

CS 4110

Programming Languages & Logics

Lecture 24

Compiling with Continuations



Continuations

We've seen continuations several times in this course already:

- As a way to implement break and continue
- As a way to make definitional translation more robust
- As an intermediate language in interpreters

Continuations

We've seen continuations several times in this course already:

- As a way to implement break and continue
- As a way to make definitional translation more robust
- As an intermediate language in interpreters

Now, we'll use them to translate a functional language down to an assembly-like language.

The translation works as a recipe for compiling any of the features we have discussed over the past few weeks all the way down to hardware.

Roadmap

CS 4120 in one lecture!

Roadmap

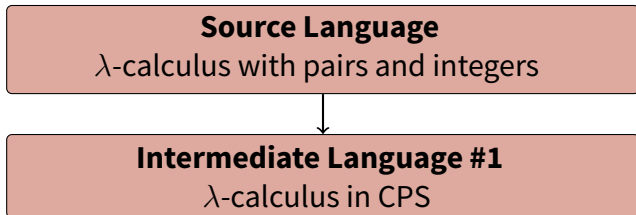
CS 4120 in one lecture!

Source Language

λ -calculus with pairs and integers

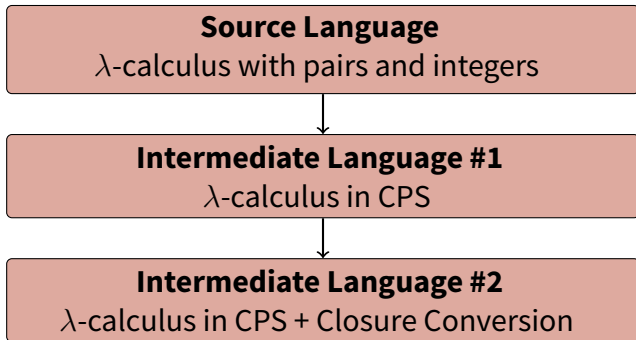
Roadmap

CS 4120 in one lecture!



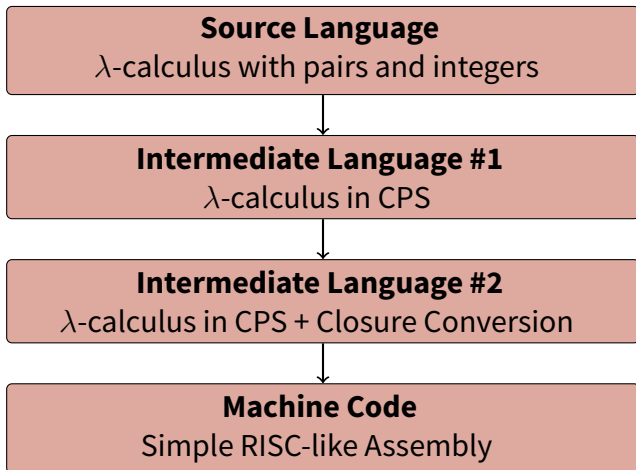
Roadmap

CS 4120 in one lecture!



Roadmap

CS 4120 in one lecture!



Source Language

We'll start from (untyped) λ -calculus with pairs and integers.

$$\begin{array}{lcl} e & ::= & x \\ & | & \lambda x. e \\ & | & e_1 e_2 \\ & | & (e_1, e_2) \\ & | & \#i e \\ & | & n \\ & | & e_1 + e_2 \end{array}$$

Target Language

$$p ::= bb_1; bb_2; \dots; bb_n$$

A program p consists of a series of *basic blocks* bb .

Target Language

$$\begin{aligned} p &::= bb_1; bb_2; \dots; bb_n \\ bb &::= lb : c_1; c_2; \dots; c_n; \text{jump } x \end{aligned}$$

A basic block has a label lb and a sequence of commands c , ending with “jump.”

Target Language

$$\begin{aligned} p &::= bb_1; bb_2; \dots; bb_n \\ bb &::= lb : c_1; c_2; \dots; c_n; \text{jump } x \\ c &::= \text{mov } x_1, x_2 \end{aligned}$$

Commands correspond to assembly language instructions and are largely self-evident.

