

Name: _____

This exam has 12 questions, for a total of 100 points.

1. 5 points What is the observable behavior of the following \mathcal{L}_{Fun} program? (e.g. does it produce an error at compile time or runtime? does it produce an integer, which one? does it diverge?)

```
def f(x : Integer) -> tuple[int]
    v = (x, x)
    return v

v1 = f(3)
v2 = f(3)
v3 = v1
if v1 is v2:
    if v1 is v3:
        print(0)
    else:
        print(1)
else:
    if v1 is v3:
        print(2)
    else:
        print(3)
```

2. 5 points What is the observable behavior of the following \mathcal{L}_{Tup} program? (e.g. does it produce an error at compile time or runtime? does it produce an integer, which one? does it diverge?)

```
v1 = (3,3)
v2 = (3,3)
if v1[0] == v2[2]:
    print(0)
else:
    print(1)
```

Name: _____

3. 8 points Given the following $\mathcal{L}_{\text{While}}$ program, apply the Explicate Control pass to translate it to \mathcal{C}_{If} .

```
sum = 0
i = input_int()
while i > 0:
    sum = sum + i
    i = i - 1
tmp = 27 + sum
print(27 + sum)
```

Name: _____

4. 13 points Apply liveness analysis to the following pseudo-x86 program to determine the set of live locations before and after every instruction. (The callee and caller saved registers are listed in the Appendix of this exam.)

```
start:
    callq read_int
    movq %rax, x57
    movq $1, y58
    callq read_int
    movq %rax, i59
    jmp loop65

loop65:
    movq i59, tmp60
    cmpq $0, tmp60
    jg block67
    jmp block66

block66:
    movq y58, tmp64
    movq x57, %rax
    addq tmp64, %rax
    jmp conclusion

block67:
    movq y58, tmp61
    movq y58, tmp62
    movq tmp61, y58
    addq tmp62, y58
    movq i59, tmp63
    movq tmp63, i59
    subq $1, i59
    jmp loop65
```

Name: _____

5. 10 points Fill in the blanks for the following `expose_alloc_tuple` auxiliary function of the Expose Allocation pass that translates from $\mathcal{L}_{\text{Tuple}}$ to $\mathcal{L}_{\text{Alloc}}$. (The grammar for $\mathcal{L}_{\text{Alloc}}$ is in the Appendix.)

```
def expose_alloc_tuple(es, tupleType, allocExp):
    n = len(es)
    num_bytes =           (a)          
    vec = generate_name('alloc')
    space_left = Compare(          (b)          , [Lt()],
                        [GlobalValue('fromspace_end')])
    xs = [Name(generate_name('init')) for e in es]
    inits = [Assign([x], e) for (x,e) in zip(xs,es)]
    initVec = []
    i = 0
    for x in xs:
        initVec += [          (c)          ]
        i += 1
    return Begin(inits \
                + [If(space_left, [], [          (d)          ])] \
                + [Assign([Name(vec)], allocExp)] \
                + initVec,
                Name(vec))

def expose_alloc_exp(e: expr) -> expr:
    match e:
        case Tuple(es, Load()):
            alloc = Allocate(len(es), e.has_type)
            return expose_alloc_tuple(          (e)          , e.has_type, alloc)
    ...
```

Name: _____

6. 4 points In the `expose_alloc_tuple` function of the previous question, why are the initializing expressions `es` assigned to the temporary variables `xs` instead of using them directly in the tuple initialization?

7. 6 points Describe the layout of the 64-bit tag at the beginning of every tuple.

