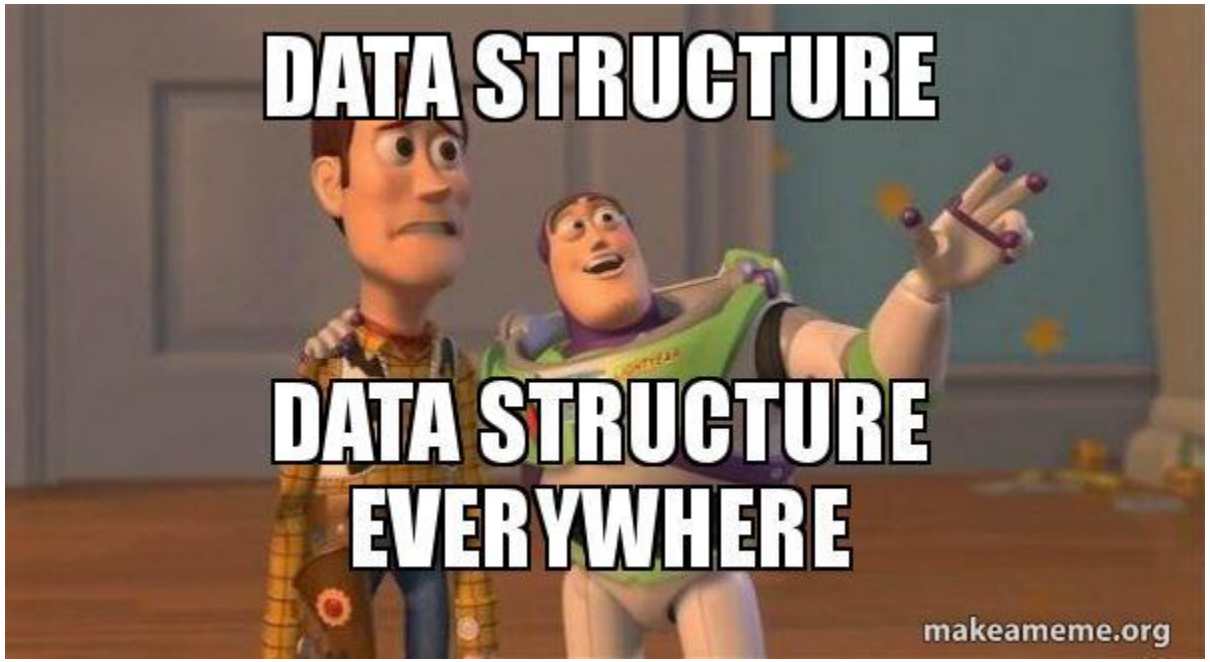
академия больших данных

mail.ru group

Базовые структуры данных

Шовкопляс Григорий

Введение в алгоритмы и структуры данных



Что такое структуры данных?

Структуры данных

- «Такой черный ящик, который хранит данные и позволяет нам удобнее пользоваться этими данными»
- Знаем какую-нибудь структуру данных?
- Массив! Простой пример, но очень корректный
- Поехали изучим что-нибудь еще!

