2. 雛形具体化

雛形具体化を用いるグラフ簡約器

2.1 雛形具体化

Implementation of Functional Programming Languageの11章、12章

- 関数プログラムは**式を評価**によって実行される
- 式は**グラフ**で表現される
- 評価は一連の簡約を行うことで実施される
- 簡約はグラフ中の簡約可能式(redex)を簡約しその結果で置き換える
- 評価は対象とする式が**正規形**(normal form)になれば終了
- 簡約系列な複数ありうるが停止したときには同じ正規形になる
- 正規形に到達する簡約系列があれば最外簡約戦略で必ず停止する

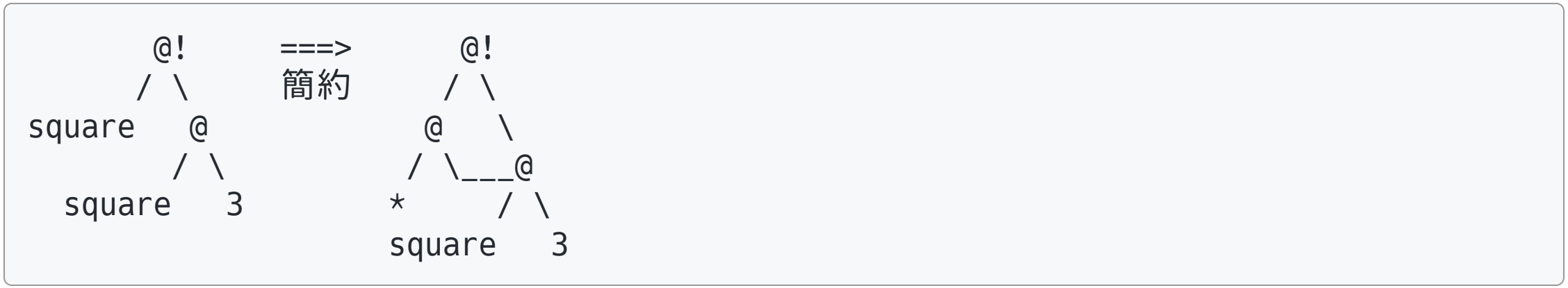
簡約例

```
main = square (square 3)
square x = x * x
```

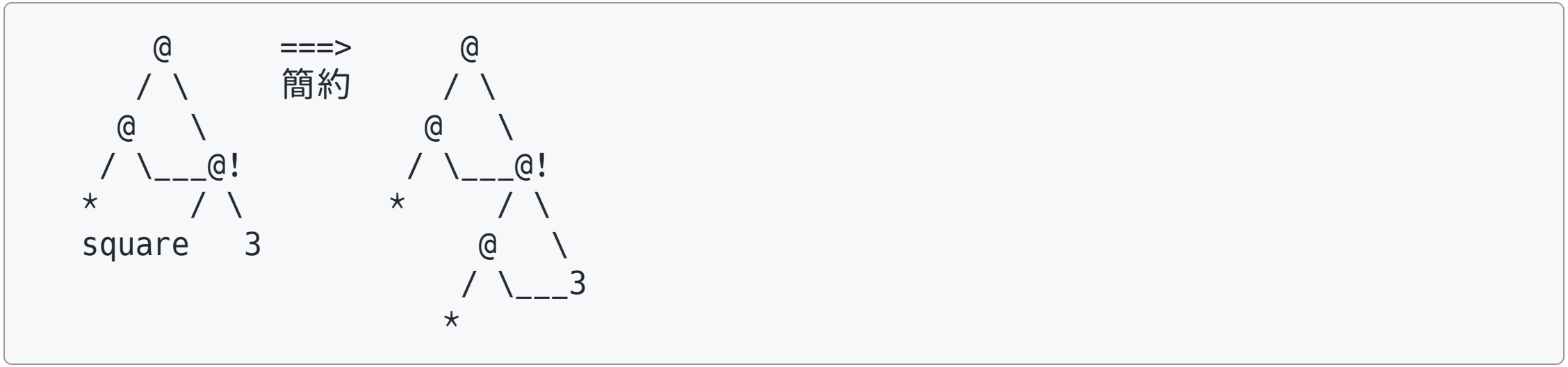
スーパーコンビネータ `main` には引数はなく、それ自身redexなので、ボディ部と置き換える。

```
main          ===>      @
              簡約     / \
                 square   @
                         / \
                   square   3
```

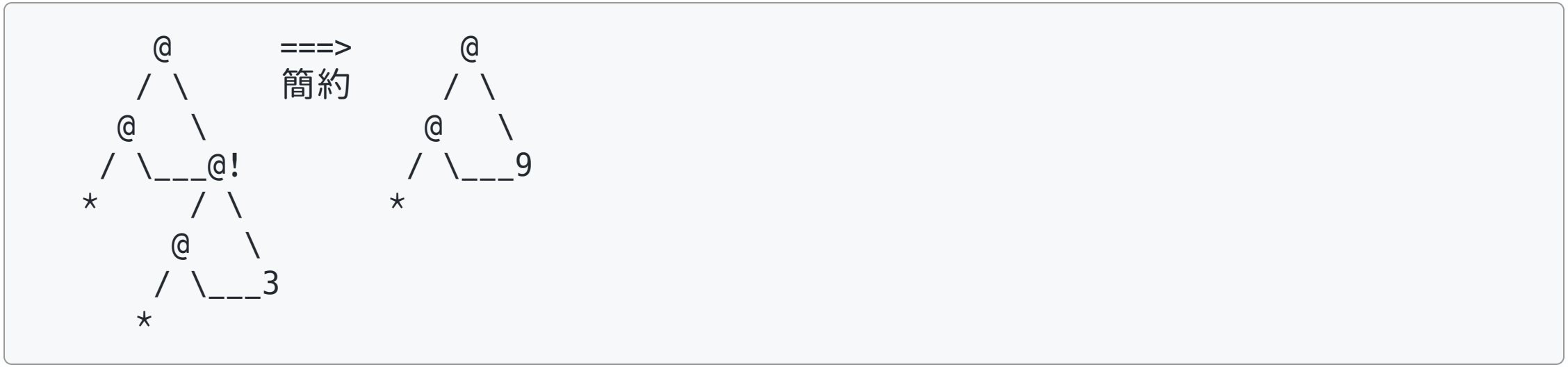
最外のredexは`square`の適用式。関数本体を具体化したものでredexを置き換える。仮引数の各出現を引数へのポインタで置き換える。



