# Extending REC with References CS496

# A New Language

- ▶ We are going to extend REC with references
- ► As a consequence we obtain a functional language that supports variables as in imperative programming
- Note that all previous programming features shall still be available, including
  - let expressions
  - procedures
  - recursion

## Implicit vs Explicit References

There are essentially two ways of doing this:

- 1. Treat every variable as a mutable reference
  - ► The resulting language is called IMPLICIT-REFS
  - ► References are implicit
- 2. Add mutable references to the non-mutable ones
  - ► The resulting language is called EXPLICIT-REFS
  - References are explicit

We are going to study both

## Implicit vs Explicit References

#### IMPLICIT-REFS

#### EXPLICIT-REFS

#### Implicit References

Concrete and Abstract Syntax

The Store
The Interpreter

**Explicit References** 

Concrete and Abstract Syntax

## Concrete Syntax

Two new productions that are added to those of  $\operatorname{REC}$ 

```
\langle \textit{Expression} \rangle ::= set \langle \textit{Identifier} \rangle = \langle \textit{Expression} \rangle
\langle \textit{Expression} \rangle ::= begin \langle \textit{Expression} \rangle^{+(;)} end
```

- ▶ The set expression is assignment
- ▶ A begin ... end expression evaluates its subexpressions in order and returns the value of the last one.

# Concrete Syntax – An Example

```
let g = let count = 0
    in proc(d)
    begin set count = -(count,-1);
    count
end
in -((g 11), (g 22))
```

# Another Example

```
1 \mid \text{let.} \quad \mathbf{x} = 0
  in letrec even(dummy)
                  = if zero?(x)
3
                    then 1
4
                    else begin
5
                            set x = -(x,1);
6
                            (odd 888)
                           end
8
               odd (dummy)
9
                  = if zero?(x)
10
                    then 0
11
                    else begin
12
                            set x = -(x,1);
13
                            (even 888)
14
15
                           end
      in begin set x = 13;
16
                  (odd -888)
17
          end
18
```

# Example (cont.)

- ► This program uses global state (the variable x) to communicate between even and odd
  - ► This is not recommended practice, it only serves the purpose of illustrating the use of references
- ► This program uses multideclaration letrec (exercise 3.32) and a begin expression (exercise 4.4).
- ► The dummy argument is necessary only because all our procedures have exactly one argument, but exercise 3.21 can fix that.

## IMPLICIT-REFS: Abstract Syntax

```
(define-datatype expression expression?
    (const-exp
      (num number?))
    (diff-exp
4
      (exp1 expression?)
5
      (exp2 expression?))
6
    (zero?-exp
7
      (exp1 expression?))
8
    (if-exp
9
      (exp1 expression?)
10
      (exp2 expression?)
      (exp3 expression?))
    (var-exp
13
      (var symbol?))
14
    (let-exp
15
      (var symbol?)
16
      (exp1 expression?)
17
      (body expression?))
18
```

#### The new variants for references

```
(proc-exp
       (var symbol?)
2
3
       (body expression?))
    (call-exp
4
       (rator expression?)
5
       (rand expression?))
6
    (letrec-exp
7
       (p-name symbol?)
8
       (b-var symbol?)
9
       (p-body expression?)
10
       (letrec-body expression?))
    (begin-exp
12
       (exp (listof expression?)))
13
    (assign-exp
14
       (id symbol?)
15
       (exp expression?)))
16
```

#### The new variants for references

```
(proc-exp
       (var symbol?)
2
3
       (body expression?))
    (call-exp
4
       (rator expression?)
5
       (rand expression?))
6
    (letrec-exp
7
       (p-name symbol?)
8
       (b-var symbol?)
9
       (p-body expression?)
10
       (letrec-body expression?))
    (begin-exp
12
       (exp (listof expression?)))
13
    (assign-exp
14
       (id symbol?)
15
       (exp expression?)))
16
```

## Example - Abstract Syntax

```
(a-program
(let-exp
'g
(let-exp
'count
(const-exp 0)
(proc-exp 'd (begin-exp (assign-exp 'count
    (diff-exp (var-exp 'count))))
(diff-exp (call-exp (var-exp 'g) (const-exp 11))
    (call-exp (var-exp 'g) (const-exp 22)))))
```

