Environments, Binding, and Scope

CS 350

Dr. Joseph Eremondi

Last updated: July 15, 2024

Overview

• Functional Programming (Plait Tutorial)

- Functional Programming (Plait Tutorial)
- Parsing (PLAI ch2)

- Functional Programming (Plait Tutorial)
- Parsing (PLAI ch2)
- Interpreters (PLAI ch3)

- Functional Programming (Plait Tutorial)
- Parsing (PLAI ch2)
- Interpreters (PLAI ch3)
- Desugaring (PLAI ch4)

- Functional Programming (Plait Tutorial)
- Parsing (PLAI ch2)
- Interpreters (PLAI ch3)
- Desugaring (PLAI ch4)
- Functions (PLAI ch5)

• Today: Environments in Curly (PLAI ch6)

- Today: Environments in Curly (PLAI ch6)
- Tues: Lambda and First-class Functions in Plait

- Today: Environments in Curly (PLAI ch6)
- Tues: Lambda and First-class Functions in Plait
- Wed: Replacements for Recursion in Plait

- Today: Environments in Curly (PLAI ch6)
- Tues: Lambda and First-class Functions in Plait
- Wed: Replacements for Recursion in Plait
- Thurs: Implementing Lambda in Curly (PLAI ch7)

- Today: Environments in Curly (PLAI ch6)
- Tues: Lambda and First-class Functions in Plait
- Wed: Replacements for Recursion in Plait
- Thurs: Implementing Lambda in Curly (PLAI ch7)
- Mon: Closures and Environments in Curly (PLAI ch7)

- Today: Environments in Curly (PLAI ch6)
- Tues: Lambda and First-class Functions in Plait
- Wed: Replacements for Recursion in Plait
- Thurs: Implementing Lambda in Curly (PLAI ch7)
- Mon: Closures and Environments in Curly (PLAI ch7)
- Tues/Wed: lectures (not included on midterm)

- Today: Environments in Curly (PLAI ch6)
- Tues: Lambda and First-class Functions in Plait
- Wed: Replacements for Recursion in Plait
- Thurs: Implementing Lambda in Curly (PLAI ch7)
- Mon: Closures and Environments in Curly (PLAI ch7)
- Tues/Wed: lectures (not included on midterm)
- Thurs: MIDTERM

- Today: Environments in Curly (PLAI ch6)
- Tues: Lambda and First-class Functions in Plait
- Wed: Replacements for Recursion in Plait
- Thurs: Implementing Lambda in Curly (PLAI ch7)
- Mon: Closures and Environments in Curly (PLAI ch7)
- Tues/Wed: lectures (not included on midterm)
- Thurs: MIDTERM

Everything up to and including Closures may appear on the midterm

Environments

• Evaluate their argument

- Evaluate their argument
- Lookup the function defn (variable, body)

- Evaluate their argument
- Lookup the function defn (variable, body)
- Replace the parameter variable with the argument value in the body

- Evaluate their argument
- Lookup the function defn (variable, body)
- Replace the parameter variable with the argument value in the body
- Evaluate the result

- Evaluate their argument
- Lookup the function defn (variable, body)
- Replace the parameter variable with the argument value in the body
- Evaluate the result
- If we ever interpret a variable, raise an error

• Each substitution is $\mathcal{O}(n)$ where n is the number of nodes in the function body AST

- Each substitution is $\mathcal{O}(n)$ where n is the number of nodes in the function body AST
- This is in addition to the cost of actually evaluating the function

- Each substitution is $\mathcal{O}(n)$ where n is the number of nodes in the function body AST
- This is in addition to the cost of actually evaluating the function
 - Very slow!

