Lecture 12 – Garbage Collection

COSE212: Programming Languages

Jihyeok Park



2023 Fall

Recall



- Mutation makes it possible to update the contents of a data structure or a variable after its creation.
 - BFAE FAE with mutable boxes
 - MFAE FAE with mutable variables
 - Evaluation with **memories**, finite maps from addresses to values:

Memories
$$M \in \mathbb{A} \xrightarrow{\text{fin}} \mathbb{V}$$

Addresses $a \in \mathbb{A}$

- In this lecture, we will learn memory management techniques to deallocate unreachable memory cells:
 - Stack and Heap
 - Manual Memory Management
 - Garbage Collection (GC)

Contents



1. Stack and Heap

Tail-Call Optimization (TCO)

2. Manual Memory Management

3. Garbage Collection

Reference Counting Mark-and-Sweep GC Copying GC (Two-Space GC) Other GC Algorithms

Contents



 Stack and Heap Tail-Call Optimization (TCO)

2. Manual Memory Management

Garbage Collection
 Reference Counting
 Mark-and-Sweep GC
 Copying GC (Two-Space GC
 Other GC Algorithms

Stack and Heap



In the previous lecture, we have seen the memory in the following MFAE expression has **unreachable** memory cells as follows:

```
/* MFAE */
var y = 1;
var f = x => {
  x = x + y;
  x * x
};
f(5);  /* 36 */
y = 3;
f(5);  /* 64 */
*
```

```
\sigma = \begin{bmatrix} y \mapsto a_0 \\ f \mapsto a_1 \end{bmatrix}

A : a_0 a_1 a_2 a_3 \dots

M = 3 v 6 8 \dots
```

where
$$v = \langle \lambda x. x = x + y; x * x, [y \mapsto a_0] \rangle$$

Then, how to detect and deallocate unreachable memory cells?

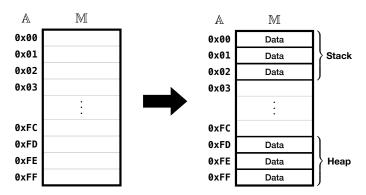
Let's delete unreachable memory cells when the program exits functions!

Stack and Heap



We can **divide** the memory into two parts:

- Stack for local variables and function parameters
- Heap for dynamically allocated memory cells



Create a **new stack frame** when the program **enters** a function, and **delete** the stack frame when it **exits** the function.





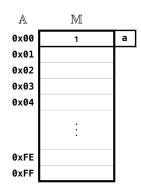
```
case class Box(var k: Int)
def f(x: Int): Int =
 var y = Box(1)
 var z = g(2)
 x + y.k + z
def g(b: Int): Int =
 var c = Box(b)
 c.k + 3
var a = 1
var d = f(42)
a + d
```

| A | \mathbb{M} |
|------|--------------|
| 0x00 | |
| 0x01 | |
| 0x02 | |
| 0x03 | |
| 0x04 | |
| | : |
| 0xFE | |
| 0xFF | |





```
case class Box(var k: Int)
def f(x: Int): Int =
 var y = Box(1)
 var z = g(2)
 x + y.k + z
def g(b: Int): Int =
 var c = Box(b)
 c.k + 3
                    /* a -> 0x00 */ *
var a = 1
var d = f(42)
a + d
```







```
case class Box(var k: Int)
def f(x: Int): Int = /* x -> 0x01 */
 var y = Box(1) /* y -> 0x02 */ *
 var z = g(2)
 x + y.k + z
def g(b: Int): Int =
 var c = Box(b)
 c.k + 3
                   /* a -> 0x00 */
var a = 1
var d = f(42)
a + d
```

| A | \mathbb{M} | | |
|------|--------------|---|---|
| 0x00 | 1 | а | |
| 0x01 | 42 | х | f |
| 0x02 | 0xFF | У | Ľ |
| 0x03 | | | |
| 0x04 | | | |
| | ÷ | | |
| 0xFE | | | |
| 0xFF | 1 | | |
| | | | |

A **new stack frame** is created when it enters the function f.