#### Implicit References

Concrete and Abstract Syntax

The Store

The Interpreter

#### **Explicit References**

Concrete and Abstract Syntax

# Motivating the Store

set! 
$$\times$$
 (+  $\times$  1)

- As we know, variables now have two readings:
  - ► An address (the blue occurrence of x)
  - ► A value (the red occurence of x)
- ▶ Environments therefore have to change their type:
  - ▶ Before: Vars → ExpVal
  - $ightharpoonup Now: Vars \longrightarrow Refs$
- Refs is a set of references or locations
- References point to expressed values
  - ► Hence we typically write *Refs*(*ExpVal*) (rather than just *Refs*)

# Motivating the Store

set! 
$$\times$$
 (+  $\times$  1)

- ► Environments therefore have to change their type:
  - ▶ Before: Vars → ExpVal
  - ▶ Now:  $Vars \longrightarrow Refs(ExpVal)$
- Revisiting the two readings of a variable
  - ▶ Blue x: we just look it up in the environment
  - ► Red x: we look it up in the environment and then use that location to get its value in a store or memory
- Our interpreter thus shall need a store

## **Environment and Store**

#### Environment



#### Store

(num-val 3)	(bool-val #t)	(num-val 7)
0	1	2

# Summing Up Our Analysis

```
ExpVal = Int + Bool + Proc

DenVal = Ref(ExpVal)
```

- Recall: Denoted values = Values assigned to variables in the environment
- References exist only as the bindings of variables.
- Two options for the type of the interpreter
  - Store is passed as argument

```
value-of::{ expression, environment, store } -> expval
```

2. Store is held in global variable (we choose this one)

```
value-of::{ expression, environment } -> expval
```

- ► There are many ways to represent the store (eg. Lists, Vectors, User-defined datatypes, etc.)
  - We choose lists since it is the simplest
  - ▶ The references will just be numbers (indices of the list)
  - An example

(num-val 3)	(bool-val #t)	(num-val 7)	(num-val 28)
0	1	2	3

- Sample operations we should support
  - Create a store
  - Lookup the value associated to a reference
  - Ask the store for a fresh reference and assign it a value (allocation)

- ► As mentioned, the store will be held in a global variable
- We name that variable the-store

```
;; the-store: a Scheme variable containing the
;; current state of the store.
3 (define the-store 'uninitialized)
```

- Initially set to a dummy value ('uninitialized)
- ▶ This definition and the code that follows is in the file store.scm

```
;; empty-store : () -> Sto
 ;; Page: 111
  (define empty-store
    (lambda () '()))
5
  ;; initialize-store! : () -> Sto
  ;; usage: (initialize-store!) sets the-store to the
     empty-store
8 ;; Page 111
9 (define initialize-store!
    (lambda ()
      (set! the-store (empty-store))))
11
```

▶ Allocation of a new association (page 111)

```
;; newref : ExpVal -> Ref
 (define newref
   (lambda (val)
3
      (let ((next-ref (length the-store)))
4
5
        (set! the-store
6
              (append the-store (list val)))
        (when (instrument-newref)
7
            (eopl:printf "newref: allocating location
8
     "s with initial contents "s" next-ref val))
       next-ref)))
9
```

(num-val 3)	(bool-val #t)	(num-val 7)
0	1	2

#### After newref (num-val 28):

(num-val 3)	(bool-val #t)	(num-val 7)	(num-val 28)
0	1	2	3

► Accessing the store (page 111)

```
;; deref :: Ref -> ExpVal
(define deref
(lambda (ref)
(list-ref the-store ref)))
```

► (list-ref aList anIndex) simply accesses the anIndex-th element in the list aList

Updating value of existing reference (page 112)

```
1 ;; setref! :: {Ref, ExpVal} -> Unit
  (define setref!
   (lambda (ref val)
3
    (set! the-store
4
      (letrec
5
        ((setref-inner
6
            ;; returns a list like store1, except that
7
      position ref1 contains val.
            (lambda (store1 ref1)
8
              (cond
9
                ((null? store1)
      (report-invalid-reference ref the-store))
                ((zero? ref1) (cons val (cdr store1)))
11
                (else
12
                  (cons (car store1)
13
                         (setref-inner
14
                            (cdr store1) (- ref1 1))))))))
15
        (setref-inner the-store ref)))))
16
```