Связный список

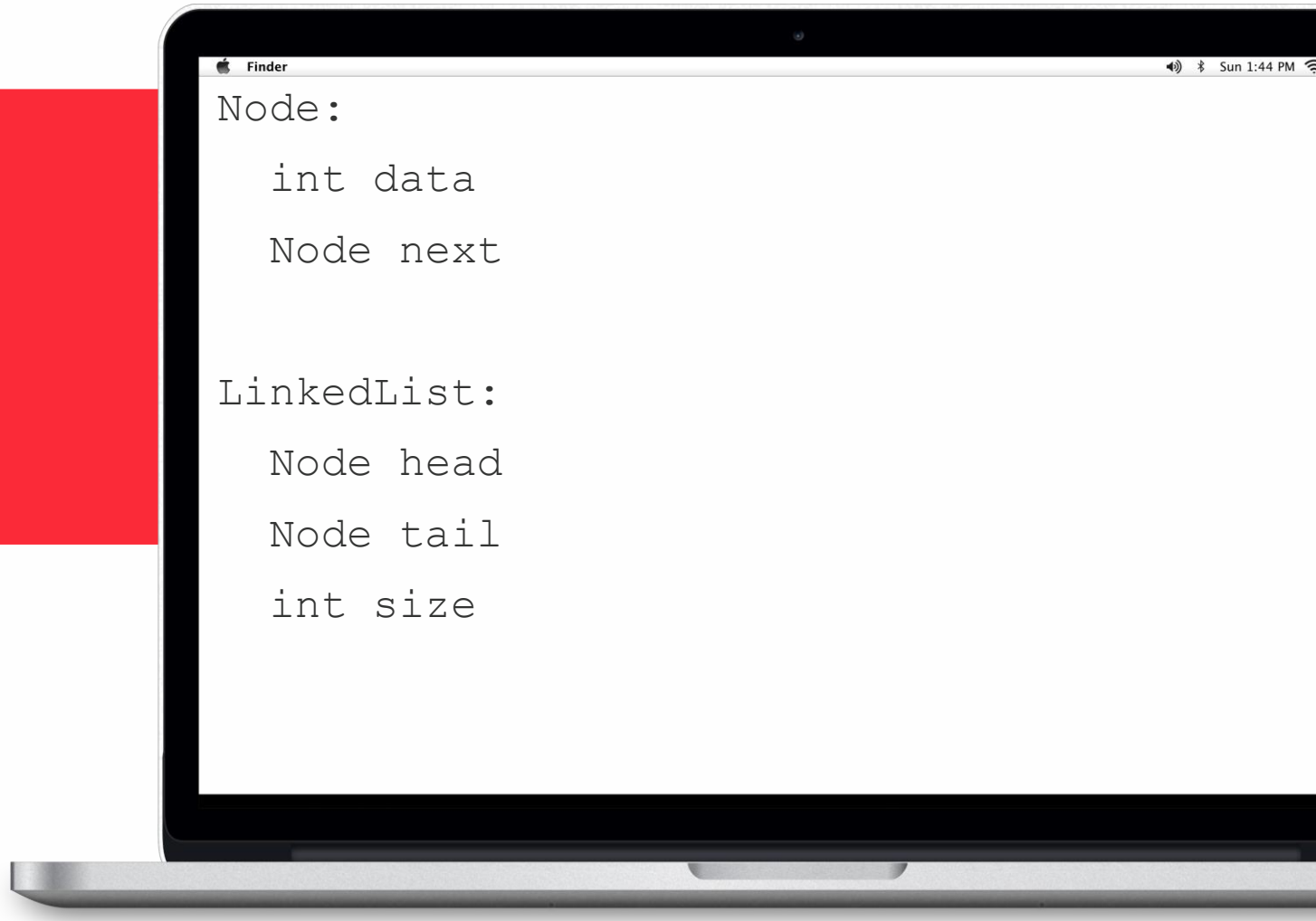
Связный список

- Каждый элемент списка узел (node)
- Хранит значение
- Знает ссылку на следующий элемент
- В чем различия с массивом?
  - O(n) памяти
  - Нельзя обратиться по индексу
  - Вставка элемента за O(1)

