

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science

1999 Final Examination and Solutions

1. Final Examination

There are *five* problems on this examination. *Make sure you don't skip over any of a problem's parts!* They are followed by an appendix that contains reference material from the course notes. The appendix contains no problems; it is just a handy reference.

You will have three hours in which to work the problems. Some problems are easier than others: read all problems before beginning to work, and use your time wisely!

This examination is open-book: you may use whatever reference books or papers you have brought to the exam. The number of points awarded for each problem is placed in brackets next to the problem number. There are 100 points total on the exam.

Do all written work in your examination booklet – we will not collect the examination handout itself; you will only be graded for what appears in your examination booklet. It will be to your advantage to show your work – we will award partial credit for incorrect solutions that make use of the right techniques.

If you feel rushed, be sure to write a brief statement indicating the key idea you expect to use in your solutions. We understand time pressure, but we can't read your mind.

This examination has text printed on only one side of each page. Rather than flipping back and forth between pages, you may find it helpful to rip pages out of the exam so that you can look at more than one page at the same time.

Contents

| | |
|--|----|
| Problem 1: Short Answer [18 points] | 2 |
| Problem 2: Operational Semantics [22 points] | 3 |
| Problem 3: Denotational Semantics [24 points] | 5 |
| Problem 4: Type Reconstruction [20 points] | 6 |
| Problem 5: Compilers [16 points] | 7 |
| Appendix A: PostFix Grammar | 9 |
| Appendix B: PostFix SOS from Chapter 3 | 10 |
| Appendix C: Standard Semantics of FLK! | 11 |
| Appendix D: Typing Rules for SCHEME/R | 14 |
| Appendix E: Type Reconstruction Algorithm for SCHEME/R | 15 |
| Appendix F: Meta-CPS Conversion Rules | 16 |

The figures in the Appendix are very similar to the ones in the course notes. Some bugs have been fixed, and some figures have been simplified to remove parts inessential for this exam. You will not be marked down if you use the corresponding figures in the course notes instead of the appendices.

Problem 1: Short Answer [18 points]

a. Explicit Polymorphism [6 points]

Rewrite the following programs in Scheme/XSP with the most flexible type possible.

(i) [3 points]

```
(lambda (f g x)
  (if (f x)
      (g x)
      (g (g x))))
```

(ii) [3 points]

```
(lambda (f)
  (if (f #t)
      (f (f 1))
      (f 2)))
```

b. Type Reconstruction [12 points]

Determine the reconstructed types of the following Scheme/R expressions. If none exists, explain why.

(i) [3 points]

```
(lambda (f x)
  (if (f #t)
      (f x)
      2))
```

(ii) [3 points]

```
(let ((f (lambda (g x) (g (g x)))))
  (if (f not #t)
      (f (lambda (x) (+ x 1)) 1)
      2))
```

(iii) [3 points]

```
(let ((f (lambda (x) x)))
  (cons (f 1)
        (cons (f #t)
              (null))))
```

(iv) [3 points]

```
(lambda (g) (g g))
```

Problem 2: Operational Semantics [22 points]

ANSWERS FOR THE FOLLOWING QUESTIONS SHOULD BE BASED ON EXTENDING THE POSTFIX GRAMMAR IN APPENDIX A AND THE POSTFIX OPERATIONAL SEMANTICS GIVEN IN APPENDIX B.

Intel has released the new 747 chip, a chip so big it can handle mind boggling amounts of number crunching. The good programmers over at Microsquish are excited because this means there is now a computer fast enough to run their dream programming language, TwoStack PostFix. In TwoStack PostFix, a program is a tuple of two command sequences, one per stack. For instance, to have one stack add the numbers 1 and 3 and have the other stack multiply 4 * 5, use the program

`<[1 3 add], [4 5 mul]>`

The meaning of a TwoStack program is also a tuple, reporting the top of each stack at the end of the program. In the previous case, `<4, 20>`.

Keep in mind that the stacks are represented as completely separate entities at different locations in memory. Note that we will model errors as stuck states. For example, the program

`<[5 1 mul], [4 0 div]>`

should get stuck. It is also a stuck state if one stack runs out of commands before the other. For instance,

`<[5 1 mul 3 add], [4 1 div]>`

should get stuck right after the transition which performs the `div` command. Note that matching commands are executed simultaneously- that is, the 5 and 4 are pushed at the same time and the `mul` and the `div` are executed at the same time.

Finally, Executives at Microsquish would like to implement a `talk` command allowing the two stacks to communicate. For now, do not worry about the transition rule for the `talk` command, but know that it requires the following domain updates:

$$\begin{aligned} C &\in \text{Command}_{\text{PostFix+Talk}} \\ C &::= \dots \text{existing PostFix commands } \dots \\ &\quad | \text{talk} \\ Q &\in \text{Commands}_{\text{PostFix+Talk}} = \text{Command}_{\text{PostFix+Talk}}^* \\ V &\in \text{Value} = \text{IntLit} + \text{Commands}_{\text{PostFix+Talk}} \end{aligned}$$

In addition, the transition relation for `PostFix+Talk` is the same as the relation for `PostFix` except it is updated to work with the new domains. It currently results in a stuck state for the `talk` command.

The important domains for TwoStack are as follows:

$$\begin{aligned} P &\in \text{Program}_{\text{TwoStack}} = \text{Commands}_{\text{TwoStack}} \\ &\quad \text{Commands}_{\text{TwoStack}} = \text{Commands}_{\text{PostFix+Talk}} \times \text{Commands}_{\text{PostFix+Talk}} \\ A &\in \text{Answer}_{\text{TwoStack}} = \text{Answer} \times \text{Answer} \end{aligned}$$

- a. [12 points] Louis Reasoner is given the job of defining the Operational Semantics of `PostFix+Talk`. He decides on the following configuration:

$$\mathcal{C}_{\text{TwoStack}} = \text{Commands}_{\text{TwoStack}} \times \text{Stack} \times \text{Stack}$$

However, he needs your help to define the rest of the 5-tuple. Define the Input, Output and Transition functions in terms of the PostFix+Talk functions. Use subscripting to express which domains the functions are over. For instance, $\Rightarrow_{\text{PostFix+Talk}}$ is the transition function for PostFix+Talk and $\Rightarrow_{\text{TwoStack}}$ is the transition function for TwoStack PostFix.