ここでredexは内側の `square` の適用式だけなので、これを簡約する。



ここで内側のかけ算が、唯一のredexとなるのでこれを簡約する。



最後の簡約は簡単。

2.1.2 簡約の3ステップ

以下を正規形が得られるまで繰り返す

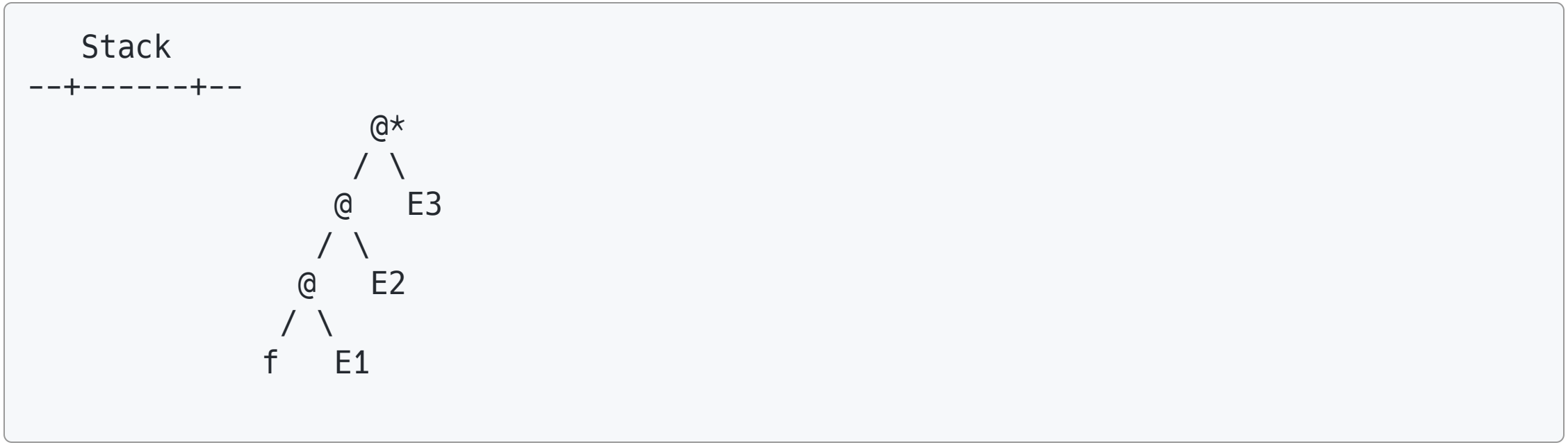
1. 次に簡約するredexを見つける
2. そのredexを簡約する
3. redexのルートを結果で更新する

最外の関数適用が

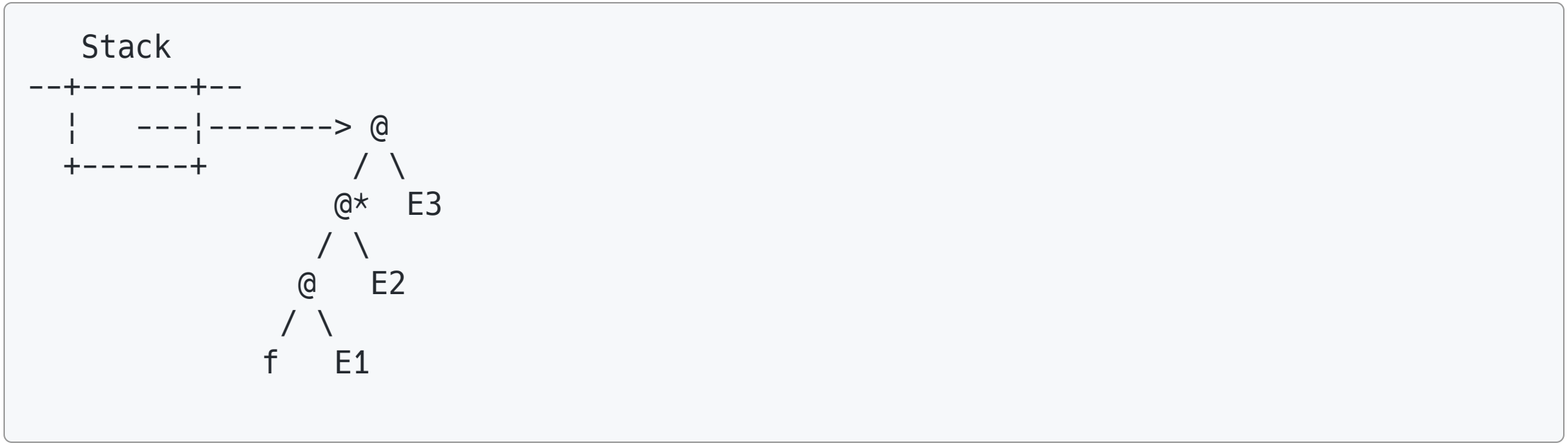
- スーパーコンビネータ適用の場合
  - この適用は必ずredexなので簡約（β簡約）
- 組込みのプリミティブ適用の場合
  - 2つの引数が共に正規形ならこの適用はredexなので簡約（δ簡約）
  - そうでないなら、引数を正規形にして（この適用がredexになって）から簡約（δ簡約）