Связный список

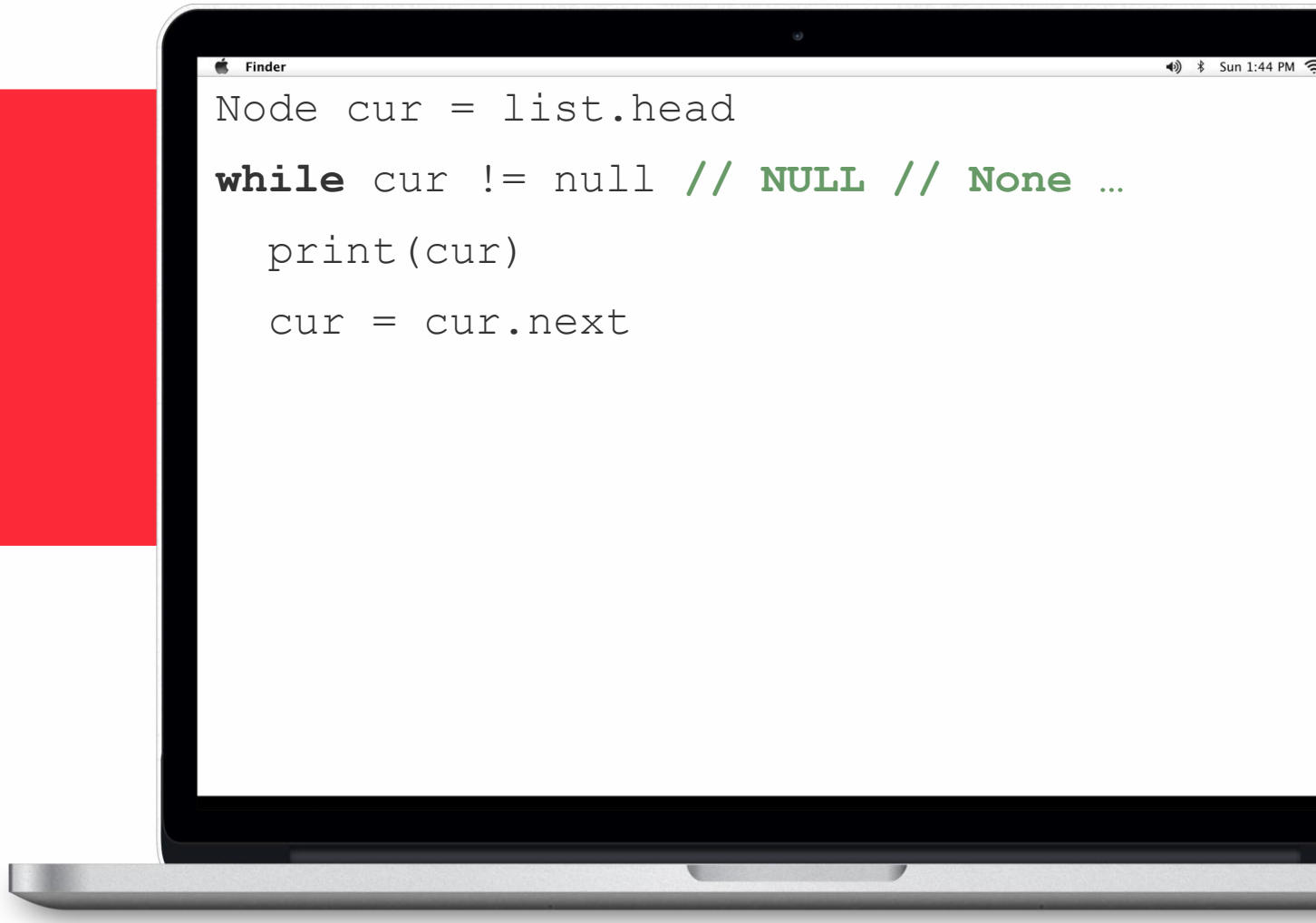
Структура Node

Список задается ссылками на начало и конец



Связный список

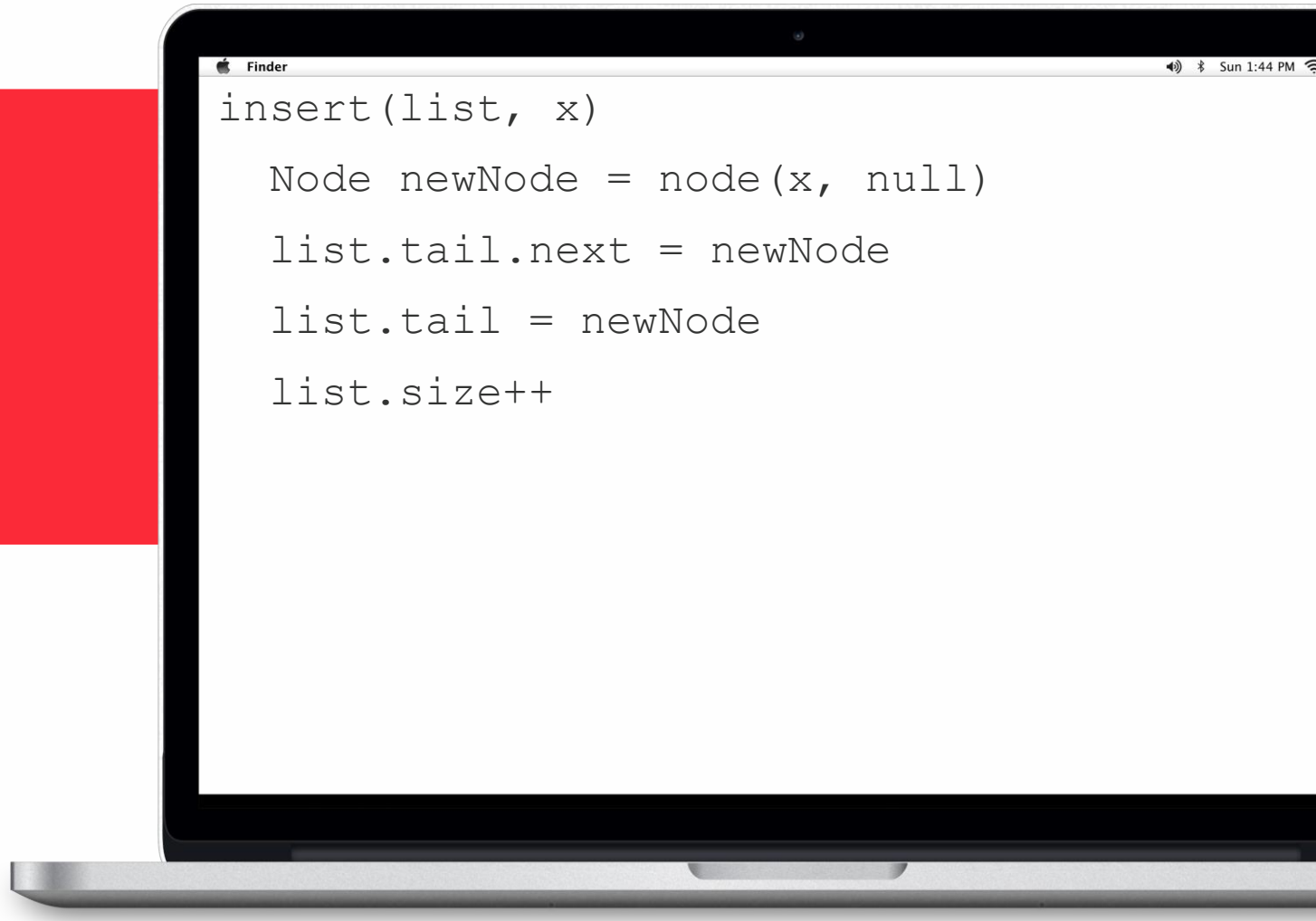
Вывод всех элементов списка



Связный список

Вставка элемента в конец

А если список пустой?