Be sure to also define the set of final configurations of TwoStack Postfix, but do not worry about reporting errors for stuck states.

It's not very exciting having two stacks run in parallel unless they can communicate. So, Louis Reasoner decides to define the `talk` command. If `talk` is at the front of both command sequences, the top value on each stack is copied to the top of the other stack. For instance,

`<[5 1 mul talk add], [4 1 div talk mul]>`

should return `<9, 20>`

- b. [6 points] Extend the transition function for TwoStack to include the `talk` command.
- c. [4 points] Mark An-treason (who is also working at Microsquish since his company was bought out) is worried that programs with `talk` may not terminate. If programs in TwoStack PostFix always terminate, set Mark's fears at rest by explaining how you would modify the energy proof to show this. If not, give an example program which does not terminate.

Problem 3: Denotational Semantics [24 points]

ANSWERS FOR THE FOLLOWING QUESTIONS SHOULD BE BASED ON MODIFYING THE FLK! STANDARD SEMANTICS GIVEN IN APPENDIX C.

Ben Bitdiddle has been called in to assist in the development of a new feature for FL!. This feature will allow procedures to call themselves recursively without having to use **rec** or **letrec**. Ben proposes adding a new form **(self E)** to FLK!. The form **(self E)** calls the containing procedure with an actual parameter that is the result of evaluating E .

The FLK! expression grammar is changed by the addition of the form **(self E)**:

$$E ::= \dots \text{ existing FLK! forms } \dots \\ | \text{ (self } E \text{)}$$

Here is an example of the use of **(self E)** written in FL! (which would be desugared into FLK! before execution):

```
(let ((fact (lambda (n) (if (= n 0) 1 (* n (self (- n 1)))))))
  (fact 4))
```

$\Rightarrow 24$

Ben further specifies that when **(self E)** is used outside of a procedure it causes the program to terminate immediately with a value that is the result of evaluating E .

Ben begins describing the denotational semantics of the **self** form by modifying the signature of the meaning function, \mathcal{E} . His new signature is:

$$\mathcal{E} : \text{Exp} \rightarrow \text{Environment} \rightarrow \text{SelfProc} \rightarrow \text{ExpCont} \rightarrow \text{CmdCont} \\ \text{SelfProc} = \text{Procedure}$$

Ben asks you to complete the denotational description of the **self** form because he is going to a conference to give a speech on why FL! is the language of the future.

- [4 points] Give the revised meaning function $\mathcal{TL}[[E]]$.
- [4 points] What is the revised $\mathcal{E}[[\text{call } E_1 \ E_2]]$?
- [4 points] What is $\mathcal{E}[[\text{self } E]]$?
- [4 points] What is the revised $\mathcal{E}[[\text{proc } I \ E]]$?
- [4 points] Prove that $\mathcal{TL}[[\text{self } (\text{self } 1)]]$ in your semantics means $(\text{Value} \rightarrow \text{Expressible } (\text{Int} \rightarrow \text{Value } 1))$.
- [4 points] Use your semantics to show **(proc x (self 1))** evaluates to a procedure that, no matter what input it is called with, loops forever.

Problem 4: Type Reconstruction [20 points]

ANSWERS FOR THE FOLLOWING QUESTIONS SHOULD BE BASED ON EXTENDING THE TYPING RULES GIVEN IN APPENDIX D AND THE TYPE RECONSTRUCTION ALGORITHM GIVEN IN APPENDIX E.

Alyssa P. Hacker has been asked to extend the type system for SCHEME/R to handle **(label I E)** and **(jump E₁ E₂)**. As introduced in class and in the book, **(label I E)** establishes a control point called *I* in *E*, and **(jump E₁ E₂)** jumps to the control point that is the value of *E*₁ and gives it the value obtained by evaluating *E*₂.

For example:

```
(label out
  (if (= x 0)
      (jump out 0)
      (/ y x)))
```

Alyssa has added a control point type to the type grammar for SCHEME/R as follows:

```
T ::= ... existing SCHEME/R types ...
    | (control-point T)
```

In the example above, the control point **out** would have type (control-point int). It is possible to have control point type errors in SCHEME/R. For example, a **label** expression must have the same type regardless of whether a jump is encountered, making the following expression not well-typed:

```
(label out
  (if (= x 0)
      (jump out #f)
      (/ y x)))
```

Your job is to complete the implementation of typed control points.

- [5 points] Give the new typing rules in SCHEME/R for **(label I E)** and **(jump E₁ E₂)**.
- [5 points] Give the type reconstruction algorithm clause for **(label I E)**.
- [5 points] Give the type reconstruction algorithm clause for **(jump E₁ E₂)**.
- [5 points] Give the reconstructed type of the following expression, or give the input parameters to the procedure that fails during the reconstruction:

```
(let ((x (label y y)))
      (jump x (label z z)))
```

Problem 5: Compilers [16 points]

ANSWERS FOR THE FOLLOWING QUESTIONS SHOULD BE BASED ON META CPS CONVERSION ALGORITHM GIVEN IN APPENDIX F.

