Programming Abstractions

Week 8-2: MiniScheme D and E and Lexical vs. Dynamic Bindings

Stephen Checkoway

What can MiniScheme do at this point?

MiniScheme C has numbers

MiniScheme C has pre-defined variables

MiniScheme C has procedure calls to built-in procedures

MiniScheme D: Conditionals

Booleans in MiniScheme

In Scheme: `#t` and `#f`

In MiniScheme: `True` and `False`

You'll need to add symbols `True` and `False` to `init-env`

- Bind them to `'True` and `'False`

New special form: if

EXP →	number	parse into `lit-exp`
|	symbol	parse into `var-exp`
|	(if *EXP EXP EXP*)	parse into `ite-exp`
|	(*EXP EXP**)	parse into `app-exp`

We need a new data type for the if-then-else expression

- `ite-exp`
- `ite-exp?`
- `ite-exp-cond`
- `ite-exp-then`
- `ite-exp-else`

The parser

MiniScheme D

```
(define (parse input)
  (cond [(number? input) (lit-exp input)]
        [(symbol? input) (var-exp input)]
        [(list? input)
         (cond [(empty? input) (error ...)]
               [(eq? (first input) 'if)
                (if (= (length input) 4)
                    (ite-exp ...)
                    (error ...))]
               [else (app-exp ...)])]
        [else (error 'parse "Invalid syntax ~s" input)]))
```

Parsing if-then-else expressions

If-then-else expressions are recursive

- E.g., *EXP* → `( if` *EXP EXP EXP* `)`

When parsing an if-then-else expression, you want to parse the sub expressions using `parse`

The input to `parse` will look like `'(if (lt? x 1) (+ y 100) z)`

The condition is `(second input)`

The then-branch is `(third input)`

The else-branch is `(fourth input)`

Evaluating `ite-exp`

Parse tree is recursive: `(parse '(if x 10 20))`

- `'(ite-exp (var-exp x) (lit-exp 10) (lit-exp 20))`

When evaluating, you should call `eval-exp` recursively

- First, call it on the conditional expression
  - If the condition is `False`, call it on the last expression
  - Otherwise, call it on the middle expression

Can you evaluate all parts of the ite-exp?

What would happen if you instead called eval-exp on all three parts of the expression before deciding which one to return?

Think about recursive procedures using if to decide if it is the base case or a recursive case

Primitive procedures returning booleans

Numeric procedures

- `number?`
- `eqv?` — like Scheme's `eqv?` so that it works with `True` and `False`
- `lt?` — like Scheme's <
- `gt?` — like Scheme's >
- `lte?` — like Scheme's <=
- `gte?` — like Scheme's >=

List procedures

- `null?`
- `list?`

MiniScheme E: let expressions

Let expressions

Consider

```
(let ([x (+ 3 4)]
      [y 5]
      [z (foo 8)])
  body)
```

To evaluate this, we need to extend the current environment with bindings for `x`, `y`, and `z` and then evaluate `body` in the extended environment

Extending environments

`(env list-of-symbols list-of-values previous-environment)`

Recall that the `env` constructor requires

- a list of symbols
- a list of values
- a previous environment

The parser doesn't know anything about environments but we can create a `let-exp` data type that stores

- the binding symbols
- the parsed binding values
- the parsed body

Parsing let expressions

```
(let ([x (+ 3 4)] [y 5] [z (foo 8)])
  body)
```

The binding list is `(second input)` where `input` is the whole let expression

The symbols are `(map first binding-list)`

The binding expressions are `(map second binding-list)`

How can we parse each of these expressions?

The body is simply `(third input)` which we can `parse`

Evaluating let expressions

Evaluating a let expressions just takes a little more work

- Evaluate each of the binding expressions in the `let-exp`

```
(map (λ (exp)
       (eval-exp exp current-env))
     (let-exp-exps tree))
```

- Bind the symbols to these values by extending the current environment
- Evaluate the body of the let expression using the extended environment

Lexical Binding

Variable usage

There are two ways a variable can be used in a program:

- As a declaration
- As a "reference" or use of the variable

Scheme has two kinds of variable "declarations"

- the bindings of a let-expression and
- the parameters of a lambda-expression

Scope of a declaration

The scope of a declaration is the portion of the expression or program to which that declaration applies

Lexical binding

- Scope of a variable is determined by textual layout of the program
- C, Java, Scheme/Racket use lexical binding

Dynamic binding

- Scope of a variable is determined by most recent *runtime* declaration
- Bash and classic Lisp use dynamic binding

Java example

What is the scope of y in this Java program?