- Each substitution is $\mathcal{O}(n)$ where n is the number of nodes in the function body AST
- This is in addition to the cost of actually evaluating the function
 - Very slow!
- Want a way to have $\mathcal{O}(1)$ function calls

- Each substitution is $\mathcal{O}(n)$ where n is the number of nodes in the function body AST
- This is in addition to the cost of actually evaluating the function
 - o Very slow!
- Want a way to have $\mathcal{O}(1)$ function calls
 - Not including the time to evaluate the function body

- Each substitution is $\mathcal{O}(n)$ where n is the number of nodes in the function body AST
- This is in addition to the cost of actually evaluating the function
 - o Very slow!
- Want a way to have $\mathcal{O}(1)$ function calls
 - o Not including the time to evaluate the function body
- Substitution is forgetful

- Each substitution is $\mathcal{O}(n)$ where n is the number of nodes in the function body AST
- This is in addition to the cost of actually evaluating the function
 - o Very slow!
- Want a way to have $\mathcal{O}(1)$ function calls
 - Not including the time to evaluate the function body
- · Substitution is forgetful
 - o Just replaces function variable with expression

- Each substitution is $\mathcal{O}(n)$ where n is the number of nodes in the function body AST
- This is in addition to the cost of actually evaluating the function
 - o Very slow!
- Want a way to have $\mathcal{O}(1)$ function calls
 - Not including the time to evaluate the function body
- · Substitution is forgetful
 - Just replaces function variable with expression
 - o Not very useful for debugging

• Data structure for deferred substitution

- Data structure for deferred substitution
 - o List of variable/value pairs

- Data structure for deferred substitution
 - o List of variable/value pairs
- Intuition:

- Data structure for deferred substitution
 - o List of variable/value pairs
- Intuition:
 - Instead of replacing all x with value v, keep a list of replacements you need to do

- Data structure for deferred substitution
 - o List of variable/value pairs
- Intuition:
 - Instead of replacing all x with value v, keep a list of replacements you need to do
 - When you interpret x, check the environment before raising an error

- Data structure for deferred substitution
 - o List of variable/value pairs
- Intuition:
 - Instead of replacing all x with value v, keep a list of replacements you need to do
 - When you interpret x, check the environment before raising an error
 - o If there's an entry for x in the environment, return that

- Data structure for deferred substitution
 - List of variable/value pairs
- Intuition:
 - Instead of replacing all x with value v, keep a list of replacements you need to do
 - When you interpret x, check the environment before raising an error
 - o If there's an entry for x in the environment, return that
 - · Error otherwise

The Solution: Environments

- Data structure for deferred substitution
 - List of variable/value pairs
- Intuition:
 - Instead of replacing all x with value v, keep a list of replacements you need to do
 - When you interpret x, check the environment before raising an error
 - o If there's an entry for x in the environment, return that
 - · Error otherwise
 - Means reference to undefined variable

```
;; Just a pair, but we get better names than fst and snd
(define-type Binding
   (bind [name : Symbol]
        [val : Number]))
```

```
;; Just a pair, but we get better names than fst and snd
(define-type Binding
  (bind [name : Symbol]
        [val : Number]))
```

```
;; Get helper functions from the type-def
(bind-name (bind 'x 3))
(bind-val (bind 'x 3))
```

```
;; Just a pair, but we get better names than fst and snd
(define-type Binding
   (bind [name : Symbol]
        [val : Number]))
```

```
;; Get helper functions from the type-def
(bind-name (bind 'x 3))
(bind-val (bind 'x 3))
```

```
'x
3
```

The Environment Data Structure: Environments

The Environment Data Structure: Environments

```
;; Lets us write Env instead of (Listof Binding)
;; But it's not defining a new type,
;; just a new name for the same type.
(define-type-alias Env (Listof Binding))
;; Environment is either empty or extended env
(define emptyEnv : Env
 empty)
(define (extendEnv [bnd : Binding]
                   [env : Envl)
        : Fnv
 (cons bnd env))
emptyEnv
(extendEnv (bind 'x 3) (extendEnv (bind 'y 4) empty))
```

The Environment Data Structure: Environments

```
;; Lets us write Env instead of (Listof Binding)
;; But it's not defining a new type,
;; just a new name for the same type.
(define-type-alias Env (Listof Binding))
;; Environment is either empty or extended env
(define emptyEnv : Env
  empty)
(define (extendEnv [bnd : Binding]
                   [env : Envl)
        : Fnv
  (cons bnd env))
emptyEnv
(extendEnv (bind 'x 3) (extendEnv (bind 'y 4) empty))
```

```
'()
(list (bind 'x 3) (bind 'y 4))
```

• Find the **first** binding in the environment

- Find the **first** binding in the environment
 - $\circ\;$ This is important for shadowing

- Find the **first** binding in the environment
 - This is important for shadowing
- Just a linear search, like we've seen lots already