► The function report-invalid-reference is used for error

(num-val 3)	(bool-val #t)	(num-val 7)	(num-val 28)
0	1	2	3

#### After setref! 1 (num-val 42):

(num-val 3)	(num-val 42)	(num-val 7)	(num-val 28)
0	1	2	3

#### Assessment

- ► This representation is extremely inefficient
- Ordinary memory operations require approximately constant time, but in our representation these operations require time proportional to the size of the store.
- ► No real implementation would ever do this, of course, but it suffices for our purposes.

#### Implicit References

Concrete and Abstract Syntax

The Interpreter

#### **Explicit References**

Concrete and Abstract Syntax

# Specification of Behavior of the Interpreter

- ▶ We now specify how the interpreter value-of behaves
- ► Input:
  - expression
  - environment
  - store
- Output:
  - expressed value
  - updated store

```
value-of::{ expression, environment, store } -> expval * store
```

▶ Note: As mentioned, in our implementation the store won't be passed as a parameter, it will be held in a global variable

## Specification of Behavior of Interpreter on Constants

 First we revisit the interpreter's behavior for the simplest cases.

```
(value-of (const-exp n) \rho \sigma) = (num-val n, \sigma)
```

- (var-exp var) is the expression to evaluate
- $\triangleright \rho$  is an environment
- $\triangleright \sigma$  is the store
- The specification of value-of denotes a pair (result, updatedStore)
- ▶ In this case the store  $\sigma$  is returned unaltered

# Specification of Behavior of Interpreter on Difference

```
(value-of (diff-exp exp1 exp2) \rho \sigma_0) =

2 let (val1,\sigma_1) = (value-of exp1 \rho \sigma_0)

3 (val2,\sigma_2) = (value-of exp2 \rho \sigma_1)

4 in (num-val ((expval->num val1) - (expval->num val2)), \sigma_2)
```

# Specification of Behavior of Interpreter on Variables

- 1. look up the identifier in the environment to find the location to which it is bound
- 2. look up in the store to find the value at that location

```
(value-of (var-exp var) \rho \sigma) = (\sigma(\rho(var)), \sigma)
```

# Specification of Behavior of Interpreter w.r.t. Assignment

```
(value-of (assign-exp var exp1) \rho \sigma_0) = let (val1,\sigma_1) = (value-of exp1 \rho \sigma_0) in (num-val 27, [\rho(\text{var}) = \text{val1}]\sigma_1)
```

- ▶ 27 is a dummy value; an assignment is evaluated to cause an effect (the value is irrelevant)
- $[\rho(var) = val1]\sigma_1$  stands for "update the reference  $\rho(var)$  in the store  $\sigma_1$  with the new value val1

## Implementation

```
1 (value-of (var-exp var) 
ho \sigma) = (\sigma(
ho(var)),\sigma)
```

#### **Implementation**

```
1 (value-of (assign-exp var exp1) \rho \sigma_0) = 2 let (val1, \sigma_1) = (value-of exp1 \rho \sigma_0) 3 in (num-val 27, [\rho(var) = val1]\sigma_1)
```

```
(define value-of
    (lambda (exp env)
2
      (cases expression exp
          (var-exp (var) ...)
4
          (assign-exp (var exp1)
5
            (begin
               (setref!
                 (apply-env env var)
8
                 (value-of exp1 env))
9
               (num-val 27)))
 )))
```

## Updating the Implementation of Extant REC Features

- ► The implementation of features that were already present in REC also have to be updated
- ▶ We've already seen the cases for constants and difference
- Here is the specification of the behavior of the interpreter for let

```
(value-of (let-exp var exp1 body) \rho \sigma_0) = let (val1,\sigma_1) = (value-of exp1 \rho \sigma_0) in (value-of body [var = I]\rho [I = val1]\sigma_1)
```