Ben Bitdiddle is a consummate compiler hacker, and has been asked by JavaHard to analyze their new Scheme compiler. JavaHard has recently realized that Java will not win in the marketplace against Scheme, and has adopted the 6.821 Scheme compiler as the cornerstone of their crash development effort. The following code was produced by the compiler after the `desugar`, `globals/wrap`, `cps-convert`, `closures/flat-all`, and `lift-convert` phases of compilation.

```
(program
  (define
    .lambda24.
    (lambda
      (.closure11. .t5.)
      (call-closure .t5. 1 (primop closure-ref .closure11. 1))))
  (define
    .lambda21.
    (lambda
      (.closure14. f g .k1.)
      (call-closure .k1. (primop closure .lambda22. g f))))
  (define
    .lambda22.
    (lambda
      (.closure13. x .k2.)
      (call-closure
        (primop closure-ref .closure13. 1)
        x
        (primop closure .lambda23. (primop closure-ref .closure13. 2) .k2.))))
  (define
    .lambda23.
    (lambda
      (.closure12. .t3.)
      (call-closure
        (primop closure-ref .closure12. 1)
        .t3.
        (primop closure-ref .closure12. 2))))
  (define .lambda20. (lambda (.closure17. x) x))
  (define
    .lambda19.
    (lambda
      (.closure16. a .k9.)
      (let ((.t10. (primop not a))) (call-closure .k9. .t10.))))
  (define
    .lambda18.
    (lambda
      (.closure15. a .k7.)
      (let ((.t8. (primop integer? a))) (call-closure .k7. .t8.))))
  (define *top* (primop closure .lambda20.))
  (define not (primop closure .lambda19.))
  (define integer? (primop closure .lambda18.))
  (let
    ((compose (primop closure .lambda21.)))
    (call-closure compose not integer? (primop closure .lambda24. *top*))))
```

- a. [6 points] What source code resulted in the incompletely compiled code above?

JavaHard soon discovers another problem with the compiler. The compiler they are using cannot handle the **begin** form because the `cps-convert` phase does not include a rule for translating **begin**.

- b. [4 points] Ben looked in the 6.821 book and could not find the meta-cps rule for **begin**. What is $\mathcal{MCPS}[(\text{begin } E1 \ E2)]$?

JavaHard has decided to not let programmers access control features such as `cwcc`, `label`, and `jump` to simplify the analysis of programs. Ben suggests to JavaHard that they include a region identifier in every procedure type:

$$T ::= \dots | (\rightarrow (T^*) T R)$$

Just as regions are assigned to cells, every procedure will be assigned a new region identifier except when two procedures are aliased together.

A procedure in E can be explicitly freed if all of the following conditions are met: (1) the procedure is in region R , (2) region R is not in the type of any free variables of E , and (3) region R is not in the type of E .

- c. [4 points] Excited about explicit freedom, Ben invents a new primitive called **(%procedure-free x)** that frees the procedure represented by value x . In the following example, the lambda expression **(lambda (y) y)** is bound to x and freed:

```
(let ((x (lambda (y) y)))
  (%procedure-free x))
```

Let E be an expression that contains a non-nested lambda expression P . From effect analysis, we know that P 's value is no longer needed after E completes. Thus, Ben would like to use **(%procedure-free x)** to free the procedure value corresponding to P . Help Ben by writing a translation function for E that will free the value of P and return the value of E . Use $[v/P]E$ to substitute the variable v for lambda expression P in E . Assume that your translation occurs before MCPS conversion and that the variable v does not appear in E .

```
T[E] = (let ((v P))
  ...fill in text here...
)
```

- d. [2 points] Ben revises his compiler to call this primitive to free all closures using your translation rule (assuming it works), and notes that certain programs slow down as a consequence. Ben had thought that reducing the work that the garbage collector had to do would make programs run faster. What is a possible explanation for this behavior?

Appendix A: PostFix Grammar

| | | | |
|-----|-------|---------------------|---------------------------------------|
| P | \in | Program | |
| Q | \in | Commands | |
| C | \in | Command | |
| A | \in | Arithmetic-operator | $= \{add, sub, mul, div\}$ |
| R | \in | Relational-operator | $= \{lt, eq, gt\}$ |
| N | \in | Intlit | $= \{\dots, -2, -1, 0, 1, 2, \dots\}$ |
| P | $::=$ | (Q) | [Program] |
| Q | $::=$ | C^* | [Command-sequence] |
| C | $::=$ | N | [Integer-literal] |
| | | pop | [Pop] |
| | | swap | [Swap] |
| | | A | [Arithmetic-op] |
| | | R | [Relational-op] |
| | | sel | [Select] |
| | | exec | [Execute] |
| | | (Q) | [Executable-sequence] |

Figure 1: The S-Expression Grammar for PostFix

Appendix B: PostFix SOS from Chapter 3

$\mathcal{C}, \mathcal{F}, \mathcal{I}$ and \mathcal{O} for our PostFix SOS are given by:

$\mathcal{C} = \text{Commands} \times \text{Stack}$

$\mathcal{F} = \{\llbracket _ \rrbracket_{\text{Command}}\} \times \text{Stack}$

$\mathcal{I} : \text{Program} \rightarrow \mathcal{C}$
 $= \lambda P . \mathbf{matching} P$
 $\triangleright \langle Q \rangle \parallel \langle Q, \llbracket _ \rrbracket_{\text{Value}} \rangle$
 $\mathbf{endmatching}$

$\mathcal{O} : \mathcal{F} \rightarrow \text{Answer}$
 $= \lambda (\llbracket _ \rrbracket_{\text{Command}}, S) . \mathbf{matching} S$
 $\triangleright V . S' \parallel (\text{Value} \mapsto \text{Answer } V)$
 $\triangleright \llbracket _ \rrbracket \parallel (\text{Error} \mapsto \text{Answer error})$
 $\mathbf{endmatching}$

| | |
|--|---------------|
| $\langle N . Q, S \rangle \Rightarrow \langle Q, N . S \rangle$ | [numeral] |
| $\langle \langle Q_{exec} \rangle . Q_{rest}, S \rangle \Rightarrow \langle Q_{rest}, Q_{exec} . S \rangle$ | [executable] |
| $\langle \text{pop} . Q, V_{top} . S \rangle \Rightarrow \langle Q, S \rangle$ | [pop] |
| $\langle \text{swap} . Q, V_1 . V_2 . S \rangle \Rightarrow \langle Q, V_2 . V_1 . S \rangle$ | [swap] |
| $\langle \text{sel} . Q_{rest}, V_{false} . V_{true} . 0 . S \rangle \Rightarrow \langle Q_{rest}, V_{false} . S \rangle$ | [sel-false] |
| $\langle \text{sel} . Q_{rest}, V_{false} . V_{true} . N_{test} . S \rangle \Rightarrow \langle Q_{rest}, V_{true} . S \rangle$ where $N_{test} \neq 0$ | [sel-true] |
| $\langle \text{exec} . Q_{rest}, Q_{exec} . S \rangle \Rightarrow \langle Q_{exec} @ Q_{rest}, S \rangle$ | [execute] |
| $\langle A . Q, N_1 . N_2 . S \rangle \Rightarrow \langle Q, N_{result} . S \rangle$ where $N_{result} \equiv (\text{calculate } A \ N_2 \ N_1)$ and $\neg((A \equiv \text{div}) \wedge (N_1 \equiv 0))$ | [arithop] |
| $\langle R . Q, N_1 . N_2 . S \rangle \Rightarrow \langle Q, 1 . S \rangle$ where $(\text{compare } R \ N_2 \ N_1)$ | [relop-true] |
| $\langle R . Q, N_1 . N_2 . S \rangle \Rightarrow \langle Q, 0 . S \rangle$ where $\neg(\text{compare } R \ N_2 \ N_1)$ | [relop-false] |