Target Language

$$\begin{aligned} p &::= bb_1; bb_2; \dots; bb_n \\ bb &::= lb : c_1; c_2; \dots; c_n; \text{jump } x \\ c &::= \text{mov } x_1, x_2 \\ &\quad | \quad \text{mov } x, n \end{aligned}$$

Commands correspond to assembly language instructions and are largely self-evident.

Target Language

$$\begin{aligned} p &::= bb_1; bb_2; \dots; bb_n \\ bb &::= lb : c_1; c_2; \dots; c_n; \text{jump } x \\ c &::= \text{mov } x_1, x_2 \\ &\quad | \text{mov } x, n \\ &\quad | \text{mov } x, lb \end{aligned}$$

Commands correspond to assembly language instructions and are largely self-evident.

Target Language

$$\begin{aligned} p &::= bb_1; bb_2; \dots; bb_n \\ bb &::= lb : c_1; c_2; \dots; c_n; \text{jump } x \\ c &::= \text{mov } x_1, x_2 \\ &\quad | \text{mov } x, n \\ &\quad | \text{mov } x, lb \\ &\quad | \text{add } x_1, x_2, x_3 \end{aligned}$$

Commands correspond to assembly language instructions and are largely self-evident.

Target Language

$$\begin{aligned} p &::= bb_1; bb_2; \dots; bb_n \\ bb &::= lb : c_1; c_2; \dots; c_n; \text{jump } x \\ c &::= \text{mov } x_1, x_2 \\ &\quad | \text{mov } x, n \\ &\quad | \text{mov } x, lb \\ &\quad | \text{add } x_1, x_2, x_3 \\ &\quad | \text{load } x_1, x_2[n] \end{aligned}$$

Commands correspond to assembly language instructions and are largely self-evident.

Target Language

$$\begin{aligned} p &::= bb_1; bb_2; \dots; bb_n \\ bb &::= lb : c_1; c_2; \dots; c_n; \text{jump } x \\ c &::= \text{mov } x_1, x_2 \\ &\quad | \text{mov } x, n \\ &\quad | \text{mov } x, lb \\ &\quad | \text{add } x_1, x_2, x_3 \\ &\quad | \text{load } x_1, x_2[n] \\ &\quad | \text{store } x_1, x_2[n] \end{aligned}$$

Commands correspond to assembly language instructions and are largely self-evident.

Target Language

$$\begin{aligned} p &::= bb_1; bb_2; \dots; bb_n \\ bb &::= lb : c_1; c_2; \dots; c_n; \text{jump } x \\ c &::= \text{mov } x_1, x_2 \\ &\quad | \text{mov } x, n \\ &\quad | \text{mov } x, lb \\ &\quad | \text{add } x_1, x_2, x_3 \\ &\quad | \text{load } x_1, x_2[n] \\ &\quad | \text{store } x_1, x_2[n] \\ &\quad | \text{malloc } n \end{aligned}$$

The only un-RISC-y command is malloc. It allocates n words of space and places its address into a special register r_0 . Ignoring garbage, it can be implemented as simply as “add $r_0, r_0, -n$.”

Intermediate Language

$$\begin{array}{lcl} c & ::= & \text{let } x = e \text{ in } c \\ & | & v_1 \ v_2 \ v_3 \\ & | & v_1 \ v_2 \end{array}$$

Commands c look like basic blocks.

Intermediate Language

$$\begin{array}{lcl} c & ::= & \text{let } x = e \text{ in } c \\ & | & v_1 \ v_2 \ v_3 \\ & | & v_1 \ v_2 \\ e & ::= & v \mid v_1 + v_2 \mid (v_1, v_2) \mid (\#i \ v) \end{array}$$

There are no subexpressions in the language!

Intermediate Language

$$\begin{aligned} c &::= \text{let } x = e \text{ in } c \\ &\quad | \quad v_1 \ v_2 \ v_3 \\ &\quad | \quad v_1 \ v_2 \\ e &::= v \mid v_1 + v_2 \mid (v_1, v_2) \mid (\#i \ v) \\ v &::= n \mid x \mid \lambda x. \lambda k. c \mid \text{halt} \mid \underline{\lambda x. c} \end{aligned}$$

Abstractions encoding continuations are marked with an underline. These are called *administrative lambdas* and can be eliminated at compile time.

