texttt{m\_tenv}(m, var) = t}{\texttt{m\_tenv}; \texttt{env} \vdash \texttt{from m take var} :: t} TFromTake
```

Type-Checking Modules

```
module m1
interface
    [u : int
    z: int]
body
    [u = 2
    v = 33-u
    z = 7]

4
```

Must compare actual and expected types of each module:

<:-decls is the operation in charge of this

```
(\texttt{m\_tenv}; [y_1 := s_1] \dots [y_{j-1} := s_{j-1}] \texttt{tenv} \vdash e_j :: s_j)_{j \in J} 
(t_i = s_{f(i)})_{i \in I} 
\underline{m[x_i : t_i]_{i \in I}} \text{ m\_tenv} \vdash \texttt{M} :: \texttt{m\_tenv} 
\underline{m[x_i : t_i]_{i \in I}[y_j = e_j]_{j \in J}} \text{ M} \text{ } :: m[x_i : t_i]_{i \in I} \text{ m\_tenv} 
\underline{m[x_i : t_i]_{i \in I}[y_j = e_j]_{j \in J}} \text{ M} \text{ } :: m[x_i : t_i]_{i \in I} \text{ m\_tenv} 
\underline{m[x_i : t_i]_{i \in I}[y_j = e_j]_{j \in J}} \text{ M} \text{ } :: m[x_i : t_i]_{i \in I} \text{ m\_tenv}
```

Type-Checking Modules

```
⊢M::m_tenv m_tenv;empty-tenv⊢e::t

⊢M e::t
```

```
⊢M::m_tenv m_tenv;empty-tenv⊢e::t

⊢M e::t
```

```
⊢M::m_tenv m_tenv;empty-tenv⊢e::t

⊢M e::t
```

```
add_module_type_definitions tenv:(mdecl list->tenv)

    function

     | [] -> tenv
     | AMod(mname, ASimpleInterface(expected_iface), mbody)::ms
         → ->
       let i_body = type_of_module_body tenv mbody
5
       in if (is_subtype i_body expected_iface)
      then add_module_type_definitions
7
                 (ExtendTEnvMod(mname, var_decls_to_tenv
                     → expected_iface,tenv))
8
                ms
       else raise (Subtype_failure(mname))
9
10
   and
     type_of_module_body en (AModBody vdefs):tenv =
11
12
     reverse_tenv @@ snd @@ List.fold_left (fun (env,renv) (
         → AValDef(var,decl)) ->
         let ty = type_of_expr env decl
13
         in (ExtendTEnv(var, ty,env),ExtendTEnv(var, ty,renv)))
14
                 (en, EmptyTEnv) vdefs
```

Subtype Checking

```
module m1
interface
    [u : int
    z: int]
body
    [u = 2
    v = 33-u
    z = 7]

4
```

Must compare actual and expected types of each module:

```
let rec is_subtype (actual:tenv) (expected: vdecl list) =
match actual, expected with
| _,[] -> true
| EmptyTEnv,_::_ -> false
| ExtendTEnv(ida,tya,tenv),(AValDecl(ide,tye))::ys ->
if ida=ide
then tya=tye && is_subtype tenv ys
else is_subtype tenv ((AValDecl(ide,tye))::ys)
| _,_ -> failwith "Case not reachable"
```

The Interpreter for SIMPLE-MODULES

- Code available in Canvas Modules/Interpreters
- Directory simple-modules
- Compile with ocamlbuild -use-menhir checker.ml
- ► Make sure the .ocamlinit file is in the folder of your sources
- Run utop

Beyond SIMPLE-MODULES

- Opaque types: modules that declare types (not just values)
- Module procedures (known as functors in OCaml): modules that take a module as argument and produce another module as result