Could we print y instead of x in the last line?

```
public static void main(String[] args) {
    int x;
    x = 1;
    while (x < 10) {
        int y = x;
        System.out.println(y);
        x += 1;
    }
    System.out.println(x);
}
```

Scope in Scheme

Scope of variables bound (declared) in a `let` is the body of the `let`
Scope of parameters in a λ is the body of the λ

```
(let ([x 5]
      [y 10])
  (* ((λ (z) (+ z y)) 7)
     x
     y))
```

Shadowing bindings

Shadowing: Declaring a new variable with the same name as an existing variable in the same scope

```
(let ([x 5]
      [y 10])
  (* ((λ (x) (+ x y)) 7)
     x
     y))
```

We say that the inner binding for x *shadows* the outer binding for x

Determining the appropriate binding

Start at the use of a variable

Search the enclosing regions starting with the innermost and working outward looking for a binding (declaration) of the variable

The first binding you find is the appropriate binding

If there are no such bindings, we say the variable is *free*

Contour diagrams

Draw the boundaries of the regions in which variable bindings are in effect

```
(λ (x)
  (λ (y)
    ((λ (x) (x y)) x)))
```

The body of a let or a lambda expression determines a contour

Each variable refers to the innermost declaration *outside* its contour

Lexical depth

The lexical depth of a variable reference is 1 less than the number of contours crossed between the reference and the declaration it refers to

```
(λ (x)
  (λ (y)
    ((λ (x) (x y)) x)))
```

In `(x y)`

- `x` has lexical depth 0
- `y` has lexical depth 1

The other `x` has lexical depth 1

Lexical addresses

(depth, position)

We can use the lexical depth of a variable along with the 0-based position of the variable in its declaration to come up with a *lexical address* of the variable

```
(let ([x 3]
      [y 4])
  (λ (a b)
    (λ (c)
      (a (+ (b x) c)))))
```

(1,0) (1,1) (2,0) (0,0)

Lexical addresses are essentially pointers to where the variable can be found on the run-time stack; can eliminate names

Dynamic binding vs. lexical binding

What is the value of y in the body of `(f 2)`

```
(let ([y 3])
  (let ([f (λ (x) (+ x y))])
    (let ([y 17])
      (f 2))))
```

With lexical (also called static) binding: `y` is 3

- The value of `y` comes from the closest lexical binding of `y`, namely `[y 3]`

With dynamic binding: `y` is 17

- The value of y comes from the most-recent *run-time* binding of `y`, namely `[y 17]`

Lambdas in a lexically-scoped language

A lambda expression evaluates to a closure which is a triple containing

- the environment at the time the lambda is evaluated
- the parameters
- the body of the lambda

When we apply the closure to argument expressions

- we evaluate the arguments in the current environment
- extend the **closure's** environment with bindings of parameters to argument values
- evaluate the closure's body in the new environment

Lexical binding example

```
(let ([y 3])
  (let ([f (λ (x) (+ x y))])
    (let ([y 17])
      (f 2))))
```

Lexical binding example

```
(let ([y 3])
  (let ([f (λ (x) (+ x y))])
    (let ([y 17])
      (f 2))))
```