Figure 2: Rewrite rules defining the transition relation for PostFix.

Appendix C: Standard Semantics of FLK!

| |
|--|
| $v \in \text{Value} = \text{Unit} + \text{Bool} + \text{Int} + \text{Sym} + \text{Pair} + \text{Procedure} + \text{Location}$ $k \in \text{Expcont} = \text{Value} \rightarrow \text{Cmdcont}$ $\gamma \in \text{Cmdcont} = \text{Store} \rightarrow \text{Expressible}$ $\text{Expressible} = (\text{Value} + \text{Error})_{\perp}$ $\text{Error} = \text{Sym}$ $p \in \text{Procedure} = \text{Denotable} \rightarrow \text{Expcont} \rightarrow \text{Cmdcont}$ $d \in \text{Denotable} = \text{Value}$ $e \in \text{Environment} = \text{Identifier} \rightarrow \text{Binding}$ $\beta \in \text{Binding} = (\text{Denotable} + \text{Unbound})_{\perp}$ $\text{Unbound} = \{\text{unbound}\}$ $s \in \text{Store} = \text{Location} \rightarrow \text{Assignment}$ $l \in \text{Location} = \text{Nat}$ $\alpha \in \text{Assignment} = (\text{Storable} + \text{Unassigned})_{\perp}$ $\sigma \in \text{Storable} = \text{Value}$ $\text{Unassigned} = \{\text{unassigned}\}$ $\text{top-level-cont} : \text{Expcont}$ $= \lambda v . \lambda s . (\text{Value} \mapsto \text{Expressible } v)$ $\text{error-cont} : \text{Error} \rightarrow \text{Cmdcont}$ $= \lambda y . \lambda s . (\text{Error} \mapsto \text{Expressible } y)$ $\text{empty-env} : \text{Environment} = \lambda I . (\text{Unbound} \mapsto \text{Binding } \text{unbound})$ $\text{test-boolean} : (\text{Bool} \rightarrow \text{Cmdcont}) \rightarrow \text{Expcont}$ $= \lambda f . (\lambda v . \mathbf{matching } v$ $\quad \triangleright (\text{Bool} \mapsto \text{Value } b) \parallel (f \ b)$ $\quad \triangleright \mathbf{else } (\text{error-cont } \text{non-boolean})$ $\quad \mathbf{endmatching })$ Similarly for: $\text{test-procedure} : (\text{Procedure} \rightarrow \text{Cmdcont}) \rightarrow \text{Expcont}$ $\text{test-location} : (\text{Location} \rightarrow \text{Cmdcont}) \rightarrow \text{Expcont}$ etc. $\text{ensure-bound} : \text{Binding} \rightarrow \text{Expcont} \rightarrow \text{Cmdcont}$ $= \lambda \beta k . \mathbf{matching } \beta$ $\quad \triangleright (\text{Denotable} \mapsto \text{Binding } v) \parallel (k \ v)$ $\quad \triangleright (\text{Unbound} \mapsto \text{Binding } \text{unbound}) \parallel (\text{error-cont } \text{unbound-variable})$ $\quad \mathbf{endmatching }$ Similarly for: $\text{ensure-assigned} : \text{Assignment} \rightarrow \text{Expcont} \rightarrow \text{Cmdcont}$ |
|--|

Figure 3: Semantic algebras for standard semantics of strict CBV FLK!.

```

same-location? : Location → Location → Bool = λl1l2 . (l1 =Nat l2)
next-location : Location → Location = λl . (l +Nat 1)
empty-store : Store = λl . (Unassigned ↦ Assignment unassigned)
fetch : Location → Store → Assignment = λls . (s l)
assign : Location → Storable → Store → Store
= λl1σs . λl2 . if (same-location? l1 l2)
    then (Storable ↦ Assignment σ)
    else (fetch l2 s)
fresh-loc : Store → Location = λs . (first-fresh s 0)
first-fresh : Store → Location → Location
= λsl . matching (fetch l s)
    ▷ (Unassigned ↦ Assignment unassigned) ∥ l
    ▷ else (first-fresh s (next-location l))
    endmatching
lookup : Environment → Identifier → Binding = λeI . (e I)

```

Figure 4: Store helper functions for standard semantics of strict CBV FLK!.

$\mathcal{TL} : \text{Exp} \rightarrow \text{Expressible}$
 $\mathcal{E} : \text{Exp} \rightarrow \text{Environment} \rightarrow \text{Expcont} \rightarrow \text{Cmdcont}$
 $\mathcal{L} : \text{Lit} \rightarrow \text{Value}$; *Defined as usual*

$\mathcal{TL}[E] = \mathcal{E}[E]$ *empty-env top-level-cont empty-store*

$\mathcal{E}[L] = \lambda ek . k \mathcal{L}[L]$

$\mathcal{E}[I] = \lambda ek . \text{ensure-bound } (\text{lookup } e \ I) \ k$

$\mathcal{E}[(\text{proc } I \ E)] = \lambda ek . k \ (\text{Procedure} \mapsto \text{Value } (\lambda dk' . \mathcal{E}[E] \ [I : d]e \ k'))$