CPS Translation

The contract of the translation is that $\llbracket e \rrbracket k$ will evaluate e and pass its result to the continuation k .

To translate an entire program, we use $k = \text{halt}$, where halt is the continuation to send the result of the entire program to.

CPS Translation

$$\llbracket x \rrbracket k = kx$$

CPS Translation

$$\llbracket x \rrbracket k = kx$$

$$\llbracket n \rrbracket k = kn$$

CPS Translation

$$\llbracket x \rrbracket k = k x$$

$$\llbracket n \rrbracket k = k n$$

$$\llbracket (e_1 + e_2) \rrbracket k = \llbracket e_1 \rrbracket (\lambda x_1. \llbracket e_2 \rrbracket (\lambda x_2. \text{let } z = x_1 + x_2 \text{ in } k z))$$

CPS Translation

$$\llbracket x \rrbracket k = k x$$

$$\llbracket n \rrbracket k = k n$$

$$\llbracket (e_1 + e_2) \rrbracket k = \llbracket e_1 \rrbracket (\lambda x_1. \llbracket e_2 \rrbracket (\lambda x_2. \text{let } z = x_1 + x_2 \text{ in } k z))$$

$$\llbracket (e_1, e_2) \rrbracket k = \llbracket e_1 \rrbracket (\lambda x_1. \llbracket e_2 \rrbracket (\lambda x_2. \text{let } t = (x_1, x_2) \text{ in } k t))$$

CPS Translation

$$\llbracket x \rrbracket k = k x$$

$$\llbracket n \rrbracket k = k n$$

$$\llbracket (e_1 + e_2) \rrbracket k = \llbracket e_1 \rrbracket (\underline{\lambda} x_1. \llbracket e_2 \rrbracket (\underline{\lambda} x_2. \text{let } z = x_1 + x_2 \text{ in } k z))$$

$$\llbracket (e_1, e_2) \rrbracket k = \llbracket e_1 \rrbracket (\underline{\lambda} x_1. \llbracket e_2 \rrbracket (\underline{\lambda} x_2. \text{let } t = (x_1, x_2) \text{ in } k t))$$

$$\llbracket \#i e \rrbracket k = \llbracket e \rrbracket (\underline{\lambda} t. \text{let } y = \#i t \text{ in } k y)$$

CPS Translation

$$\llbracket x \rrbracket k = k x$$

$$\llbracket n \rrbracket k = k n$$

$$\llbracket (e_1 + e_2) \rrbracket k = \llbracket e_1 \rrbracket (\underline{\lambda} x_1. \llbracket e_2 \rrbracket (\underline{\lambda} x_2. \text{let } z = x_1 + x_2 \text{ in } k z))$$

$$\llbracket (e_1, e_2) \rrbracket k = \llbracket e_1 \rrbracket (\underline{\lambda} x_1. \llbracket e_2 \rrbracket (\underline{\lambda} x_2. \text{let } t = (x_1, x_2) \text{ in } k t))$$

$$\llbracket \#i e \rrbracket k = \llbracket e \rrbracket (\underline{\lambda} t. \text{let } y = \#i t \text{ in } k y)$$

$$\llbracket \lambda x. e \rrbracket k = k (\lambda x. \lambda k'. \llbracket e \rrbracket k')$$

CPS Translation

$$\llbracket x \rrbracket k = k x$$

$$\llbracket n \rrbracket k = k n$$

$$\llbracket (e_1 + e_2) \rrbracket k = \llbracket e_1 \rrbracket (\underline{\lambda} x_1. \llbracket e_2 \rrbracket (\underline{\lambda} x_2. \text{let } z = x_1 + x_2 \text{ in } k z))$$

$$\llbracket (e_1, e_2) \rrbracket k = \llbracket e_1 \rrbracket (\underline{\lambda} x_1. \llbracket e_2 \rrbracket (\underline{\lambda} x_2. \text{let } t = (x_1, x_2) \text{ in } k t))$$

$$\llbracket \#i e \rrbracket k = \llbracket e \rrbracket (\underline{\lambda} t. \text{let } y = \#i t \text{ in } k y)$$

$$\llbracket \lambda x. e \rrbracket k = k (\lambda x. \lambda k'. \llbracket e \rrbracket k')$$

$$\llbracket e_1 e_2 \rrbracket k = \llbracket e_1 \rrbracket (\underline{\lambda} f. \llbracket e_2 \rrbracket (\underline{\lambda} v. f v k))$$

Example

Let's translate the expression $\llbracket (\lambda a. \#1\ a) (3, 4) \rrbracket k$, using $k = \text{halt}$.