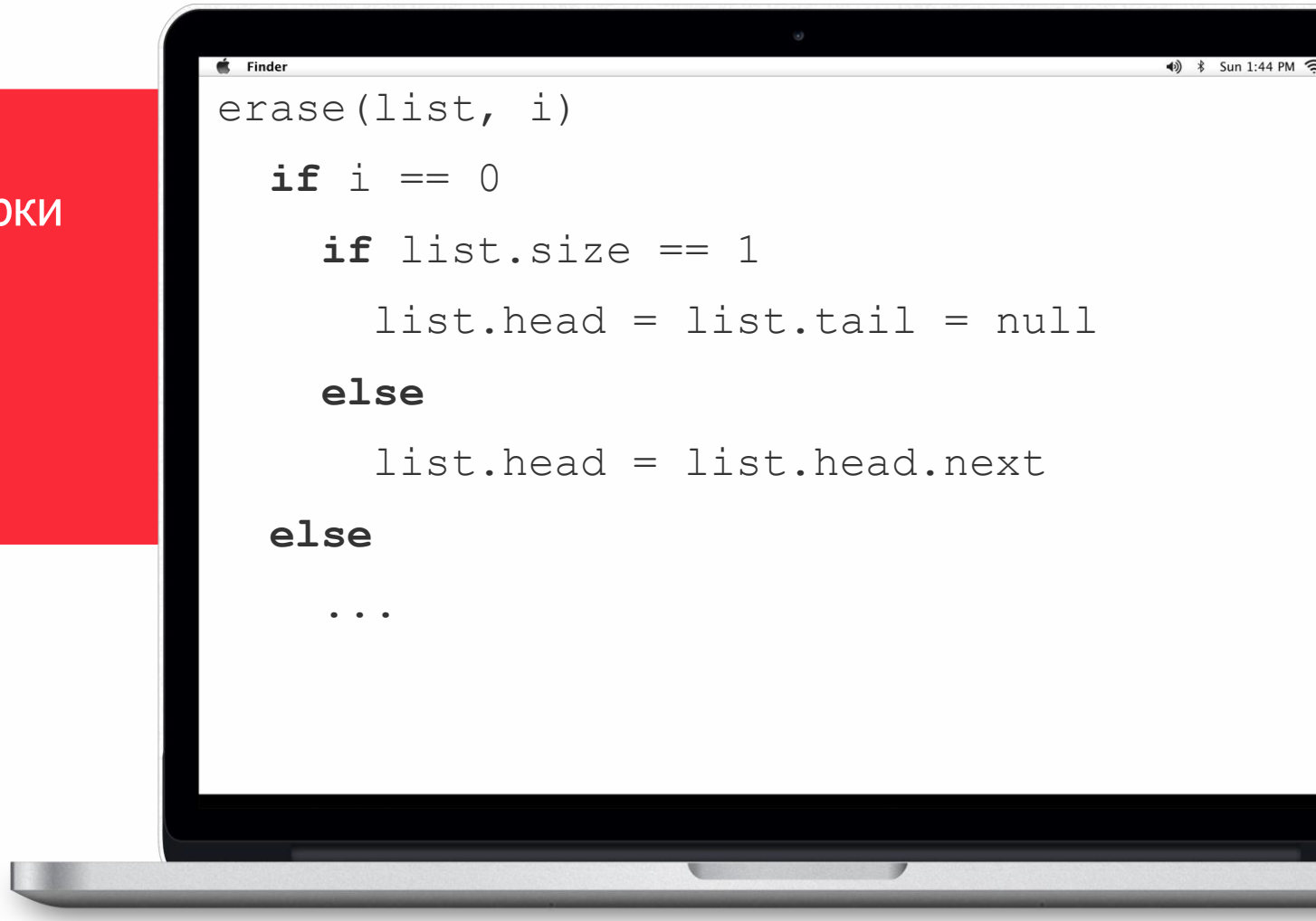


Связный список

Вставка элемента в конец

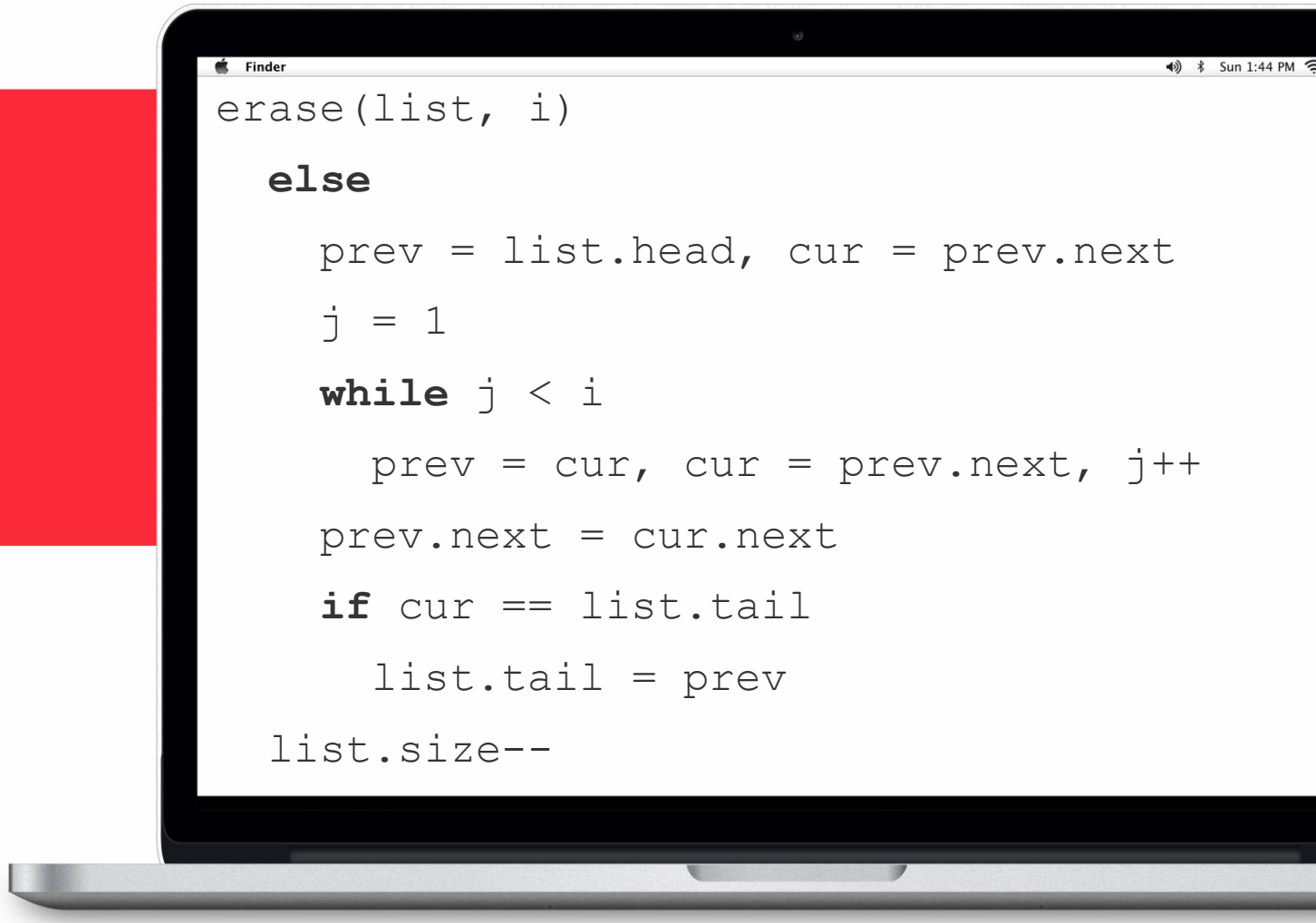
Связный список

Удаление элемента без проверки на дурака



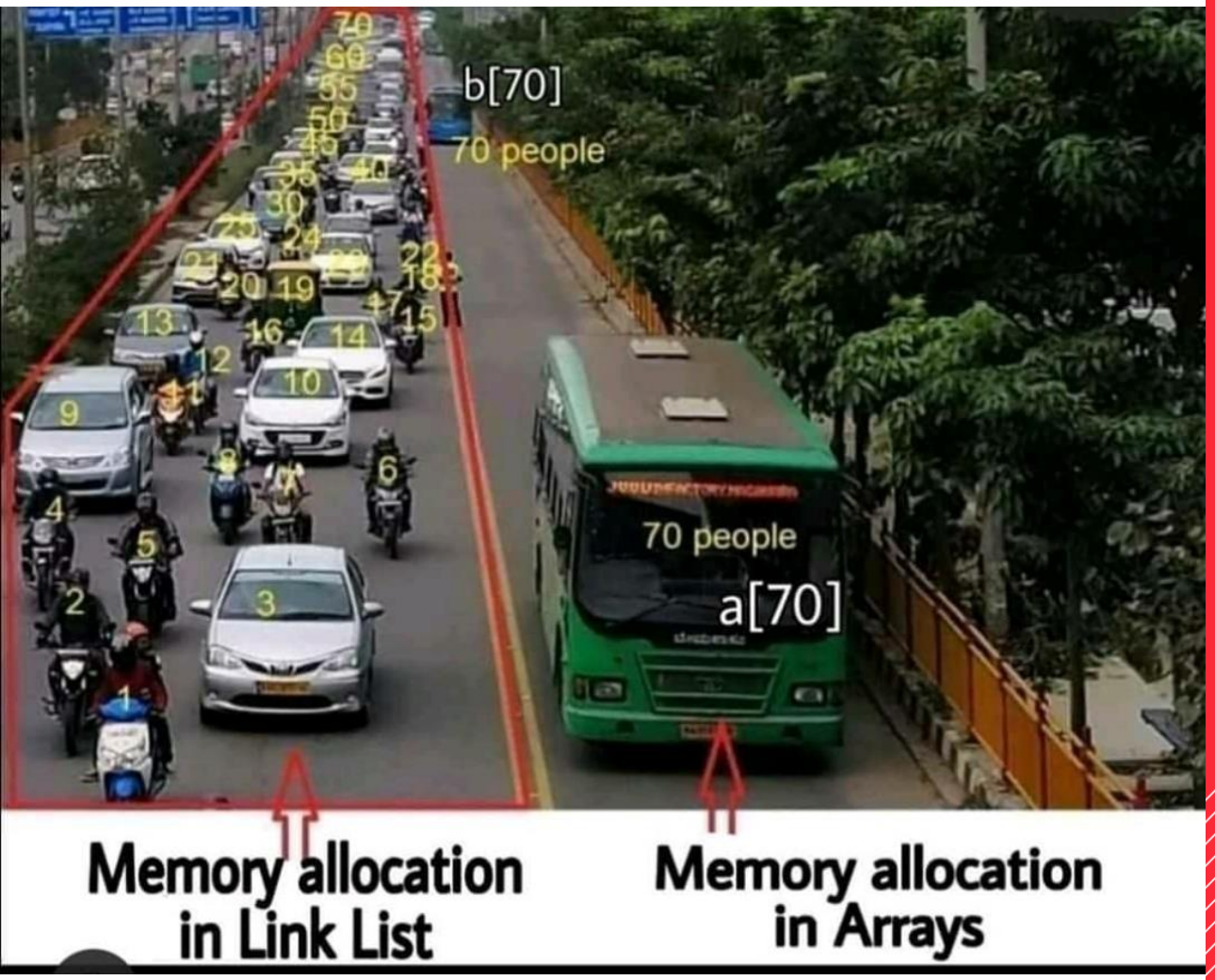
Связный список