2.1.3 背骨を巻き戻して次のredexをみつける

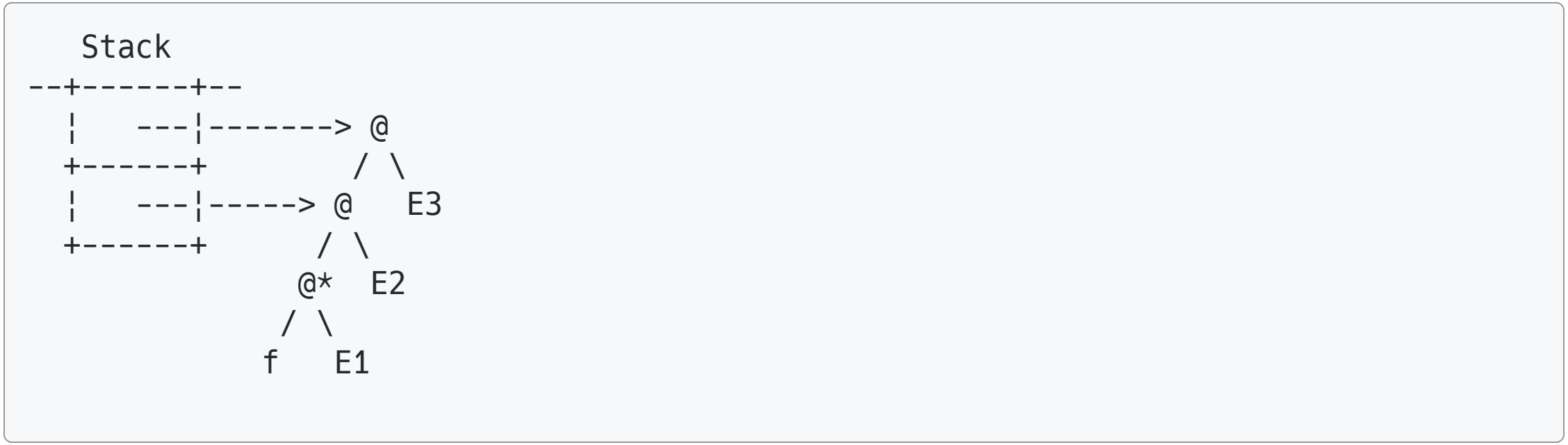
ルートから適用ノードの左側を辿る



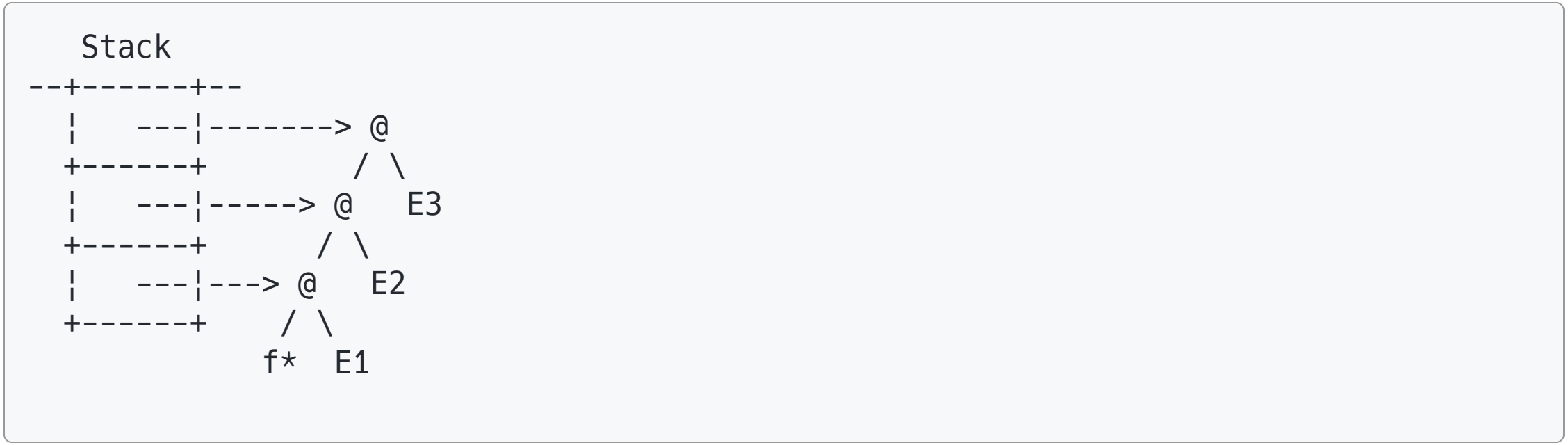
ルートの適用ノードをスタックに積んで、左へ降りる



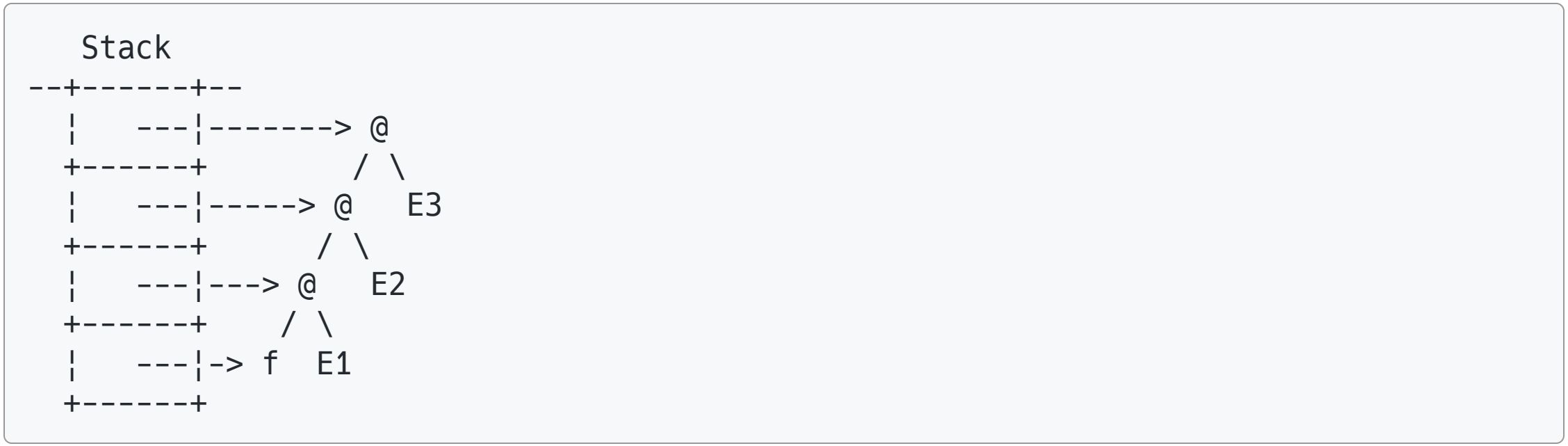
適用ノードをスタックに積んで、左に降りる



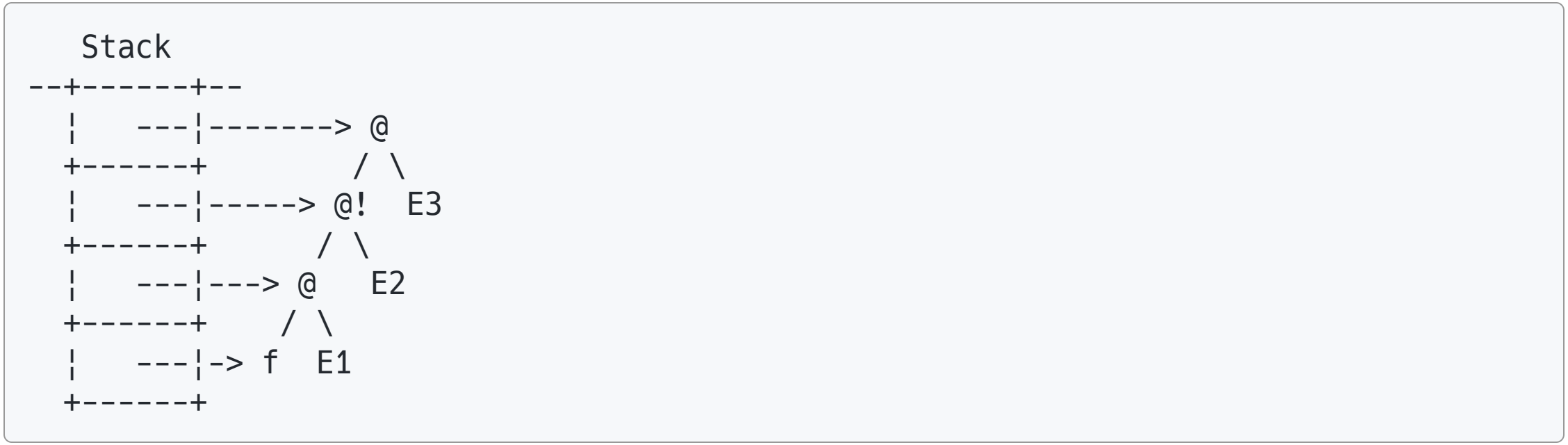
適用ノードをスタックに積んで、左に降りる



`f`がスーパーコンビネータの場合： `f` をスタックに積んで、 `f` のアリティ（ここでは2とする）を確認



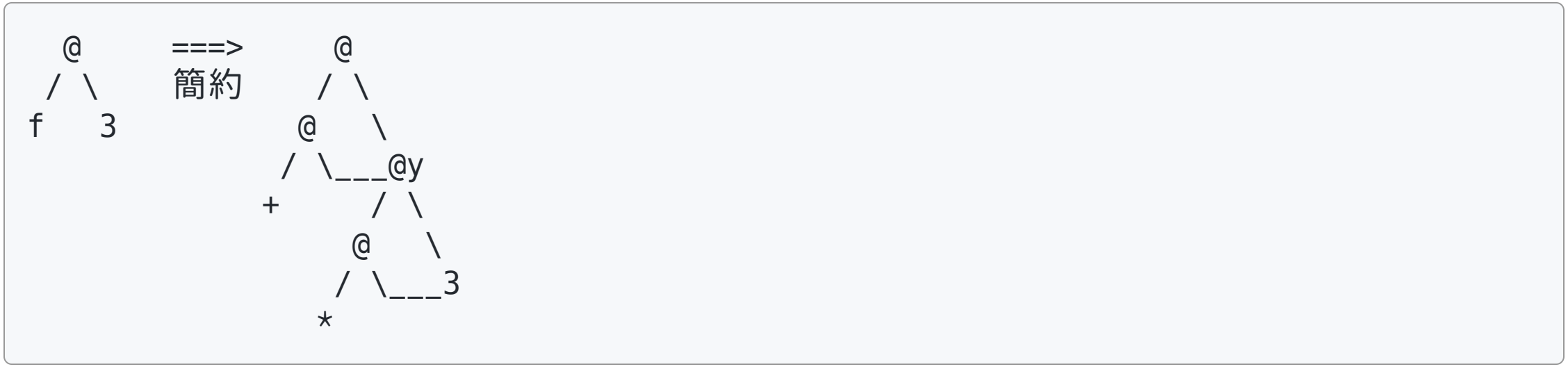
`f` がスーパーコンビネータの場合： アリティ（ここでは2とする）の分だけノードを上へもどったところが、最外の簡約可能項のルートノード



`f` がプリミティブ（算術2項演算子）の場合： 被演算子 `E1` および `E2` が共に正規形なら `!` のついたノードが最外の簡約可能項のルートノード。さもなければ、先に被演算子項を正規形にまで簡約する。