$\mathcal{E}[(\text{call } E_1 \ E_2)] = \lambda ek . \mathcal{E}[E_1] \ e \ (\text{test-procedure } (\lambda p . \mathcal{E}[E_2] \ e \ (\lambda v . p \ v \ k)))$

$\mathcal{E}[(\text{if } E_1 \ E_2 \ E_3)] =$
 $\lambda ek . \mathcal{E}[E_1] \ e \ (\text{test-boolean } (\lambda b . \text{if } b \ \text{then } \mathcal{E}[E_2] \ e \ k \ \text{else } \mathcal{E}[E_3] \ e \ k))$

$\mathcal{E}[(\text{pair } E_1 \ E_2)] = \lambda ek . \mathcal{E}[E_1] \ e \ (\lambda v_1 . \mathcal{E}[E_2] \ e \ (\lambda v_2 . k \ (\text{Pair} \mapsto \text{Value } \langle v_1, v_2 \rangle)))$

$\mathcal{E}[(\text{cell } E)] = \lambda ek . \mathcal{E}[E] \ e \ (\lambda vs . k \ (\text{Location} \mapsto \text{Value } (\text{fresh-loc } s))$
 $\quad \quad \quad (\text{assign } (\text{fresh-loc } s) \ v \ s))$

$\mathcal{E}[(\text{begin } E_1 \ E_2)] = \lambda ek . \mathcal{E}[E_1] \ e \ (\lambda v_{\text{ignore}} . \mathcal{E}[E_2] \ e \ k)$

$\mathcal{E}[(\text{primop cell-ref } E)] = \lambda ek . \mathcal{E}[E] \ e \ (\text{test-location } (\lambda ls . \text{ensure-assigned } (\text{fetch } l \ s) \ k \ s))$

$\mathcal{E}[(\text{primop cell-set! } E_1 \ E_2)]$
 $= \lambda ek . \mathcal{E}[E_1] \ e \ (\text{test-location } (\lambda l . \mathcal{E}[E_2] \ e \ (\lambda vs . k \ (\text{Unit} \mapsto \text{Value } \text{unit}) \ (\text{assign } l \ v \ s))))$

$\mathcal{E}[(\text{rec } I \ E)] = \lambda eks . \text{let } f = \mathbf{fix}_{\text{Expressible}} (\lambda a . \mathcal{E}[E] \ [I : (\text{extract-value } a)] \ e \ \text{top-level-cont } s)$
 $\quad \quad \quad \mathbf{matching} \ f$
 $\quad \quad \quad \triangleright (\text{Value} \mapsto \text{Expressible } v) \parallel \mathcal{E}[E] \ [I : v] \ e \ k \ s$
 $\quad \quad \quad \triangleright \text{else } f$
 $\quad \quad \quad \mathbf{endmatching}$

$\text{extract-value} : \text{Expressible} \rightarrow \text{Binding}$
 $= \lambda a . \mathbf{matching} \ a$
 $\quad \triangleright (\text{Value} \mapsto \text{Expressible } v) \parallel (\text{Denotable} \mapsto \text{Binding } v)$
 $\quad \triangleright \text{else } \perp_{\text{Binding}}$
 $\quad \mathbf{endmatching}$

Figure 5: Valuation clauses for standard semantics of strict CBV FLK!

Appendix D: Typing Rules for SCHEME/R

| | |
|---|--------------------|
| $\vdash \#u : \text{unit}$ | [<i>unit</i>] |
| $\vdash B : \text{bool}$ | [<i>bool</i>] |
| $\vdash N : \text{int}$ | [<i>int</i>] |
| $\vdash (\text{symbol } I) : \text{sym}$ | [<i>symbol</i>] |
| $[\dots, I : T, \dots] \vdash I : T$ | [<i>var</i>] |
| $[\dots, I : (\text{generic } (I_1 \dots I_n) T_{body}), \dots] \vdash I : (\forall i [T_i/I_i])T_{body}$ | [<i>genvar</i>] |
| $\frac{A \vdash E_{test} : \text{bool} ; A \vdash E_{con} : T ; A \vdash E_{alt} : T}{A \vdash (\text{if } E_{test} E_{con} E_{alt}) : T}$ | [<i>if</i>] |
| $\frac{A[I_1 : T_1, \dots, I_n : T_n] \vdash E_{body} : T_{body}}{A \vdash (\text{lambda } (I_1 \dots I_n) E_{body}) : (-> (T_1 \dots T_n) T_{body})}$ | [λ] |
| $\frac{A \vdash E_{rator} : (-> (T_1 \dots T_n) T_{body}) \quad \forall i . (A \vdash E_i : T_i)}{A \vdash (E_{rator} E_1 \dots E_n) : T_{body}}$ | [<i>apply</i>] |
| $\frac{\forall i . (A \vdash E_i : T_i) \quad A[I_1 : \text{Gen}(T_1, A), \dots, I_n : \text{Gen}(T_n, A)] \vdash E_{body} : T_{body}}{A \vdash (\text{let } ((I_1 E_1) \dots (I_n E_n)) E_{body}) : T_{body}}$ | [<i>let</i>] |
| $\frac{\forall i . (A[I_1 : T_1, \dots, I_n : T_n] \vdash E_i : T_i) \quad A[I_1 : \text{Gen}(T_1, A), \dots, I_n : \text{Gen}(T_n, A)] \vdash E_{body} : T_{body}}{A \vdash (\text{letrec } ((I_1 E_1) \dots (I_n E_n)) E_{body}) : T_{body}}$ | [<i>letrec</i>] |
| $\frac{\forall i . (A \vdash E_i : T_i)}{A \vdash (\text{record } (I_1 E_1) \dots (I_n E_n)) : (\text{recordof } (I_1 T_1) \dots (I_n T_n))}$ | [<i>record</i>] |
| $\frac{A \vdash E_r : (\text{recordof } (I_1 T_1) \dots (I_n T_n)) \quad A[I_1 : T_1, \dots, I_n : T_n] \vdash E_b : T}{A \vdash (\text{with } (I_1 \dots I_n) E_r E_b) : T}$ | [<i>with</i>] |
| $\frac{A \vdash (\text{letrec } ((I_1 E_1) \dots (I_n E_n)) E_{body}) : T}{A \vdash (\text{program } (\text{define } I_1 E_1) \dots (\text{define } I_n E_n) E_{body}) : T}$ | [<i>program</i>] |