Stack and Heap



For example, consider the following Scala program:

```
case class Box(var k: Int)
def f(x: Int): Int = /* x -> 0x01 */
 var y = Box(1) /* y -> 0x02 */
 var z = g(2)
 x + y.k + z
def g(b: Int): Int = /* b -> 0x03 */
 var c = Box(b) /* c -> 0x04 */
 c.k + 3 /* 5 */
                /* a -> 0x00 */
var a = 1
var d = f(42)
a + d
```

| A | \mathbb{M} | | |
|------|--------------|---|---|
| 0x00 | 1 | а | |
| 0x01 | 42 | х | f |
| 0x02 | 0xFF | у | |
| 0x03 | 2 | b | _ |
| 0x04 | 0xFE | С | g |
| | : | | |
| 0xFE | 2 | | |
| 0xFF | 1 | | |
| | | - | |

A new stack frame is created when it enters the function g.





```
case class Box(var k: Int)
def f(x: Int): Int = /* x -> 0x01 */
 var y = Box(1) /* y -> 0x02 */
 var z = g(2) /* z -> 0x03 */
 x + y.k + z /* 48 */
def g(b: Int): Int =
 var c = Box(b)
 c.k + 3
var a = 1
                /* a -> 0x00 */
var d = f(42)
a + d
```

| A | \mathbb{M} | | |
|------|--------------|---|---|
| 0x00 | 1 | а | |
| 0x01 | 42 | х | |
| 0x02 | 0xFF | у | f |
| 0x03 | 5 | z | |
| 0x04 | | | |
| | | | |
| 0xFE | 2 | | |
| 0xFF | 1 | | |
| | | _ | |

After exiting the function g, its stack frame is **deleted**. The memory cells allocated for b and c in the stack frame are **deallocated**.





```
case class Box(var k: Int)
def f(x: Int): Int =
 var y = Box(1)
 var z = g(2)
 x + y.k + z
def g(b: Int): Int =
 var c = Box(b)
 c.k + 3
                 /* a -> 0x00 */
var a = 1
                /* d -> 0x01 */
var d = f(42)
                   /* 49 */
a + d
```

```
Α
           M
0x00
0x01
           48
                     d
0x02
0x03
0x04
0xFE
0xFF
```

After exiting the function f, its stack frame is **deleted**. The memory cells allocated for x, y, and z in the stack frame are **deallocated**.





```
def sum(x: Int, acc: Int): Int =
  if (x < 1) acc
  else sum(x - 1, x + acc) // tail-call
  sum(1000, 0)</pre>
```

| A | \mathbb{M} | | |
|------|--------------|-----|-------|
| 0x00 | 1000 | х | sum |
| 0x01 | 0 | acc | Suiii |
| 0x02 | | | |
| 0x03 | | | |
| 0x04 | | | |
| 0x05 | | | |
| | : | | |
| 0xFE | | | |
| 0xFF | | | |





```
def sum(x: Int, acc: Int): Int =
  if (x < 1) acc
  else sum(x - 1, x + acc) // tail-call
  sum(1000, 0)</pre>
```

| \mathbb{A} | \mathbb{M} | | |
|--------------|--------------|-----|-------|
| 0x00 | 1000 | х | sum |
| 0x01 | 0 | acc | Juiii |
| 0x02 | 999 | х | sum |
| 0x03 | 1000 | acc | Suiii |
| 0x04 | | | , |
| 0x05 | | | |
| | • | | |
| 0xFE | | | |
| 0xFF | | | |





```
def sum(x: Int, acc: Int): Int =
  if (x < 1) acc
  else sum(x - 1, x + acc) // tail-call
  sum(1000, 0)</pre>
```

| \mathbb{A} | \mathbb{M} | | |
|--------------|--------------|-----|-------|
| 0x00 | 1000 | х | sum |
| 0x01 | 0 | acc | Sum |
| 0x02 | 999 | х | sum |
| 0x03 | 1000 | acc | Sum |
| 0x04 | 998 | х | sum |
| 0x05 | 1999 | acc | Suiii |
| | ••• | | |
| 0xFE | 873 | х | sum |
| 0xFF | 118999 | acc | Suiii |
| | | | |

It fails with a stack overflow error.