8. 10 points Fill in the blanks for the following `explicate_control` that translates $\mathcal{L}_{\text{FunRef}}^{\text{mon}}$ programs into \mathcal{C}_{Fun} programs.

```
def explicate_tail(e : expr, blocks: Dict[str, List[stmt]]) -> List[stmt]:
    match e:
        case Call(Name(f), args) if f in builtin_functions:
            return [ (a) ]
        case Call(func, args):
            return [ (b) ]
        ...

def explicate_pred(cnd: expr, thn: List[stmt], els: List[stmt],
                  basic_blocks: Dict[str, List[stmt]]) -> List[stmt]:
    match cnd:
        case Call(func, args):
            tmp = generate_name('call')
            return [ (c),
                    If(Compare(Name(tmp), [Eq()]), [Constant(False)]),
                      create_block(els, basic_blocks),
                      create_block(thn, basic_blocks))]
        ...

def explicate_stmt(s: stmt, cont: List[stmt],
                  blocks: Dict[str, List[stmt]]) -> List[stmt]:
    match s:
        case Return(value):
            return (d)
        ...

def explicate_def(d) -> stmt:
    match d:
        case FunctionDef(name, params, body, _, returns, _):
            new_body = []
            blocks = {}
            if isinstance(returns, VoidType):
                body = body + [Return(Constant(None))]
            for s in reversed(body):
                new_body = (e)
            blocks[label_name(name + '_start')] = new_body
            return FunctionDef(name, params, blocks, None, returns, None)
        ...
```

Name: _____

9. 6 points What is the purpose of the Reveal Functions pass? How does the output of Reveal Functions facilitate decisions made in later passes of the compiler?
10. 8 points Describe the general layout of the procedure call frame that your compiler uses.

11. 13 points Apply Instruction Selection to the following two functions, translating them from \mathcal{C}_{Fun} to $\text{x86}_{\text{callq}*}^{\text{Def}}$. (The definitions of \mathcal{C}_{Fun} and $\text{x86}_{\text{callq}*}^{\text{Def}}$ are in the Appendix, as is the list of argument-passing registers.) (The function `even` calls `odd`, but you do not need to translate the `odd` function for this exam question, so its definition is omitted.) (The below functions are presented in concrete syntax except for the `FunRef` nodes.)

```
def even(x : int) -> bool:
    even_start:
        if x == 0:
            goto block.146
        else:
            goto block.147
    block.146:
        return True
    block.147:
        fun.138 = FunRef(odd, 1)
        tmp.139 = (x - 1)
        tail fun.138(tmp.139)

def main() -> int:
    main_start:
        fun.142 = FunRef(even, 1)
        tmp.143 = input_int()
        call.152 = fun.142(tmp.143)
        if call.152 == False:
            goto block.153
        else:
            goto block.154
    block.153:
        tmp.144 = 0
        goto block.151
    block.154:
        tmp.144 = 42
        goto block.151
    block.151:
        print(tmp.144)
        return 0
```


12. 12 points Draw the interference graph for the following program fragment by adding edges between the nodes below. You do not need to include edges between two registers. The live-after set for each instruction is given to the right of each instruction and the types of each variable is listed below. (The callee and caller saved registers are listed in the Appendix of this exam.)

```
def main() -> int:
    var_types:
        tmp73 : int, tmp72 : int, tmp75 : Callable[[int], int],
        tmp74 : Callable[Callable[[int], int], tuple[int, int]], Void), tmp71 : int,
        _65 : int, _64 : int, _66 : int, alloc63 : tuple[int, int],
        vec62 : tuple[int,int]

    mainstart:
        {}
    movq free_ptr(%rip), tmp71
        {tmp71}
    movq tmp71, tmp72
        {tmp72}
    addq $24, tmp72
        {tmp72}
    movq fromspace_end(%rip), tmp73
        {tmp72 tmp73}
    cmpq tmp73, tmp72
        {}
    jl block77
        {}
    jmp block78
        {}

    block77:
        {}
    movq $0, _66
        {}
    jmp block76
        {}

    block78:
        {}
    movq %r15, %rdi
        {rdi}
    movq $24, %rsi
        {rdi rsi}
    callq collect
        {}
    jmp block76
        {}

    block76:
        {}
    movq free_ptr(%rip), %r11
        {}
    addq $24, free_ptr(%rip)
        {}
    movq $5, 0(%r11)
        {r11}
    movq %r11, alloc63
        {alloc63}
    movq alloc63, %r11
        {alloc63}
    movq $0, 8(%r11)
        {alloc63}
    movq $0, _65
        {alloc63}
    movq alloc63, %r11
        {alloc63}
    movq $41, 16(%r11)
        {alloc63}
    movq $0, _64
        {alloc63}
    movq alloc63, vec62
        {vec62}
    leaq map_vec_57(%rip), tmp74
        {vec62 tmp74}
    leaq add158(%rip), tmp75
        {vec62 tmp74 tmp75}
    movq tmp75, %rdi
        {tmp74 rdi vec62}
    movq vec62, %rsi
        {rsi tmp74 rdi vec62}
    callq *tmp74
        {vec62}
    movq vec62, %r11
        {r11}
    movq 16(%r11), %rax
        {rax}
    jmp mainconclusion
        {rax}
)
```