/ denotes a fresh store location

## Updating the Implementation of Extant REC Features

► The implementation for let

```
(define value-of
    (lambda (exp env)
      (cases expression exp
3
          (var-exp (var) ...)
4
          (assign-exp (var exp1) ...)
          (let-exp (var exp1 body)
6
              (let ((val1 (value-of exp1 env)))
7
                (value-of body
8
                   (extend-env var (newref val1) env))))
9
 )))
```

# Specification of Behavior of Interpreter w.r.t. Procedures

```
1 (value-of (proc-exp var body) \rho \sigma) = 2 ((proc-val (closure var body \rho)), \sigma)
```

▶ No changes w.r.t. REC here

# Specification of Behavior of Interpreter w.r.t. Procedure Calls

```
(value-of (call-exp rator rand) \rho \sigma_0)

= let (proc,\sigma_1) = (value-of rator \rho \sigma_0)

(arg,\sigma_2) = (value-of rand \rho \sigma_1))

((apply-procedure (expval->proc proc) arg), \sigma_2)
```

▶ No changes w.r.t. REC here but must update apply-procedure

# Specifying the Behavior of Procedure Application

```
(apply-procedure (procedure var body \rho) val \sigma) = (value-of body [var = I]\rho [I = val]\sigma)
```

► Here / denotes a fresh store location

#### Implementing Procedure Application

```
;; Proc * ExpVal --> ExpVal
(define apply-procedure
(lambda (proc1 val)
(cases proc proc1
(procedure (var body saved-env)
(value-of body
(extend-env var (newref val) saved-env))))))
```

# **Environment Lookup**

- Last we have to update environment lookup
- ► This operation was called apply-env
- Only the case for letrec has to be updated
- Before doing so we recall its definition

#### apply-env as Implemented in REC

```
(define apply-env
    (lambda (env search-var)
2
      (cases environment env
3
        (empty-env ()
4
          (report-no-binding-found search-var))
5
        (extend-env (saved-var saved-val saved-env)
6
          (if (eqv? saved-var search-var)
             saved-val
8
             (apply-env saved-env search-var)))
9
        (extend-env-rec (p-name
                          b-var
                          p-body
                           saved-env)
          (if (eqv? search-var p-name)
14
             (proc-val (procedure b-var p-body env))
15
             (apply-env saved-env search-var))))))
16
```

#### apply-env as Implemented in REC

```
(define apply-env
    (lambda (env search-var)
2
      (cases environment env
3
         (empty-env ()
4
           (report-no-binding-found search-var))
5
         (extend-env (saved-var saved-val saved-env)
6
           (if (eqv? saved-var search-var)
             saved-val
8
             (apply-env saved-env search-var)))
9
         (extend-env-rec (p-name
                           b-var
                           p-body
                           saved-env)
13
           (if (eqv? search-var p-name)
14
             (proc-val (procedure b-var p-body env))
15
             (apply-env saved-env search-var))))))
16
```

#### Updating apply-env

- Before:
  - ▶ apply-env would return a closure
- ► Now:
  - apply-env should return a reference to a location in the store containing the appropriate closure
- ▶ The reason: how the interpreter manages variable lookup

```
1 (value-of (var-exp var) \rho \sigma) = (\sigma(\rho(var)), \sigma)
```

 $\rho(var)$  is environment lookup of var in  $\rho$ 

# Updating apply-env

```
(extend-env-rec (p-names b-vars p-bodies saved-env)
    (let ((n (location search-var p-names)))
      (if n
3
        (newref
4
           (proc-val
5
             (procedure
6
               (list-ref b-vars n)
               (list-ref p-bodies n)
8
               env)))
9
        (apply-env saved-env search-var))))
10
```

- Manages multideclaration letrec.
- ▶ location takes a variable and a list of variables and returns either the position of the variable in the list, or #f if it is not present.
- ► This completes the implementation of IMPLICIT-REFS.