However, is it really necessary to keep **all the stack frames? No!** Scala supports **tail-call optimization** (TCO).





```
def sum(x: Int, acc: Int): Int =
  if (x < 1) acc
  else sum(x - 1, x + acc) // tail-call
  sum(1000, 0)</pre>
```

| \mathbb{A} | \mathbb{M} | | |
|--------------|--------------|-----|-----|
| 0x00 | 1000 | х | sum |
| 0x01 | 0 | acc | Sum |
| 0x02 | | | |
| 0x03 | | | |
| 0x04 | | | |
| 0x05 | | | |
| | : | | |
| 0xFE | | | |
| 0xFF | | | |

Why? the function call is in **tail-call position** (i.e., the final action in the function). It means that it directly returns the result without any further computation.

Thus, we can safely **discard** the current stack frame **before** calling the function, and it is called **tail-call optimization** (TCO).





```
def sum(x: Int, acc: Int): Int =
  if (x < 1) acc
  else sum(x - 1, x + acc) // tail-call
  sum(1000, 0)</pre>
```

| \mathbb{A} | \mathbb{M} | | |
|--------------|--------------|-----|-------|
| 0x00 | 999 | х | sum |
| 0x01 | 1000 | acc | Juiii |
| 0x02 | | | |
| 0x03 | | | |
| 0x04 | | | |
| 0x05 | | | |
| | : | | |
| 0xFE | | | |
| 0xFF | | | |

Why? the function call is in **tail-call position** (i.e., the final action in the function). It means that it directly returns the result without any further computation.

Thus, we can safely **discard** the current stack frame **before** calling the function, and it is called **tail-call optimization** (TCO).





```
def sum(x: Int, acc: Int): Int =
  if (x < 1) acc
  else sum(x - 1, x + acc) // tail-call
  sum(1000, 0) // 500500</pre>
```

```
    A
    M

    0x00
    0
    x

    0x01
    500500
    acc

    0x02
    acc

    0x03
    acc

    0x04
    acc

    0x7E
    acc

    0xFE
    acc
```

Why? the function call is in **tail-call position** (i.e., the final action in the function). It means that it directly returns the result without any further computation.

Thus, we can safely **discard** the current stack frame **before** calling the function, and it is called **tail-call optimization** (TCO).

Contents



1. Stack and Heap

Tail-Call Optimization (TCO)

2. Manual Memory Management

3. Garbage Collection

Reference Counting Mark-and-Sweep GC Copying GC (Two-Space GC) Other GC Algorithms





Let's see the previous example again:

```
case class Box(var k: Int.)
def f(x: Int): Int =
 var y = Box(1)
 var z = g(2)
 x + y.k + z
def g(b: Int): Int =
 var c = Box(b)
 c.k + 3
                 /* a -> 0x00 */
var a = 1
var b = f(42)
                /* b -> 0x01 */
a + b
                   /* 49 */
```

```
\mathbb{M}
 Α
0x00
                         а
0x01
              48
                         b
0x02
0x03
0x04
0xFE
                          Heap
0xFF
```

Unfortunately, we still cannot deallocate memory cells (e.g., 0xFE and 0xFF) dynamically allocated in the **heap** rather than the **stack**.





One way to resolve this is using the **manual memory management**, and C++ is an example language that supports it with special keywords for memory allocation (new) and deallocation (delete) in heap, respectively:

```
\mathbb{M}
struct Box { int k; Box(int k): k(k_{0x00})
                                                             а
int f(int \times) {
                                            0x01
                                                     48
                                                             b
  Box* y = @ GL \{0.07\} @ new Box (1);
                                            0x02
// alloc 0xFF
                                            0x03
  int z = g(2);
                                            0x04
  int k = y - > k;
  return x + k + z:
                                            0xFE
                                                              Heap
int g(int b) {
                                            0xFF
  Box* c = @\backslash GL\{0.07\} @new Box(b);
// alloc 0xFE
  int k = c -> k:
  return k + 3:
```