$\text{Gen}(T, A) = (\text{generic } (I_1 \dots I_n) T)$, where $\{I_i\} = \text{FTV}(T) - \text{FTE}(A)$

Appendix E: Type Reconstruction Algorithm for SCHEME/R

$R[\#\mathbf{u}] A S = \langle \mathbf{unit}, S \rangle$

$R[B] A S = \langle \mathbf{bool}, S \rangle$

$R[N] A S = \langle \mathbf{int}, S \rangle$

$R[(\mathbf{symbol} I)] A S = \langle \mathbf{sym}, S \rangle$

$R[I] A[I : T] S = \langle T, S \rangle$

$R[I] A[I : (\mathbf{generic} (I_1 \dots I_n) T)] S = \langle T[?v_i/I_i], S \rangle$ ($?v_i$ are new)

$R[I] A S = \mathbf{fail}$ (when I is unbound)

$R[(\mathbf{if} E_t E_c E_a)] A S = \mathbf{let} \langle T_t, S_t \rangle = R[E_t] A S$
 $\mathbf{in} \mathbf{let} S'_t = U(T_t, \mathbf{bool}, S_t)$
 $\mathbf{in} \mathbf{let} \langle T_c, S_c \rangle = R[E_c] A S'_t$
 $\mathbf{in} \mathbf{let} \langle T_a, S_a \rangle = R[E_a] A S_c$
 $\mathbf{in} \mathbf{let} S'_a = U(T_c, T_a, S_a)$
 $\mathbf{in} \langle T_a, S'_a \rangle$

$R[(\mathbf{lambda} (I_1 \dots I_n) E_b)] A S = \mathbf{let} \langle T_b, S_b \rangle = R[E_b] A[I_i : ?v_i] S$
 $\mathbf{in} \langle (-> (?v_1 \dots ?v_n) T_b), S_b \rangle$ ($?v_i$ are new)

$R[(E_0 E_1 \dots E_n)] A S = \mathbf{let} \langle T_0, S_0 \rangle = R[E_0] A S$
 $\mathbf{in} \dots$
 $\mathbf{let} \langle T_n, S_n \rangle = R[E_n] A S_{n-1}$
 $\mathbf{in} \mathbf{let} S_f = U(T_0, (-> (T_1 \dots T_n) ?v_f), S_n)$
 $\mathbf{in} \langle ?v_f, S_f \rangle$ ($?v_f$ is new)

$R[(\mathbf{let} ((I_1 E_1) \dots (I_n E_n)) E_b)] A S = \mathbf{let} \langle T_1, S_1 \rangle = R[E_1] A S$
 $\mathbf{in} \dots$
 $\mathbf{let} \langle T_n, S_n \rangle = R[E_n] A S_{n-1}$
 $\mathbf{in} R[E_b] A[I_i : Rgen(T_i, A, S_n)] S_n$

$R[(\mathbf{letrec} ((I_1 E_1) \dots (I_n E_n)) E_b)] A S = \mathbf{let} A_1 = A[I_i : ?v_i]$ ($?v_i$ are new)
 $\mathbf{in} \mathbf{let} \langle T_1, S_1 \rangle = R[E_1] A_1 S$
 $\mathbf{in} \dots$
 $\mathbf{let} \langle T_n, S_n \rangle = R[E_n] A_1 S_{n-1}$
 $\mathbf{in} \mathbf{let} S_b = U(?v_i, T_i, S_n)$
 $\mathbf{in} R[E_b] A[I_i : Rgen(T_i, A, S_b)] S_b$

$R[(\mathbf{record} (I_1 E_1) \dots (I_n E_n))] A S = \mathbf{let} \langle T_1, S_1 \rangle = R[E_1] A S$
 $\mathbf{in} \dots$
 $\mathbf{let} \langle T_n, S_n \rangle = R[E_n] A S_{n-1}$
 $\mathbf{in} \langle \mathbf{recordof} (I_1 T_1) \dots (I_n T_n), S_n \rangle$

$R[(\mathbf{with} (I_1 \dots I_n) E_r E_b)] A S = \mathbf{let} \langle T_r, S_r \rangle = R[E_r] A S$
 $\mathbf{in} \mathbf{let} S_b = U(T_r, (\mathbf{recordof} (I_1 ?v_i) \dots (I_n ?v_n)), S_r)$ ($?v_i$ are new)
 $\mathbf{in} R[E_b] A[I_i : ?v_i] S_b$

$Rgen(T, A, S) = Gen((S T), (\mathbf{subst-in-type-env} S A))$

Appendix F: Meta-CPS Conversion Rules

In the following rules, grey mathematical notation (like λv) and square brackets $[]$ are used for “meta-application”, which is evaluated *as part of meta-CPS conversion*. Code in BLACK TYPEWRITER FONT is part of the output program; meta-CPS conversion does *not* evaluate any of this code. Therefore, you can think of meta-CPS-converting an expression E as rewriting $\mathcal{MCP}S[E]$ until no grey is left.

$E \in \text{Exp}$
 $m \in \text{Meta-Continuation} = \text{Exp} \rightarrow \text{Exp}$

$meta\text{-}cont \rightarrow exp : (\text{Exp} \rightarrow \text{Exp}) \rightarrow \text{Exp} = [\lambda m . (\text{LAMBDA } (t) [m t])]$
 $exp \rightarrow meta\text{-}cont : \text{Exp} \rightarrow (\text{Exp} \rightarrow \text{Exp}) = [\lambda E . [\lambda V . (\text{CALL } E V)]]$

$meta\text{-}cont \rightarrow exp [\lambda V . (\text{CALL } K V)] = K$

$\mathcal{MCP}S : \text{Exp} \rightarrow \text{Meta-Continuation} \rightarrow \text{Exp}$