Name: _____

Appendix

The caller-saved registers are:

```
rax rcx rdx rsi rdi r8 r9 r10 r11
```

and the callee-saved registers are:

```
rsp rbp rbx r12 r13 r14 r15
```

The argument-passing registers are:

```
rdi rsi rdx rcx r8 r9
```

Grammar for $\mathcal{L}_{\text{While}}$

```

exp ::= Constant(int) | Call(Name('input_int'), [])
      | UnaryOp(USub(), exp) | BinOp(exp, Add(), exp)
      | BinOp(exp, Sub(), exp)
stmt ::= Expr(Call(Name('print'), [exp])) | Expr(exp)
-----
exp ::= Name(var)
stmt ::= Assign([Name(var)], exp)
-----
boolop ::= And() | Or()
unaryop ::= Not()
cmp ::= Eq() | NotEq() | Lt() | LtE() | Gt() | GtE()
bool ::= True | False
exp ::= Constant(bool) | BoolOp(boolop, [exp, exp])
      | Compare(exp, [cmp], [exp]) | IfExp(exp, exp, exp)
stmt ::= If(exp, stmt+, stmt+)
-----
stmt ::= While(exp, stmt+, [])
 $\mathcal{L}_{\text{While}}$  ::= Module(stmt*)

```

Grammar for \mathcal{C}_{if}

```

atm ::= Constant(int) | Name(var) | Constant(bool)
exp ::= atm | Call(Name('input_int'), []) | UnaryOp(USub(), atm)
      | BinOp(atm, Sub(), atm) | BinOp(atm, Add(), atm)
      | Compare(atm, [cmp], [atm])
stmt ::= Expr(Call(Name('print'), [atm])) | Expr(exp)
      | Assign([Name(var)], exp)
tail ::= Return(exp) | Goto(label)
      | If(Compare(atm, [cmp], [atm]), [Goto(label)], [Goto(label)])
 $\mathcal{C}_{\text{if}}$  ::= CProgram({label: [stmt, ..., tail], ...})

```

Grammar for $\mathcal{L}_{\text{Tuple}}$

```

exp ::= Constant(int) | Call(Name('input_int'), [])
    | UnaryOp(USub(), exp) | BinOp(exp, Add(), exp)
    | BinOp(exp, Sub(), exp)
stmt ::= Expr(Call(Name('print'), [exp])) | Expr(exp)
-----
exp ::= Name(var)
stmt ::= Assign([Name(var)], exp)
-----
boolop ::= And() | Or()
unaryop ::= Not()
cmp ::= Eq() | NotEq() | Lt() | LtE() | Gt() | GtE()
bool ::= True | False
exp ::= Constant(bool) | BoolOp(boolop, [exp, exp])
    | Compare(exp, [cmp], [exp]) | IfExp(exp, exp, exp)
stmt ::= If(exp, stmt+, stmt+)
-----
stmt ::= While(exp, stmt+, [])
-----
cmp ::= Is()
exp ::= Tuple(exp+, Load()) | Subscript(exp, Constant(int), Load())
    | Call(Name('len'), [exp])
 $\mathcal{L}_{\text{Tuple}}$  ::= Module(stmt*)

```

Grammar for $\mathcal{L}_{\text{Alloc}}$

The $\mathcal{L}_{\text{Alloc}}$ language extends $\mathcal{L}_{\text{Tuple}}$ with the following grammar rules:

```

exp ::= Collect(int) | Allocate(int, type) | GlobalValue(name)
stmt ::= Assign([Subscript(exp, int, Store())], exp)

