

CS 242 Midterm Exam

This is a *closed*-book exam. The maximum possible score is 100 points. Make sure you print your name legibly and sign the honor code below. All of the intended answers may be written within the space provided. You may use the back of the preceding page for scratch work. If you need to use the back side of a page to write part of your answer, be sure to mark your answer clearly.

The following is a statement of the Stanford University Honor Code:

- A. The Honor Code is an undertaking of the students, individually and collectively:
- (1) that they will not give or receive aid in examinations; that they will not give or receive unpermitted aid in class work, in the preparation of reports, or in any other work that is to be used by the instructor as the basis of grading;
 - (2) that they will do their share and take an active part in seeing to it that others as well as themselves uphold the spirit and letter of the Honor Code.
- B. The faculty on its part manifests its confidence in the honor of its students by refraining from proctoring examinations and from taking unusual and unreasonable precautions to prevent the forms of dishonesty mentioned above. The faculty will also avoid, as far as practicable, academic procedures that create temptations to violate the Honor Code.
- C. While the faculty alone has the right and obligation to set academic requirements, the students and faculty will work together to establish optimal conditions for honorable academic work.

I acknowledge and accept the Honor Code.

(Signature)

(Print your name)

Prob	# 1	# 2	# 3	# 4	# 5	# 6	# 7	# 8	# 9	# 10	Total
Score											
Max	10	9	12	7	6	10	12	10	16	8	100

1. (10 points) True or False

Mark each statement *true* or *false*, as appropriate.

- _____ (a) Allocation and deallocation of activation records is always done using a stack (LIFO) structure.
- _____ (b) Two function parameters passed by value are never aliased inside the function body.
- _____ (c) Two function parameters passed by reference are never aliased inside the function body.
- _____ (d) Two function parameters, one passed by value and the other by reference, are never aliased inside the function body.
- _____ (e) A dangling pointer refers to a region of memory that has been deallocated.
- _____ (f) Determining whether a C function has a side effect is undecidable.
- _____ (g) The tree reduce function (from homework) is a higher-order function.
- _____ (h) A value on the left-hand-side of an operator is called an L-value.
- _____ (i) In principle, it is possible to parallelize C code by first translating a C program to a pure functional language and then evaluating independent parts of the functional program in parallel.
- _____ (j) Java, C, and Lisp programming languages are Turing complete but ML is not.

2. (9 points) Terminology

Define the following terms.

- (a) (3 points) alpha-conversion

- (b) (3 points) polymorphic function

- (c) (3 points) garbage

3. (12 points) Short Answer

(a) (3 points) Write the “dotted pair” Lisp expression for the list (A B C D). Your answer might look something like ((X . Y) . (Z . W)).

(b) (3 points) Some variables in the following lambda expression are free and some are bound. Circle each bound variable and draw a line from the bound variable to the corresponding λ that binds it. For example, if you were asked to do this for $\lambda x.x+y$, you would circle x and draw a line from this circle to the letter λ at the beginning of the expression.

$(\lambda a. \lambda b. \lambda c. c (a b)) (\lambda a. \lambda c. c a c) a$

(c) (6 points)

In a language that does not provide automatic storage management, a *memory leak* occurs when a program no longer has a pointer to a heap location that has been allocated to it. The reason this is called a memory leak is that a program is supposed to free locations allocated to it from the heap. But if the program no longer has a pointer to a location, it may be impossible to deallocate the location.