Example

Let's translate the expression $\llbracket (\lambda a. \#1\ a) (3, 4) \rrbracket k$, using $k = \text{halt}$.

$$\llbracket (\lambda a. \#1\ a) (3, 4) \rrbracket k$$

Example

Let's translate the expression $\llbracket (\lambda a. \#1\ a) (3, 4) \rrbracket k$, using $k = \text{halt}$.

$$\begin{aligned} & \llbracket (\lambda a. \#1\ a) (3, 4) \rrbracket k \\ = & \llbracket \lambda a. \#1\ a \rrbracket (\underline{\lambda} f. \llbracket (3, 4) \rrbracket (\underline{\lambda} v. f\ v\ k)) \end{aligned}$$

Example

Let's translate the expression $\llbracket (\lambda a. \#1\ a) (3, 4) \rrbracket k$, using $k = \text{halt}$.

$$\begin{aligned} & \llbracket (\lambda a. \#1\ a) (3, 4) \rrbracket k \\ = & \llbracket \lambda a. \#1\ a \rrbracket (\underline{\lambda} f. \llbracket (3, 4) \rrbracket (\underline{\lambda} v. f\ v\ k)) \\ = & (\underline{\lambda} f. \llbracket (3, 4) \rrbracket (\underline{\lambda} v. f\ v\ k)) (\lambda a. \lambda k'. \llbracket \#1\ a \rrbracket k') \end{aligned}$$

Example

Let's translate the expression $\llbracket (\lambda a. \#1\ a) (3, 4) \rrbracket k$, using $k = \text{halt}$.

$$\begin{aligned} & \llbracket (\lambda a. \#1\ a) (3, 4) \rrbracket k \\ = & \llbracket \lambda a. \#1\ a \rrbracket (\underline{\lambda} f. \llbracket (3, 4) \rrbracket (\underline{\lambda} v. f\ v\ k)) \\ = & (\underline{\lambda} f. \llbracket (3, 4) \rrbracket (\underline{\lambda} v. f\ v\ k)) (\lambda a. \lambda k'. \llbracket \#1\ a \rrbracket k') \\ = & (\underline{\lambda} f. \llbracket 3 \rrbracket (\underline{\lambda} x_1. \llbracket 4 \rrbracket (\underline{\lambda} x_2. \text{let } b = (x_1, x_2) \text{ in } (\underline{\lambda} v. f\ v\ k)\ b))) \\ & (\lambda a. \lambda k'. \llbracket \#1\ a \rrbracket k') \end{aligned}$$

Example

Let's translate the expression $\llbracket (\lambda a. \#1\ a) (3, 4) \rrbracket k$, using $k = \text{halt}$.

$$\begin{aligned} & \llbracket (\lambda a. \#1\ a) (3, 4) \rrbracket k \\ = & \llbracket \lambda a. \#1\ a \rrbracket (\underline{\lambda} f. \llbracket (3, 4) \rrbracket (\underline{\lambda} v. f\ v\ k)) \\ = & (\underline{\lambda} f. \llbracket (3, 4) \rrbracket (\underline{\lambda} v. f\ v\ k)) (\lambda a. \lambda k'. \llbracket \#1\ a \rrbracket k') \\ = & (\underline{\lambda} f. \llbracket 3 \rrbracket \left(\underline{\lambda} x_1. \llbracket 4 \rrbracket (\underline{\lambda} x_2. \text{let } b = (x_1, x_2) \text{ in } (\underline{\lambda} v. f\ v\ k)\ b) \right)) \\ & (\lambda a. \lambda k'. \llbracket \#1\ a \rrbracket k') \\ = & (\underline{\lambda} f. \left(\underline{\lambda} x_1. (\underline{\lambda} x_2. \text{let } b = (x_1, x_2) \text{ in } (\underline{\lambda} v. f\ v\ k)\ b)\ 4 \right) 3) \\ & (\lambda a. \lambda k'. \llbracket \#1\ a \rrbracket k') \end{aligned}$$

