Garbage Collection

Today: various garbage collection strategies; basic ideas:

- Allocate until we run out of space; then try to free stuff
- **Invariant:** only the PL implementation (runtime system) knows about pointers so we can tag everything and find all reachable data
- Unlike C, C++, asm;
- Like ruby, python, perl, java, racket, ... everything else really

Reference Counting

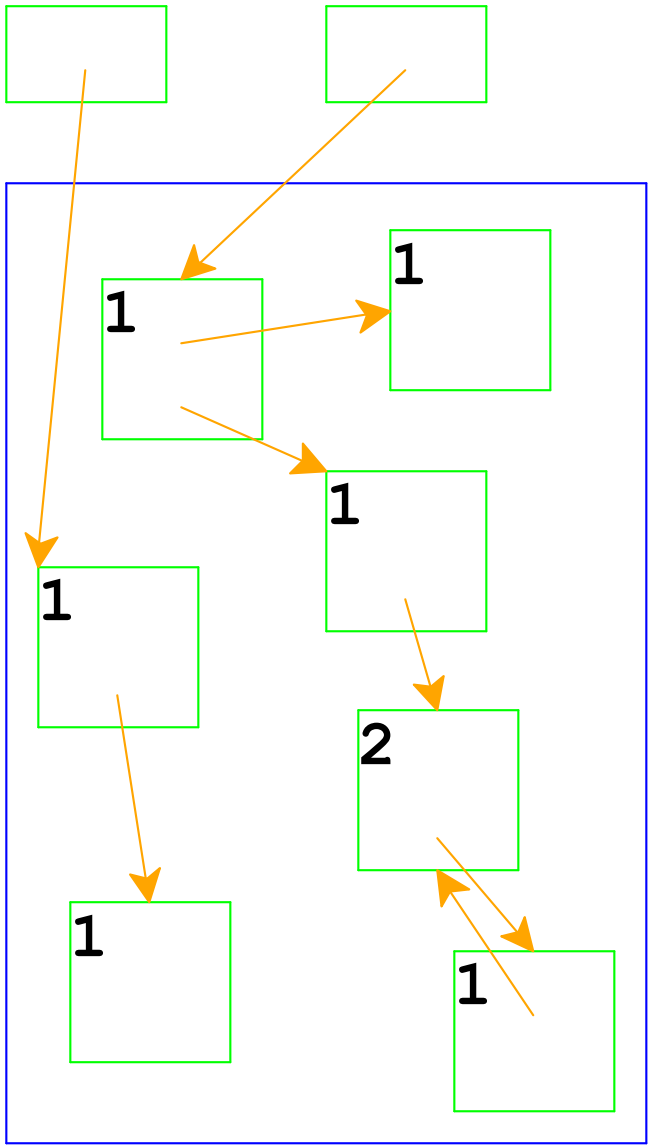
Reference counting: a way to know whether a record has other users

Reference Counting

Reference counting: a way to know whether a record has other users

- Attach a count to every record, starting at 0
- When installing a pointer to a record increment its count
- When replacing a pointer to a record, decrement its count
- When a count reaches 0, decrement counts for other records referenced by the record, then free it

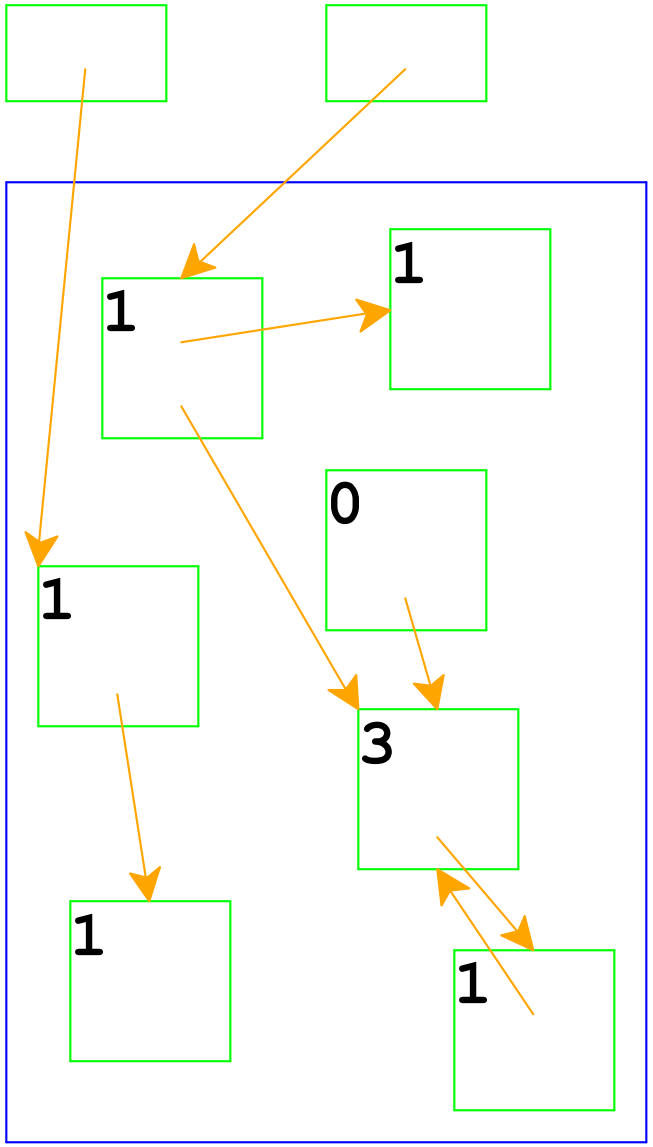
Reference Counting



Top boxes are the roots

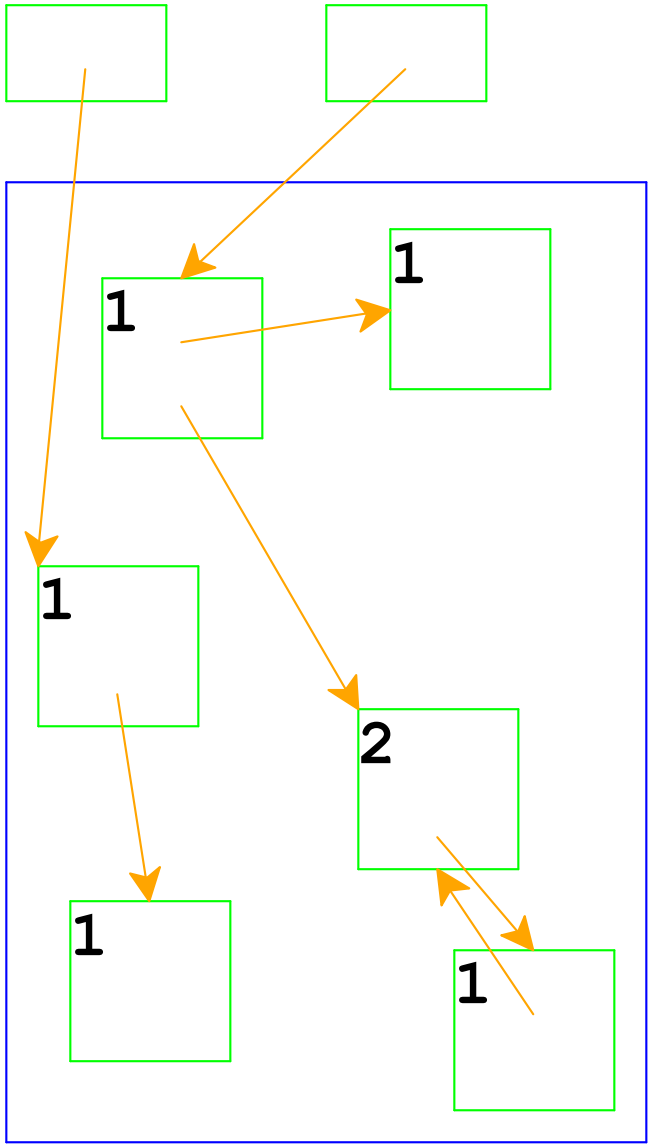
Boxes in the blue area are allocated memory

Reference Counting



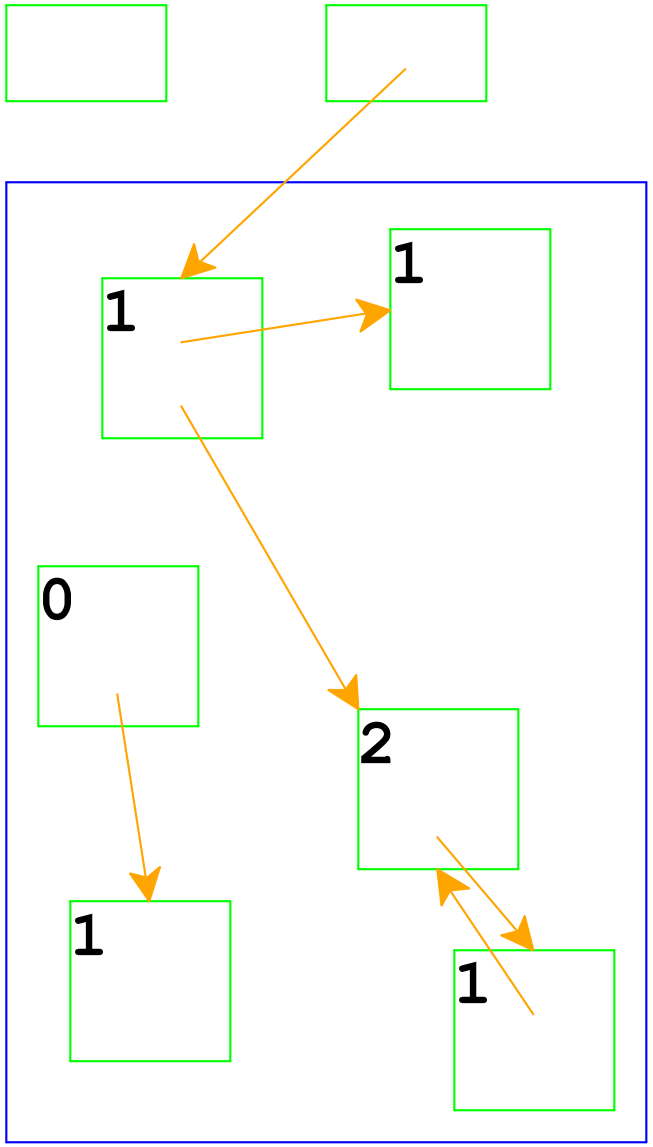
Adjust counts when a pointer is changed...

Reference Counting



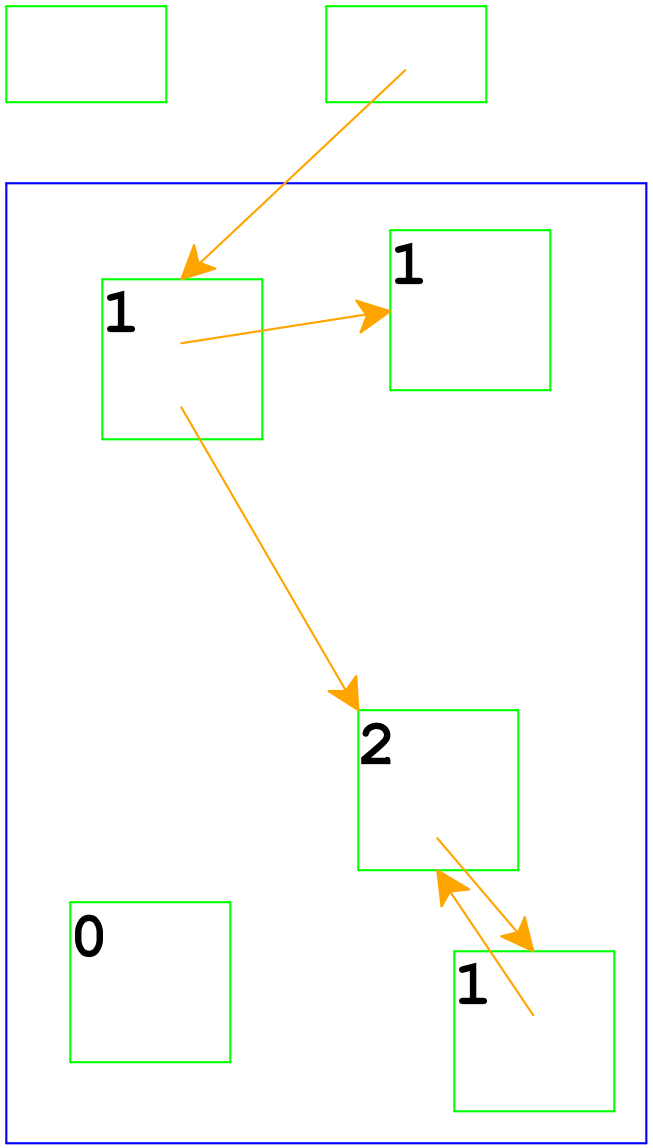
... freeing a record if its count goes to 0