Suppose you are asked to design a “leak detector” tool for C. Explain how you might use a garbage collection algorithm to detect potential memory leaks. Describe briefly:

- any additional data structures you require

- whether you can identify all leaks

- two features of the C language make this task difficult

4. (7 points) Symbolic evaluation of Lisp expression

Lisp expressions can be evaluated symbolically, using β -reduction and simple calculational properties of conditional, cond, and arithmetic. Evaluate the following expression to a number, showing your work. The first two steps are done for you.

```
((lambda (fact)           // function with one argument
  (fact fact 2))
 (lambda (ft k)          // this lambda function is the argument
  (cond ((= k 0) 1)
        (t (* k (ft ft (- k 1))))))
) )
```

= // after substituting lambda function for both fact references

```
((lambda (ft k)          // lambda function with 2 arguments
  (cond ((= k 0) 1)
        (t (* k (ft ft (- k 1))))))
)
(lambda (ft k)          // this lambda will be the first argument
  (cond ((= k 0) 1)
        (t (* k (ft ft (- k 1))))))
)
2 // this is the second argument
)
```

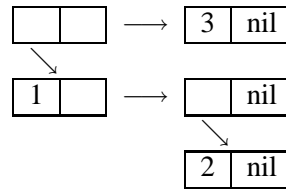
= // after applying the first & second arguments and resolving cond

```
(* 2
 (lambda (ft k)          // lambda function with 2 arguments
  (cond ((= k 0) 1)
        (t (* k (ft ft (- k 1))))))
)
(lambda (ft k)          // first argument
  (cond ((= k 0) 1)
        (t (* k (ft ft (- k 1))))))
)
1 // second argument
)
)
```

=

5. (6 points) Nested Lists in ML program

Lisp supports arbitrarily nested lists, like ((1 (2)) 3), which are represented internally by cons cells:



To support such nested lists in ML, we can define the following datatype:

```
datatype 'a lisplist = nil | atom of 'a | cons of 'a lisplist * 'a lisplist;
```

This allows us to simulate Lisp representations of nested lists:

```
(1) is cons(atom(1),nil)
(1 2) is cons(atom(1),cons(atom2,nil))
```

Flattening a list means removing all internal nesting, while presenting the atoms in the same order. For example, flattening the list ((1 (2)) 3) yields: (1 2 3).

Below is a skeleton for a flatten function in ML. Fill in the missing code. As an example of how the flatten function should behave, flatten(cons(cons(atom(1),nil),nil)) should produce: cons(atom(1), nil) That is, ((1)) is flattened to (1).

To simplify your task, you can assume that an concat function is provided, which takes a pair of lisplist as arguments and concatenates them, producing a single lisplist. That is, concat applied to (1 2) and (3 4) gives (1 2 3 4).

```
fun flatten (nil) =

| flatten (atom(X)) =

| flatten (cons(Y,Z)) =
```


7. (12 points) Denotational Semantics

This problem asks about a nonstandard semantics that characterizes a form of type analysis for expressions given by the following grammar:

$$e ::= 0 \mid 1 \mid \text{true} \mid \text{false} \mid x \mid e + e \mid e = e \mid \text{if } e \text{ then } e \text{ else } e \mid \text{fun } f(x:\sigma):\tau=e \text{ in } e \mid f(e)$$

In the fun...in... expression, which declares a function and provides a scope where the function can be called, the types σ and τ may be either *int* or *bool*. The meaning $\mathcal{V}[[e]](s)$ of an expression e depends on the state s . For the purpose of this problem, a state s is a mapping from variables to values, with function names considered as variables. In type analysis, we will use three basic type values and function types involving *int* and *bool*:

$$\text{Values} = \{ \text{int}, \text{bool}, \text{type_error}, \text{int} \rightarrow \text{int}, \text{int} \rightarrow \text{bool}, \text{bool} \rightarrow \text{int}, \text{bool} \rightarrow \text{bool} \}$$

Intuitively, $\mathcal{V}[[e]](s) = \text{int}$ means that the value of expression e is an integer (in state s) and $\mathcal{V}[[e]](s) = \text{type_error}$ means that evaluation of e may involve a type error.