Удаление элемента



Связный список

- На что нужно обратить внимание в зависимости от языка?
  - Нужно чистить память!
- Как удалять из середины, зная ссылку?
  - Двусвязный список



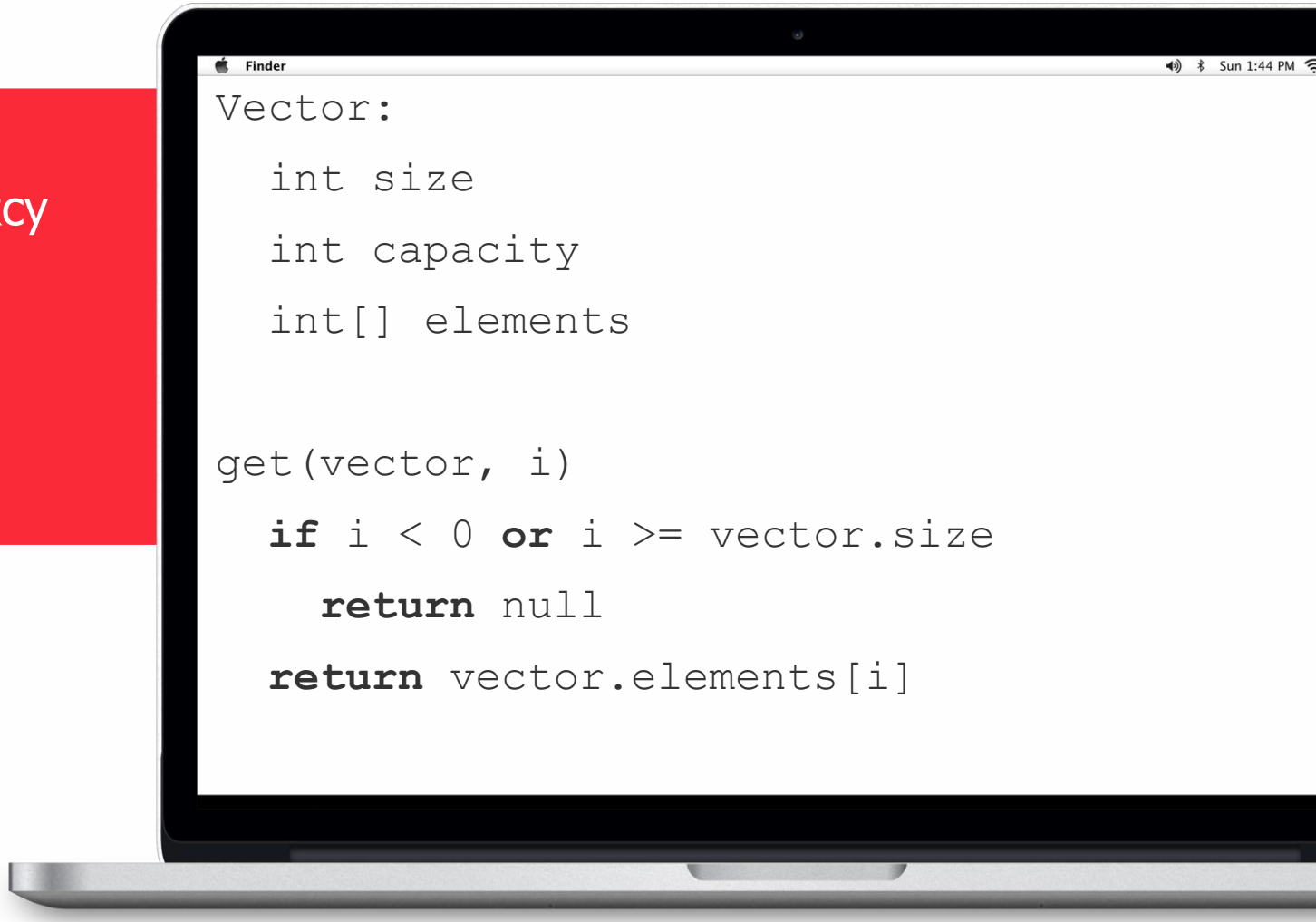
Саморасширяющийся массив (Вектор)