Example

Let's translate the expression $\llbracket (\lambda a. \#1 a) (3, 4) \rrbracket k$, using $k = \text{halt}$.

$$\begin{aligned} & \llbracket (\lambda a. \#1 a) (3, 4) \rrbracket k \\ = & \llbracket \lambda a. \#1 a \rrbracket (\underline{\lambda} f. \llbracket (3, 4) \rrbracket (\underline{\lambda} v. f v k)) \\ = & (\underline{\lambda} f. \llbracket (3, 4) \rrbracket (\underline{\lambda} v. f v k)) (\lambda a. \lambda k'. \llbracket \#1 a \rrbracket k') \\ = & (\underline{\lambda} f. \llbracket 3 \rrbracket (\underline{\lambda} x_1. \llbracket 4 \rrbracket (\underline{\lambda} x_2. \text{let } b = (x_1, x_2) \text{ in } (\underline{\lambda} v. f v k) b))) \\ & (\lambda a. \lambda k'. \llbracket \#1 a \rrbracket k') \\ = & (\underline{\lambda} f. (\underline{\lambda} x_1. (\underline{\lambda} x_2. \text{let } b = (x_1, x_2) \text{ in } (\underline{\lambda} v. f v k) b) 4) 3) \\ & (\lambda a. \lambda k'. \llbracket \#1 a \rrbracket k') \\ = & (\underline{\lambda} f. (\underline{\lambda} x_1. (\underline{\lambda} x_2. \text{let } b = (x_1, x_2) \text{ in } (\underline{\lambda} v. f v k) b) 4) 3) \\ & (\lambda a. \lambda k'. \llbracket a \rrbracket (\underline{\lambda} t. \text{let } y = \#1 t \text{ in } k' y))) \end{aligned}$$

Optimization

Clearly, the translation generates a lot of administrative λ s!

Optimization

Clearly, the translation generates a lot of administrative λ s!

To make the code more efficient and compact, we will optimize using some simple rewriting rules to eliminate administrative λ s

Optimization

Clearly, the translation generates a lot of administrative λ s!

To make the code more efficient and compact, we will optimize using some simple rewriting rules to eliminate administrative λ s

We can eliminate applications to variables by copy propagation:

$$(\underline{\lambda}x.e) y \rightarrow e\{y/x\}$$

Optimization

Clearly, the translation generates a lot of administrative λ s!

To make the code more efficient and compact, we will optimize using some simple rewriting rules to eliminate administrative λ s

We can eliminate applications to variables by copy propagation:

$$(\underline{\lambda}x.e) y \rightarrow e\{y/x\}$$

Other lambdas can be converted into lets:

$$(\underline{\lambda}x.c)v \rightarrow \text{let } x = v \text{ in } c$$

Optimization

Clearly, the translation generates a lot of administrative λ s!

To make the code more efficient and compact, we will optimize using some simple rewriting rules to eliminate administrative λ s

We can eliminate applications to variables by copy propagation:

$$(\underline{\lambda}x.e) y \rightarrow e\{y/x\}$$

Other lambdas can be converted into lets:

$$(\underline{\lambda}x.c)v \rightarrow \text{let } x = v \text{ in } c$$

We can also perform administrative η -reductions:

$$\underline{\lambda}x.k x \rightarrow k$$

Example, Redux

After applying these rewrite rules to the expression we had previously, we obtain:

```
let  $f = \lambda a. \lambda k'. \text{let } y = \#1 \ a \text{ in } k' \ y \text{ in}$   
let  $x_1 = 3 \text{ in}$   
let  $x_2 = 4 \text{ in}$   
let  $b = (x_1, x_2) \text{ in}$   
 $f \ b \ k$ 
```

This is starting to look a lot more like our target language!