Reference Counting



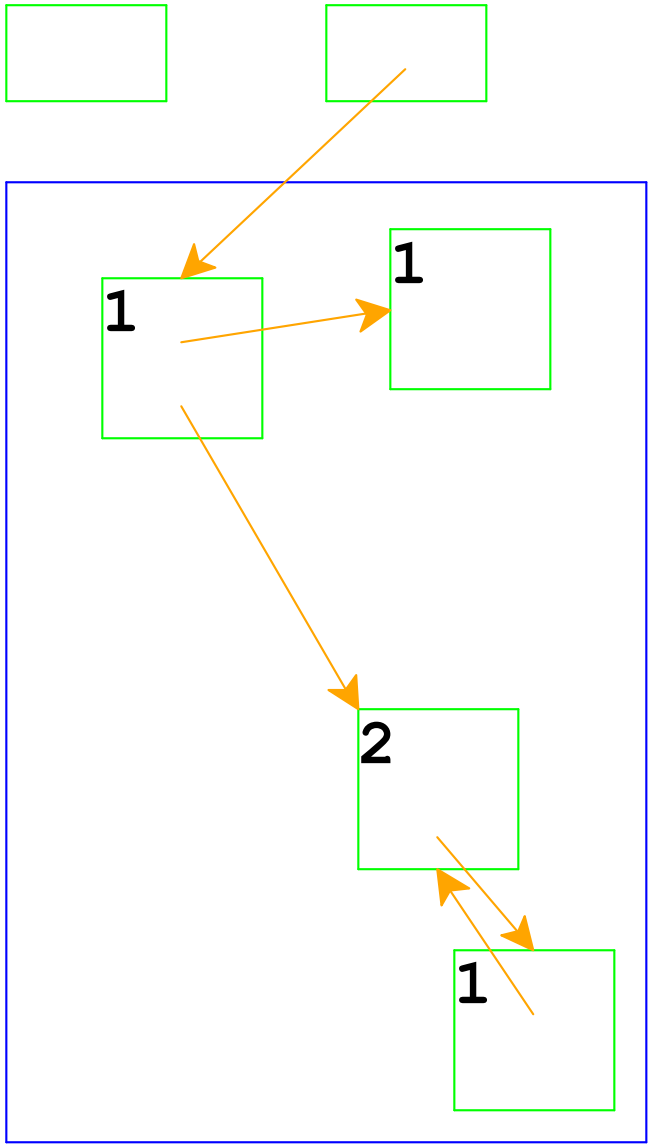
Same if the pointer is a root

Reference Counting



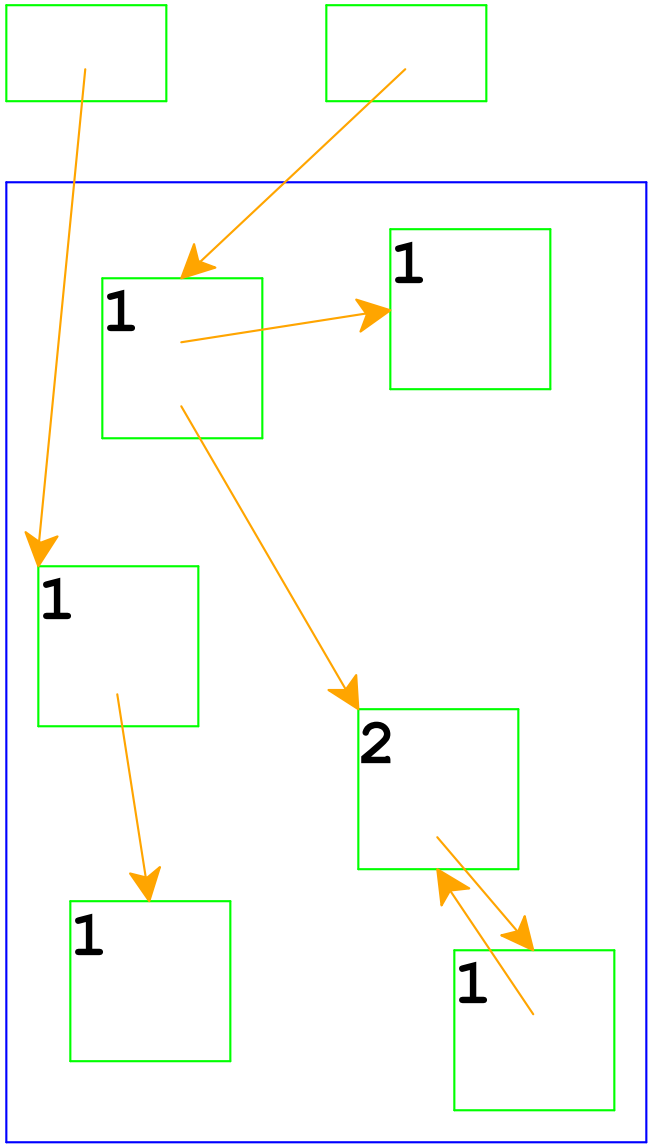
Adjust counts after frees, too...

Reference Counting



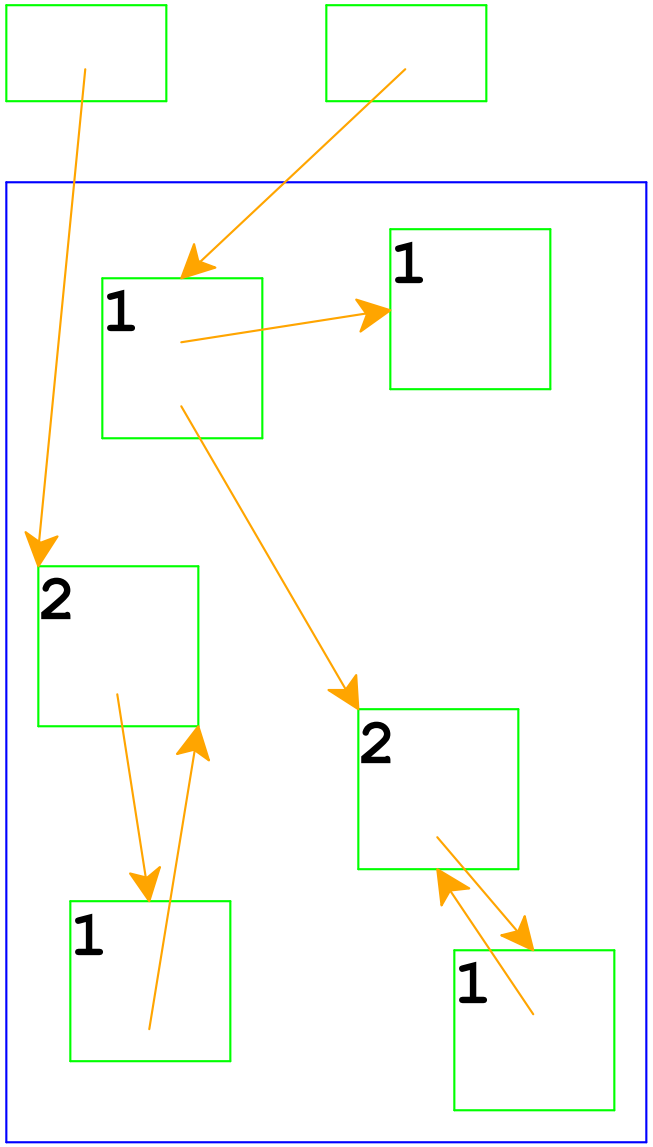
... which can trigger more frees

Reference Counting And Cycles



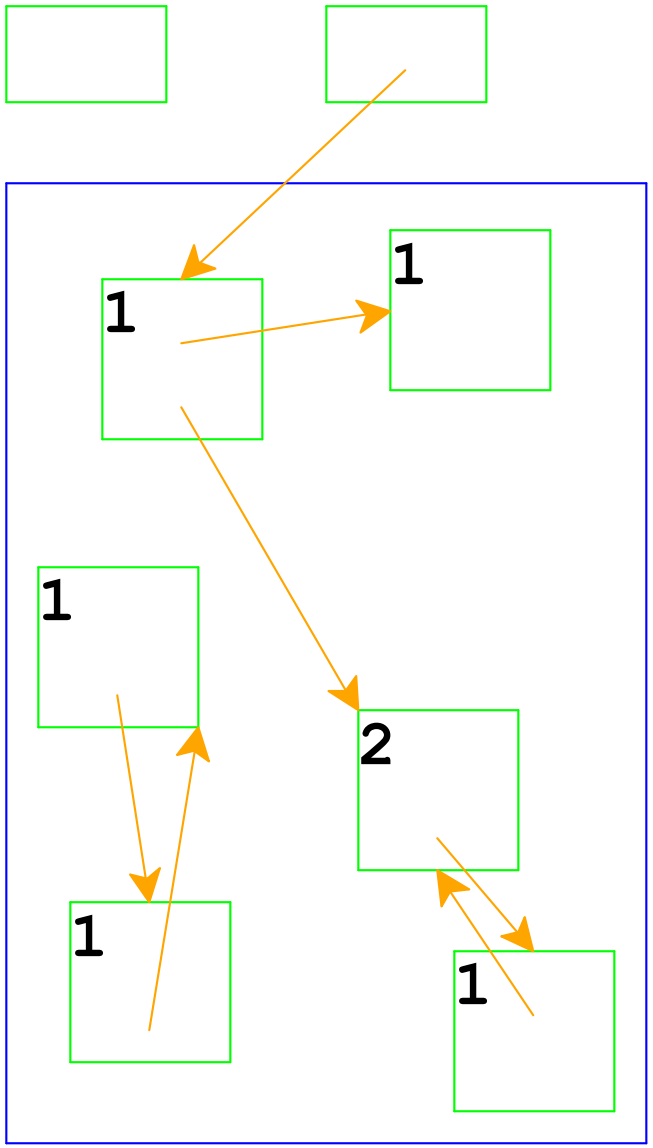
An assignment can create a cycle...

Reference Counting And Cycles



Adding a reference increments a count

Reference Counting And Cycles



Lower-left records are inaccessible, but not deallocated

In general, cycles break reference counting

Reference counting problems

- Cycles
- Maintaining counts wastes time & space
- Need to use free lists to track available memory

Reference counting

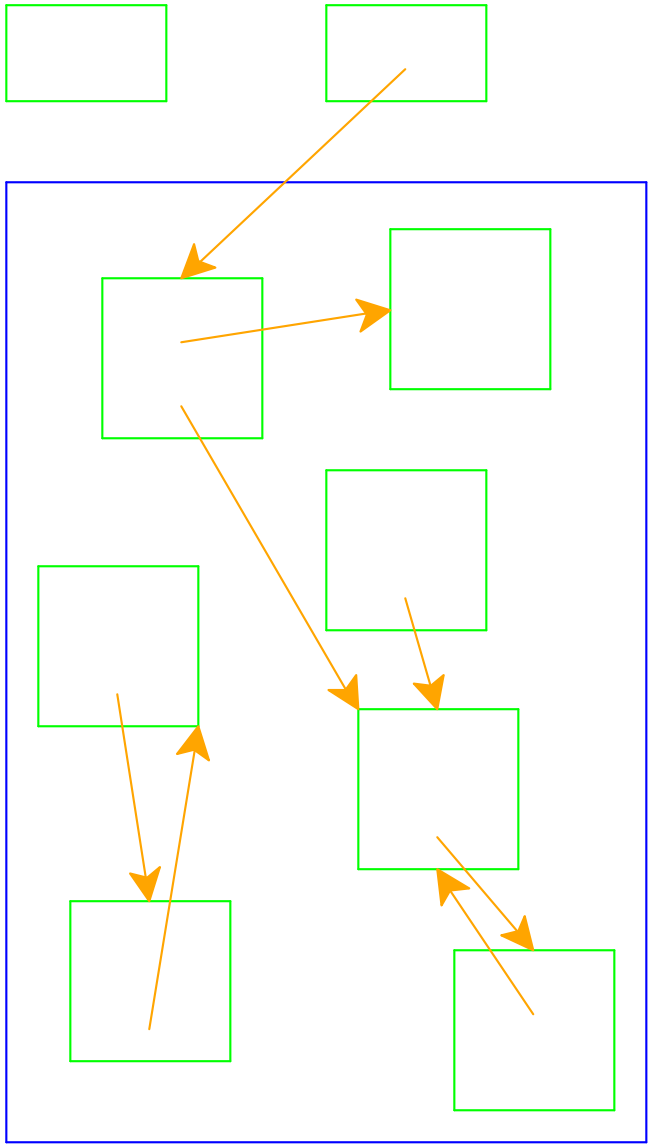
But there are times when this is a good choice:

- No pauses
- Interop with non-GC languages (but see the Boehm collector)

Mark & Sweep Garbage Collection Algorithm

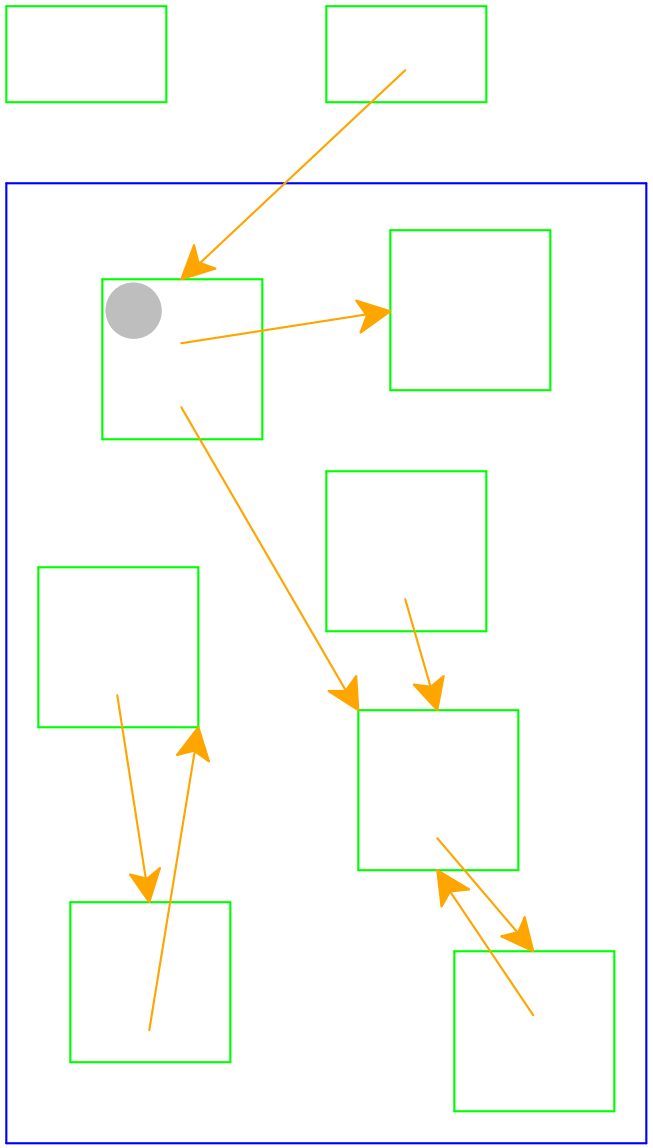
- Color all records ***white***
- Color records referenced by roots ***gray***
- Repeat until there are no gray records:
  - Pick a gray record, *r*
  - For each white record that *r* points to, make it gray
  - Color *r* ***black***
- Deallocate all white records