```

Name: _____

Grammar for \mathcal{L}_{Fun}

```

exp ::= Constant(int) | Call(Name('input_int'), [])
      | UnaryOp(USub(), exp) | BinOp(exp, Add(), exp)
      | BinOp(exp, Sub(), exp)
stmt ::= Expr(Call(Name('print'), [exp])) | Expr(exp)
-----
exp ::= Name(var)
stmt ::= Assign([Name(var)], exp)
-----
boolop ::= And() | Or()
unaryop ::= Not()
cmp ::= Eq() | NotEq() | Lt() | LtE() | Gt() | GtE()
bool ::= True | False
exp ::= Constant(bool) | BoolOp(boolop, [exp, exp])
      | Compare(exp, [cmp], [exp]) | IfExp(exp, exp, exp)
stmt ::= If(exp, stmt+, stmt+)
-----
stmt ::= While(exp, stmt+, [])
-----
cmp ::= Is()
exp ::= Tuple(exp+, Load()) | Subscript(exp, Constant(int), Load())
      | Call(Name('len'), [exp])
-----
type ::= IntType() | BoolType() | VoidType() | TupleType[type+]
      | FunctionType(type*, type)
exp ::= Call(exp, exp*)
stmt ::= Return(exp)
params ::= (var, type)*
def ::= FunctionDef(var, params, stmt+, None, type, None)
 $\mathcal{L}_{\text{Fun}}$  ::= Module([def ... stmt ...])

```

Name: _____

Grammar for $\mathcal{L}_{\text{FunRef}}^{\text{mon}}$

```

atm ::= Constant(int) | Name(var)
exp ::= atm | Call(Name('input_int'), [])
      | UnaryOp(USub(), atm) | BinOp(atm, Add(), atm)
      | BinOp(atm, Sub(), atm)
stmt ::= Expr(Call(Name('print'), [atm])) | Expr(exp)
       | Assign([Name(var)], exp)
-----
atm ::= Constant(bool)
exp ::= Compare(atm, [cmp], [atm]) | IfExp(exp, exp, exp)
      | Begin(stmt*, exp)
stmt ::= If(exp, stmt*, stmt*)
-----
stmt ::= While(exp, stmt+, [])
-----
exp ::= Subscript(atm, atm, Load())
      | Call(Name('len'), [atm])
      | Allocate(int, type) | GlobalValue(var)
stmt ::= Assign([Subscript(atm, atm, Store())], atm)
       | Collect(int)
-----
type ::= IntType() | BoolType() | VoidType() | TupleType[type+]
       | FunctionType(type*, type)
exp ::= FunRef(label, int) | Call(atm, atm*)
stmt ::= Return(exp)
params ::= (var, type)*
def ::= FunctionDef(var, params, stmt+, None, type, None)
 $\mathcal{L}_{\text{FunRef}}^{\text{mon}}$  ::= Module([def ... stmt ...])

```

Grammar for \mathcal{C}_{Fun}

```

atm ::= Constant(int) | Name(var) | Constant(bool)
exp ::= atm | Call(Name('input_int'), []) | UnaryOp(USub(), atm)
      | BinOp(atm, Sub(), atm) | BinOp(atm, Add(), atm)
      | Compare(atm, [cmp], [atm])
stmt ::= Expr(Call(Name('print'), [atm])) | Expr(exp)
       | Assign([Name(var)], exp)
tail ::= Return(exp) | Goto(label)
       | If(Compare(atm, [cmp], [atm]), [Goto(label)], [Goto(label)])
-----
exp ::= Subscript(atm, atm, Load()) | Allocate(int, type)
      | GlobalValue(var) | Call(Name('len'), [atm])
stmt ::= Collect(int)
       | Assign([Subscript(atm, atm, Store())], atm)
-----
exp ::= FunRef(label, int) | Call(atm, atm*)
tail ::= TailCall(atm, atm*)
params ::= [(var, type), ...]
block ::= label: stmt* tail
blocks ::= {block, ...}
def ::= FunctionDef(label, params, blocks, None, type, None)
 $\mathcal{C}_{\text{Fun}}$  ::= CProgramDefs([def, ...])

```

Name: _____

Grammar for $x86_{callq}^{Def}$

```
arg ::= Constant(int) | Reg(reg) | Deref(reg,int) | ByteReg(reg)
      | Global(label) | FunRef(label, int)
instr ::= ... | IndirectCallq(arg, int) | TailJump(arg, int)
      | Instr('leaq', [arg, Reg(reg)])
block ::= label: instr*
blocks ::= {block, ...}
def ::= FunctionDef(label, [], blocks, -, type, -)
 $x86_{callq}^{Def}$  ::= X86ProgramDefs([def, ...])
```