Variable	Value
y	3

Lexical binding example

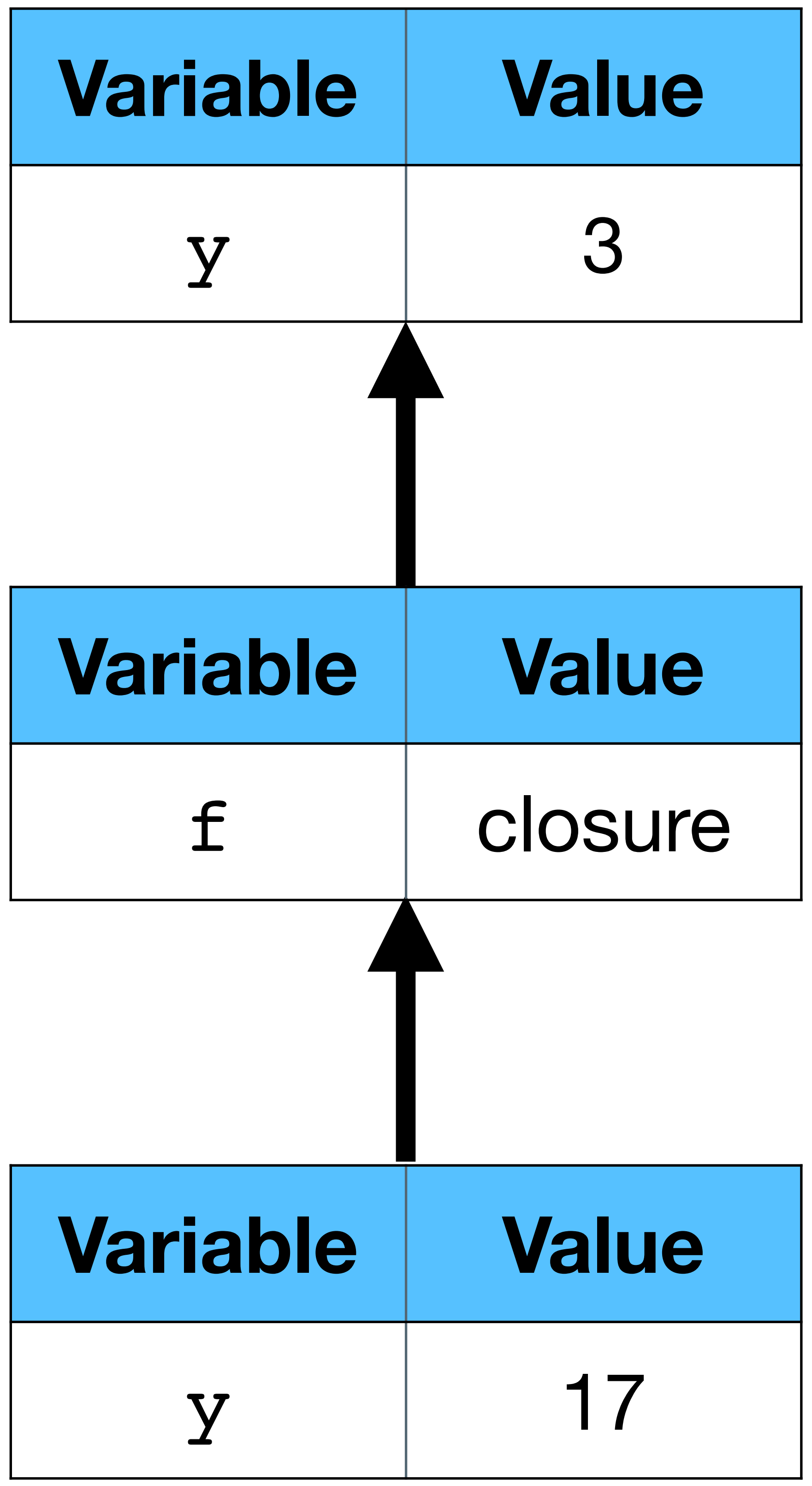
```
(let ([y 3])
  (let ([f (λ (x) (+ x y))])
    (let ([y 17])
      (f 2))))
```