Mark & Sweep Garbage Collection



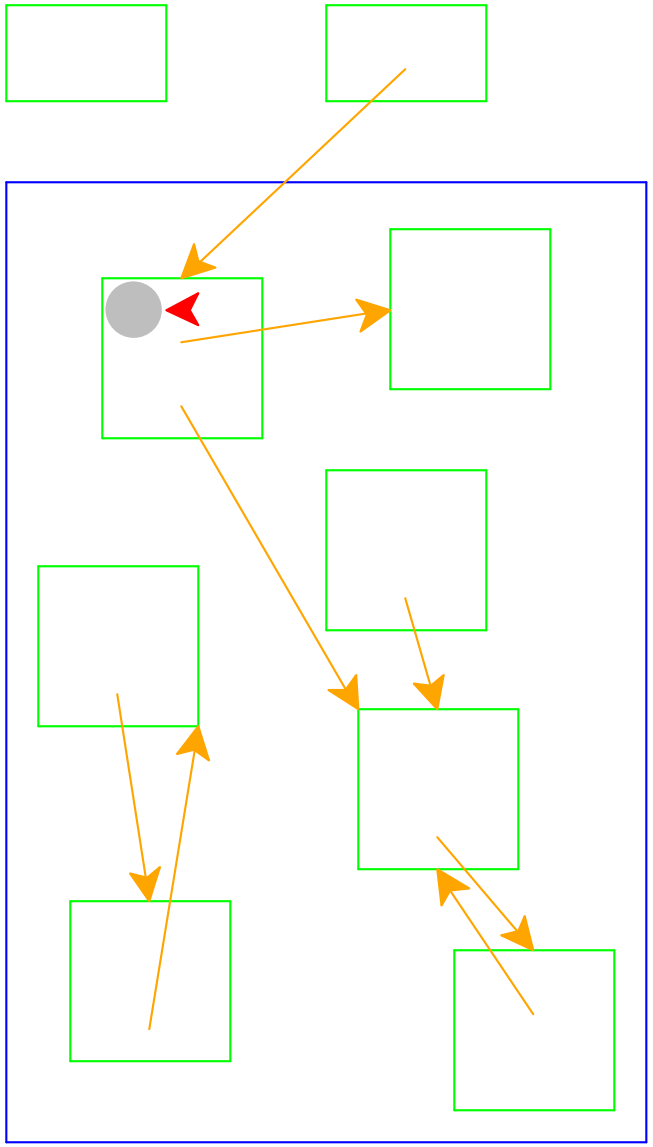
All records are marked white

Mark & Sweep Garbage Collection



Mark records referenced by roots as gray

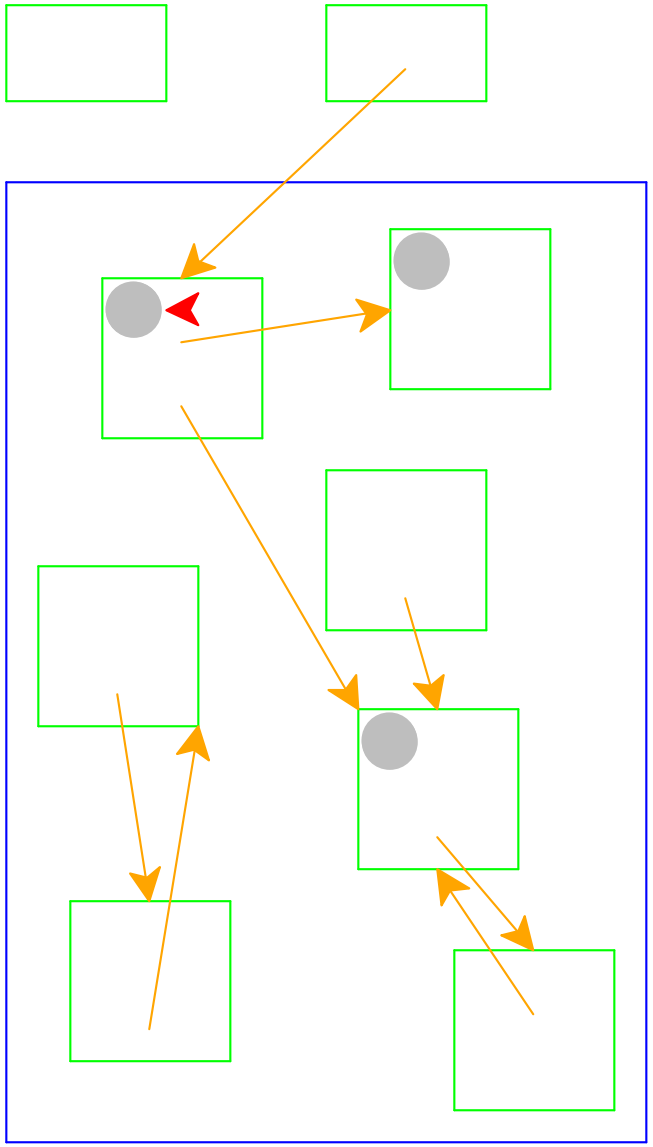
Mark & Sweep Garbage Collection



Need to pick a gray record

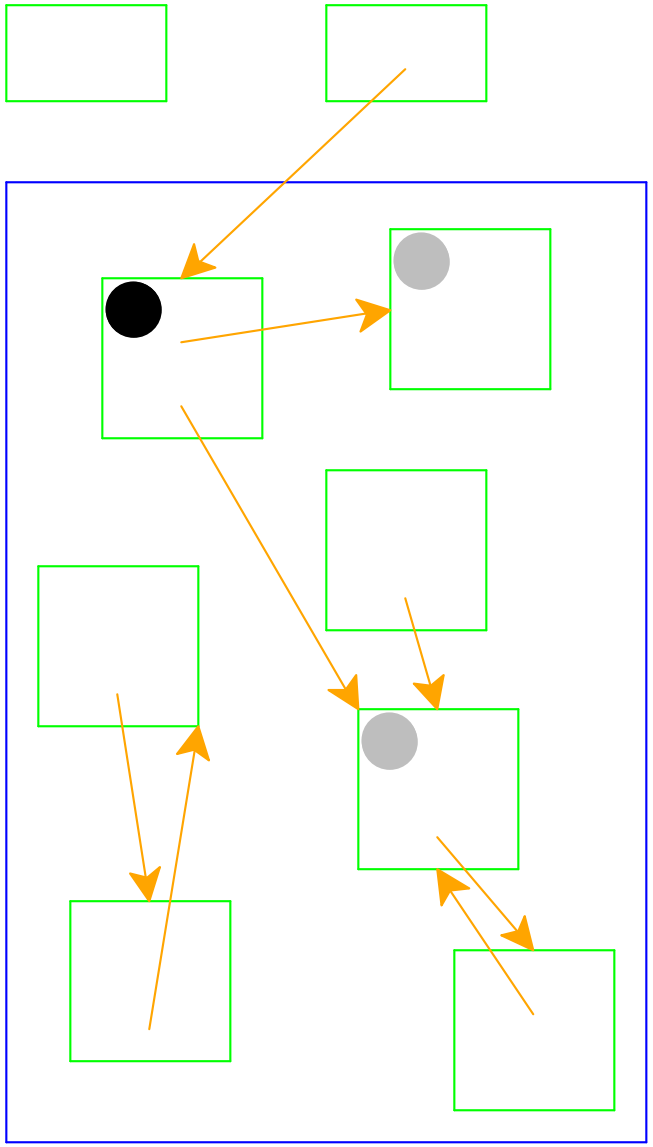
Red arrow indicates the chosen record

Mark & Sweep Garbage Collection



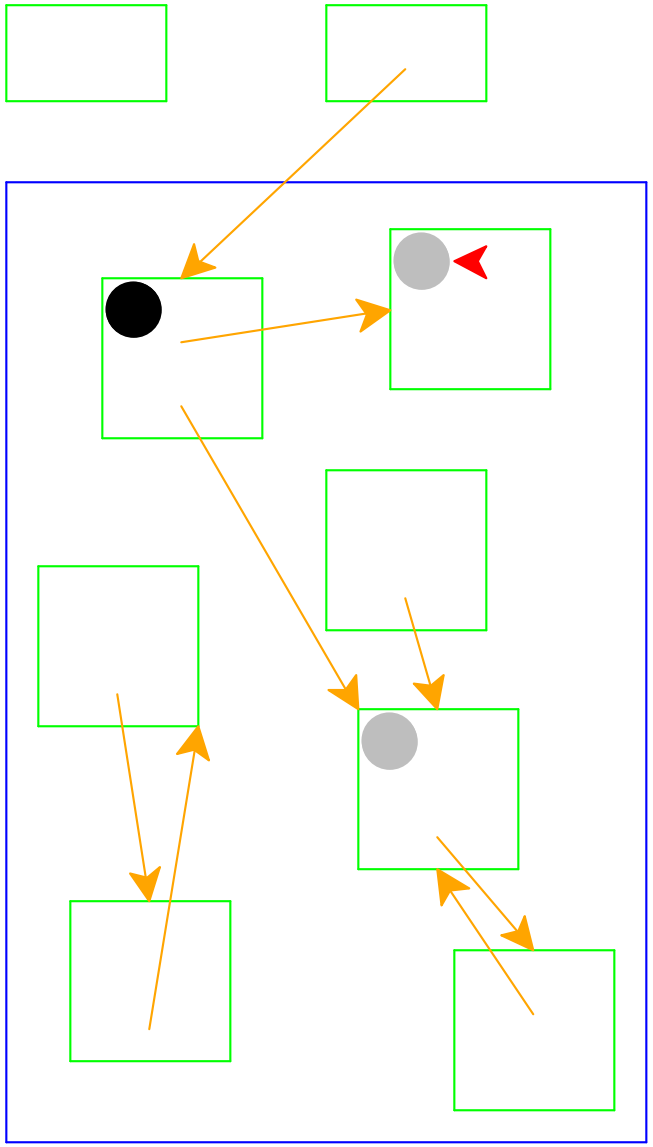
Mark white records referenced by chosen record as gray