- Find the **first** binding in the environment
 - This is important for shadowing
- Just a linear search, like we've seen lots already

- Find the **first** binding in the environment
 - o This is important for shadowing
- Just a linear search, like we've seen lots already

```
(define (lookup [n : Symbol] [env : Env]) : Number
  (type-case (Listof Binding) env
   ;; Can't find a variable in an empty env
   [empty (error 'lookup "undefined variable")]
   ;; Cons: check if the first binding is the var
   ;; we're looking for.
   ;; Return its value if it is, otherwise
   ;; keep looking in the rest of the list
   [(cons b rst-env) (cond
                       [(symbol=? n (bind-name b))
                        (bind-val b)]
                       [else (lookup n rst-env)])))
```

• We can change the implementation without changing the surface language

- We can change the implementation without changing the surface language
- Programs should run the exact same in both interpreters

- We can change the implementation without changing the surface language
- Programs should run the exact same in both interpreters
- Strategy: add an extra context argument for Environment

- We can change the implementation without changing the surface language
- Programs should run the exact same in both interpreters
- Strategy: add an extra context argument for Environment
 - o Unlike fundefs, this will change across recursive calls

- We can change the implementation without changing the surface language
- Programs should run the exact same in both interpreters
- Strategy: add an extra context argument for Environment
 - o Unlike fundefs, this will change across recursive calls

- We can change the implementation without changing the surface language
- Programs should run the exact same in both interpreters
- Strategy: add an extra context argument for Environment
 - o Unlike fundefs, this will change across recursive calls

 Exactly like before, except we have to pass the environment in the recursive call

- Exactly like before, except we have to pass the environment in the recursive call
- Other operations are similar

- Exactly like before, except we have to pass the environment in the recursive call
- Other operations are similar

- Exactly like before, except we have to pass the environment in the recursive call
- Other operations are similar

 Can't return an error, because we might have added a deferred substitution to the environment

- Can't return an error, because we might have added a deferred substitution to the environment
- So we look in the environment and see if there's a value bound to x

- Can't return an error, because we might have added a deferred substitution to the environment
- So we look in the environment and see if there's a value bound to x
- If there is return it

- Can't return an error, because we might have added a deferred substitution to the environment
- So we look in the environment and see if there's a value bound to x
- If there is return it
 - o Otherwise, variable not found error

- Can't return an error, because we might have added a deferred substitution to the environment
- So we look in the environment and see if there's a value bound to x
- If there is return it
 - o Otherwise, variable not found error

- Can't return an error, because we might have added a deferred substitution to the environment
- So we look in the environment and see if there's a value bound to x
- If there is return it
 - Otherwise, variable not found error

```
[(Var x)
(lookup x env)]
```

• Just like before, we get the function body + variable, and value for argument

- Just like before, we get the function body + variable, and value for argument
- Still interpret body

- Just like before, we get the function body + variable, and value for argument
- Still interpret body
- Instead of replacing in body, put a variable-value pair in the environment

- Just like before, we get the function body + variable, and value for argument
- Still interpret body
- Instead of replacing in body, put a variable-value pair in the environment
 - $\circ \;$ Called $\emph{binding}$ the variable to its value

- Just like before, we get the function body + variable, and value for argument
- Still interpret body
- Instead of replacing in body, put a variable-value pair in the environment
 - $\circ \;$ Called $\emph{binding}$ the variable to its value

- Just like before, we get the function body + variable, and value for argument
- Still interpret body
- Instead of replacing in body, put a variable-value pair in the environment
 - o Called binding the variable to its value

Static Scoping

• We evaluated the body of the function in the empty environment, plus a value for its variable

- We evaluated the body of the function in the empty environment, plus a value for its variable
 - o Didn't extend the additional environment

- We evaluated the body of the function in the empty environment, plus a value for its variable
 - Didn't extend the additional environment
- Want functions to be abstractions

- We evaluated the body of the function in the empty environment, plus a value for its variable
 - Didn't extend the additional environment
- Want functions to be abstractions
 - Should be able to predict how a function behaves from how it's called

- We evaluated the body of the function in the empty environment, plus a value for its variable
 - Didn't extend the additional environment
- Want functions to be abstractions
 - Should be able to predict how a function behaves from how it's called
 - Don't want result to depend on context, just arguments

- We evaluated the body of the function in the empty environment, plus a value for its variable
 - o Didn't extend the additional environment
- Want functions to be abstractions
 - Should be able to predict how a function behaves from how it's called
 - Don't want result to depend on context, just arguments
- Gives the same results as substitution.