Variable	Value
y	3

Variable	Value
f	closure

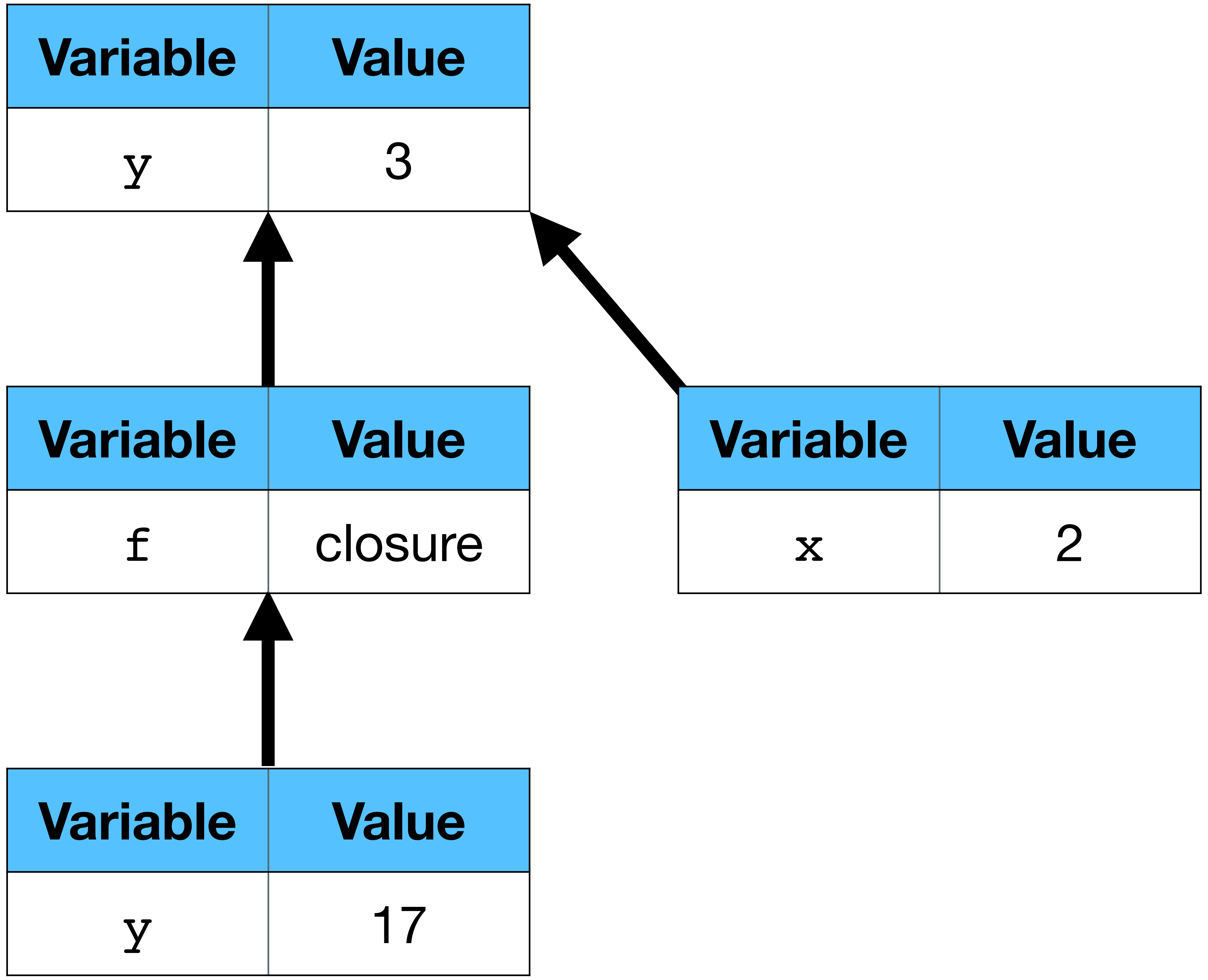
Lexical binding example

```
(let ([y 3])
  (let ([f (λ (x) (+ x y))])
    (let ([y 17])
      (f 2))))
```



Lexical binding example

```
(let ([y 3])
  (let ([f (λ (x) (+ x y))])
    (let ([y 17])
      (f 2))))
```



Lambdas in a dynamically-scoped language

A lambda expression evaluates to a procedure which is just a pair containing

- the parameters
- the body of the lambda

When we apply the procedure to argument expressions

- we evaluate the arguments in the current environment
- extend the **current** environment with bindings of parameters to argument values
- evaluate the lambda's body in the new environment

Dynamic binding example

```
(let ([y 3])
  (let ([f (λ (x) (+ x y))])
    (let ([y 17])
      (f 2))))
```

Dynamic binding example

Variable	Value
y	3

```
(let ([y 3])
  (let ([f (λ (x) (+ x y))])
    (let ([y 17])
      (f 2))))
```

Dynamic binding example

```
(let ([y 3])
  (let ([f (λ (x) (+ x y))])
    (let ([y 17])
      (f 2))))
```

Variable	Value
y	3

Variable	Value
f	procedure

Dynamic binding example

```
(let ([y 3])
  (let ([f (λ (x) (+ x y))])
    (let ([y 17])
      (f 2))))
```