Manual Memory Management



One way to resolve this is using the **manual memory management**, and C++ is an example language that supports it with special keywords for memory allocation (new) and deallocation (delete) in heap, respectively:

```
\mathbb{M}
struct Box { int k; Box(int k): k(k_{0x00})
                                                            а
int f(int \times) {
                                           0x01
                                                    48
                                                            b
  Box* y = @\GL\{0.07\}@new Box(1);
                                           0x02
// alloc 0xFF
                                           0x03
  int z = g(2);
                                           0x04
  int k = y->k; @\GL\{0.21\} @delete y
  return x + k + z:
                                           0xFE
int g(int b) {
                                           0xFF
  Box* c = Q\backslash GL\{0.07\} @new Box(b);
// alloc 0xFE
  int k = c \rightarrow k; Q \subseteq \{0.21\} Q = c \in C; C \in C
  return k + 3:
```

Manual Memory Management



Pros:

• **Efficient** – Users can **explicitly** deallocate memory cells allocated in heap whenever they want.

Cons:

- **Error-prone** Users have all the **responsibility** to deallocate memory cells allocated in heap:
 - Memory leak occurs if users forget to deallocate memory cells.

```
b = new Box(42); ...
```

• Dangling pointer occurs if users deallocate memory cells too early.

```
b = new Box(42); \dots delete b; \dots b\rightarrow k;
```

• **Double free** occurs if users deallocate memory cells more than once.

```
b = new Box(42); \dots delete b; \dots delete b;
```

Contents



 Stack and Heap Tail-Call Optimization (TCO)

Manual Memory Management

3. Garbage Collection

Reference Counting Mark-and-Sweep GC Copying GC (Two-Space GC) Other GC Algorithms



Is there any way to automatically deallocate memory cells in heap? Yes!

Garbage collection (GC) is a representative technique for **automatic memory management**.

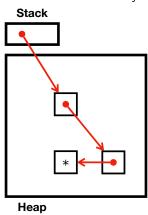
Let's learn several GC algorithms:

- Reference counting
- Mark-and-sweep GC
- Copying GC (Two-space GC)
- Others

Before explaining them, let's represent memory cells in heap in a **graphical** way without actual addresses.

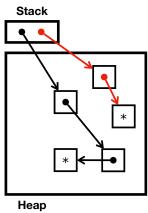


```
case class A(var x: A)
var a = A(A(A(null))) *
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```



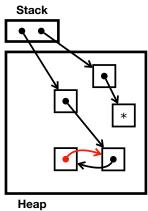


```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```



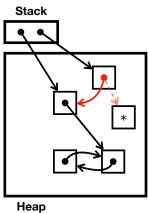


```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```



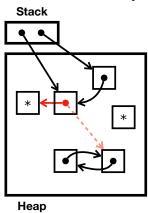


```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```



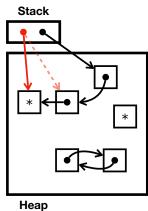


```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```



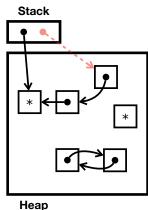


```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```





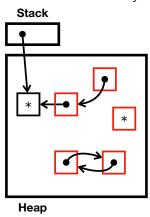
```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```





From now on, we will use the **graphical representation** of memory cells:

```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```



We need to deallocate five unreachable memory cells in heap.

Reference Counting



Reference counting is a simple GC algorithm that keeps track of the **number of references** to each memory cell in heap.

- 1 Initialize the reference count of each cell to 0.
- When a reference to a cell is created, increment its reference count.
- **3** When a reference to a cell is deleted, **decrement** its reference count.
- When the reference count of a cell reaches 0, deallocate the cell.

Many programming languages use reference counting to implement GC:

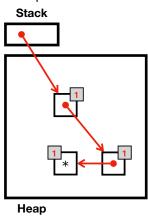
Python, Swift, Perl, Objective-C, etc.