Mark & Sweep Garbage Collection



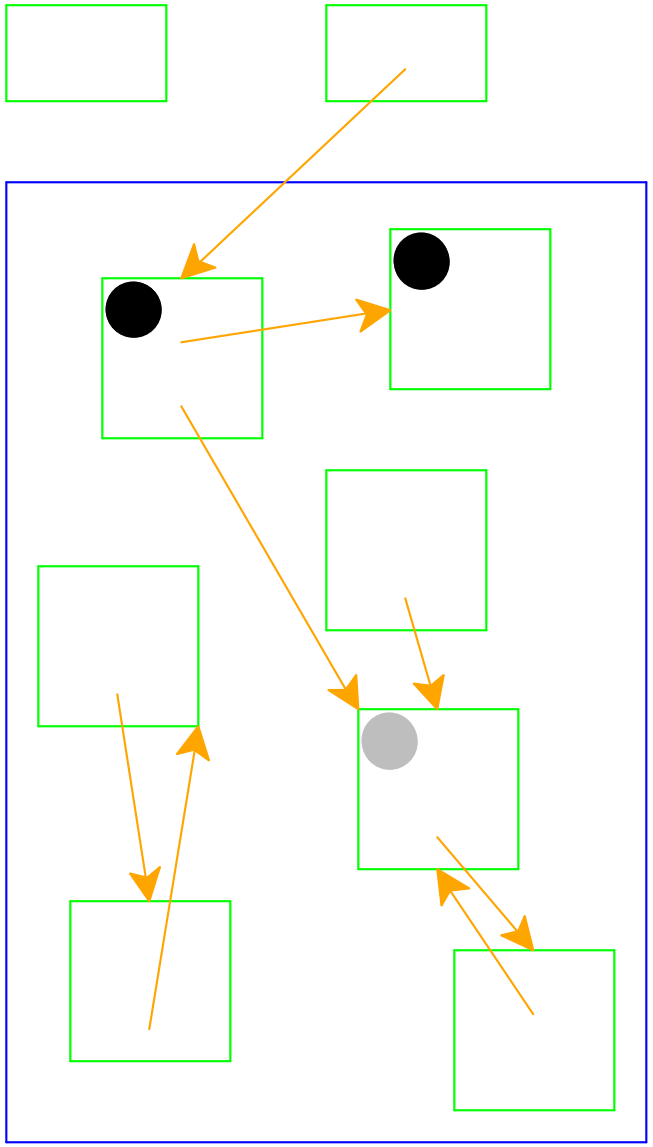
Mark chosen record black

Mark & Sweep Garbage Collection



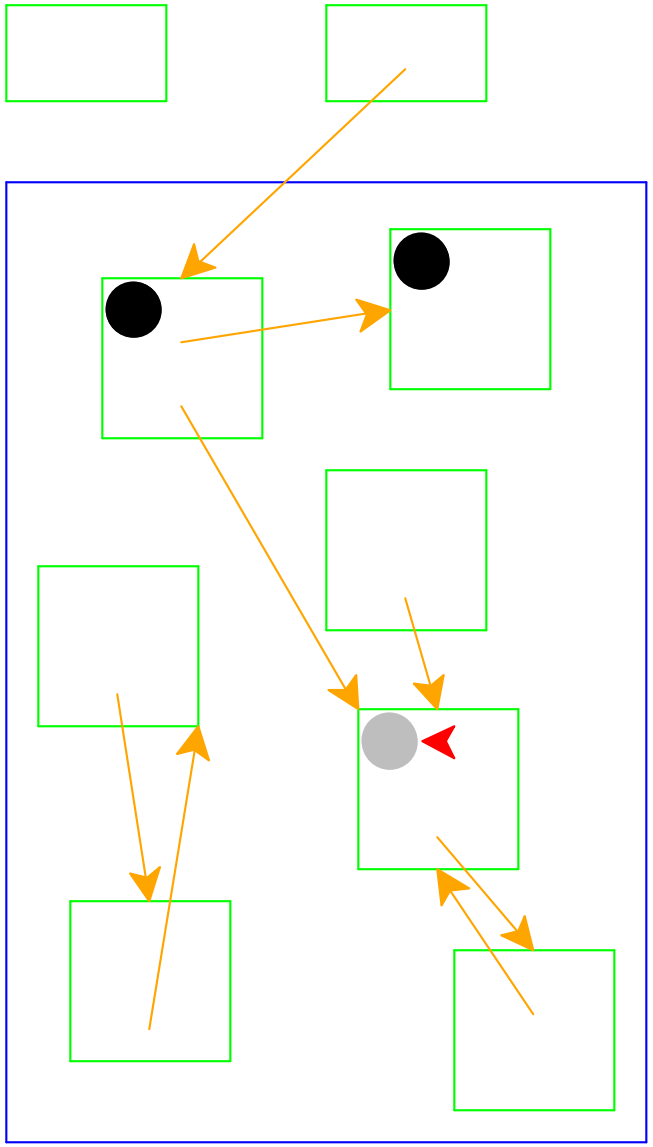
Start again: pick a gray record

Mark & Sweep Garbage Collection



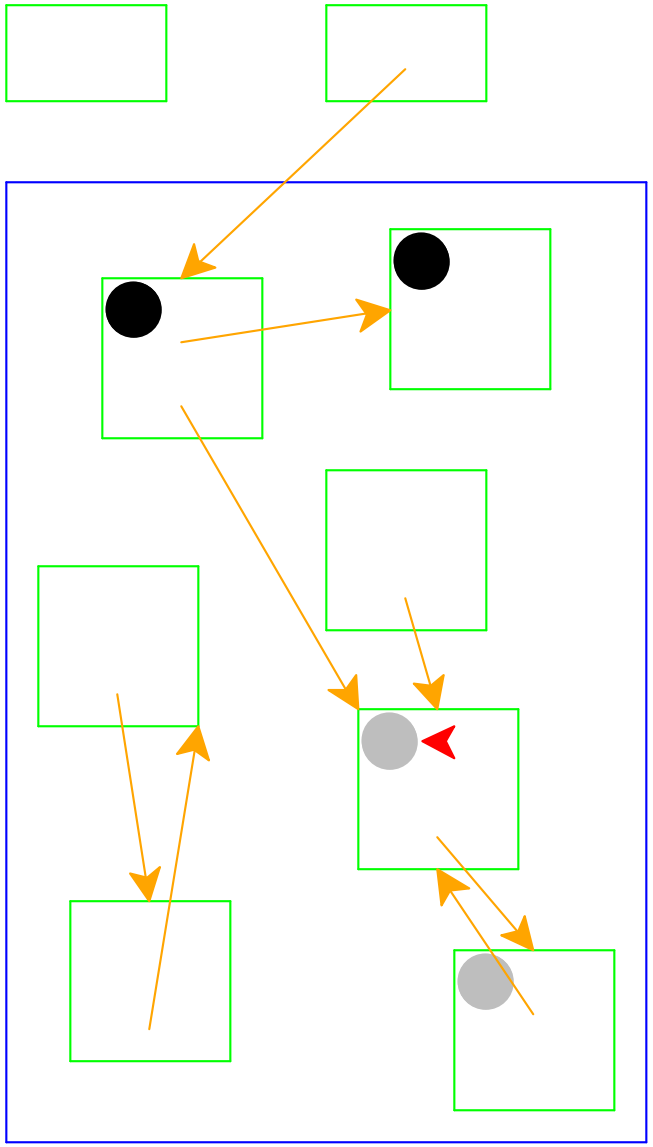
No referenced records; mark black

Mark & Sweep Garbage Collection



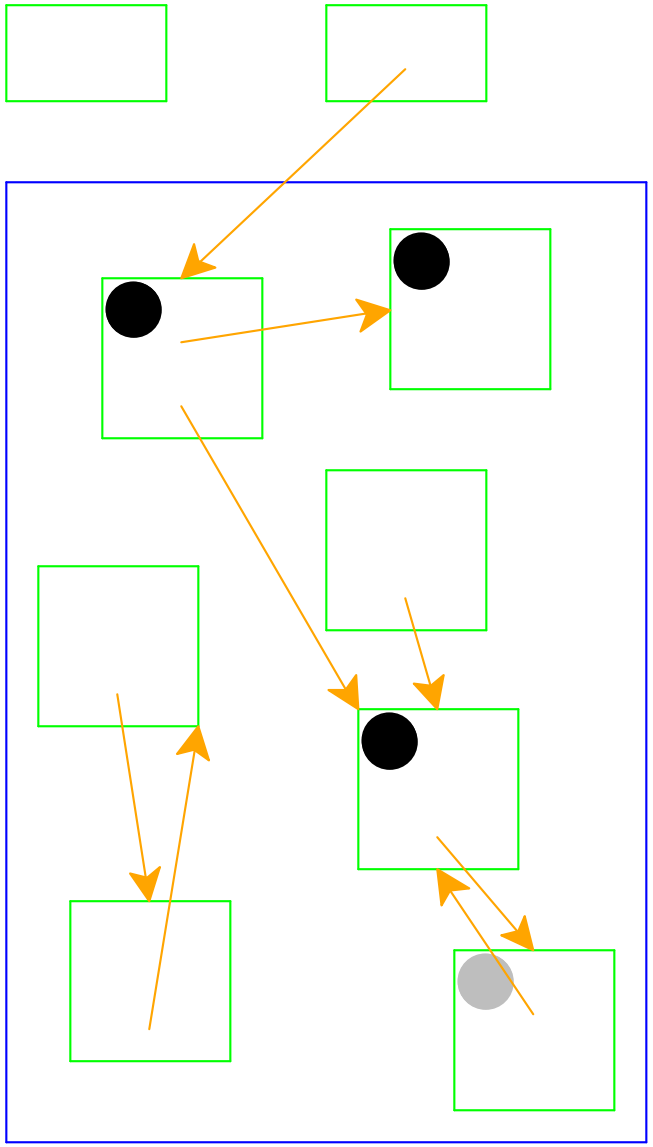
Start again: pick a gray record

Mark & Sweep Garbage Collection



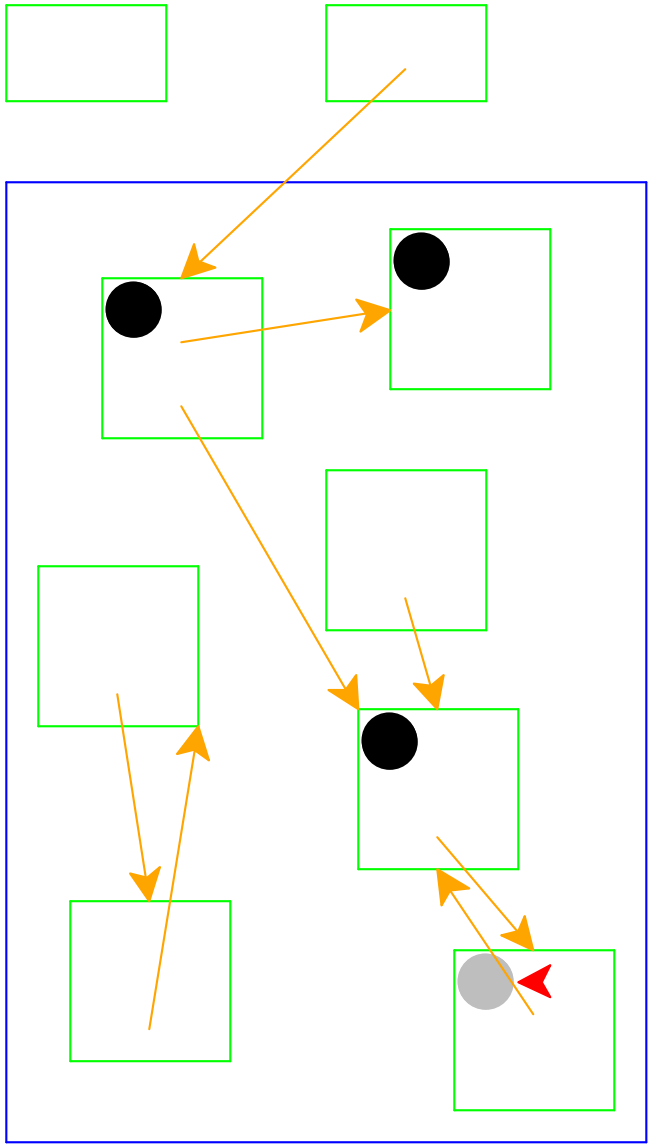
Mark white records referenced by chosen record as gray

Mark & Sweep Garbage Collection



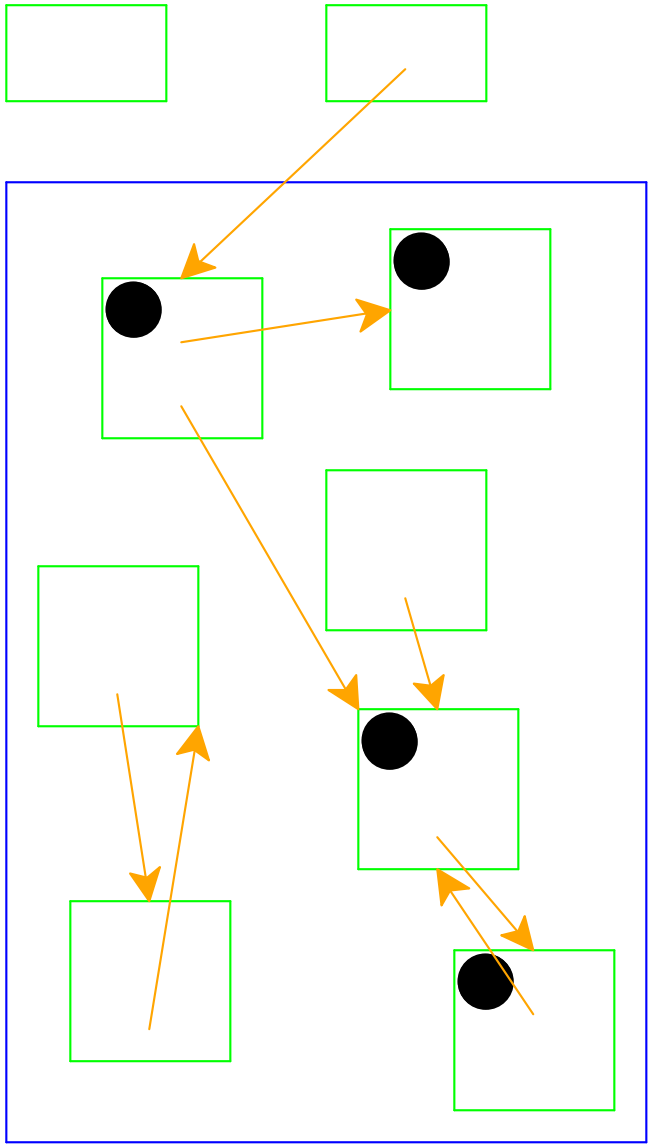
Mark chosen record black

Mark & Sweep Garbage Collection



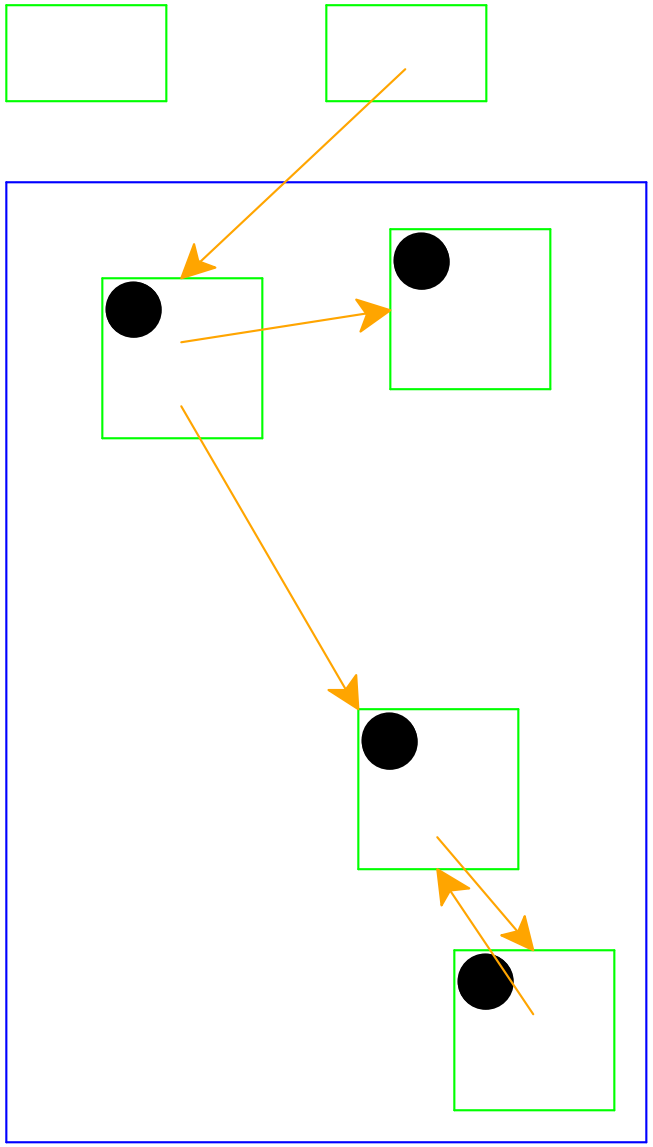
Start again: pick a gray record

Mark & Sweep Garbage Collection



No referenced white records; mark black

Mark & Sweep Garbage Collection



No more gray records; deallocate white records

Cycles ***do not*** break garbage collection

Mark & Sweep Problems

- Cost of collection proportional to (entire) heap
- Bad locality
- Need to use free lists to track available memory

(But there are times when this is a good choice)

Two-Space Copying Collectors

A ***two-space*** copying collector compacts memory as it collects, making allocation easier.