2.1.4 スーパーコンビネータの簡約可能項

- スーパーコンビネータの簡約可能項は、仮パラメータの出現位置を対応する実引数項へのポインタに置き換えたスーパーコンビネータ本体で置き換える
- 実引数項をコピーせず、ポインタを使い、共有していることに注意



2.1.5 更新

```
id x = x
f p = (id p) * p
main = f (sqrt 4)

f の簡約後

     @       ===>       @
    / \                 / \
   @   \               @   \
  / \   \             / \   \
 *   @   \           *   #   \
    / \___@             \___@
  id     / \             / \
      sqrt   4         sqrt   4
```

`#` は間接参照ノード

2.1.6 CAF

```
fac20 = factorial 20
```

スーパーコンビネータ `fac20` は CAF でかつ redex のルートなので、 `fac20` の簡約結果で上書きする

2.2 状態遷移システム

かけ算機械

```haskell
type MultState = (Int, Int, Int, Int) -- ^ (n, m, d, t)

evalMult :: MultState -> [MultState]
evalMult state = if multFinal state
  then [state]
  else state : evalMult (stepMult state)

stepMult (n, m, d, t)
  | d >  0 = (n,   m, d-1, t+1)  -- ^ 規則 1
  | d == 0 = (n, m-1,   n, t  )  -- ^ 規則 2
```

練習問題 2.1

かけ算マシンを手で走らせよ。初期状態 (2,3,0,0) からはじめ、各ステップで発火する規則を特定し、最終状態が (2,0,0,6) であることを確かめよ。

練習問題 2.2

状態列の不変条件とは、すべての状態で真となる述語である。 n および m の初期値 N および M と現在の n、m、d、t の値のとの関係を見つけよ。これにより、このかけ算機械が、かけ算を実行するものであることを証明せよ。すなわち、以下を示せ。

1. 初期状態で不変条件が成り立つ
2. ある状態で、不変条件が成り立てば、次の状態でも不変状態が成り立つ
3. 不変条件と停止条件が成り立てば、$t = N \times M$ である
4. このかけ算機械は停止する

状態遷移システムは、以下の点で便利

- 低レベルの詳細にわずらわされない程度に抽象的
- 隠れた詳細に依存していないことが確認できる程度に具体的
- 状態遷移システムは直截に実行可能なHaskellのコードに変換できる

練習問題 2.3

状態が最終状態であるかを判断する述語 `multFinal :: multState -> Bool` を定義し、初期状態 `(2,3,0,0)` からかけ算機械を走らせると、最終状態が `(2,0,0,6)` になることを示せ。

2.3 Mark 1: 最小雛形具体化グラフ簡約器

マシン状態： $(stack, dump, heap, globals)$ の4つ組

- $stack$: ヒープ上のノードを特定するアドレスのスタック
  - $a_1 : s$ という記法は、a_1 がスタックトップ、s がのこりのスタックであることを示す
- $dump$: 正格なプリミティブ演算の引数評価に先立ち、スパインのスタックを記録
- $heap$: タグ付きノードを集めたもの
  - $h[a : node]$ という記法は、ヒープ h において、a はノード$node$のアドレスであることを示す
- $globals$: スーパーコンビネータおよびプリミティブを表すノードへのアドレス

ノードの表現

```
data Node
  = NAp Addr Addr
  | NSupercomb Name [Name] CoreExpr
  | NNum Int
```

- `NAp` a_1 a_2 はアドレスa_1にあるノードのアドレスa_2にあるノードへの適用を表す
- `NSupercomb` $args$ $body$ は引数$args$と本体$body$をもつスーパーコンビネータを表す
- `NNum` n は整数nを表す

状態遷移規則

(2.1)

$$\begin{array}{rcccc} & a : s & d & h[a : \mathtt{NAp}\ a_1\ a_2] & f \\ \Longrightarrow \quad a_1 : a : s & d & h & & f \end{array}$$

状態遷移規則

(2.2)

$$\begin{array}{lrlll} & a_0 : a_1 : \cdots : a_n : s & d & h[a_0 : \texttt{NSupercomb}\ [x_1, \ldots, x_n]\ body] & f \\ \Longrightarrow & a_r : s & d & h' & f \end{array}$$

ここで、$(h', a_r) = instantiate\ body\ h\ f[x_1 \mapsto a_1, \ldots, x_n \mapsto a_n]$

関数 $instatiate$ の引数は、

1. 具体化する式
2. ヒープ
3. 名前からヒープ上のアドレスへのグローバルマッピング f をスタックにある引数名からヒープアドレスへのマッピングで拡張したもの

返り値は、

- 新しいヒープと、新しく構成されたインスタンスの（ルートノードの）アドレス

2.3.2 実装の構造

```
run :: String -> String
run = showResult . eval . compile . parse
```

1. `parse` はソースコード(`:: String`)を構文解析して `CoreProgram` を構成する関数（教科書とは意味を変更）

```
parse :: String -> CoreProgram
```

2. `compile` は `CoreProgram` を雛形具体化機械の初期状態に変換

```
compile :: CoreProgram -> TiState
```

3. `eval` はプログラム実行関数、初期状態から状態遷移を繰り返し、最終状態にまで遷移させ、結果は通過したすべての状態のリスト

```
eval :: TiState -> [TiState]
```

4. `showResult` は最終結果を整形して表示する

```
showResult :: [TiState] -> String
```

2.3.3 パーザ

`Language` モジュールをインポートする

2.3.4 コンパイラ

`TiState`

```
type TiState = (TiStack, TiDump, TiHeap, TiGlobals, TiStats)
```

1. `TiStack` スパインスタック、ヒープアドレス `Addr` のスタック

```
type TiStack = [Addr]
```

2. `TiDump` ダンプ、2.6節までは不要。ダミー

```
data TiDump = DummyTiDump
initalTiDump :: TiDump
initalTiDump = DummyTiDump
```

3. `TiHeap` は `Node` を格納するヒープ

```
type TiHeap = Heap Node
type Heap a = (Int, [Addr], [(Addr, a)])
```

`Heap` は使用アドレス数、未使用アドレス集合、アドレスと内容の2つ組のリスト、の3つを組にしたもの

4. `TiGlobal` はスーパーコンビネータ名とその定義が納められているヒープ上のアドレスの連想リスト

```
type TiGlobals = Assoc Name Addr
type Assoc a b = [(a, b)]
```

5. `TiStats` 実行時性能統計のためのデータ収集用、ひとまずステップカウント

```
type TiStats = Int

tiStatInitial :: TiStats
tiStatInitial = 0

tiStatIncSteps :: TiStats -> TiStats
tiStatIncSteps s = s + 1
tiStatGetSteps :: TiStats -> Int
tiStatGetSteps s = s

applyToStats :: (TiStats -> TiStats) -> TiState -> TiState
applyToStats f (stack, dump, heap, scDefs, stats)
  = (stack, dump, heap, scDefs, f stats)
```

`compile`

```
compile program
  = (initialStack, initalTiDump, initialHeap, globals, tiStatInitial)
    where
      scDefs = program ++ preludeDefs ++ extraPreludeDefs
      (initialHeap, globals) = buildInitialHeap scDefs
      initialStack = [addressOfMain]
      addressOfMain = aLookup globals "main" (error "main is not defined")

extraPreludeDefs :: CoreProgram
extraPreludeDefs = []
```

`buildInitialHeap` プログラムから `NSupercomb` ノードを含むヒープと、スーパーコンビネータ名とヒープ上のアドレスの対応を示す連想リストを構成する。

```
buildInitialHeap :: [CoreScDefn] -> (TiHeap, TiGlobals)
buildInitialHeap scDefs = mapAccumL allocateSc hInitial scDefs

allocateSc :: TiHeap -> CoreScDefn -> (TiHeap, (Name, Addr))
allocateSc heap scDefn = case scDefn of
  (name, args, body) -> (heap', (name, addr))
    where
      (heap', addr) = hAlloc heap (NSupercomb name args body)
```