Саморасширяющийся массив

- Вот бы можно было пользоваться массивами, но чтобы памяти было O(n)
- А так можно сделать!
- Пусть добавляем по элементу в конец массива
- Выделим массив, на константное число элементов
- Если нам этого мало будем увеличивать размер в два раза
- Говорят, что будет работать быстро

Саморасширяющийся массив

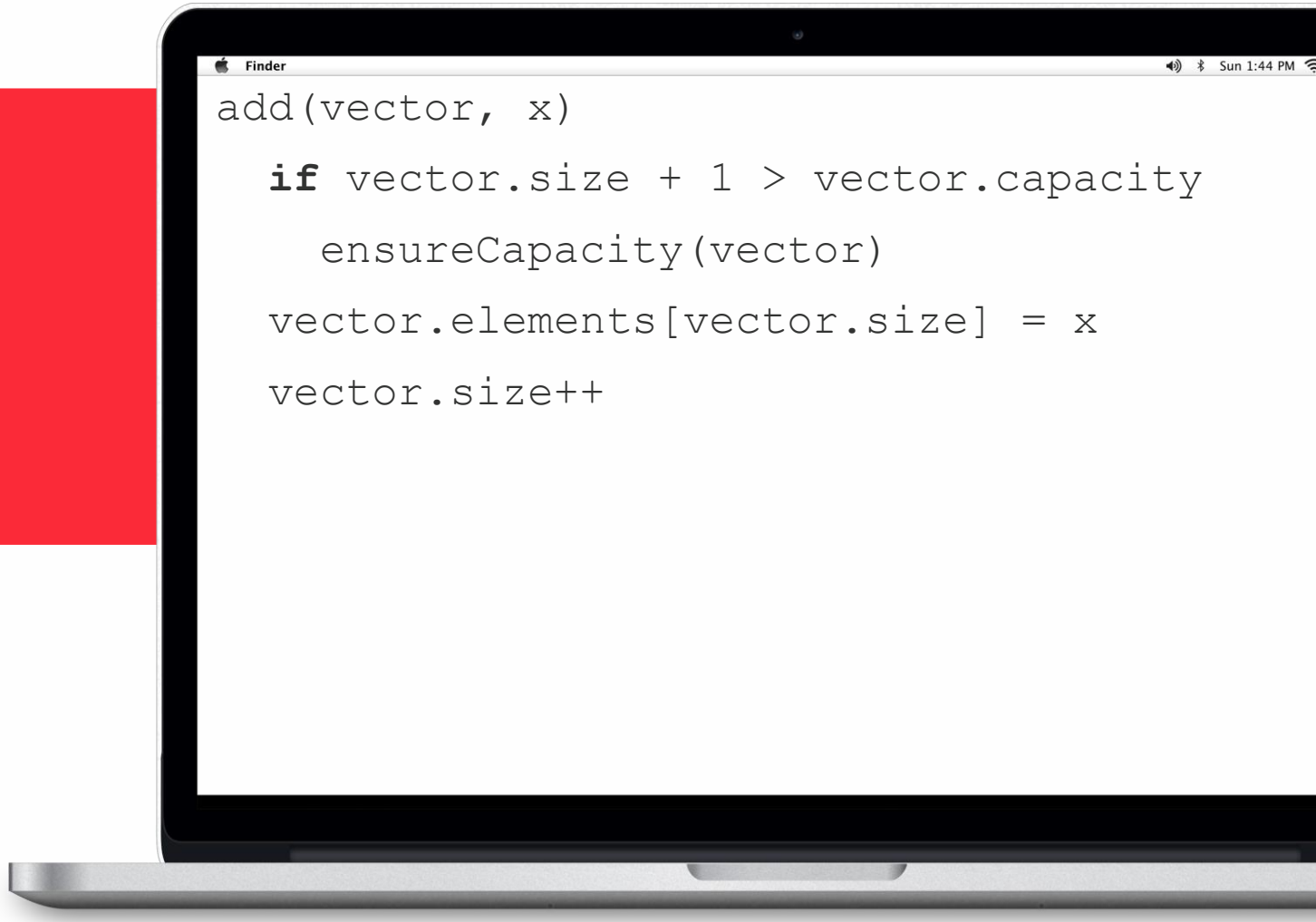
Структура + элемент по индексу



Саморасширяющийся массив

Добавление в конец

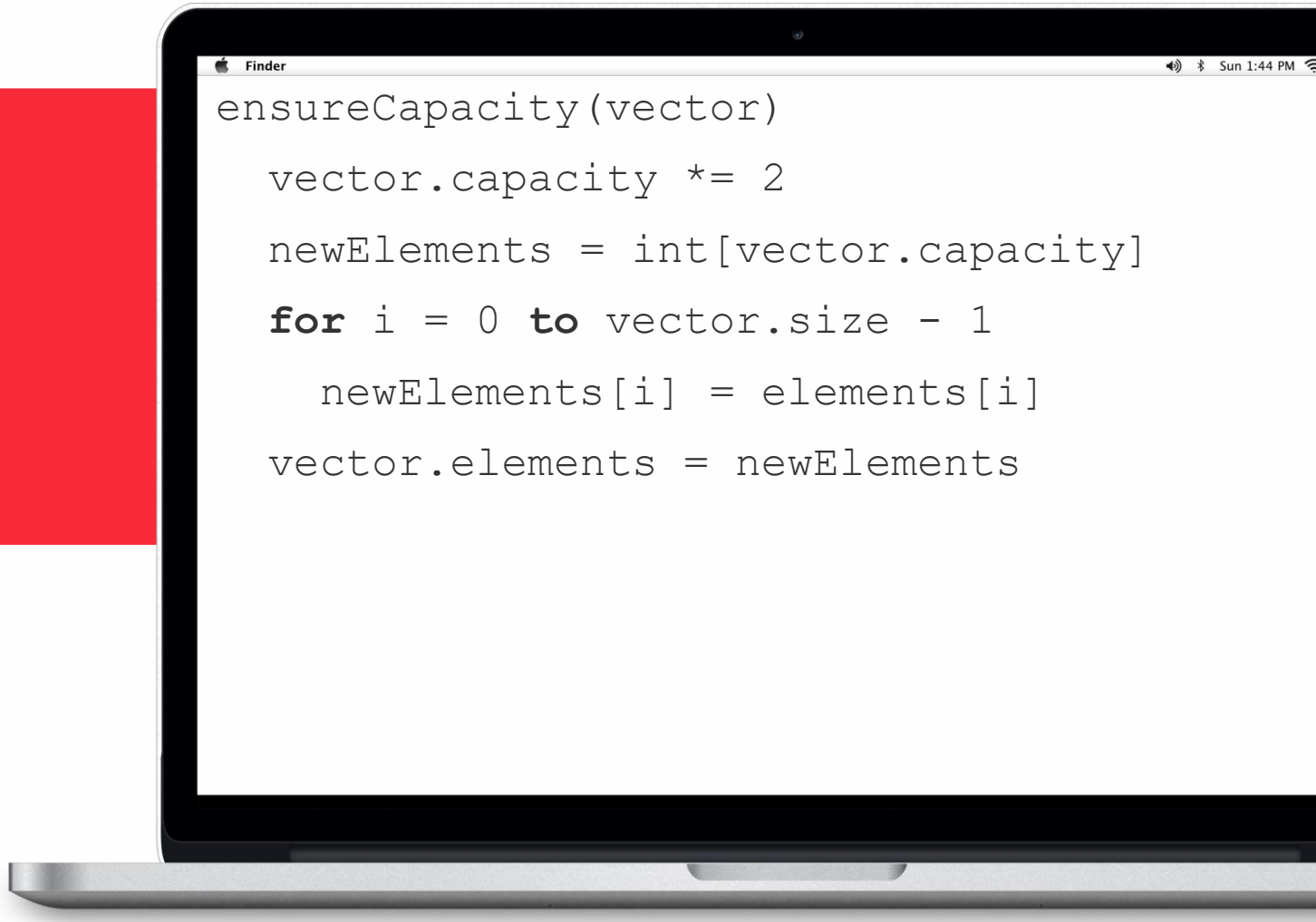
А если массив закончился?