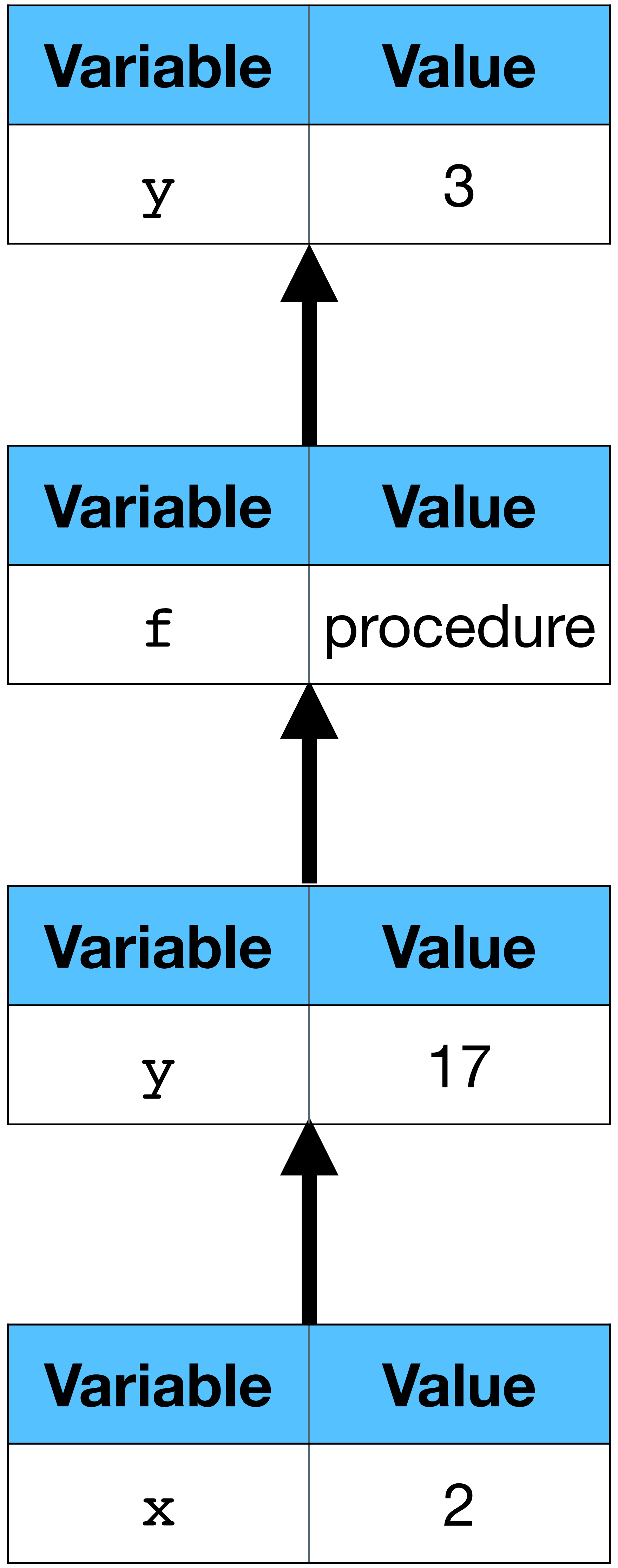
Variable	Value
y	3

Variable	Value
f	procedure

Variable	Value
y	17

Dynamic binding example

```
(let ([y 3])
  (let ([f (λ (x) (+ x y))])
    (let ([y 17])
      (f 2))))
```



Why was dynamic binding ever used?

It's easy to implement

- Dynamic binding was understood several years before static binding

It made sense to some people that `(λ (x) (+ x y))` should use whatever the latest version of y is

Why do we now use lexical binding?

Most languages are derived from Algol-60 which used lexical binding

Compilers can use lexical addresses known at compile time for all variable references

Code from lexically-bound languages is easier to verify

- E.g., in Racket, we can ensure a variable is declared before it is used *before* we run the program
- It makes more sense to most people

Python example

```python
def fun(x):
    return lambda y: x + y

def main():
    f = fun(10)
    print(f(7))      # Prints 17
    x = 20
    print(f(7))      # Prints 17

main()
```

Bash example

```bash
#!/bin/bash

x=0

setx() {
  x=$1
}

printx() {
  echo "${x}"
}

main() {
  printx  # prints 0
  setx 10
  printx  # prints 10
  local x=25
  printx  # prints 25!
  setx 100
  printx  # prints 100!
}

main
printx    # prints 10
```