# The Interpreter for IMPLICIT-REFS

- Code available from http://www.eopl3.com
- ▶ Directory chapter4/implicit-refs
- Open top.scm in Racket
- ▶ There are a number of tests in tests.scm
- You can run them with run-one. Eg.

```
[(run-one 'gensym-test)
2 (num-val -1)
```

#### Implicit References

Concrete and Abstract Syntax
The Store

The Interpreter

#### **Explicit References**

Concrete and Abstract Syntax

# EXPLICIT-REFS: A language with explicit references

- 1. We now define a new language EXPLICIT-REFS, which adds references as expressed values to our language.
- 2. Concrete and Abstract Syntax
- 3. Specification
- 4. Implementation

#### Implicit vs Explicit References

#### IMPLICIT-REFS

#### EXPLICIT-REFS

## Implicit vs Explicit Store

- ▶ In the implicit store design, every variable is mutable.
  - Allocation, dereferencing and mutation are built into the language.
- ► The explicit reference design gives a clear account of allocation, dereferencing, and mutation
  - ▶ All these operations are explicit in the programmer's code.

#### Expressed and Denoted Values

#### Before (IMPLICIT-REFS)

```
ExpVal = Int + Bool + Proc

DenVal = Ref(ExpVal)
```

#### Now (EXPLICIT-REFS):

```
ExpVal = Int + Bool + Proc + Ref(ExpVal)

DenVal = ExpVal
```

- Before: programs could only produce numbers, booleans or closures as a result
- Now: programs can also produce references as a result (or even store them)

#### Expressed Values for EXPLICIT-REFS

#### Expressed values before

#### Now

#### **Environment and Store**

#### Environment

(num-val 3)	(num-val 3)	(ref-val 0)
х	у	z

#### Store

(num-val 7)	(ref-val 9)	(bool-val #t)
0	1	2

# Yet Another Example – References to References

```
let x = newref(newref(0))
in begin
setref(deref(x), 11);
deref(deref(x))
end
```

- Allocates a new reference containing 0.
- ▶ Then binds x to a reference containing the above reference.
- ► The value of deref(x) is a reference to the first reference.
- ▶ So when it evaluates the setref, it is the first reference that is modified, and the entire program returns 11.

#### **EXPLICIT-REFS:** Concrete Syntax

```
 \begin{array}{lll} \langle \textit{Expression} \rangle & ::= & \langle \textit{Number} \rangle \\ \langle \textit{Expression} \rangle & ::= & -(\langle \textit{Expression} \rangle \,,\, \langle \textit{Expression} \rangle) \\ \langle \textit{Expression} \rangle & ::= & \textit{zero?} \, (\langle \textit{Expression} \rangle) \\ \langle \textit{Expression} \rangle & ::= & \textit{if} \, \langle \textit{Expression} \rangle \\ & & & & \text{then} \, \langle \textit{Expression} \rangle \, & \text{else} \, \langle \textit{Expression} \rangle \\ \langle \textit{Expression} \rangle & ::= & \langle \textit{Identifier} \rangle \\ \langle \textit{Expression} \rangle & ::= & \text{let} \, \langle \textit{Identifier} \rangle \, & \langle \textit{Expression} \rangle \, & \text{in} \, \langle \textit{Expression} \rangle \\ \langle \textit{Expression} \rangle & ::= & (\langle \textit{Expression} \rangle \, \langle \textit{Expression} \rangle) \\ \langle \textit{Expression} \rangle & ::= & \text{letrec} \, \langle \textit{Identifier} \rangle (\langle \textit{Identifier} \rangle) \, & \langle \textit{Expression} \rangle \\ & & & \text{in} \, \langle \textit{Expression} \rangle \\ & & & & \text{in} \, \langle \textit{Expression} \rangle \end{array}
```

#### **EXPLICIT-REFS:** Concrete Syntax

We add three new operations to create and use references.

- newref: allocates a new location and returns a reference to it.
- deref: dereferences a reference: that is, it returns the contents of the location that the reference represents.
- setref: changes the contents of the location that the reference represents.