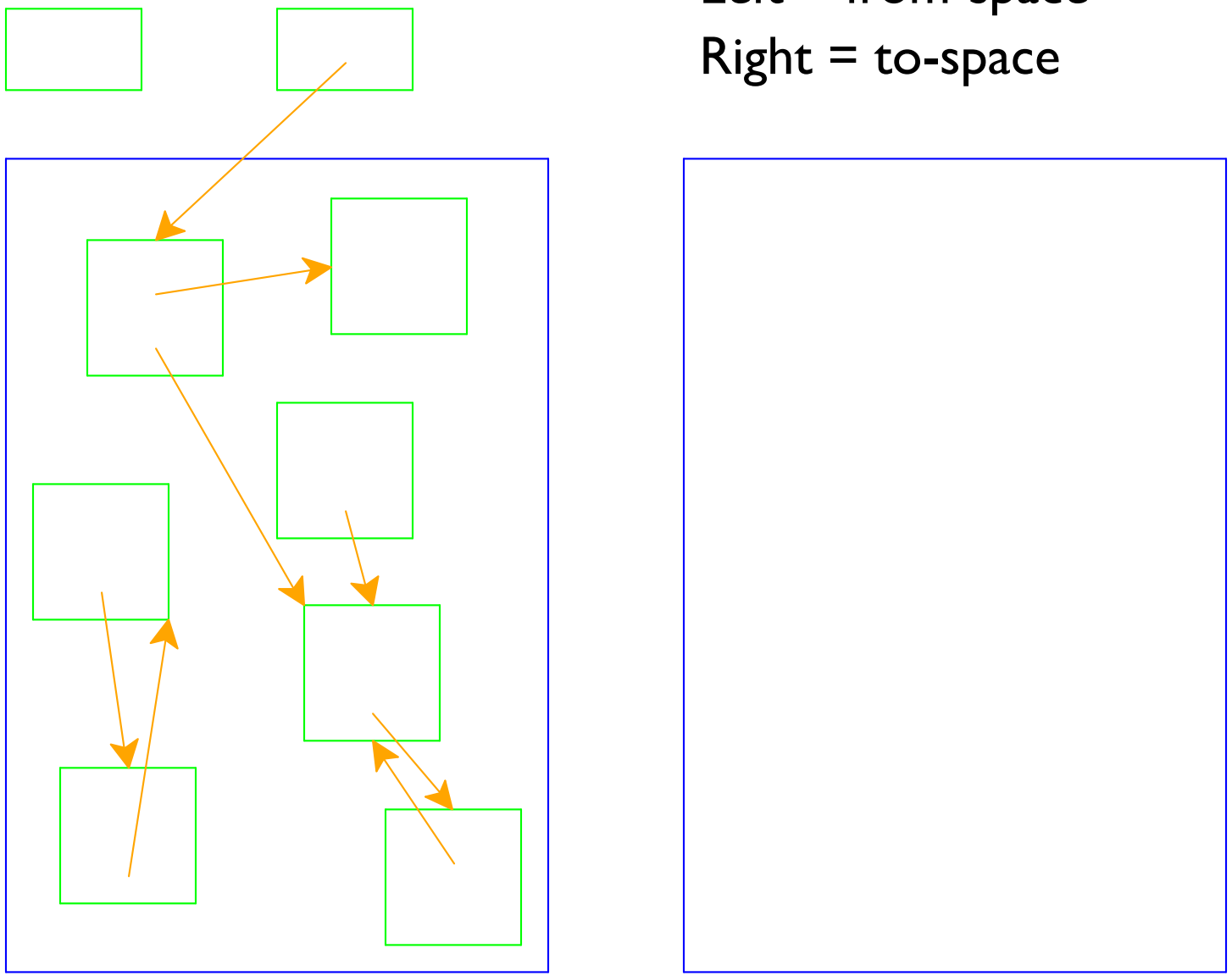
Allocator:

- Partitions memory into ***to-space*** and ***from-space***
- Allocates only in ***to-space***

Collector:

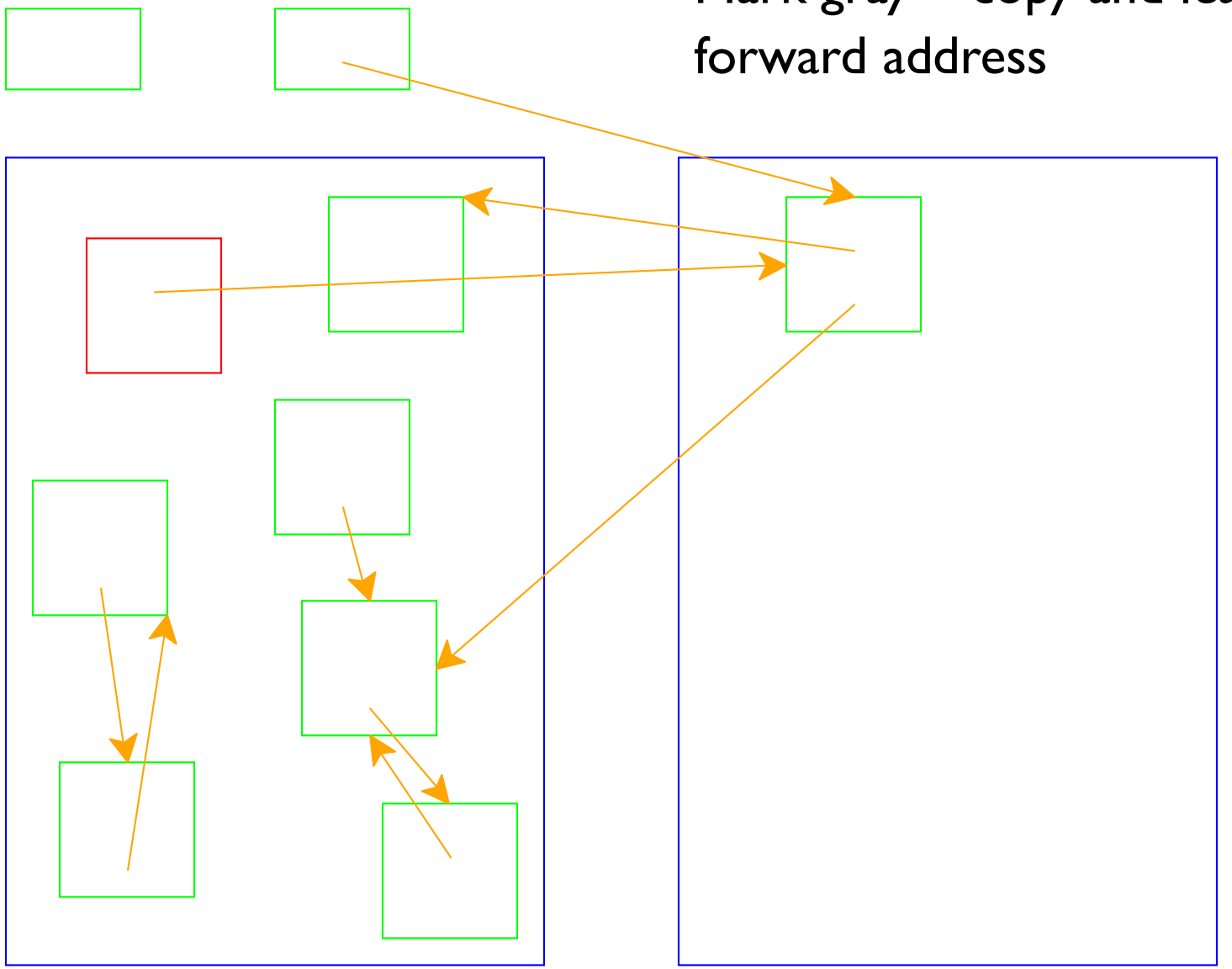
- Starts by swapping ***to-space*** and ***from-space***
- Coloring gray ⇒ copy from ***from-space*** to ***to-space***
- Choosing a gray record ⇒ walk once though the new ***to-space***, update pointers