2.3.5 評価器

```
eval state = state : restStates
  where
    restStates
      ¦ tiFinal state = []
      ¦ otherwise     = eval nextState
    nextState = doAdmin (step state)

doAdmin :: TiState -> TiState
doAdmin state = applyToStats tiStatIncSteps state
```

最終状態の判定

- 計算はスタックに単一の数またはデータオブジェクトが含まれる状態になったときにのみ停止する

```
tiFinal :: TiState -> Bool
tiFinal state = case state of
  ([soleAddr], _, heap, _, _) -> isDataNode (hLookup heap soleAddr)
  ([], _, _, _, _)            -> error "Empty stack!"
  _                           -> False

isDataNode :: Node -> Bool
isDataNode node = case node of
  NNum _ -> True
  _      -> False
```

`step` ある状態から1つ次の状態への遷移

```
step :: TiState -> TiState
step state = case state of
  (stack, dump, heap, globals, stats) -> dispatch (hLookup heap (head stack))
  where
    dispatch (NNum n)                  = numStep state n
    dispatch (NAp a1 a2)               = apStep  state a1 a2
    dispatch (NSupercomb sc args body) = scStep  state sc args body
```

`NNum` ノードと `NAp` ノードはやさしい

```
numStep :: TiState -> Int -> TiState
numStep state n = error "Number applied as a function"

apStep :: TiState -> Addr -> Addr -> TiState
apStep state a1 a2 = case state of
  (stack, dump, heap, globals, stats) -> (a1:stack, dump, heap, globals, stats)
```

スーパーコンビネータの適用：

1. 本体を具体化、引数名をスタックにあるアドレスと結びつける（規則2.2）
2. 簡約可能項のルートを含む引数をスタックから除去、簡約結果をスタックにプッシュ

（Mark 1 では更新は行わない）

```
scStep :: TiState -> Name -> [Name] -> CoreExpr -> TiState
scStep state scName argNames body = case state of
  (stack, dump, heap, globals, stats)
    -> (stack', dump, heap', globals, stats)
    where
      stack' = resultAddr : drop (length argNames + 1) stack
      (heap', resultAddr) = instantiate body heap env
      env = argBindings ++ globals
      argBindings = zip argNames (getargs heap stack)
```

`getargs` スタックにある引数名に対応するヒープ上の実引数ノードのアドレスを取得する

```
getargs :: TiHeap -> TiStack -> [Addr]
getargs heap stack = case stack of
  sc:stack' -> map getarg stack'
    where
      getarg addr = arg
        where
          NAp fun arg = hLookup heap addr
  []        -> error "Empty stack"
```

`instantiate`

```
instantiate :: CoreExpr          -- Body of suprercombinator
            -> TiHeap            -- Heap before instatiation
            -> Assoc Name Addr   -- Association of names to address
            -> (TiHeap, Addr)    -- Heap after instatiation, and address of root of instance
instantiate expr heap env = case expr of
  ENum n                -> hAlloc heap  (NNum n)
  EAp e1 e2             -> hAlloc heap2 (NAp a1 a2)
    where
      (heap1, a1) = instantiate e1 heap  env
      (heap2, a2) = instantiate e2 heap1 env
  EVar v                -> (heap, aLookup env v (error ("Undefined name " ++ show v)))
  EConstr tag arity     -> instantiateConstr tag arity heap env
  ELet isrec defs body -> instantiateLet isrec defs body heap env
  ECase e alts          -> error "Can't instantiate case exprs"
  ELam vs e             -> error "Can't instantiate lambda abstractions"
```

2.3.6 結果の整形

`showResults`

```
showResults :: [TiState] -> String
showResults states
  = iDisplay (iConcat [ iLayn (map showState states)
                      , showStats (last states)
                      ])
```

`showState`

```
showState :: TiState -> IseqRep
showState (stack, dump, heap, globals, stats)
  = iConcat [ showStack heap stack, iNewline ]
```

`showStack`

```
showStack :: TiHeap -> TiStack -> IseqRep
showStack heap stack
  = iConcat
    [ iStr "Stack ["
    , iIndent (iInterleave iNewline (map showStackItem stack))
    , iStr " ]"
    ]
    where
      showStackItem addr
        = iConcat [ showFWAddr addr, iStr ": "
                  , showStkNode heap (hLookup heap addr)
                  ]
```

`showStkNode`

```
showStkNode :: TiHeap -> Node -> IseqRep
showStkNode heap (NAp funAddr argAddr)
  = iConcat [ iStr "NAp ", showFWAddr funAddr
            , iStr " ", showFWAddr argAddr, iStr " ("
            , showNode (hLookup heap argAddr), iStr ")"
            ]
showStkNode heap node = showNode node
```

`showNode`

```
showNode :: Node -> IseqRep
showNode node = case node of
  NAp a1 a2 -> iConcat [ iStr "NAp ", showAddr a1
                       , iStr " ",    showAddr a2
                       ]
  NSupercomb name args body
            -> iStr ("NSupercomb " ++ name)
  NNum n    -> iStr "NNum " `iAppend` iNum n
```

`showAddr`

```
showAddr :: Addr -> IseqRep
showAddr addr = iStr (showaddr addr)

showFWAddr :: Addr -> IseqRep
showFWAddr addr = iStr (space (4 - length str) ++ str)
  where
    str = show addr
```

`showStats`

```
showStats :: TiState -> IseqRep
showStats (stack, dump, heap, globals, stats)
  = iConcat [ iNewline, iNewline, iStr "Total number of steps = "
            , iNum (tiStatGetSteps stats)
            ]
```

練習問題 2.4

ここまでの実装をテストせよ

```
testProg0, testProg1, testProg2 :: String
testProg0 = "main = S K K 3"
testProg1 = "main = S K K" -- wrong (not saturated)
testProg2 = "id = S K K ;\n\
            \main = twice twice twice id 3"

test :: String -> IO ()
test = putStrLn . showResults . eval . compile . parse
```

練習問題 2.5

`showState` を改変してヒープの内容をすべて表示するようにせよ。

```
showState :: TiState -> IseqRep
showState (stack, dump, heap, globals, stats)
  = iConcat [ showStack heap stack, iNewline
            , showHeap heap, iNewline
            ]

showHeap :: TiHeap -> IseqRep
showHeap heap = case heap of
  (_,_,contents) -> iConcat
    [ iStr "Heap  ["
    , iIndent (iInterleave iNewline (map showHeapItem contents))
    , iStr " ]"
    ]
  where
    showHeapItem (addr, node)
      = iConcat [ showFWAddr addr, iStr ": "
                , showNode node
                ]
```

練習問題 2.6

`scStep` が引数が充満でない場合に適切なエラーを表示するようにせよ。

```
scStep :: TiState -> Name -> [Name] -> CoreExpr -> TiState
scStep state scName argNames body = case state of
  (stack, dump, heap, globals, stats)
    | length stack < length argNames + 1
      -> error "Too few argments given"
    | otherwise
      -> (stack', dump, heap', globals, stats)
    where
      stack' = resultAddr : drop (length argNames + 1) stack
      (heap', resultAddr) = instantiate body heap env
      env = argBindings ++ globals
      argBindings = zip argNames (getargs heap stack)
```

練習問題 2.7

より多くの実行時情報を収集するようにせよ。

- プリミティブの簡約とスーパーコンビネータの簡約を分けて計数
- ヒープ操作（とくにアロケーション）の回数
- スタックの最大深さ

練習問題 2.8

`instantiate`に渡す環境`env`は、

```
env = argBindings ++ globals
```

と定義されているが、これを

```
env = globals ++ argBindings
```

とするとどうなるか

練習問題 2.9

`eval` の定義を

```
eval state
  ¦ tiFinal state = [state]
  ¦ otherwise     = state : eval nextState
```