Here are the semantic clauses for the first few expression forms.

$$\begin{aligned} \mathcal{V}[[0]](s) &= \text{int} \\ \mathcal{V}[[1]](s) &= \text{int} \\ \mathcal{V}[[\text{true}]](s) &= \text{bool} \\ \mathcal{V}[[\text{false}]](s) &= \text{bool} \end{aligned}$$

The value of a variable in some state is simply the value the state gives to that variable.

$$\mathcal{V}[[x]](s) = s(x)$$

For addition, the value will be *type_error* unless both operands are integers.

$$\mathcal{V}[[e_1 + e_2]](s) = \begin{cases} \text{int} & \text{if } \mathcal{V}[[e_1]](s) = \text{int} \text{ and } \mathcal{V}[[e_2]](s) = \text{int} \\ \text{type_error} & \text{otherwise} \end{cases}$$

For boolean equality, the value will be *type_error* unless both operands are integers or booleans.

$$\mathcal{V}[[e_1 = e_2]](s) = \begin{cases} \text{bool} & \text{if } \mathcal{V}[[e_1]](s) = \text{int} \text{ and } \mathcal{V}[[e_2]](s) = \text{int} \\ \text{bool} & \text{if } \mathcal{V}[[e_1]](s) = \text{bool} \text{ and } \mathcal{V}[[e_2]](s) = \text{bool} \\ \text{type_error} & \text{otherwise} \end{cases}$$

Since function declarations involve declaring the types of parameters and return values, and these types are recorded in a state, the semantics of declarations involves changing the state. Before giving the clause for function declaration, we define the notation $s[x \mapsto \sigma]$ for a state that is similar to s , except that it must map x to type σ . More precisely,

$$s[x \mapsto \sigma] = \lambda y \in \text{Variables}. \text{if } y = x \text{ then } \sigma \text{ else } s(y)$$

Using this notation, we can define the semantics of function declarations by

$$\mathcal{V}[[\text{fun } f(x:\sigma):\tau=e_1 \text{ in } e_2]](s) = \begin{cases} \mathcal{V}[[e_2]](s[f \mapsto (\sigma \rightarrow \tau)]) & \text{if } \mathcal{V}[[e_1]](s[x \mapsto \sigma]) = \tau \\ \text{type_error} & \text{otherwise} \end{cases}$$

Similarly, the semantics of function application are

$$\mathcal{V}[[f(e_1)]](s) = \begin{cases} \tau & \text{if } s(f) = \sigma \rightarrow \tau \text{ and } \mathcal{V}[[e_1]](s) = \sigma \\ \text{type_error} & \text{otherwise} \end{cases}$$

The clause for conditional is

$$\mathcal{V}[[\text{if } e_1 \text{ then } e_2 \text{ else } e_3]](s) = \begin{cases} \text{bool} & \text{if } \mathcal{V}[[e_1]](s) = \text{bool} \text{ and } \mathcal{V}[[e_2]](s) = \mathcal{V}[[e_3]](s) = \text{bool} \\ \text{int} & \text{if } \mathcal{V}[[e_1]](s) = \text{bool} \text{ and } \mathcal{V}[[e_2]](s) = \mathcal{V}[[e_3]](s) = \text{int} \\ \text{type_error} & \text{otherwise} \end{cases}$$

For example, using $s_0 = \lambda y \in \text{Variables}. \text{type_error}$,

$$\mathcal{V}[[\text{if true then } 0 + 1 \text{ else } x + 1]](s_0) = \text{type_error}$$

since the expression $x + 1$ produces a type error if evaluating x produces a type error.

Question:

(12 points) Show how to calculate the meaning of the expression

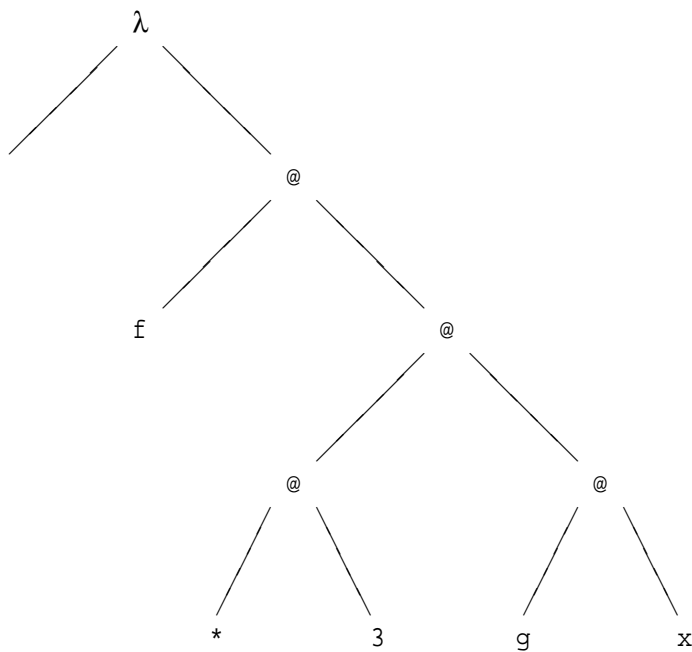
`fun f(x:int):bool=(x=0) in (if f(0) then true else false)`

in state $s_0 = \lambda y \in \text{Variables}. \text{type_error}$.

8. (10 points) Type inference on parse graph

Use the parse graph to perform type inference for the following ML function:

```
fun arith (f, g, x) = f(3 * g(x));
```



9. (16 points) Activation Records and Scope

This question asks about memory management in the evaluation of the following statically-scoped ML expression.

```

fun sum_compose(f, g) =
  let val x : int ref = ref 0
  in
    fn (y) => (x := (!x) + f(g(y)); !x)
  end;
val x = 2;
fun add_two(y) = y + x;
fun square(x) = x * x;
val comp = sum_compose(add_two, square);
comp(2);

```

(a) (13 points) Fill in the missing information in the following depiction of the run-time stack at the time when the call on line 10 of this code fragment is made (`comp(2)`). Remember that function values are represented by closures, and that a closure is a pair consisting of an environment (pointer to an activation record) and compiled code. Remember also that in ML, function arguments are evaluated before the function is called.

In this drawing, a bullet (●) indicates that a pointer should be drawn from this slot to the appropriate closure, compiled code or heap cell. Since the pointers to activation records cross and could become difficult to read, each activation record is numbered at the far left. In each activation record, place the number of the activation record of the statically enclosing scope in the slot labelled “access link.” Closures and heap cells are labeled with capital letters and compiled code with lowercase letters. Use activation record numbers for the environment pointer part of each closure pair, and write the lowercase letter corresponding to the compiled code for the second part of the pair. Write the values of local variables and function parameters in the activation records. The first two activation records are done for you.

	Activation Records	Closures and Cells	Compiled Code
(1)	access link (0) sum_compose A ●	(A)((1), i ●)	(i)sum_compose
(2)	access link (1) x 2		(j)add_two
(3)	access link () add_two ●	(B)((), ●)	(k)square
(4)	access link () square ●	(C)((), ●)	(l) $\lambda y \dots$
(5) comp = (...)	access link () comp ●	(D)((), ●)	
(6)	access link () f ● g ● x ●	(E) <input type="text"/>	
(7) comp(2)	access link () y		

(b) (3 points) What is the value of the expression `comp(2)`? Why?

10. (8 points) **Implementing Exceptions**

One way of implementing exceptions is to make a table mapping exception names to code for handlers, for each scope, and store this table on the run-time stack. This has little performance impact at run-time, unless an exception is raised, since the tables can be determined at compile time. However, when an exception is raised, there is some cost. Specifically, if the current activation record contains a handler, control is transferred to this handler. If not, then the exception will have to be “re-raised” in another scope.

- (a) (2 points) If an exception is raised and there is no handler in the current scope, which pointer in the activation record should be used to find the next scope?

- (b) (2 points) Can the compiler determine the number of pointers to follow, for a given exception and scope, at compile time? Explain why or why not.

- (c) (2 points) Optimizing compilers often change the order of instructions for various reasons. Why do languages with exceptions make this kind of optimization more difficult?

- (d) (2 points) What information would a compiler like to know, at compile time, about a given expression such as a function call, in order to reorder instructions?