Two-Space Collection



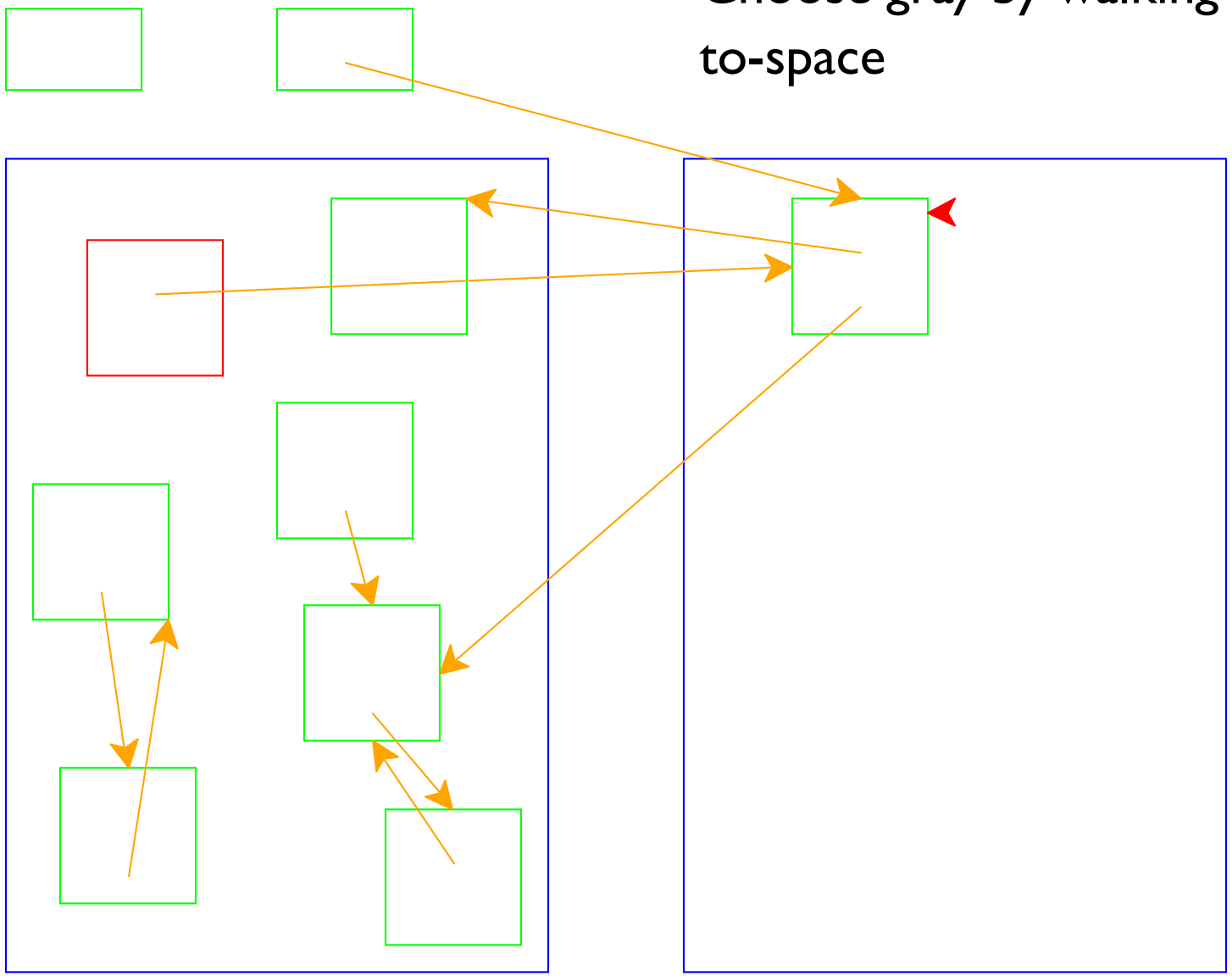
Two-Space Collection

Mark gray = copy and leave forward address



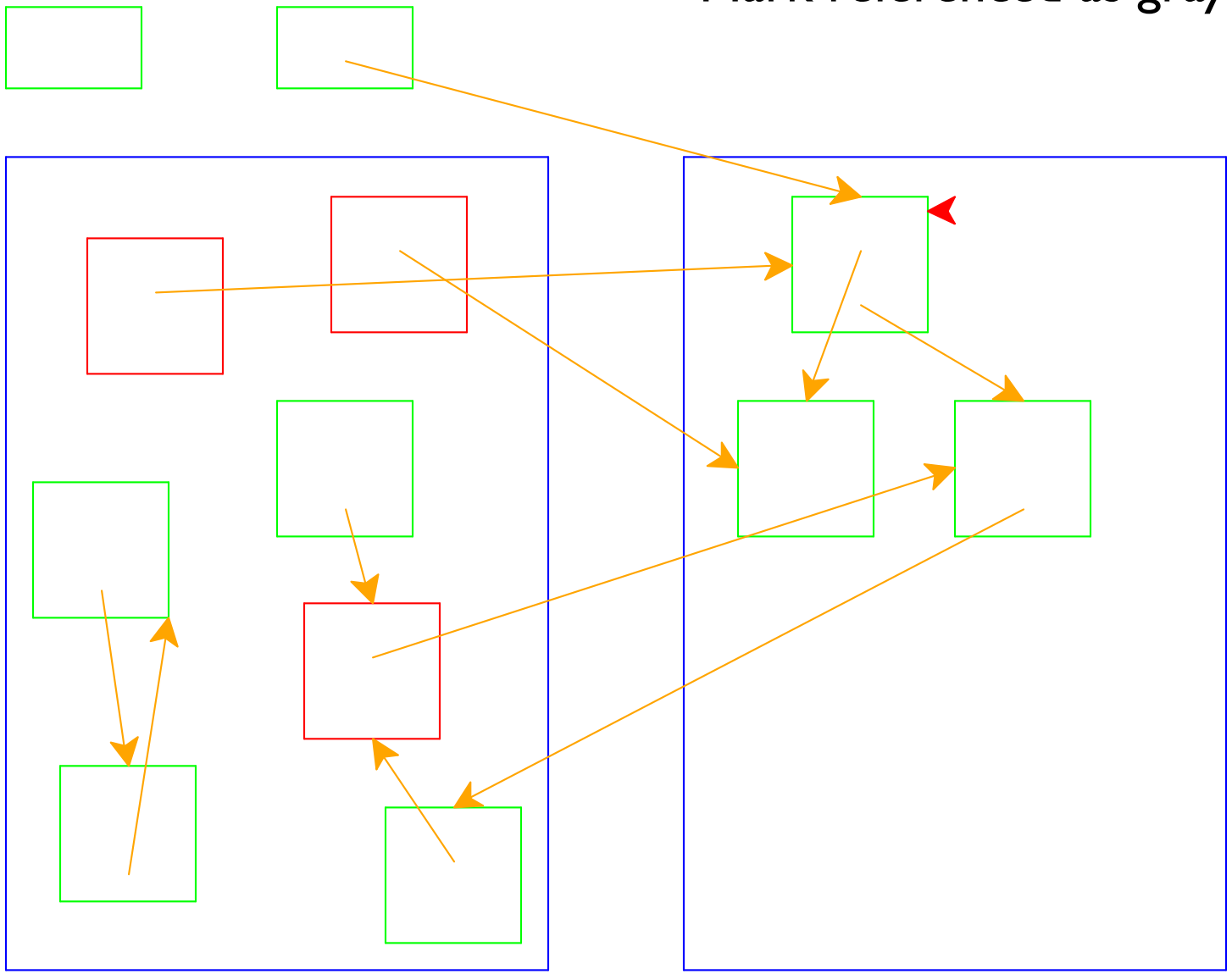
Two-Space Collection

Choose gray by walking through to-space



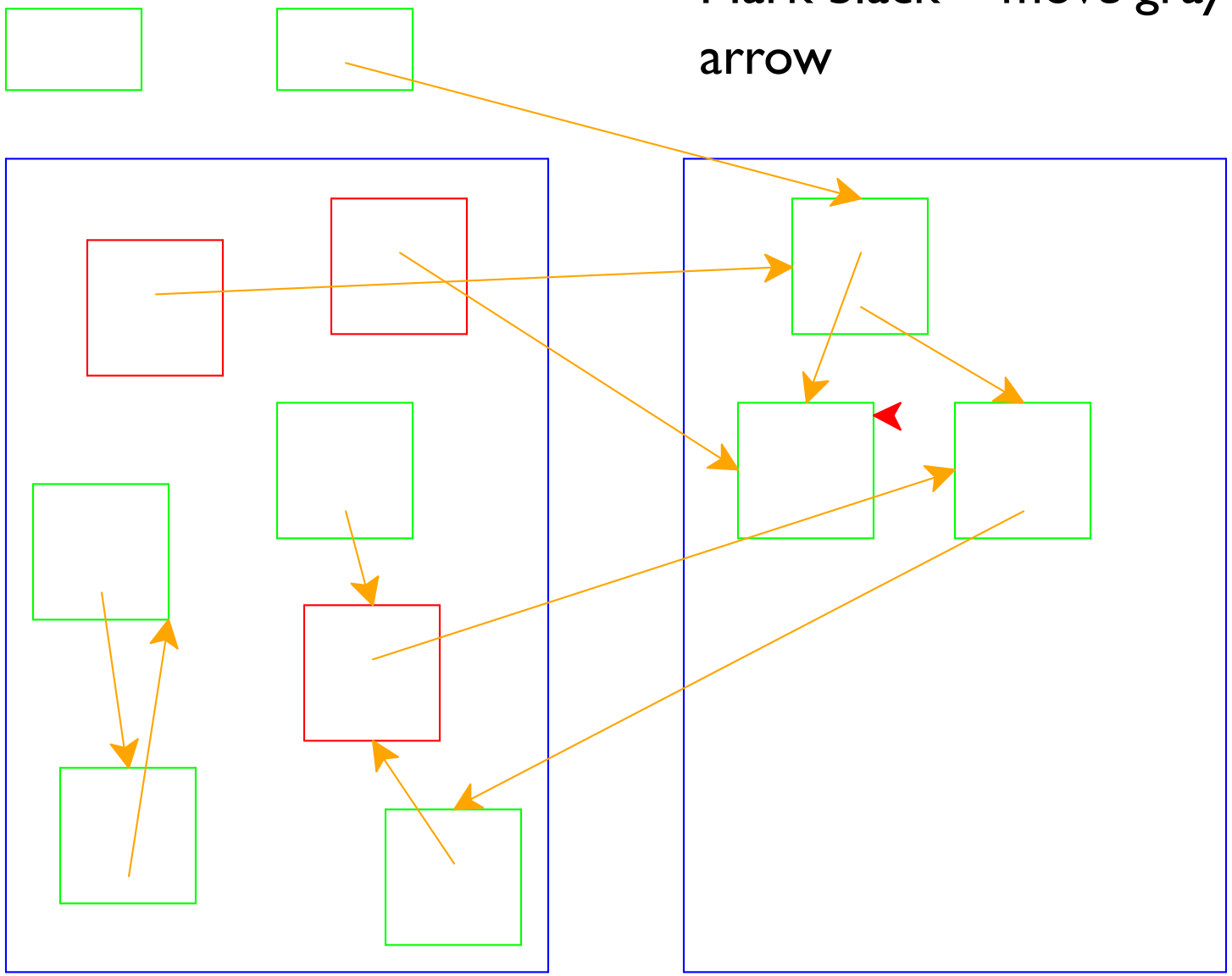
Two-Space Collection

Mark referenced as gray



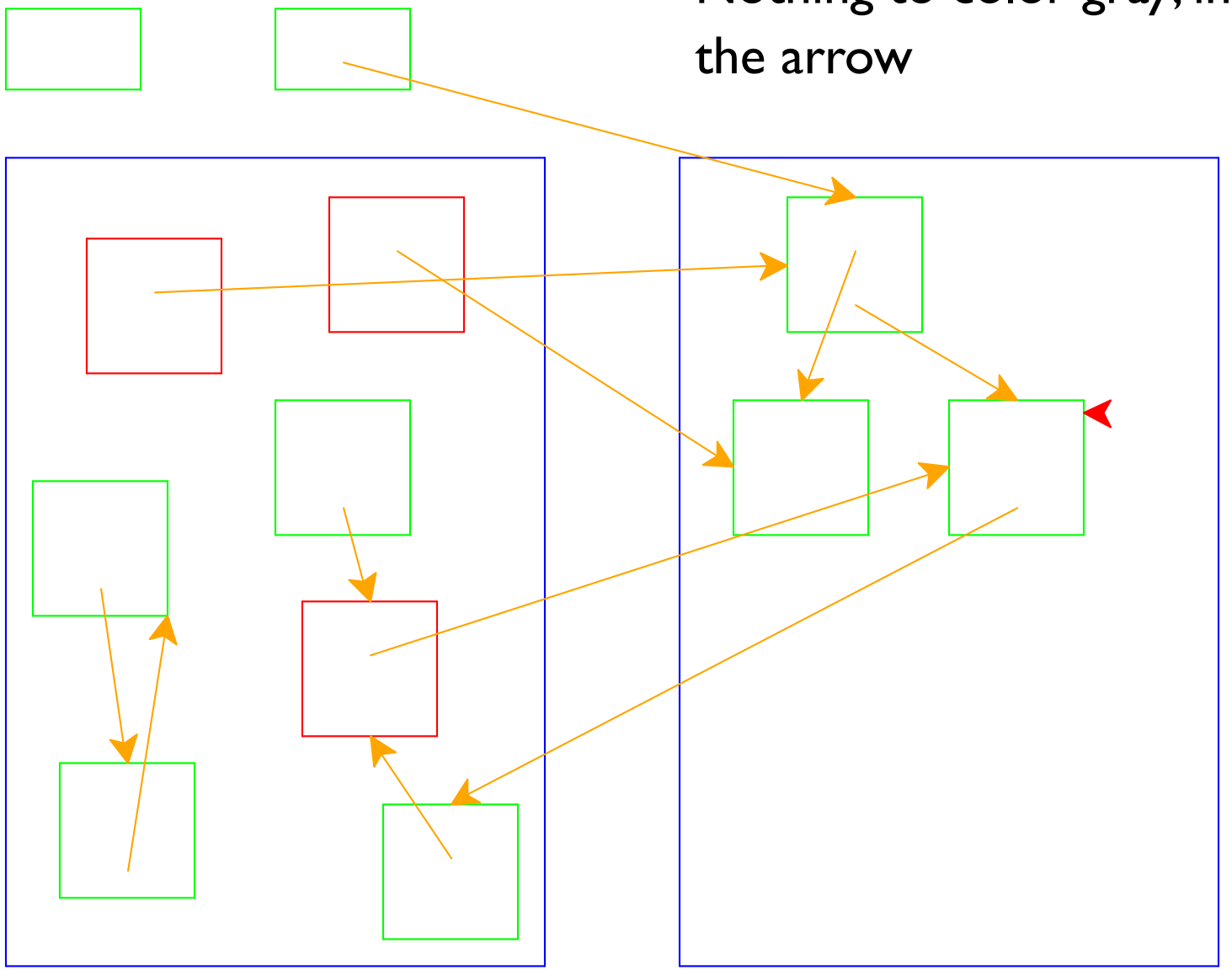
Two-Space Collection

Mark black = move gray-choosing arrow



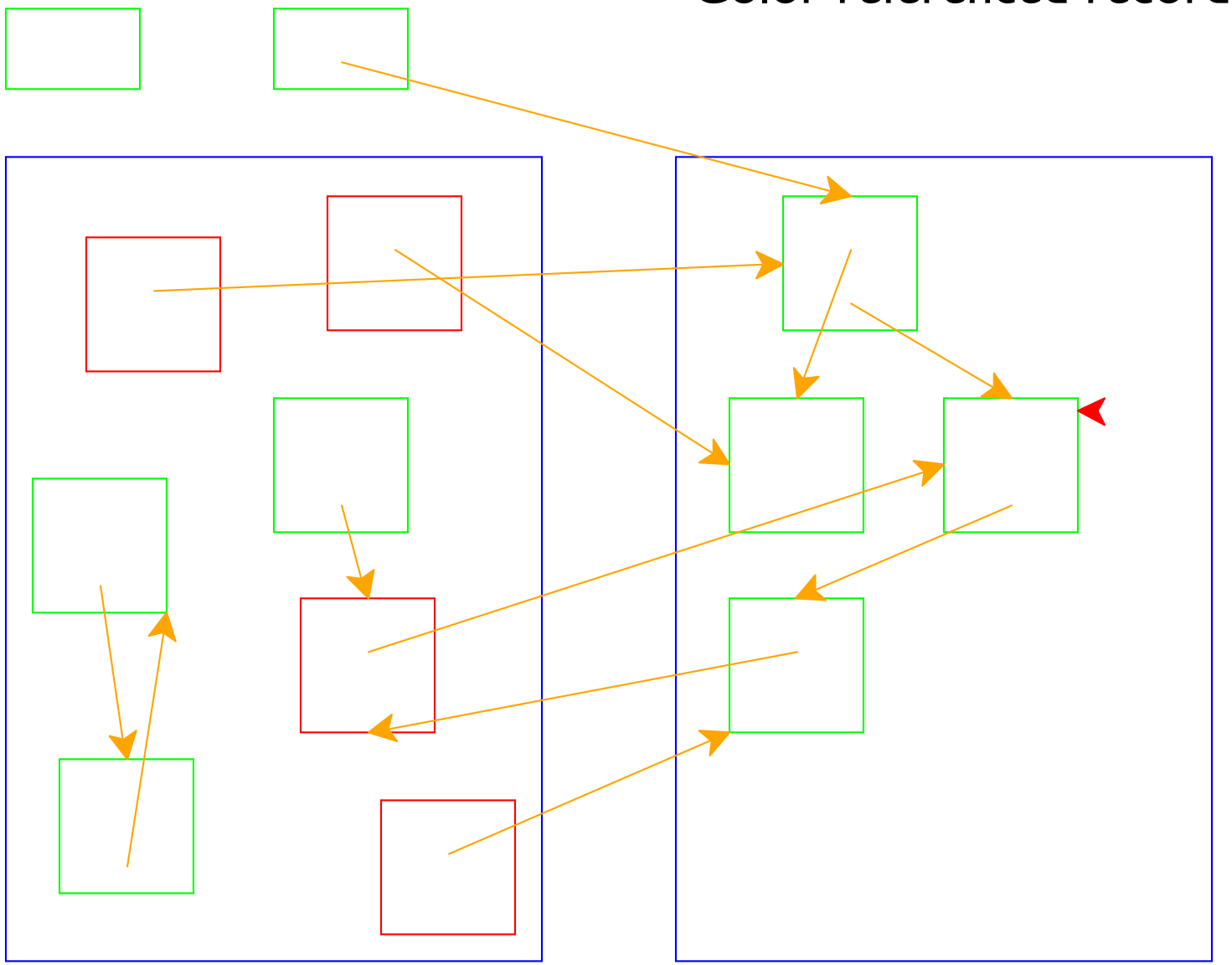
Two-Space Collection

Nothing to color gray; increment the arrow



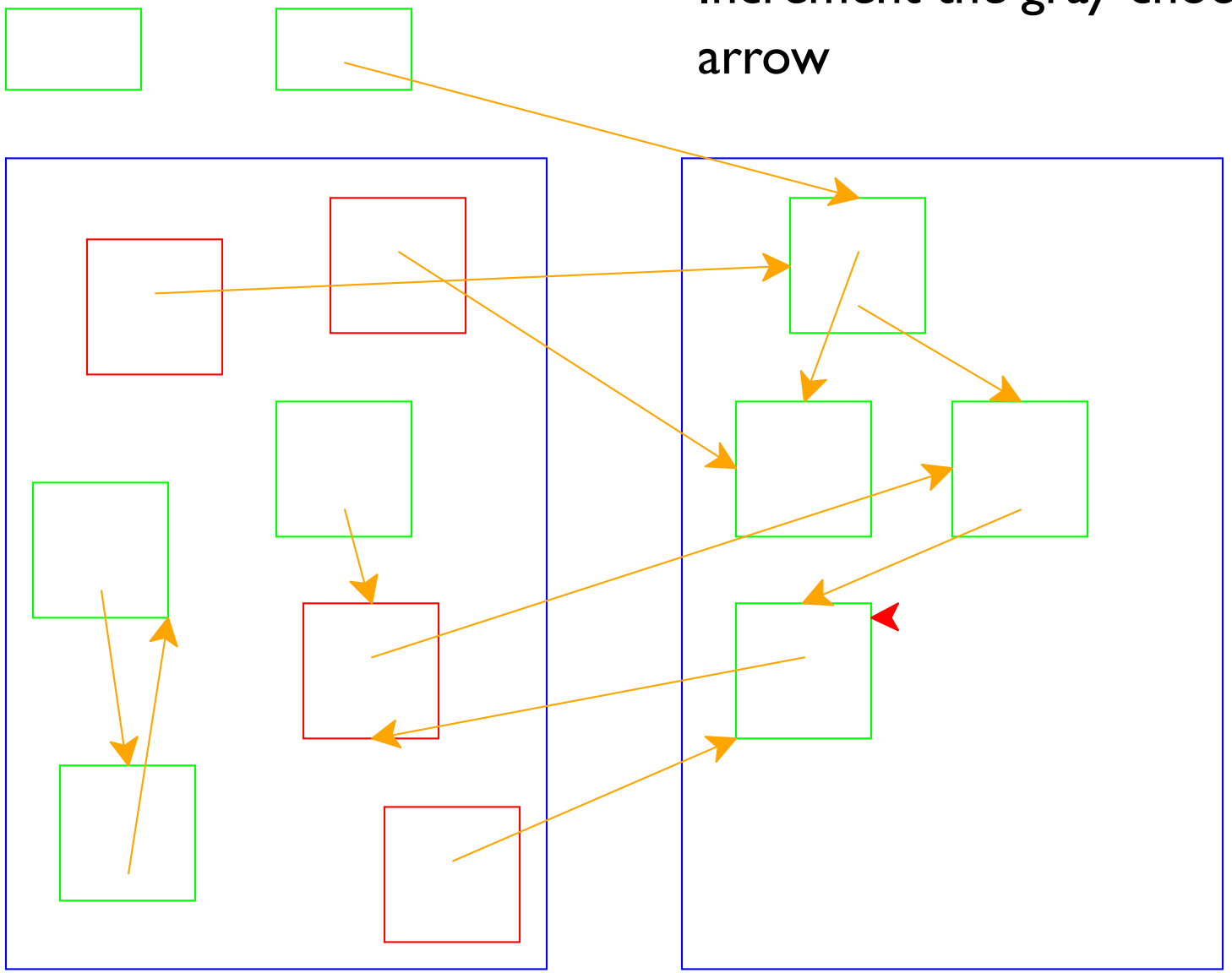
Two-Space Collection

Color referenced record gray



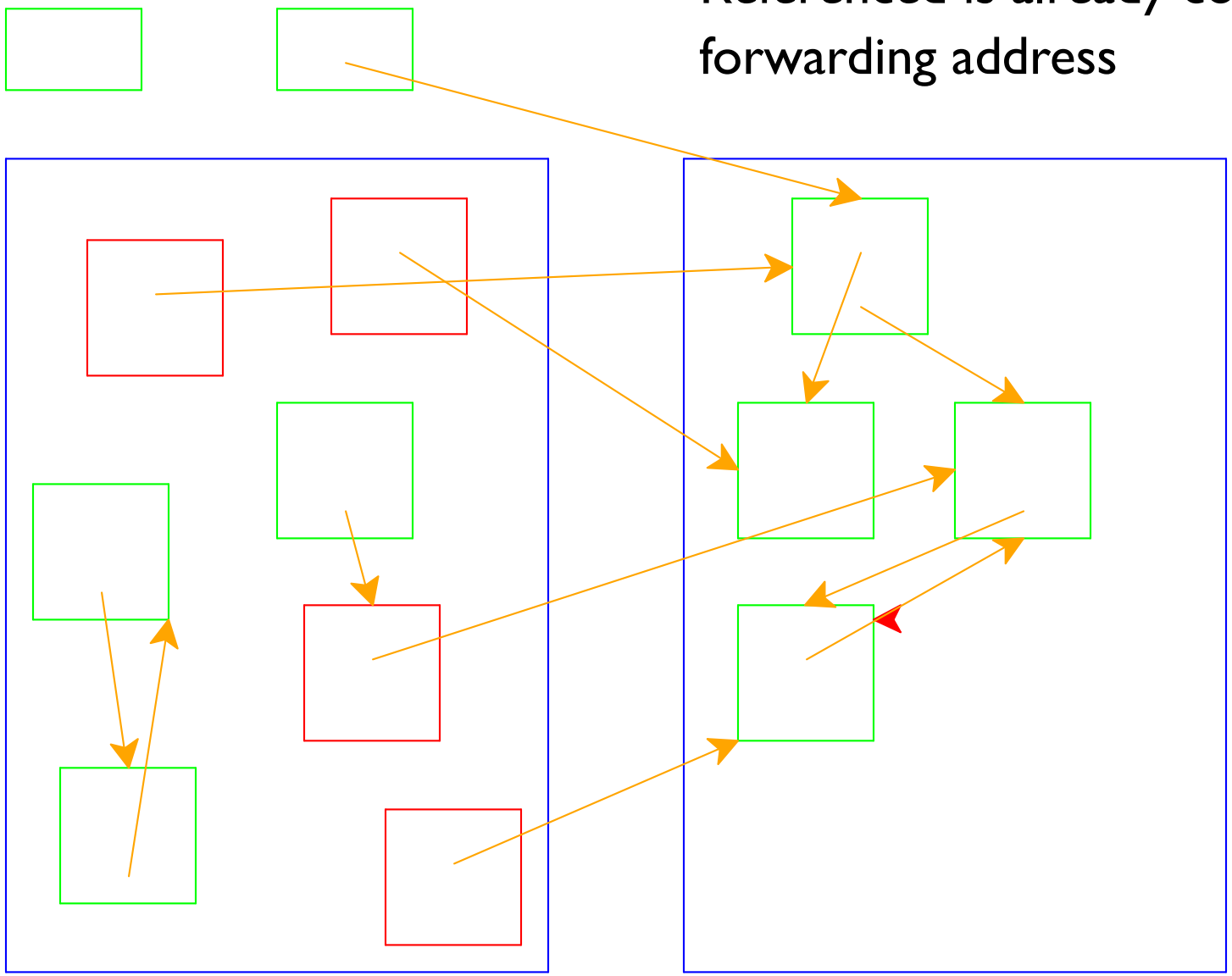
Two-Space Collection

Increment the gray-choosing arrow



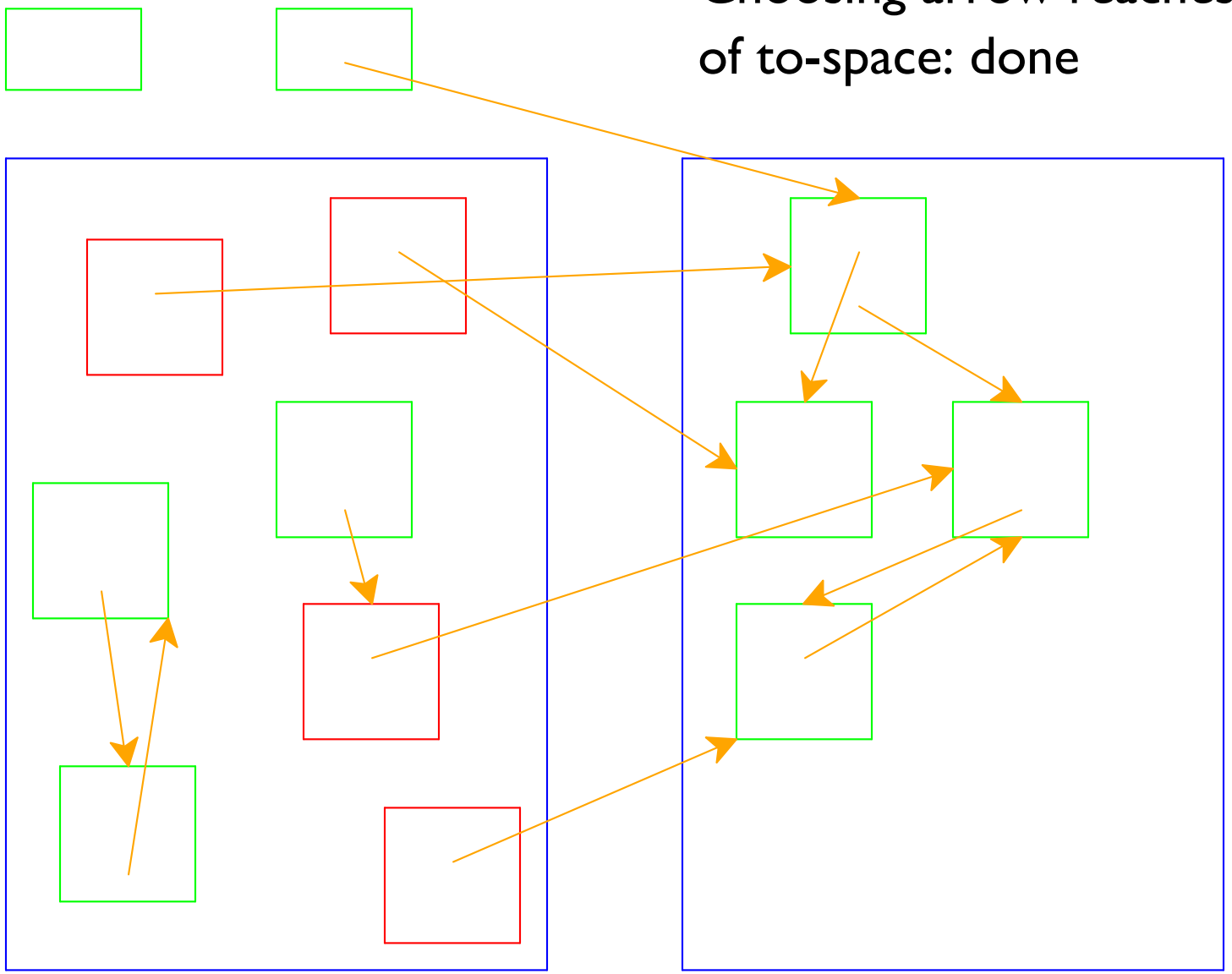
Two-Space Collection