- We evaluated the body of the function in the empty environment, plus a value for its variable
 - Didn't extend the additional environment
- Want functions to be abstractions
 - Should be able to predict how a function behaves from how it's called
 - o Don't want result to depend on context, just arguments
- Gives the same results as substitution.
- We call this static scoping

- We evaluated the body of the function in the empty environment, plus a value for its variable
 - Didn't extend the additional environment
- Want functions to be abstractions
 - Should be able to predict how a function behaves from how it's called
 - o Don't want result to depend on context, just arguments
- Gives the same results as substitution.
- We call this static scoping
- If we extend the environment from the call site, we get dynamic scoping

 A language has static scoping if undefined variables in a term get their values from the environment where the function is defined

- A language has static scoping if undefined variables in a term get their values from the environment where the function is defined
 - Right now, variables come from top-level functions, so undefined variables are always an error

- A language has static scoping if undefined variables in a term get their values from the environment where the function is defined
 - Right now, variables come from top-level functions, so undefined variables are always an error
 - We'll see more of this later

- A language has static scoping if undefined variables in a term get their values from the environment where the function is defined
 - Right now, variables come from top-level functions, so undefined variables are always an error
 - We'll see more of this later
- A language has dynamic scoping if undefined variables get their values from the point where the function is called

```
{define {f x} {+ x y}}
{define {g y} {f y}}
{g 3}
```

• Static scoping says this is an error

```
{define {f x} {+ x y}}
{define {g y} {f y}}
{g 3}
```

- Static scoping says this is an error
 - \circ No value for y in body of f

```
{define {f x} {+ x y}}
{define {g y} {f y}}
{g 3}
```

- Static scoping says this is an error
 - o No value for y in body of f
- Dynamic scoping produces 6

```
{define {f x} {+ x y}}
{define {g y} {f y}}
{g 3}
```

- · Static scoping says this is an error
 - o No value for y in body of f
- Dynamic scoping produces 6
 - Looks up y from g when evaluating f

```
{define {f x} {+ x y}}
{define {g y} {f y}}
{g 3}
```

- · Static scoping says this is an error
 - o No value for y in body of f
- Dynamic scoping produces 6
 - Looks up y from g when evaluating f
- Dynamic scoping is WRONG

```
{define {f x} {+ x y}}
{define {g y} {f y}}
{g 3}
```

- · Static scoping says this is an error
 - o No value for y in body of f
- Dynamic scoping produces 6
 - o Looks up y from g when evaluating f
- Dynamic scoping is WRONG
 - You should understand it, but know that static scoping is what we want

Implementing Let

• New language: Curly-Let

- New language: Curly-Let
- Curly-Fundef, but with one new feature

- New language: Curly-Let
- Curly-Fundef, but with one new feature
 - ∘ {letvar x e1 e2}

- New language: Curly-Let
- Curly-Fundef, but with one new feature
 - o {letvar x e1 e2}
 - Gives x the value e1 in the expression e2

- New language: Curly-Let
- Curly-Fundef, but with one new feature
 - o {letvar x e1 e2}
 - Gives x the value e1 in the expression e2
 - · Called letvar so we don't confuse with plait

- New language: Curly-Let
- Curly-Fundef, but with one new feature
 - o {letvar x e1 e2}
 - Gives x the value e1 in the expression e2
 - · Called letvar so we don't confuse with plait
- We'll implement with both substitution and environments

Abstract Syntax

Abstract Syntax

Parsing and Desugaring are the same as usual

Abstract Syntax

- Parsing and Desugaring are the same as usual
 - ∘ See Curly-Let.rkt

• Want variable to have the given value in the body

- Want variable to have the given value in the body
 - $\circ\;$ So just substitute the value for the variable in the body

- Want variable to have the given value in the body
 - o So just substitute the value for the variable in the body
- Key detail: expression only evaluated once

- Want variable to have the given value in the body
 - o So just substitute the value for the variable in the body
- Key detail: expression only evaluated once
 - Can have exponential speedup in some algorithms