としたほうが、わかりやすそうにみえるが、この定義には欠点がある。それはどのようなものか。

Mark 2 : let(rec) 式

`instantiate` を拡張して、 `ELet` 項に対応する

練習問題 2.10

`instantiate` の定義に非再帰的 `let` 式に対応する等式を加えよ。

`ELet nonRecursive defs body` を具体化する

1. `defs` の各定義の右辺を具体化する
2. `defs` の各定義の左辺（名前）と新たに具体化されたものとを結びつけて環境を拡張する
3. 拡張された環境と式本体を渡して `instantiate` を呼ぶ

練習問題 2.11

再帰的 `let` に `instantiate` が対応できるようにせよ

（ヒント）前問のステップ 1 で `instantiate` に既存の環境を渡していたが、代りにステップ 2 で拡張した環境を渡すようにする

```
instantiateLet :: IsRec -> Assoc Name CoreExpr -> CoreExpr -> TiHeap -> Assoc Name Addr -> (TiHeap, Addr)
instantiateLet isrec defs body heap env
  = instantiate body heap' env'
    where
        (heap', extraBindings) = mapAccumL instantiateRhs heap defs
        env' = extraBindings ++ env
        rhsEnv ¦ isrec     = env'
               ¦ otherwise = env
        instantiateRhs heap (name, rhs)
          = (heap1, (name, addr))
            where
                (heap1, addr) =instantiate rhs heap rhsEnv
```

テストプログラム

```
pair x y f = f x y ;
fst p = p K ;
snd p = p K1 ;
f x y = letrec
            a = pair x b ;
            b = pair y a
        in
        fst (snd (snd (snd a))) ;
main = f 3 4
```

結果は `4` である

練習問題 2.12

以下のプログラムを実行するとどうなるか？

```
main = letrec f = f x in f
```

Haskellのような強い型付けを行う言語でも同じ問題が起きるか？

Mark 3: 更新の追加

遷移規則(2.2)の代わりに(2.3)を使う

(2.3)

$$\begin{array}{rrll} & a_0 : a_1 : \cdots : a_n : s \quad d & h[a_0 : \texttt{NSupercomb}\ [x_1, \ldots, x_n]\ body] & f \\ \Longrightarrow & a_r : s \quad d & h'[a_n : \texttt{NInd}\ a_r] & f \end{array}$$

ここで、$(h', a_r) = instantiate\ body\ h\ f[x_1 \mapsto a_1, \ldots, x_n \mapsto a_n]$である。
$instantiate$が返す結果のルートへの間接参照a_rで、ノードa_nを上書きする。

スパインをアンワインドするときに間接参照に出会す可能性があるので、その場合に対応する規則(2.4)を追加する。

(2.4)

$$\begin{array}{rcccc} & a : s & d & h[a : \mathtt{NInd}\ a_1] & f \\ \Longrightarrow & a_1 : s & d & h & f \end{array}$$

新しい規則を実装するには以下を行う

1. `Node` 型に新しいデータ構成子 `NInd` を追加。それに伴い `showNode` もこれに対応する必要がある

```
data Node = NAp Addr Addr                          -- Application
          | NSupercomb Name [Name] CoreExpr        -- Supercombinator
          | NNum Int                               -- Number
          | NInd Addr                              -- Indirection
```

2. 規則(2.3)対応：リデックスのルートを `hUpdate` をつかって、結果の間接参照に置き換えるように、 `scStep` を変更する
3. 規則(2.4)対応： `dispatch` の定義に間接参照を扱う等式を追加

練習問題 2.13

間接参照による更新を実装せよ。以下のプログラムの実行を、Mark 1 と Mark 3 で走らせ効果を比較せよ。

```
id x = x ;
main = twice twice id 3
```

まず、手で簡約を行ってどう計算されるか確認せよ

```
main = twice twice twice id 3
```

と定義した場合どうなるか。

2.5.1 間接参照の低減

あきらかに不要な間接参照を構成しないようにする。

```
instantiateAndUpdate
  :: CoreExpr          -- ^ スーパーコンビネータの本体
  -> Addr              -- ^ 更新するノードのアドレス
  -> TiHeap            -- ^ 具体化前のヒープ
  -> Assoc Name Addr   -- ^ 仮引数からアドレスへの連想リスト
  -> TiHeap            -- ^ 具体化語のヒープ
```

式が関数適用であったときの `instantiateAndUpdate` の定義

```
instantiateAndUpdate (EAp e1 e2) updAddr heap env
  = hUpdate heap2 updAddr (NAp a a2)
    where
      (heap1, a1) = instantiate e1 heap  env
      (heap2, a2) = instantiate e2 heap1 env
```

更新の必要があるのはルートノードだけなので、`e1` および `e2` に対しては元の `instantiate` にする

練習問題 2.14

`instantiateAndUpdate`の定義を完成せよ。以下の点に留意すること

- 具体化する式が単なる変数なら、間接参照が必要である。（なぜか）
- `let(rec)`式に対する再帰的な具体化の部分は注意深く考えること。

`scStep`を`instantiateAndUpdate`を呼ぶように変更せよ

- `scStep`でおこなっていた更新のコードは取り除け。

簡約段数とヒープアロケーション数を計測し、効果を確かめよ。

2.6 Mark 4: 算術演算の追加

2.6.1 算術に対する遷移規則

符号反転の遷移規則(2.5) 引数が評価済みの場合

$$
\begin{array}{rcccl}
 & a : a_1 : [] & d & h \left[\begin{array}{lcl} a & : & \texttt{NPrim Neg} \\ a_1 & : & \texttt{NAp}\ a\ b \\ b & : & \texttt{NNum}\ n \end{array}\right] & f \\
\Longrightarrow & a_1 : [] & d & h\,[a_1 : \texttt{NNum}\ (-n)] & f
\end{array}
$$

符号反転の遷移規則(2.6) 引数未評価の場合

$$\begin{array}{lllll} a : a_1 : [] & & d & h & \left[\begin{array}{lll} a & : & \texttt{NPrim Neg} \\ a_1 & : & \texttt{NAp}\ a\ b \end{array}\right] \quad f \\ \Longrightarrow \quad b : [] & (a : a_1 : []) : d & h & & f \end{array}$$

スタックをダンプに積んでから、引数の評価にはいる。評価が済んだら、ダンプに積んだスタックを復帰する(2.7)

$$\begin{array}{llll} a : [] & s : d & h\,[a : \texttt{NNum}\ n] & f \\ \Longrightarrow \quad s & d & h & f \end{array}$$

引数のルートノードが間接参照になっている場合がある。 (2.1)の特殊な場合として、(2.8)を導入する

(2.8)

$$\begin{array}{rlll} & a : s \quad d & h \left[\begin{array}{lll} a & : & \texttt{NAp}\ a_1\ a_2 \\ a_2 & : & \texttt{NInd}\ a_3 \end{array} \right] & f \\ \Longrightarrow & a : s \quad d & h \left[a : \texttt{NAp}\ a_1\ a_3 \right] & f \end{array}$$

(2.8)が機能するよう、(2.6)を変更して(2.9)にする必要がある

(2.9)

$$\begin{array}{rlll} & a : a_1 : [] \quad\quad\quad d & h \left[\begin{array}{lll} a & : & \texttt{NPrim Neg} \\ a_1 & : & \texttt{NAp}\ a\ b \end{array} \right] & f \\ \Longrightarrow & b : [] \quad (a_1 : []) : d & h & f \end{array}$$

練習問題

加算に対する遷移規則を書け。（ほかの二項算術演算子も実質同じである。）

2.6.2 算術演算式の実装

1. `TiDump` をスタックのスタックとして再定義
2. 名前 `n` 値 `p` のプリミティブを表すヒープノード `NPrim Name Primitive` を追加
3. `showNode` を拡張
4. `Primitive` を定義
5. `NPrim` ノードを初期ヒープにわりあてるよう `buildInitialHeap` を拡張
6. プリミティブの名前と値の連想リスト `primitives`
7. `allocatePrim` を `allocateSc` を参考に実装
8. `step` の `dispatch` を `NPrim` に対応して、 `primStep` を呼ぶものとして拡張