Reference Counting



Reference counting is a simple GC algorithm that keeps track of the **number of references** to each memory cell in heap.

```
case class A(var x: A)
var a = A(A(A(null))) *
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```

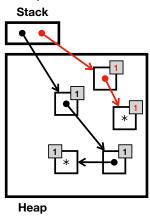


Reference Counting

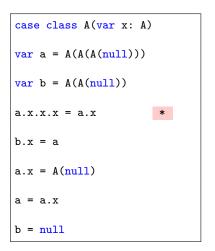


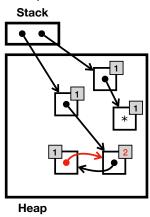
Reference counting is a simple GC algorithm that keeps track of the **number of references** to each memory cell in heap.

```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```

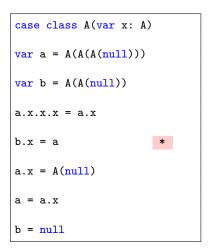


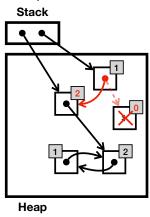






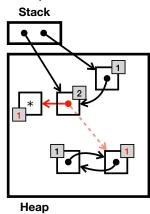




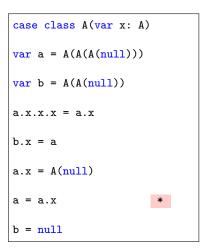


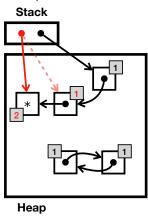


```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```



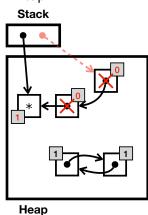








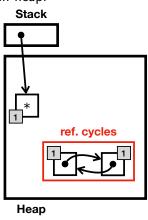
```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```





Reference counting is a simple GC algorithm that keeps track of the **number of references** to each memory cell in heap.

```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```



Unfortunately, we cannot deallocate unreachable reference cycles.



Reference counting is a simple GC algorithm that keeps track of the **number of references** to each memory cell in heap.

Pros:

- **Easy to implement** Simply increment and decrement the reference count when a reference is created and deleted.
- Low overhead Deallocation is immediate and takes a short time.

Cons:

- **Reference cycles** It cannot deallocate unreachable reference cycles.
- **Reference count cost** It requires space to store reference counts.
- Free List and Fragmentation It requires a free list to keep track
 of available free memory cells in heap, and it also suffers from
 fragmentation making it difficult to allocate large objects.

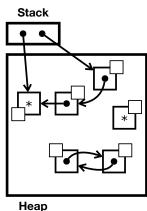


Mark-and-Sweep GC is one of tracing GC algorithms that traverses the heap to find unreachable objects when it is triggered under some conditions.

- **1** Mark all memory cells as unreachable (white).
- Mark all memory cells referenced by roots as unscanned (gray).
- 3 Repeat until there are no unscanned (gray) memory cells:
 - 1 Pick an unscanned (gray) memory cell.
 - 2 Mark memory cells referenced by the picked one as unscanned (gray).
 - 3 Mark the picked memory cell as scanned (black).
- Oeallocate (sweep) all memory cells that are still marked as unreachable (white).

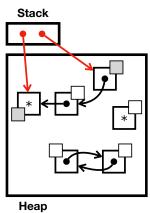


```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```



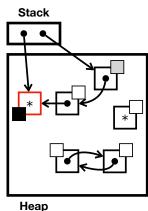


```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```



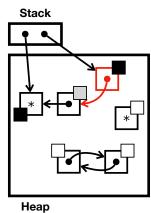


```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```



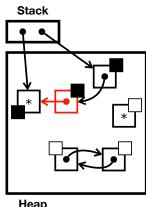


```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```





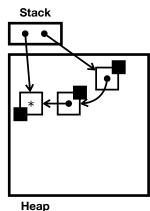
```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```



Heap



```
case class A(var x: A)
var a = A(A(A(null)))
var b = A(A(null))
a.x.x.x = a.x
b.x = a
a.x = A(null)
a = a.x
b = null
```





Mark-and-Sweep GC is one of tracing GC algorithms that traverses the heap to find unreachable objects when it is triggered under some conditions.

Pros:

• Reference cycles – It can deallocate unreachable reference cycles.