- Want variable to have the given value in the body
 - o So just substitute the value for the variable in the body
- Key detail: expression only evaluated once
 - Can have exponential speedup in some algorithms

- Want variable to have the given value in the body
 - o So just substitute the value for the variable in the body
- Key detail: expression only evaluated once
 - o Can have exponential speedup in some algorithms

Substituting in a Let Expression

• {letvar x e1 e2} binds x in e2

Substituting in a Let Expression

- {letvar x e1 e2} binds x in e2
- So when substituting in e2 we don't ever replace x

Substituting in a Let Expression

- {letvar x e1 e2} binds x in e2
- So when substituting in e2 we don't ever replace x
- Implements shadowing

Substituting in a Let Expression

- {letvar x e1 e2} binds x in e2
- So when substituting in e2 we don't ever replace x
- Implements shadowing

Substituting in a Let Expression

- {letvar x e1 e2} binds x in e2
- So when substituting in e2 we don't ever replace x
- · Implements shadowing

• Interpret the variable's value in the current environment

- Interpret the variable's value in the current environment
- Interpret the let body in the extended environment

- Interpret the variable's value in the current environment
- Interpret the let body in the extended environment
- When we hit x we'll look in the env

- Interpret the variable's value in the current environment
- Interpret the let body in the extended environment
- When we hit x we'll look in the env

- Interpret the variable's value in the current environment
- Interpret the let body in the extended environment
- When we hit x we'll look in the env

Notice that we only add to the env to interp body, not xexp

- Notice that we only add to the env to interp body, not xexp
- This is because x is in scope for body but not xexp

- Notice that we only add to the env to interp body, not xexp
- This is because x is in scope for body but not xexp
- We say the **scope of a variable** is the part of the program in which its value is either substituted or bound

- Notice that we only add to the env to interp body, not xexp
- This is because x is in scope for body but not xexp
- We say the **scope of a variable** is the part of the program in which its value is either substituted or bound
- Let extends scope by adding a variable, while calls transfer scope to the function

- Notice that we only add to the env to interp body, not xexp
- This is because x is in scope for body but not xexp
- We say the **scope of a variable** is the part of the program in which its value is either substituted or bound
- Let extends scope by adding a variable, while calls transfer scope to the function
- Later we'll see more complex examples of scope

- Notice that we only add to the env to interp body, not xexp
- This is because x is in scope for body but not xexp
- We say the **scope of a variable** is the part of the program in which its value is either substituted or bound
- Let extends scope by adding a variable, while calls transfer scope to the function
- Later we'll see more complex examples of scope
 - e.g. in Plait, let and let* have different rules for what's in scope

• Environments have a stack structure

- Environments have a stack structure
 - $\circ\,$ Push on new bindings when variables are defined

- Environments have a stack structure
 - Push on new bindings when variables are defined
 - Don't directly pop, but will sometimes interpret in the unextended environment

- Environments have a stack structure
 - o Push on new bindings when variables are defined
 - Don't directly pop, but will sometimes interpret in the unextended environment
- If we bind a value to a variable that's already in the environment, we say we shadow the old binding

- Environments have a stack structure
 - Push on new bindings when variables are defined
 - Don't directly pop, but will sometimes interpret in the unextended environment
- If we bind a value to a variable that's already in the environment, we say we shadow the old binding
 - lookup always takes the most recent binding

- Environments have a stack structure
 - o Push on new bindings when variables are defined
 - Don't directly pop, but will sometimes interpret in the unextended environment
- If we bind a value to a variable that's already in the environment, we say we shadow the old binding
 - lookup always takes the most recent binding
- Part of the call stack

- Environments have a stack structure
 - o Push on new bindings when variables are defined
 - Don't directly pop, but will sometimes interpret in the unextended environment
- If we bind a value to a variable that's already in the environment, we say we shadow the old binding
 - lookup always takes the most recent binding
- Part of the call stack
 - Theoretical in Curly

- Environments have a stack structure
 - o Push on new bindings when variables are defined
 - Don't directly pop, but will sometimes interpret in the unextended environment
- If we bind a value to a variable that's already in the environment, we say we shadow the old binding
 - lookup always takes the most recent binding
- Part of the call stack
 - Theoretical in Curly
 - Actually implemented for most languages