Referenced is already copied, use forwarding address

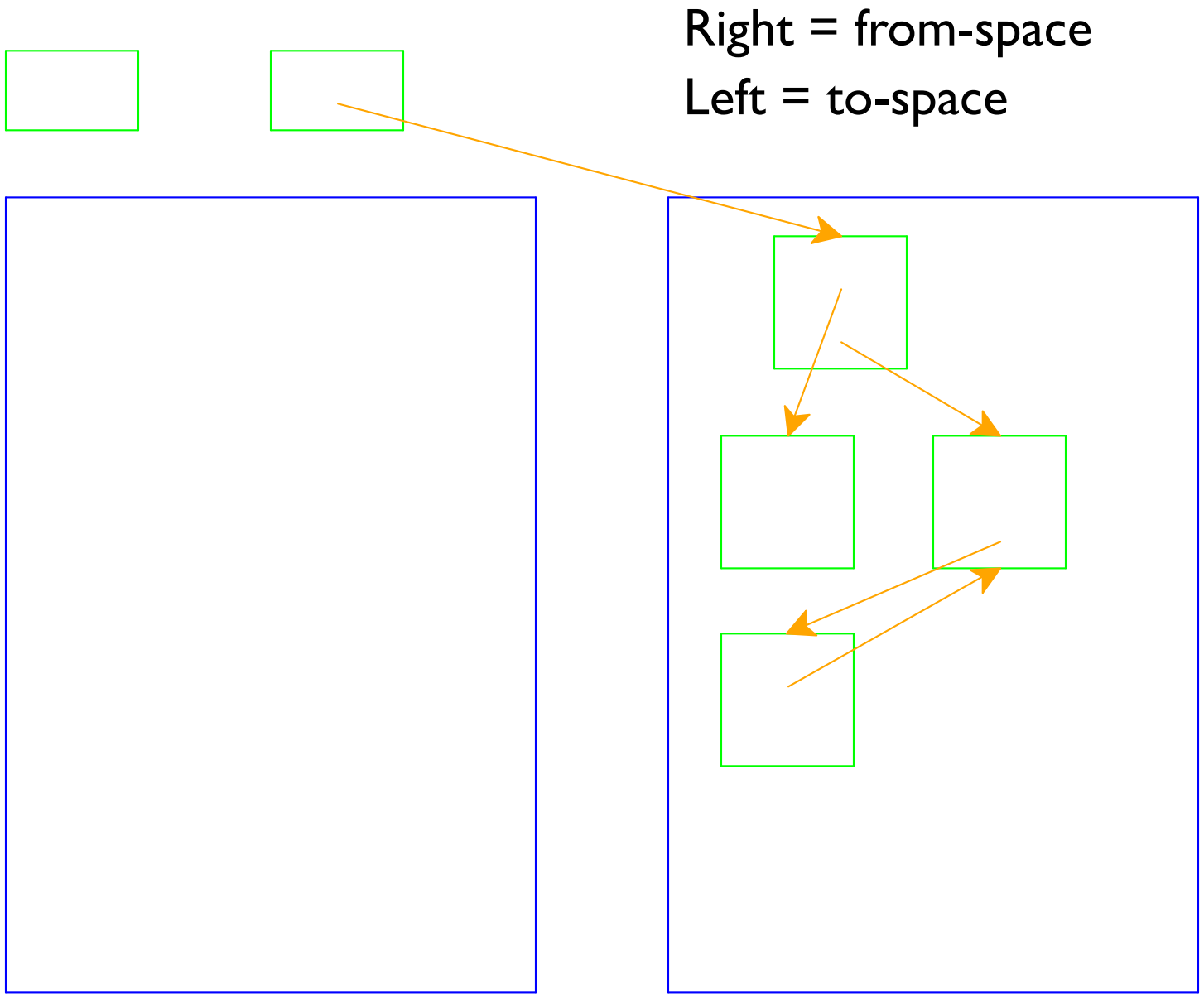


Two-Space Collection

Choosing arrow reaches the end of to-space: done



Two-Space Collection



Two-Space Collection on Vectors

- Everything is a number:
  - Some numbers are immediate integers
  - Some numbers are pointers
- An allocated record in memory starts with a tag, followed by a sequence of pointers and immediate integers
  - The tag describes the shape

Two-Space Collection on Vectors

Use breadth-first search; as you traverse the heap, move objects from one space to the other. Maintain a queue in the to-space. Invariant: objects in the queue have pointers to the from-space; objects outside the queue have pointers to new locations of the objects, i.e., in the to-space.