9. `primStep` を実装
    - `primNeg` を実装
10. 遷移規則(2.7)を実装するために `numStep` を変更
11. 同様に遷移規則(2.8)を実装するために `apStep` を変更する
12. `tiFinal` を変更する

練習問題 2.16

符号反転演算を実装し以下のプログラムでテストせよ。

算術演算の実装には 2項算術演算汎用の

```
primArith :: TiState -> (Int -> Int -> Int) -> TiState
```

を用意して、これを使う

練習問題 2.17

`primArith` を実装し、テストせよ

2.7 Mark 5: 構造をもつデータ

`case` 式が実装できるといいのだが、雛形具体化機械の枠組みでは難しい。 代りに `if` 、 `casePair` 、 `caseList` などの組み込み関数を使い特定の構造をもつデータ型を扱えるようにする。

練習問題 2.18

なぜ雛形具体化機械で `case` 式を導入するのは難しいのか

(ヒント) `case` 式に対して `instantiate` がすべきことを考えよ

構造をもつデータの構築

状態遷移規則(2.10)

$$
\begin{array}{lllll}
 & a : a_1 : \cdots : a_n : [] & d & h \left[\begin{array}{lll} a & : & \texttt{NPrim (PrimConstr } t\ n\texttt{)} \\ a_1 & : & \texttt{NAp } a\ b_1 \\ \cdots & & \\ a_n & : & \texttt{NAp } a_{n-1}\ b_n \end{array}\right] & f \\
\Longrightarrow & a_n : [] & d & h\,[a_n : \texttt{NData } t [b_1, \ldots, b_n]] & f
\end{array}
$$

2.7.2 条件式

練習問題 2.19

条件式に関する状態遷移規則を書け

練習問題 2.20

コア言語で、`or`、`xor`、`not` の定義を与えよ。これらの定義を `extraPreludeDefs` に追加せよ。

2.7.3 構造をもつデータの実装

練習問題 2.21

これまでの変更をすべて行え。これで意味のある再帰定義関数がかける。

```
fac n = if (n == 0) 1 (n * fac (n - 1));
main = fac 3
```

2.7.4 対（ペア）

練習問題 2.22

`casePair` に対応する状態遷移規則を書け。

必要な規則は

- 第1引数が評価済みの場合、簡約を実行するための規則
- 第1引数が未評価の場合、評価を開始するための規則

実装した上で以下のプログラムでテストせよ。

```
main = fst (snd (fst (MkPair (MkPair 1 (MkPair 2 3)) 4)))
```

2.7.5 リスト

練習問題 2.23

コア言語で `Cons`、`Nil`、`head`、`tail` の定義を与えよ。 空リストに対して `head` や `tail` を適用したときに返るものとして、新たなプリミティブ `abort` を導入し、Haskell の `error` を呼び出して、プログラムを停止するようにせよ。

練習問題 2.24

`caseList` に対する遷移規則を書き、実装せよ。 `abort` に対する遷移規則を書き、実装せよ。 `preludeDefs` に `Nil`、 `Cons`、 `head`、 `tail` の定義を追加せよ。

練習問題 2.25

`case` 式を実装するのに代えて、構造をもつデータ型ごとに `case` プリミティブを考えるとき、主たる利点と欠点はなにか。

2.7.6 リストの表示

プログラムの結果が、数値のリストであった場合、

- 空リストなら、印字は終了
- 先頭が、未評価なら先に評価
- 先頭が、評価済ならそれを印字してから、末尾部リストを再帰的に印字

状態遷移システムにおいて数値の出力をどうモデル化するか

- `TiState` に「出力（数値のリスト）」という構成要素を追加、これに数値を追加することで、「数値の出力」をモデル化する
- プリミティブ `Print` と `Stop` を追加

`Stop` に対応する遷移規則

状態遷移規則(2.11)

$$\begin{array}{rcccccc} & o & a : [] & [] & h & [\ a & : \ \texttt{NPrim Stop}\ ] & f \\ \Longrightarrow & o & [] & [] & h & & f \end{array}$$

`Print` に対応する遷移規則（リストの先頭が評価済みの場合）

状態遷移規則(2.12)

$$
\begin{array}{lllll}
 & o & a : a_1 : a_2 : [] \quad [] & h & \left[\begin{array}{lll} a & : & \texttt{NPrim Print} \\ a_1 & : & \texttt{NAp}\ a\ b_1 \\ a_2 & : & \texttt{NAp}\ a_1\ b_2 \\ b_1 & : & \texttt{NNum}\ n \end{array}\right] \quad f \\
\Longrightarrow & o \sharp\sharp [n] & b_2 : [] \quad [] & h & \quad f
\end{array}
$$

`Print` に対応する遷移規則（リストの先頭が未評価の場合）

状態遷移規則(2.13)

$$
\begin{array}{lllllll}
 & o & a : a_1 : a_2 : [] & [] & h \left[\begin{array}{lll} a & : & \texttt{NPrim Print} \\ a_1 & : & \texttt{NAp}\ a\ b_1 \\ a_2 & : & \texttt{NAp}\ a_1\ b_2 \end{array} \right] & f \\
\Longrightarrow & o & b_1 : [] & (a_2 : []) : [] & h & f
\end{array}
$$

`extraPreludeDefs` に以下を追加

```
printList xs = caseList xs stop printCons
printCons h t = print h (printList t)
```

ここで、`print` は `primitives` でプリミティブ `Print` に、`stop` は `Stop` に束縛する

練習問題 2.26

ここまでの変更を実装し、数値のリストを返すプログラムを作成して実装をテストせよ。

2.8 別の実装

別実装の演習

2.8.1 プリミティブの別表現

プリミティブに対してやることは、そのプリミティブを実行するだけなので、`Primitive`型は単に`TiState`から`TiState`への関数と表現できる

```
type Primitive = TiState -> TiState
```

`Add`や`Sub`といったデータ構成子は消え、case分岐はなくなり、`primStep`は単なる関数適用になる

```
primStep state prim = prim state
primitive = [ ("negatate", primNeg)
            , ("+", primArith (+))
            , ("-", primArith (-))
            , ("*", primArith (*)
            , ("/", primArith div))
            ]
```

練習問題 2.27

この変更を実装しテストせよ。

2.8.2 ダンプの別表現

新しいスタックを古いスタックの上にそのまま作っていくという実装が普通。 ダンプはスパインスタックの底からのオフセットを積むスタックとして表現する

```
type TiDump = Stack Int
```

練習問題 2.28

この変更を実装しテストせよ。引数評価の始まりと終わりの部分の定義に変更が必要になる。また `tiFinal` も変更が必要

2.8.3 データ値の別表現

ブール値

`if` 式の簡約規則は、第2、第3、引数のどちらかを選択するだけなので、 ブール値を構造を持つデータではなく、2つの引数の一方を選択するという関数として表現する。

```
True  t f = t
False t f = f
```

こうすると、各ブール値演算の定義は以下のとおり

```
if = I
and b1 b2 t f = b1 (b2 t f) f
or b1 b2 t f = b1 t (b2 t f)
not b t f = b f t
```

ペア

おなじ方法をペアに対しても使える

```
pair a b f = f a b
casePair = I
fst p = p K
snd p = p K1
```

リスト

またリストにも使えるが、`Cons`用に追加の引数を使う

```
cons a b cn cc = cc a b
nil cn cc      = cn
caseList = I
```