Optimization

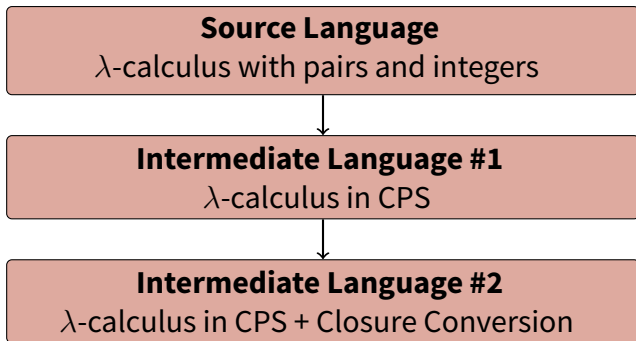
Writing these optimizations separately makes it easier to define the CPS conversion uniformly, without worrying about efficiency.

Optimization

Writing these optimizations separately makes it easier to define the CPS conversion uniformly, without worrying about efficiency.

We may not be able to remove all administrative lambdas. Any that cannot be eliminated using the rules above are converted into “real” lambdas.

Roadmap



Closure Conversion

The next step is to bring all λ s to the top level, with no nesting.

$$\begin{aligned} P &::= \text{let } x_f = \lambda x_1. \dots \lambda x_n. \lambda k. c \text{ in } P \\ &\quad | \text{let } x_c = \lambda x_1. \dots \lambda x_n. c \text{ in } P \\ &\quad | c \\ c &::= \text{let } x = e \text{ in } c \mid x_1 x_2 \dots x_n \\ e &::= n \mid x \mid \text{halt} \mid x_1 + x_2 \mid (x_1, x_2) \mid \#i x \end{aligned}$$

This translation requires the construction of *closures* that capture the free variables of the lambda abstractions and is known as *closure conversion*.

Closure Conversion

The main part of the translation is:

$$\begin{aligned} \llbracket \lambda x. \lambda k. c \rrbracket \sigma = & \\ \text{let } (c', \sigma') = \llbracket c \rrbracket \sigma \text{ in} & \\ \text{let } y_1, \dots, y_n = \text{fvs}(\lambda x. \lambda k. c') \text{ in} & \\ (f y_1 \dots y_n, \sigma' [f \mapsto \lambda y_1. \dots \lambda y_n. \lambda x. \lambda k. c']) & \text{ where } f \text{ fresh} \end{aligned}$$

Closure Conversion

The main part of the translation is:

$$\begin{aligned} \llbracket \lambda x. \lambda k. c \rrbracket \sigma = & \\ \text{let } (c', \sigma') = \llbracket c \rrbracket \sigma \text{ in} & \\ \text{let } y_1, \dots, y_n = \text{fvs}(\lambda x. \lambda k. c') \text{ in} & \\ (f y_1 \dots y_n, \sigma'[f \mapsto \lambda y_1. \dots \lambda y_n. \lambda x. \lambda k. c']) & \text{ where } f \text{ fresh} \end{aligned}$$

The translation of $\lambda x. \lambda k. c$ above first translates the body c , then creates a new function f parameterized on x as well as the free variables y_1 to y_n of the translated body.

Closure Conversion

The main part of the translation is:

$$\begin{aligned} \llbracket \lambda x. \lambda k. c \rrbracket \sigma = & \\ \text{let } (c', \sigma') = \llbracket c \rrbracket \sigma \text{ in} & \\ \text{let } y_1, \dots, y_n = \text{fvs}(\lambda x. \lambda k. c') \text{ in} & \\ (f y_1 \dots y_n, \sigma' [f \mapsto \lambda y_1. \dots \lambda y_n. \lambda x. \lambda k. c']) & \text{ where } f \text{ fresh} \end{aligned}$$

The translation of $\lambda x. \lambda k. c$ above first translates the body c , then creates a new function f parameterized on x as well as the free variables y_1 to y_n of the translated body.

It then adds f to the environment σ replaces the entire lambda with $(f y_n \dots y_n)$.