Two-Space Collection on Vectors

- Initialize a queue with the objects pointed at by the roots
- For each object in the queue, for each pointer in the object
  - If the pointer points to an as-yet uncopied object, copy it into queue; update the old version of the object to record where it got copied; update the pointers in the original object, remove it from the queue
  - If the pointer points to an object already copied, just update the pointer so it points to the new location of the object

Two-Space Collection on Vectors

- Use two pointers into the ***to-space*** to maintain a queue for a breadth-first traversal
- Inc the end pointer to add to the queue, increment the front pointer to remove from the queue; when the pointers come together, terminate

Two-Space Vector Example

- 26-byte memory (13 bytes per space), 2 roots
  - Tag 1: one integer
  - Tag 2: one pointer
  - Tag 3: one integer, then one pointer

Root 1: 7 Root 2: 0

From: `1 75 2 0 3 2 10 3 2 2 3 1 4`

Two-Space Vector Example

- 26-byte memory (13 bytes per space), 2 roots
  - Tag 1: one integer
  - Tag 2: one pointer
  - Tag 3: one integer, then one pointer

Root 1: 7 Root 2: 0

From:	1	75	2	0	3	2	10	3	2	2	3	1	4
Addr:	00	01	02	03	04	05	06	07	08	09	10	11	12

Two-Space Vector Example

- 26-byte memory (13 bytes per space), 2 roots
  - Tag 1: one integer
  - Tag 2: one pointer
  - Tag 3: one integer, then one pointer

Root 1: 7 Root 2: 0

From:	1	75	2	0	3	2	10	3	2	2	3	1	4
Addr:	00	01	02	03	04	05	06	07	08	09	10	11	12
	^		^		^			^			^		

Two-Space Vector Example

- 26-byte memory (13 bytes per space), 2 roots
  - Tag 1: one integer
  - Tag 2: one pointer
  - Tag 3: one integer, then one pointer

Root 1: 7 Root 2: 0

From:	1	75	2	0	3	2	10	3	2	2	3	1	4
Addr:	00	01	02	03	04	05	06	07	08	09	10	11	12
	^		^		^			^			^		
To:	0	0	0	0	0	0	0	0	0	0	0	0	0
Q:													
Addr:	13	14	15	16	17	18	19	20	21	22	23	24	25

Two-Space Vector Example

- 26-byte memory (13 bytes per space), 2 roots
  - Tag 1: one integer
  - Tag 2: one pointer
  - Tag 3: one integer, then one pointer
  - Tag 99: forwarding pointer (to to-space)

Root 1: 13 Root 2: 0

From:	1	75	2	0	3	2	10	99	13	2	3	1	4
Addr:	00	01	02	03	04	05	06	07	08	09	10	11	12
	^		^		^			^			^		
To:	3	2	2	0	0	0	0	0	0	0	0	0	0
Q:	^			^									
Addr:	13	14	15	16	17	18	19	20	21	22	23	24	25

Two-Space Vector Example

- 26-byte memory (13 bytes per space), 2 roots
  - Tag 1: one integer
  - Tag 2: one pointer
  - Tag 3: one integer, then one pointer
  - Tag 99: forwarding pointer (to to-space)

Root 1: 13 Root 2: 16

From:	99	16	2	0	3	2	10	99	13	2	3	1	4
Addr:	00	01	02	03	04	05	06	07	08	09	10	11	12
	^		^		^			^			^		
To:	3	2	2	1	75	0	0	0	0	0	0	0	0
Q:	^					^							
Addr:	13	14	15	16	17	18	19	20	21	22	23	24	25

Two-Space Vector Example

- 26-byte memory (13 bytes per space), 2 roots
  - Tag 1: one integer
  - Tag 2: one pointer
  - Tag 3: one integer, then one pointer
  - Tag 99: forwarding pointer (to to-space)

Root 1: 13 Root 2: 16

From:	99	16	99	18	3	2	10	99	13	2	3	1	4
Addr:	00	01	02	03	04	05	06	07	08	09	10	11	12
	^		^		^			^			^		
To:	3	2	18	1	75	2	0	0	0	0	0	0	0
Q:				^				^					
Addr:	13	14	15	16	17	18	19	20	21	22	23	24	25

Two-Space Vector Example

- 26-byte memory (13 bytes per space), 2 roots
  - Tag 1: one integer
  - Tag 2: one pointer
  - Tag 3: one integer, then one pointer
  - Tag 99: forwarding pointer (to to-space)

Root 1: 13 Root 2: 16

From:	99	16	99	18	3	2	10	99	13	2	3	1	4
Addr:	00	01	02	03	04	05	06	07	08	09	10	11	12
	^		^		^			^			^		
To:	3	2	18	1	75	2	0	0	0	0	0	0	0
Q:						^		^					
Addr:	13	14	15	16	17	18	19	20	21	22	23	24	25

Two-Space Vector Example

- 26-byte memory (13 bytes per space), 2 roots
  - Tag 1: one integer
  - Tag 2: one pointer
  - Tag 3: one integer, then one pointer
  - Tag 99: forwarding pointer (to to-space)

Root 1: 13 Root 2: 16

```
From:  99 16 99 18  3  2 10 99 13  2  3  1  4
Addr:  00 01 02 03 04 05 06 07 08 09 10 11 12
        ^     ^     ^        ^        ^
To:     3  2 18  1 75  2 16  0  0  0  0  0  0
Q:                          ^^
Addr:  13 14 15 16 17 18 19 20 21 22 23 24 25
```

Further reading

Uniprocessor Garbage Collection Techniques, by Wilson

`ftp://ftp.cs.utexas.edu/pub/garbage/gcsurvey.ps`

Mark and sweep implementation, with linear-time allocator

```
;; init-allocator : -> void
(define (init-allocator)
  (for ([i (in-range 0 (heap-size))])
    (heap-set! i 'free)))

;; gc:flat? : location? -> flat-value?
(define (gc:flat? loc)
  (equal? (heap-ref loc) 'flat))

;; gc:deref : location? -> void?
(define (gc:deref loc)
  (cond
    [(equal? (heap-ref loc) 'flat)
     (heap-ref (+ loc 1))]
    [else
     (error 'gc:deref
            "attempted to deref a non flat value, loc ~s"
            loc)]))
```

Mark and sweep implementation, with linear-time allocator

```
;; gc:cons? : location? -> boolean?
(define (gc:cons? loc)
  (equal? (heap-ref loc) 'pair))

;; gc:first : location? -> location?
(define (gc:first pr-ptr)
  (if (equal? (heap-ref pr-ptr) 'pair)
      (heap-ref (+ pr-ptr 1))
      (error 'first "non pair")))

;; gc:rest : location? -> location?
(define (gc:rest pr-ptr)
  (if (equal? (heap-ref pr-ptr) 'pair)
      (heap-ref (+ pr-ptr 2))
      (error 'rest "non pair")))
```

Mark and sweep implementation, with linear-time allocator

```
;; gc:set-first! : location? location? -> void?
(define (gc:set-first! pr-ptr new)
  (if (equal? (heap-ref pr-ptr) 'pair)
      (heap-set! (+ pr-ptr 1) new)
      (error 'set-first! "non pair")))

;; gc:set-rest! : location? location? -> void?
(define (gc:set-rest! pr-ptr new)
  (if (equal? (heap-ref pr-ptr) 'pair)
      (heap-set! (+ pr-ptr 2) new)
      (error 'set-first! "non pair")))
```

Mark and sweep implementation, with linear-time allocator

```
;; gc:closure-code-ptr : loc -> heap-value
(define (gc:closure-code-ptr loc)
  (if (gc:closure? loc)
      (heap-ref (+ loc 1))
      (error 'gc:closure-code "non closure @ ~a, got ~s"
             loc (heap-ref loc))) )

;; gc:closure-env-ref : loc number -> loc
(define (gc:closure-env-ref loc i)
  (if (gc:closure? loc)
      (if (< i (heap-ref (+ loc 2)))
          (heap-ref (+ loc 3 i))
          (error 'closure-env-ref
                 "closure-env-ref out of bounds"))
      (error 'closure-env-ref "non closure")))

;; gc:closure? : loc -> boolean
(define (gc:closure? loc)
  (equal? (heap-ref loc) 'proc))
```

Mark and sweep implementation, with linear-time allocator

```
;; gc:alloc-flat : flat-values? -> location?
(define (gc:alloc-flat fv)
  (let ([ptr (alloc 2 #f #f)])
    (heap-set! ptr 'flat)
    (heap-set! (+ ptr 1) fv)
    ptr))

;; gc:cons : root? root? -> location?
(define (gc:cons hd tl)
  (define ptr (alloc 3 hd tl))
  (heap-set! ptr 'pair)
  (heap-set! (+ ptr 1) (read-root hd))
  (heap-set! (+ ptr 2) (read-root tl))
  ptr)
```

Mark and sweep implementation, with linear-time allocator

```
;; gc:closure : heap-value (vectorof loc) -> loc
(define (gc:closure code-ptr free-vars)
  (define free-vars-count (length free-vars))
  (define next (alloc (+ free-vars-count 3) free-vars '()))
  (heap-set! next 'proc)
  (heap-set! (+ next 1) code-ptr)
  (heap-set! (+ next 2) free-vars-count)
  (for ([x (in-range 0 free-vars-count)]
        [r (in-list free-vars)])
    (heap-set! (+ next 3 x) (read-root r)))
  next)
```

Mark and sweep implementation, with linear-time allocator

```
a roots is either:
   - root
   - loc
   - (listof roots)

;; alloc : number[size] roots roots -> loc
(define (alloc n some-roots some-more-roots)
  (let ([next (find-free-space 0 n)])
    (cond
      [next
       next]
      [else
       (collect-garbage some-roots some-more-roots)
       (let ([next (find-free-space 0 n)])
         (unless next
           (error 'alloc "out of space"))
         next)])))
```

Mark and sweep implementation, with linear-time allocator

```
;; find-free-space : loc number -> loc or #f
;; start must be a valid pointer
;; (not to the middle of an object)
(define (find-free-space start size)
  (cond
    [(= start (heap-size)) #f]
    [else
     (case (heap-ref start)
       [(free)
        (if (n-free-blocks? start size)
            start
            (find-free-space (+ start 1) size))]
       [(flat) (find-free-space (+ start 2) size)]
       [(pair) (find-free-space (+ start 3) size)]
       [(proc)
        (find-free-space (+ start 3 (heap-ref (+ start 2)))
                         size)]
       [else
        (error 'find-free-space "ack ~s" start)])]))
```

Mark and sweep implementation, with linear-time allocator

```
;; n-free-blocks? : location? nat -> boolean?
(define (n-free-blocks? start size)
  (cond
    [(= size 0) #t]
    [(= start (heap-size)) #f]
    [else
     (and (eq? 'free (heap-ref start))
          (n-free-blocks? (+ start 1) (- size 1)))]))
```

Mark and sweep implementation, with linear-time allocator

```
;; collect-garbage : roots roots -> void?
(define (collect-garbage some-roots some-more-roots)
  (validate-heap)
  (mark-white! 0)
  (traverse/roots (get-root-set))
  (traverse/roots some-roots)
  (traverse/roots some-more-roots)
  (free-white! 0)
  (validate-heap))
```

Mark and sweep implementation, with linear-time allocator

```
;; validate-heap : -> void?
(define (validate-heap)
  (define (valid-pointer? i)
    (unless (memq (heap-ref i) '(flat pair proc))
      (error 'validate-heap "found bad pointer @ ~a" i)))
  (let loop ([i 0])
    (when (< i (heap-size))
      (case (heap-ref i)
        [(flat)
         (loop (+ i 2))]
        [(pair)
         (valid-pointer? (heap-ref (+ i 1)))
         (valid-pointer? (heap-ref (+ i 2)))
         (loop (+ i 3))]
        [(proc)
         (for ([x (in-range 0 (heap-ref (+ i 2)))])
           (valid-pointer? (heap-ref (+ i 3 x))))
         (loop (+ i 3 (heap-ref (+ i 2))))]
        [(free)
         (loop (+ i 1))]
        [else (error 'validate-heap "corruption! @ ~a" i)])))))
```

Mark and sweep implementation, with linear-time allocator

```
;; mark-white! : location -> void?
(define (mark-white! i)
  (when (< i (heap-size))
    (case (heap-ref i)
      [(pair) (heap-set! i 'white-pair)
              (mark-white! (+ i 3))]
      [(flat) (heap-set! i 'white-flat)
              (mark-white! (+ i 2))]
      [(proc) (heap-set! i 'white-proc)
              (mark-white! (+ i 3 (heap-ref (+ i 2))))]
      [(free) (mark-white! (+ i 1))]
      [else
       (error 'mark-white! "unknown tag @ ~a" i)])))
```

Mark and sweep implementation, with linear-time allocator

```
;; free-white! : location? -> void?
(define (free-white! i)
  (when (< i (heap-size))
    (case (heap-ref i)
      [(pair)       (free-white! (+ i 3))]
      [(flat)       (free-white! (+ i 2))]
      [(proc)       (free-white! (+ i 3 (heap-ref (+ i 2))))]
      [(white-pair) (heap-set! i 'free)
                    (heap-set! (+ i 1) 'free)
                    (heap-set! (+ i 2) 'free)
                    (free-white! (+ i 3))]
      [(white-flat) (heap-set! i 'free)
                    (heap-set! (+ i 1) 'free)
                    (free-white! (+ i 2))]
      [(white-proc) (define closure-size (heap-ref (+ i 2)))
                    (for ([dx (in-range 0 (+ closure-size 3))])
                      (heap-set! (+ i dx) 'free))
                    (free-white! (+ i 3 closure-size))]
      [(free)       (free-white! (+ i 1))]
      [else (error 'free-white! "unknown tag ~s"
                   (heap-ref i))])))
```

Mark and sweep implementation, with linear-time allocator

```
;; traverse/roots : roots -> void
(define (traverse/roots thing)
  (cond
    [(list? thing) (for-each traverse/roots thing)]
    [(root? thing) (traverse/roots (read-root thing))]
    [(number? thing) (traverse/loc thing)]))
```

Mark and sweep implementation, with linear-time allocator

```
;; traverse/loc : location -> void
(define (traverse/loc loc)
  (case (heap-ref loc)
    [(white-pair)
     (heap-set! loc 'pair)
     (traverse/loc (heap-ref (+ loc 1)))
     (traverse/loc (heap-ref (+ loc 2)))]
    [(white-flat)
     (heap-set! loc 'flat)]
    [(white-proc)
     (heap-set! loc 'proc)
     (for ([i (in-range (heap-ref (+ loc 2)))])
       (traverse/loc (heap-ref (+ loc 3 i))))]
    [(pair) (void)]
    [(flat) (void)]
    [(proc) (void)]
    [else
     (error 'traverse/loc "crash ~s" loc)]))
```