Саморсаширяющийся массив

ensureCapacity

Не забыть почистить память!



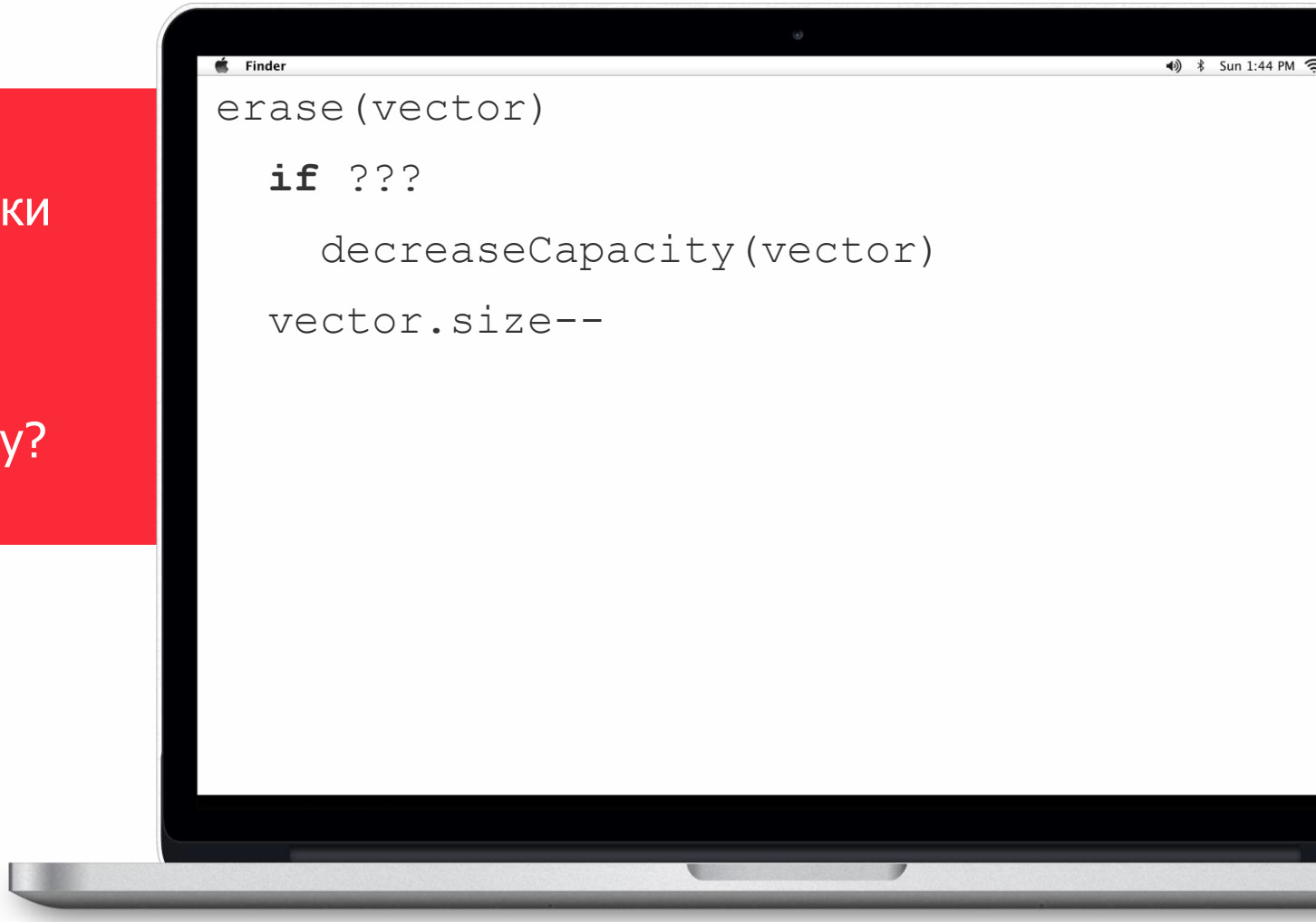
Саморасширяющийся массив

- Почему памяти O(n)?
- За сколько работают операции?
- Стандартные реализации
  - vector (C++)
  - ArrayList (Java)
  - list (Python)

Саморасширяющийся массив

Удаление из конца без проверки на дурака

Какое условие decreaseCapacity?