練習問題 2.29

ブール値、ペア、リストを前述の方法で実装し、性能を測定せよ。従来の方法と比べてどのような長所と短所があるか。

2.9 ガーベッジコレクション

実行が進むにつれ、ヒープに割り当てられるノードが増え、そのうちヒープは満杯になるので、GCが必要になる。

```
gc :: TiState -> TiState
```

簡約1ステップごとに `doAdmin` でヒープサイズをチェックし、所定のサイズを超えていたら GC を呼ぶ。

2.9.1 マークスキャンGC

1. ルートアドレスを特定する

```
findStackRoots ::  TiStack   -> [Addr]
findDumpRoots  ::  TiDump    -> [Addr]
findGlobalRoots :: TiGlobals -> [Addr]
```

2. マークフェーズでルートからたどれるすべてのノードに再帰的にマークする

```
markFrom :: TiHeap -> Addr -> TiHeap
```

3. スキャンフェーズでマークされていないノードを解放したのち、マークされたノードのマークを外す

```
scanHeap :: TiHeap -> TiHeap
```

練習問題 2.30

`gc` の定義を `findRoots`、`markFrom`、`scanHeap` を使って書き、その上で、`doAdmin` から適切に呼び出すようにせよ。

練習問題 2.31

`findRoots` の定義を書け。

マークの方法は、本物の実装のときは、1ビットをマークビットとして使うなどがあるが、ここでは `Node` を拡張する

```
data Node
    = NAp Addr Addr
    | NSupercomb Name [Name] CoreExpr
    | NNum Int
    | NInd Addr
    | NPrim Name Primitive
    | NData Tag [Addr]
    | NMarked Node
```

`markFrom :: TiHeap -> Addr -> TiHeap` はヒープ h とアドレス a を取り以下をおこなう

1. h 中で a のノードを見る。マーク済みなら直ぐに返る。
2. ノード n をマークするには `hUpdate` をつかって n を **NMarked** n で置き換える
3. ノート n 中にあるすべてのアドレスを取り出して、それぞれに `markFrom` を適用する

```
scanHeap :: TiHeap -> TiHeap
```

1. `hAddresses` でヒープ内で使われているアドレスとノードの連想リストを取り出す。
2. 連想リストを走査して、ノードがマークされていなければ、`hFree` でそのアドレスを解放する
3. ノードがマークされていれば、`NMarked` 構成しを外したものに更新する

練習問題 2.32

`markFrom` と `scanHeap` の定義を書け

2.9.2 間接参照の除去

GCの最適化。図 2.2 にあるように間接参照を除いて、`NInd`ノードを回収する。そのために`markFrom`をすこし変更する。

```
markFrom :: TiHeap -> Addr -> (TiHeap, Addr)
```

`markFrom`の返り値を新しいヒープと新しいアドレスの対とする。スタック、ダンプ、グローバル環境も変更されることになるので、以下を用意し、それぞれ`markFrom`を使うようにする

```
markFromStack   :: TiHeap -> TiStack   -> (TiHeap, TiStack)
markFromDump    :: TiHeap -> TiDump    -> (TiHeap, TiDump)
markFromGlobals :: TiHeap -> TiGlobals -> (TiHeap, TiGlobals)
```

練習問題 2.33

改定版 `markFrom` を実装せよ。改良前とのヒープサイズを比べて、どのくらい改良されたか計測せよ。

2.9.3 ポインタ反転法

ヒープ中のN個のノードがすべて単一の長いリストで連結されことがあれば、`markFrom`はN回呼び出されることになる。こうなると、作業スタックはヒープと同じ大きさになりうる。毎回、ほとんどありえない状況のために、ヒープサイズと同じ作業スタックを用意するのは。。。

そこで、`NMarked`構成子を少々変更して、ポインタ反転法を使う。

```
data Node
    = NAp Addr Addr
    | NSupercomb Name [Name] CoreExpr
    | NNum Int
    | NInd Addr
    | NPrim Name Primitive
    | NData Tag [Addr]
    | NMarked MarkState Node

data MarkState
    = Done          -- ^ Marking on this node finished
    | Visits Int    -- ^ Node visited n times so far
```

ポインタ反転法は、マーク用状態遷移システムを別に用意して説明する。

0. 状態は、フォワードポインタ、バックワードポインタ、ヒープの三つ組である

$$(f, b, h)$$

1. `markFrom` がヒープ h_{init}、アドレス a で呼ばれると、マークプロセスが以下の初期状態からスタートする。

$$(a, \mathtt{hNull}, h_{init})$$

2. この状態遷移系は、以下の状態が終了状態である

$$(f, \mathtt{hNull}, h\,[f : \mathtt{NMarked\ Done\ n}])$$

3. フォワードポインタが未だマークされていないノードを指している場合
    i. `NAp` 場合

$$\begin{array}{rcll} & f & b & h\,[f : \mathtt{NAp}\ a_1\ a_2] \\ \Longrightarrow & a_1 & f & h\,[f : \mathtt{NMarked\ (Visits\ 1)\ (NAp}\ b\ a_2\mathtt{)}] \end{array}$$

    ii. フォワードポインタが未だマークされていない `NPrim` ノードを指していた場合

$$\begin{array}{rcll} & f & b & h\,[f : \mathtt{NPrim}\ p] \\ \Longrightarrow & f & b & h\,[f : \mathtt{NMarked\ Done\ (NPrim}\ p\mathtt{)}] \end{array}$$

    iii. `NSuperCombinator` と `NNum` ノードはアドレスを含まないので、同じ扱いでよい。
    iv. `NInd` の場合はフォワードポインタを間接参照先のアドレスに変更

$$\begin{array}{rcll} & f & b & h\,[f : \mathtt{NInd}\ a] \\ \Longrightarrow & a & b & h \end{array}$$

1. フォワードポインタがマーク済ノードを指す場合
    i. バックワードポインタが `hNull` なら終了
    ii. そうでないなら、バックワードポインタが指すのはマークされた `NAp` のはず
        a. その `MarkState` が `(Visits 1)` の場合（左側は操作済み）

$$f \quad b \quad h \left[\begin{array}{lll} f & : & \texttt{NMarked Done}\ n \\ b & : & \texttt{NMarked (Visits 1) (Nap}\ b'\ a_2\texttt{)} \end{array} \right]$$
$$\Longrightarrow \quad a_2 \quad b \quad h \left[\begin{array}{lll} b & : & \texttt{NMarked (Visits 2) (Nap}\ f\ b'\texttt{)} \end{array} \right]$$

        a. その `MarkState` が `(Visits 2)` の場合

$$f \quad b \quad h \left[\begin{array}{lll} f & : & \texttt{NMarked Done}\ n \\ b & : & \texttt{NMarked (Visits 2) (Nap}\ a_1\ b'\texttt{)} \end{array} \right]$$
$$\Longrightarrow \quad b \quad b' \quad h \left[\begin{array}{lll} b & : & \texttt{NMarked Done (Nap}\ a_1\ f\texttt{)} \end{array} \right]$$

2.9.4 2スペースGC

生きているノードだけ、別のヒープ空間にコピーする

1. 生きてるノードを、新ヒープにアロケートし、旧ヒープの当該ノードは移動先のアドレスを持たせる。
2. 新しいヒープにアロケートしたノードに含まれるアドレスを更新する。

実装

1. `Node` 型に `NForward` 構成子を追加（`NMarked` は不要）
2. `markFromXXXX` の代りに `evacuateXXXX` を使う

```
evacuateStack   :: TiHeap -> TiHeap -> TiStack -> (TiHeap, TiStack)
evacuateDump    :: TiHeap -> TiHeap -> TiDump -> (TiHeap, TiDump)
evacuateGlobals :: :: TiHeap -> TiHeap -> TiGlobals -> (TiHeap, TiGlobals)
```

3. evacuate後、 `scavengeHeap` を使う

```
scavengeHeap :: TiHeap -> TiHeap -> TiHeap
```