$\mathcal{MCP}S[I] = [\lambda m . [m I]]$

$\mathcal{MCP}S[L] = [\lambda m . [m L]]$

$\mathcal{MCP}S[(\text{LAMBDA } (I_1 \dots I_n) E)]$
 $= [\lambda m . [m (\text{LAMBDA } (I_1 \dots I_n) . Ki.)$
 $\quad [\mathcal{MCP}S[E] [exp \rightarrow meta\text{-}cont . Ki.]]]]$

$\mathcal{MCP}S[(\text{CALL } E_1 E_2)]$
 $= [\lambda m . [\mathcal{MCP}S[E_1] [\lambda v_1 .$
 $\quad [\mathcal{MCP}S[E_2] [\lambda v_2 .$
 $\quad \quad (\text{CALL } v_1 v_2 [meta\text{-}cont \rightarrow exp m])]]]]]$

$\mathcal{MCP}S[(\text{PRIMOP } P E_1 E_2)]$
 $= [\lambda m . [\mathcal{MCP}S[E_1] [\lambda v_1 .$
 $\quad [\mathcal{MCP}S[E_2] [\lambda v_2 .$
 $\quad \quad (\text{LET } ((.Ti. (\text{PRIMOP } P v_1 v_2)))$
 $\quad \quad [m . .Ti.])]]]]]$

$\mathcal{MCP}S[(\text{IF } E_c E_t E_f)]$
 $= [\lambda m . [\mathcal{MCP}S[E_c] [\lambda v_1 .$
 $\quad (\text{LET } ((K [meta\text{-}cont \rightarrow exp m]))$
 $\quad (\text{IF } v_1$
 $\quad \quad [\mathcal{MCP}S[E_t] [exp \rightarrow meta\text{-}cont K]]$
 $\quad \quad [\mathcal{MCP}S[E_f] [exp \rightarrow meta\text{-}cont K]]))]]]$

$\mathcal{MCP}S[(\text{LET } ((I E_{\text{def}})) E_{\text{body}})]$
 $= [\lambda m . [\mathcal{MCP}S[E_{\text{def}}] [\lambda v .$
 $\quad (\text{LET } ((I v)) [\mathcal{MCP}S[E_{\text{body}}] m])]]]$

2. Final Examination Solutions

Problem 1: Short Answer

a. Explicit Polymorphism

- (i) `(plambda (t)
 (lambda ((f (-> (t) bool)) (g (-> (t) t) (x t)))
 (if (f x)
 (g x)
 (g (g x))))))`
- (ii) `(lambda ((f (poly (t) (-> (t) t))))
 (if ((proj f bool) #t)
 ((proj f int) ((proj f int) 1))
 ((proj f int) 2)))`

b. Type Reconstruction

- (i) No reconstructed type exists. There is no first class polymorphism allowed in Scheme/R.
- (ii) `int`
- (iii) No reconstructed type exists. No heterogeneous lists in Scheme/R.
- (iv) No reconstructed type exists. Hindley-Milner cannot reconstruct the type of an identifier involved in a self application.

Problem 2: Operational Semantics

a.

$$\begin{aligned}
 \mathcal{I}_{\text{TwoStack}} &: \text{Program}_{\text{TwoStack}} \rightarrow \mathcal{C}_{\text{TwoStack}} \\
 \mathcal{I}_{\text{TwoStack}} &= \lambda \langle Q_1, Q_2 \rangle . \langle \langle Q_1, Q_2 \rangle, \llbracket \text{Value} \rrbracket, \llbracket \text{Value} \rrbracket \rangle \\
 \mathcal{F}_{\text{TwoStack}} &= \{ \llbracket \text{Command}_{\text{PostFix+Talk}} \rrbracket \} \times \{ \llbracket \text{Command}_{\text{PostFix+Talk}} \rrbracket \} \times \text{Stack} \times \text{Stack} \\
 \mathcal{O}_{\text{TwoStack}} &: \mathcal{F} \rightarrow \text{Answer}_{\text{TwoStack}} \\
 \mathcal{O}_{\text{TwoStack}} &= \lambda \langle \langle \llbracket \text{Command}_{\text{PostFix+Talk}} \rrbracket, \llbracket \text{Command}_{\text{PostFix+Talk}} \rrbracket \rangle, S_1, S_2 \rangle . \\
 &\quad \langle \mathcal{O}_{\text{PostFix}} \langle \llbracket \text{Command}_{\text{PostFix+Talk}} \rrbracket, S_1 \rangle, \mathcal{O}_{\text{PostFix}} \langle \llbracket \text{Command}_{\text{PostFix+Talk}} \rrbracket, S_2 \rangle \rangle \\
 \Rightarrow_{\text{TwoStack}} &: \\
 &\quad \frac{\langle Q_1, S_1 \rangle \Rightarrow_{\text{PostFix+Talk}} \langle Q'_1, S'_1 \rangle \wedge \langle Q_2, S_2 \rangle \Rightarrow_{\text{PostFix+Talk}} \langle Q'_2, S'_2 \rangle}{\langle \langle Q_1, Q_2 \rangle, S_1, S_2 \rangle \Rightarrow_{\text{TwoStack}} \langle \langle Q'_1, Q'_2 \rangle, S'_1, S'_2 \rangle}
 \end{aligned}$$

b.

$$\langle \langle \text{talk}.Q_1, \text{talk}.Q_2 \rangle, V_1.S_1, V_2.S_2 \rangle \Rightarrow_{\text{TwoStack}} \langle \langle Q_1, Q_2 \rangle, V_2.V_1.S_1, V_1.V_2.S_2 \rangle$$

c. The following program does not terminate:

`<<(talk exec) talk exec>>, <<(talk exec) talk exec>>`

Problem 3: Denotational Semantics

- a. $\mathcal{TL}\llbracket E \rrbracket = (\mathcal{E}\llbracket E \rrbracket \text{ empty-env } (\lambda dk . (\text{top-level-cont } d)) \text{ top-level-cont empty-store})$
- b. $\mathcal{E}\llbracket (\text{call } E_1 E_2) \rrbracket = \lambda ep_1k . (\mathcal{E}\llbracket E_1 \rrbracket e p_1 (\text{test-procedure } (\lambda p_2 . (\mathcal{E}\llbracket E_2 \rrbracket e p_1 (\lambda v . (p_2 v k))))))$
- c. $\mathcal{E}\llbracket (\text{self } E) \rrbracket = \lambda epk . (\mathcal{E}\llbracket E \rrbracket e p \lambda v . (p v k))$
- d. $\mathcal{E}\llbracket (\text{proc } I E) \rrbracket =$
 $\lambda ep_1k_1 . (k_1 (\text{Procedure} \mapsto \text{Value } (\mathbf{fix}_{\text{Procedure}}(\lambda p_2 . (\lambda dk_2 . (\mathcal{E}\llbracket E \rrbracket [I : d]e p_2 k_2))))))$
- e.

$$\begin{aligned}
& \mathcal{TL}\llbracket (\text{self } (\text{self } 1)) \rrbracket = \\
& = (\mathcal{E}\llbracket (\text{self } (\text{self } 1)) \rrbracket \text{ empty-env } (\lambda dk . (\text{top-level-cont } d)) \text{ top-level-cont empty-store}) \\
& = (\mathcal{E}\llbracket (\text{self } 1) \rrbracket \text{ empty-env } (\lambda dk . (\text{top-level-cont } d)) \text{ top-level-cont empty-store}) \\
& \quad \text{because } \lambda v . (\lambda dk . (\text{top-level-cont } d) v k) = \text{top-level-cont} \\
& = (\mathcal{E}\llbracket 1 \rrbracket \text{ empty-env } (\lambda dk . (\text{top-level-cont } d)) \text{ top-level-cont empty-store}) \\
& = (\text{top-level-cont } \mathcal{L}\llbracket 1 \rrbracket) \\
& = (\text{Value} \mapsto \text{Expressible } (\text{Int} \mapsto \text{Value } 1))
\end{aligned}$$

f.

$$\begin{aligned}
& \mathcal{E}\llbracket (\text{proc } x (\text{self } 1)) \rrbracket \\
& = \lambda ep_1k_1 . (k_1 (\text{Procedure} \mapsto \text{Value } (\mathbf{fix}_{\text{Procedure}} \lambda p_2 . (\lambda dk_2 . (\mathcal{E}\llbracket (\text{self } 1) \rrbracket [x : d]e p_2 k_2)))))) \\
& = \lambda ep_1k_1 . (k_1 (\text{Procedure} \mapsto \text{Value } (\mathbf{fix}_{\text{Procedure}} \lambda p_2 . (\lambda dk_2 . (\mathcal{E}1) [x : d]e p_2 \lambda v . (p_2 v k_2)))))) \\
& = \lambda ep_1k_1 . (k_1 (\text{Procedure} \mapsto \text{Value } (\mathbf{fix}_{\text{Procedure}} \lambda p_2 . (\lambda dk_2 . (\lambda v . (p_2 v k_2) \mathcal{L}\llbracket 1 \rrbracket)))))) \\
& = \lambda ep_1k_1 . (k_1 (\text{Procedure} \mapsto \text{Value } (\mathbf{fix}_{\text{Procedure}} \lambda p_2 . (\lambda dk_2 . (p_2 \mathcal{L}\llbracket 1 \rrbracket k_2)))))) \\
& \quad \text{but } \perp_{\text{Procedure}} \text{ is a fixed point of } \lambda p_2 . (\lambda dk_2 . (p_2 \mathcal{L}\llbracket 1 \rrbracket k_2))
\end{aligned}$$

It should be clear that $\perp_{\text{Procedure}}$ must be the procedure-generating function's least fixed point. This means that the procedural value that the expression **(proc x (self 1))** computes is $\perp_{\text{Procedure}}$, a procedure that given any denotable and expression continuation returns \perp_{CmdCont} i.e. a procedure that loops forever regardless of its input. Since this procedural value is the value computed by **(proc x (self 1))**, we have completed the demonstration.

Problem 4: Type Reconstruction

a.

$$\frac{A[I : (\text{control-point } T)] \vdash E : T}{A \vdash (\text{label } I E) : T} \quad [\text{label}]$$

$$\frac{A \vdash E_1 : (\text{control-point } T) \quad A \vdash E_2 : T}{A \vdash (\text{jump } E_1 E_2) : T_{\text{any}}} \quad [\text{jump}]$$

- b. $R\llbracket (\text{label } I E) \rrbracket A S = \mathbf{let} \langle T_1, S_1 \rangle = R\llbracket E \rrbracket A [I : (\text{control-point } ?v_1)] S$
 $\mathbf{in} \langle T_1, U(T_1, ?v_1, S_1) \rangle$

where $?v_1$ is a fresh type variable

c. $R[(\text{jump } E_1 \ E_2)] \ A \ S = \text{let } \langle T_1, S_1 \rangle = R[E_1] \ A \ S$
 $\text{in } \text{let } \langle T_2, S_2 \rangle = R[E_2] \ A \ S_1$
 $\text{in } \langle ?v_1, U(T_1, (\text{control-point } T_2), S_2) \rangle$

where $?v_1$ is a fresh type variable

d. Unification fails while reconstructing the type of **(label y y)**. The unification call that fails is $U((\text{control-point } ?v_1), ?v_1, S_1)$.

Problem 5: Compilers

a. $(\text{let } ((\text{compose } (\text{lambda } (f \ g) (\text{lambda } (x) (f \ (g \ x))))))$
 $((\text{compose not integer?}) \ 1))$

b. $\mathcal{MCP}S[(\text{begin } E_1 \ E_2)] \ m =$
 $(\text{begin } (\mathcal{MCP}S[E_1] \ (\lambda V . V)) (\mathcal{MCP}S[E_2] \ m))$

c. $T[E] =$

```
(let ((v P))
  (let ((r [v/P]E))
    (begin
      (%procedure-free v)
      r)))
```

d. One possible explanation is that the programs do not exhaust memory, and thus the garbage collector is never called. In that case, explicitly freeing unused closures is extra work that has no benefit. Another possibility is that in some programs, explicitly freeing closures releases enough memory so the garbage collector is no longer invoked, leading to storage fragmentation that slows the program down.