Closure Conversion

The main part of the translation is:

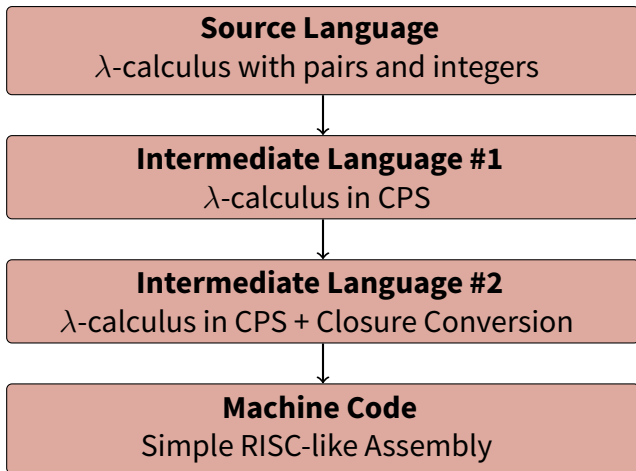
$$\begin{aligned} \llbracket \lambda x. \lambda k. c \rrbracket \sigma = & \\ \text{let } (c', \sigma') = \llbracket c \rrbracket \sigma \text{ in} & \\ \text{let } y_1, \dots, y_n = \text{fvs}(\lambda x. \lambda k. c') \text{ in} & \\ (f y_1 \dots y_n, \sigma' [f \mapsto \lambda y_1. \dots \lambda y_n. \lambda x. \lambda k. c']) & \text{ where } f \text{ fresh} \end{aligned}$$

The translation of $\lambda x. \lambda k. c$ above first translates the body c , then creates a new function f parameterized on x as well as the free variables y_1 to y_n of the translated body.

It then adds f to the environment σ replaces the entire lambda with $(f y_n \dots y_n)$.

When applied to an entire program, this has the effect of eliminating all nested λ s.

Roadmap



Code Generation

$$\mathcal{P}[[c]] = \text{main} : \mathcal{C}[[c]];$$

halt :

Code Generation

$$\mathcal{P}[\text{let } x_f = \lambda x_1. \dots \lambda x_n. \lambda k. c \text{ in } p] = x_f : \text{mov } x_1, a_1; \\ \vdots \\ \text{mov } x_n, a_n; \\ \text{mov } k, ra; \\ \mathcal{C}[c]; \\ \mathcal{P}[p]$$

Code Generation

$$\mathcal{P}[\text{let } x_c = \lambda x_1. \dots \lambda x_n. c \text{ in } p] = x_c : \text{mov } x_1, a_1;$$

\vdots

$\text{mov } x_n, a_n;$

$\mathcal{C}[c];$

$\mathcal{P}[p]$

Code Generation

$$\mathcal{C}[\text{let } x = n \text{ in } c] = \text{mov } x, n; \\ \mathcal{C}[c]$$

Code Generation

$$\mathcal{C}[\text{let } x_1 = x_2 \text{ in } c] = \text{mov } x_1, x_2; \\ \mathcal{C}[c]$$

Code Generation

$$\mathcal{C}[\text{let } x = x_1 + x_2 \text{ in } c] = \text{add } x_1, x_2, x; \\ \mathcal{C}[c]$$

Code Generation

$\mathcal{C}[\text{let } x = (x_1, x_2) \text{ in } c] =$ malloc 2;
mov x, r_0 ;
store $x_1, x[0]$;
store $x_2, x[1]$;
 $\mathcal{C}[c]$

Code Generation

$$\mathcal{C}[\text{let } x = \#i x_1 \text{ in } c] = \text{load } x, x_1[i - 1]; \\ \mathcal{C}[c]$$

Code Generation

$$\mathcal{C}[[x\ k\ x_1\ \dots\ x_n]] = \begin{array}{l} \text{mov } a_1, x_1; \\ \qquad \vdots \\ \text{mov } a_n, x_n; \\ \text{mov } ra, k; \\ \text{jump } x \end{array}$$

Final Thoughts

Note that we assume an infinite supply of registers. We would need to do register allocation and spill registers to a stack.

Also, while this translation is very simple, it is not particularly efficient. For example, we are doing a lot of register moves when calling functions and when starting the function body, which could be optimized.