- Environments have a stack structure
 - o Push on new bindings when variables are defined
 - Don't directly pop, but will sometimes interpret in the unextended environment
- If we bind a value to a variable that's already in the environment, we say we shadow the old binding
 - lookup always takes the most recent binding
- Part of the call stack
 - Theoretical in Curly
 - o Actually implemented for most languages
 - Every time we call a function or define a variable we push onto the call stack

- Environments have a stack structure
 - o Push on new bindings when variables are defined
 - Don't directly pop, but will sometimes interpret in the unextended environment
- If we bind a value to a variable that's already in the environment, we say we shadow the old binding
 - lookup always takes the most recent binding
- Part of the call stack
 - Theoretical in Curly
 - o Actually implemented for most languages
 - Every time we call a function or define a variable we push onto the call stack
- Curly-Let and others have an implicit call stack

- Environments have a stack structure
 - o Push on new bindings when variables are defined
 - Don't directly pop, but will sometimes interpret in the unextended environment
- If we bind a value to a variable that's already in the environment, we say we shadow the old binding
 - lookup always takes the most recent binding
- Part of the call stack
 - Theoretical in Curly
 - o Actually implemented for most languages
 - Every time we call a function or define a variable we push onto the call stack
- Curly-Let and others have an implicit call stack
 - We don't keep the data structure ourselves

- Environments have a stack structure
 - o Push on new bindings when variables are defined
 - Don't directly pop, but will sometimes interpret in the unextended environment
- If we bind a value to a variable that's already in the environment, we say we shadow the old binding
 - lookup always takes the most recent binding
- Part of the call stack
 - Theoretical in Curly
 - o Actually implemented for most languages
 - Every time we call a function or define a variable we push onto the call stack
- Curly-Let and others have an implicit call stack
 - We don't keep the data structure ourselves
 - Instead, call interp recursively to add to the Plait call stack

- Environments have a stack structure
 - o Push on new bindings when variables are defined
 - Don't directly pop, but will sometimes interpret in the unextended environment
- If we bind a value to a variable that's already in the environment, we say we shadow the old binding
 - lookup always takes the most recent binding
- Part of the call stack
 - Theoretical in Curly
 - Actually implemented for most languages
 - Every time we call a function or define a variable we push onto the call stack
- Curly-Let and others have an implicit call stack
 - We don't keep the data structure ourselves
 - Instead, call interp recursively to add to the Plait call stack
 - When finished eval, plait returns is to part waiting for the result

• {letvar x 3 {letvar x 4 {+ 3 x}}}

- {letvar x 3 {letvar x 4 {+ 3 x}}}
 Looks at most recent definition
 - Looks at most recent definition

- {letvar x 3 {letvar x 4 {+ 3 x}}}
 - o Looks at most recent definition
 - o So should be 7

- {letvar x 3 {letvar x 4 {+ 3 x}}}
 - Looks at most recent definition
 - So should be 7
- Substitution: subst doesn't replace x in {+ 3 x}
 because it is bound

- {letvar x 3 {letvar x 4 {+ 3 x}}}
 - Looks at most recent definition
 - So should be 7
- Substitution: subst doesn't replace x in {+ 3 x}
 because it is bound
- Environments: (x,4) is at the top of the environment, so interp of x finds 4

• There are *high level design choices* for programming languages

- There are *high level design choices* for programming languages
 - How to deal with variable name collisions (shadowing)

- There are high level design choices for programming languages
 - How to deal with variable name collisions (shadowing)
 - How to deal with undefined variables (static vs. dynamic scope)

- There are high level design choices for programming languages
 - How to deal with variable name collisions (shadowing)
 - How to deal with undefined variables (static vs. dynamic scope)
- Decisions are made concrete in the implementation

- There are high level design choices for programming languages
 - How to deal with variable name collisions (shadowing)
 - How to deal with undefined variables (static vs. dynamic scope)
- Decisions are made concrete in the implementation
 - Behavior of subst on bound variables

- There are high level design choices for programming languages
 - How to deal with variable name collisions (shadowing)
 - How to deal with undefined variables (static vs. dynamic scope)
- Decisions are made concrete in the implementation
 - Behavior of subst on bound variables
 - What environment interp is passed for function bodies