#### EXPLICIT-REFS: Abstract Syntax

```
(define-datatype expression expression?
    (const-exp
      (num number?))
    (diff-exp
4
      (exp1 expression?)
5
      (exp2 expression?))
6
    (zero?-exp
7
      (exp1 expression?))
8
    (if-exp
9
      (exp1 expression?)
10
      (exp2 expression?)
      (exp3 expression?))
    (var-exp
13
      (var identifier?))
14
    (let-exp
15
      (var identifier?)
16
      (exp1 expression?)
17
      (body expression?))
18
```

#### The new variants for references

```
(proc-exp
      (var identifier?)
      (body expression?))
    (call-exp
4
      (rator expression?)
5
      (rand expression?))
6
    (letrec-exp
      (p-name identifier?)
8
      (b-var identifier?)
9
      (p-body expression?)
10
      (letrec-body expression?))
    (newref-exp
12
13
      (exp1 expression?))
    (deref-exp
14
      (exp1 expression?))
15
    (setref-exp
16
      (ref expression?)
      (exp expression?))
18
```

#### The new variants for references

```
(proc-exp
      (var identifier?)
      (body expression?))
    (call-exp
4
      (rator expression?)
5
      (rand expression?))
6
    (letrec-exp
      (p-name identifier?)
8
      (b-var identifier?)
9
      (p-body expression?)
10
      (letrec-body expression?))
    (newref-exp
      (exp1 expression?))
13
    (deref-exp
14
      (exp1 expression?))
15
    (setref-exp
16
      (ref expression?)
      (exp expression?))
18
```

#### Example

```
let x = newref(newref(0))
in begin
setref(deref(x), 11);
deref(deref(x))
end
```

```
(a-program
(let-exp
'x
(newref-exp (newref-exp (const-exp 0)))
(begin-exp
(setref-exp (deref-exp (var-exp 'x)) (const-exp
11))
(list (deref-exp (deref-exp (var-exp 'x))))))
```

#### Specification - newref

```
(value-of (newref-exp exp) 
ho \sigma_0) = 
2 let (val,\sigma_1) = (value-of exp 
ho \sigma_0)
3 in (1, [l=val]\sigma_1)
```

- ▶ *I* is fresh, that is  $I \notin dom(\sigma_1)$
- ref-val is the tag that indicates that the result is a reference

#### Implementation - newref

```
1 (value-of (newref-exp exp) \rho \sigma_0) = 2 let (val,\sigma_1) = (value-of exp \rho \sigma_0) 3 in (1, [1=val]\sigma_1)
```

## Specification - deref and setref

```
(value-of (deref-exp exp) \rho \sigma_0) = 
2 let (1,\sigma_1) = (value-of exp \rho \sigma_0)
3 in (\sigma_1(1),\sigma_1)
```

```
1 (value-of (setref-exp exp1 exp2) \rho \sigma_0) =
2 let (1,\sigma_1) = (value-of exp1 \rho \sigma_0)
3 (val,\sigma_2) = (value-of exp2 \rho \sigma_1)
4 in (23,[l=val]\sigma_2)
```

Because in our language every expression returns a value, setref returns a dummy value: 23.

#### Implementation - deref

```
1 (value-of (deref-exp exp) \rho \sigma_0) =
2 let (1,\sigma_1) = (value-of exp \rho \sigma_0)
3 in (\sigma_1(1), \sigma_1)
```

```
(define value-of
(lambda (exp env)
(cases expression exp
(var-exp (var) ...)
(newref-exp (exp1) ...)
(deref-exp (exp1)
(let ((v1 (value-of exp1 env)))
(let ((ref1 (expval->ref v1)))
(deref ref1))))
```

#### Implementation - setref

```
(value-of (setref-exp exp1 exp2) \rho \sigma_0) =

let (1,\sigma_1) = (value-of exp1 \rho \sigma_0)

(val,\sigma_2) = (value-of exp2 \rho \sigma_1)

in (23,[1=val]\sigma_2)
```

```
(define value-of
    (lambda (exp env)
      (cases expression exp
        (var-exp (var) ...)
4
        (newref-exp (exp1) ...)
5
        (deref-exp (exp1) ...)
6
        (setref-exp (exp1 exp2)
          (let ((ref (expval->ref (value-of exp1 env))))
8
             (let ((val2 (value-of exp2 env)))
               (begin
10
                 (setref! ref val2)
                 (num-val 23)))))
12
```

## The Interpreter for EXPLICIT-REFS

- Code available from http://www.eopl3.com
- Directory chapter4/explicit-refs
- Open top.scm in Racket
- ▶ There are a number of tests in tests.scm
- You can run them with run-one. Eg.

```
1 (run-one 'gensym-test-1)
2 (num-val -1)
```