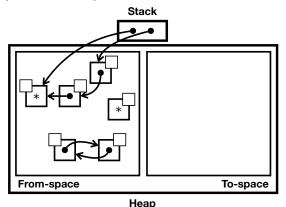
Cons:

- Stop-the-world It stops the program execution during GC.
- Free List and Fragmentation It requires a free list to keep track
 of available free memory cells in heap, and it also suffers from
 fragmentation making it difficult to allocate large objects.

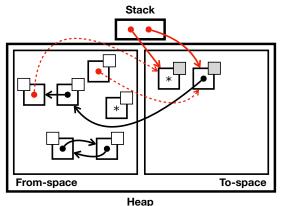


- Allocation It allocates memory cells only in from-space.
- **Deallocation** It deallocates all the unreachable objects as follows:
 - Mark all memory cells as unreachable (white).
 - 2 Copy all memory cells referenced by roots as unscanned (gray) and copy them from the from-space to the to-space
 - **3 Update** the data of the original memory cell to point to the copied one.
 - 4 Repeat until there are no unscanned (gray) memory cells
 - 1 Pick an unscanned (gray) memory cell in the from-space.
 - **2** Copy memory cells referenced by the picked one as unscanned (gray).
 - **3 Update** the data of the original memory cell to point to the copied one.
 - Mark the picked memory cell as scanned (black).
 - **5 Swap** from-space and to-space.

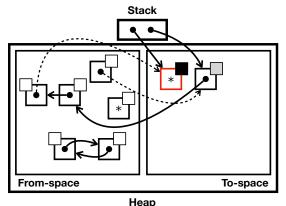




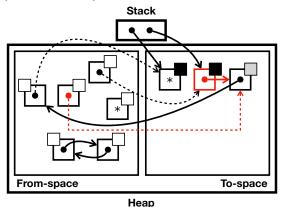




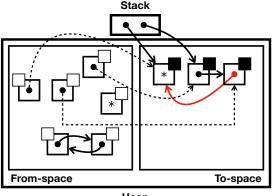






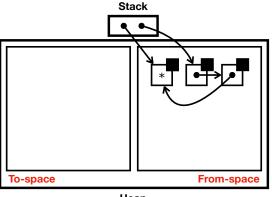






Heap





Heap



Similar to mark-and-sweep GC, **copying GC** (Two-space GC) is another **tracing GC** algorithm. However, it **copies** all the reachable objects and reorganizes them in a **compact** layout by splitting the heap into two spaces: **from-space** and **to-space**.

Pros:

- Reference cycles It can deallocate unreachable reference cycles.
- No more Free List and Fragmentation After deallocation process, the heap is always contiguous. Thus, it is enough to keep track of the first free memory cell for allocation.
- Fast Allocation It does not require any extra work to find free memory cells in the free list for allocation.

Cons:

- Stop-the-world It stops the program execution during GC.
- Only half of the heap (from-space) is used for allocation.
- Expensive copying process It copies all the reachable objects from the from-space to the to-space.

Other GC Algorithms



Existing real-world programming languages utilize more sophisticated GC algorithms, mix diverse GC algorithms, or even provide options to choose different GC algorithms:

- **Generational GC** e.g, Java¹, Python²
- Concurrent GC e.g., Java³, Golang⁴
- Ownership, Borrowing, and Lifetimes e.g., Rust⁵
- Escape Analysis e.g., Java⁶
- etc.

https://www.oracle.com/webfolder/technetwork/tutorials/obe/java/gc01/index.html

² https://devguide.python.org/internals/garbage-collector/

https://docs.oracle.com/javase/8/docs/technotes/guides/vm/gctuning/cms.html

⁴ https://tip.golang.org/doc/gc-guide

⁵ https://doc.rust-lang.org/book/ch04-01-what-is-ownership.html

https://blogs.oracle.com/javamagazine/post/escape-analysis-in-the-hotspot-jit-compiler

Summary



1. Stack and Heap

Tail-Call Optimization (TCO)

2. Manual Memory Management

3. Garbage Collection

Reference Counting Mark-and-Sweep GC Copying GC (Two-Space GC) Other GC Algorithms

Next Lecture



Lazy Evaluation

Jihyeok Park
 jihyeok_park@korea.ac.kr
https://plrg.korea.ac.kr