Стек, Очередь, Дек

Стек

- Представляет собой список элементов, организованных по принципу LIFO (last in — first out)
- Основные методы:
  - push
  - pop
  - size
  - back (top, peek)…
- **Абстракция, существуют различные реализации!**

Стек

- Будем рассматривать те реализации, в которых каждая операция за работает за O(1) и потребляют O(n) памяти
- Стек на саморасширяющемся массиве
- Стек на связном списке

Очередь

- Представляет собой список элементов, организованных по принципу FIFO (first in — first out)
- Основные методы:
  - push
  - pop
  - size
  - front
- **Абстракция, существуют различные реализации!**

Очередь

- Очередь на списке
- Очередь на саморасширяющемся массиве
  - какие проблемы в отличии от стека?
- Очередь на саморасширяющемся циклическом массиве

Дэк

- Представляет собой список элементов, который позволяет добавлять\удалять элементы, как из начала, так и из конца
- Основные методы:
  - push_back, push_front
  - pop_back, pop_front
  - size
  - front, back
- **Абстракция, существуют различные реализации!**

Куча

Приоритетная очередь

- Абстракция, которая хранит множество элементов
- Позволяет добавить элемент множество
- Удаляет элементы в порядке приоритета
  - Например, минимальный элемент множества

Куча

- Способ реализации приоритетной очереди
- Подвешенное двоичное дерево
  - Значение в любой вершине небольше значения в детях (потомках)
  - На i-ом слое 2^i вершин, кроме последнего
  - Последний слой заполнен «слева направо»
- Как хранить такую структуру?
  - Ссылки
  - Массив

Куча

- Куча на массиве
- Корень имеет номер 0
- У вершины с номером i
  - Левый ребенок 2i + 1
  - Правый ребенок 2i + 2
- Почему не будет пропусков?
- Как получить номер родителя, зная номер ребенка?
- Тогда i-й элемент кучи – i-й элемент массива

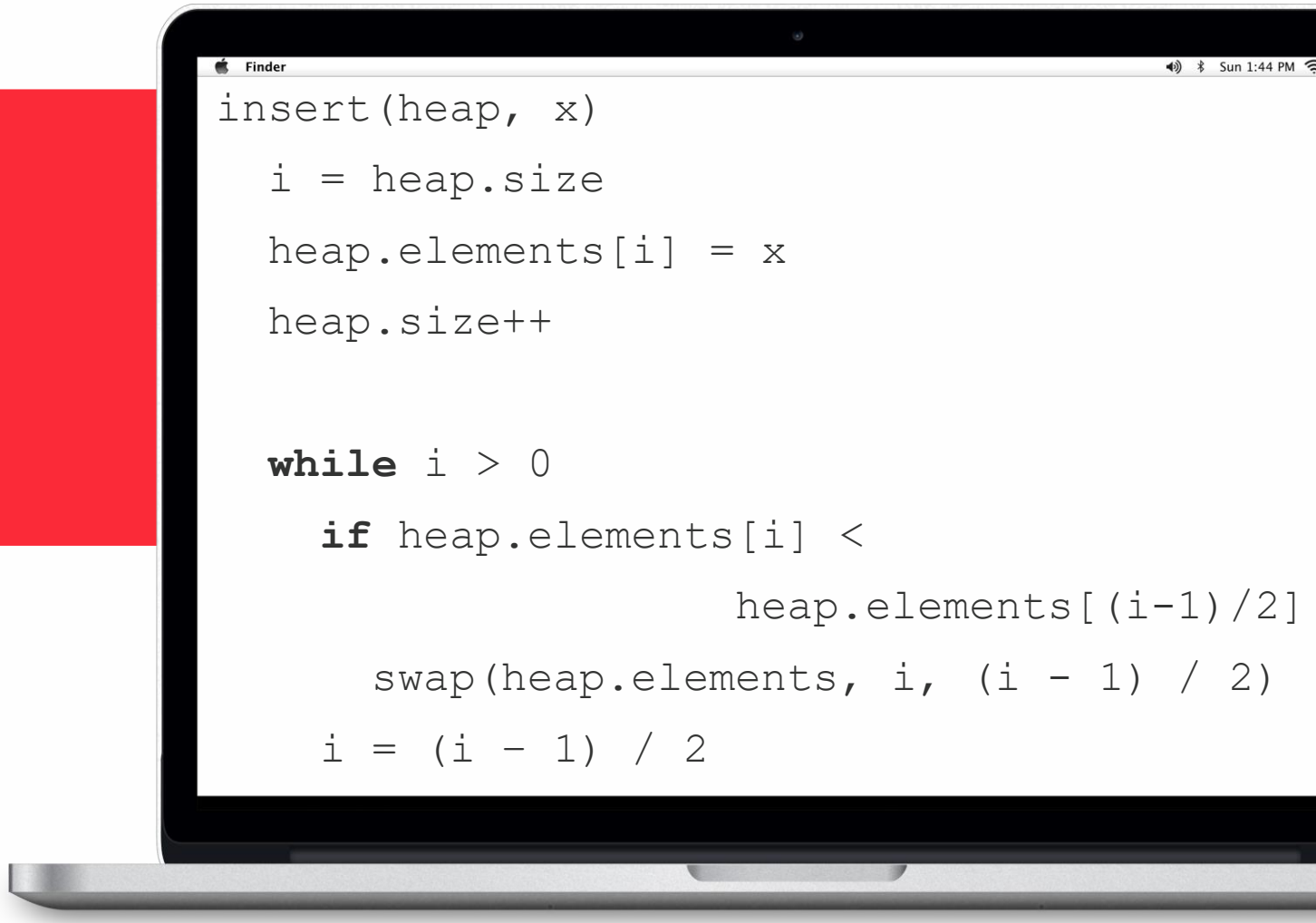
Куча

- Если куча – приоритетная очередь, должны быть операции:
  - insert
  - removeMin
- Как можно их реализовать?
- За сколько они будут работать?

Куча

Вставка элемента

Можно if заменить на and



Куча

Удаление минимального элемента

Что делать если правого сына нет?

Куча

Удаление минимального элемента

Что делать если правого сына нет?

Сортировка кучей

- Вспомним сортировку выбором
- Вставили все элементы в кучу, последовательно выбираем минимум
- O(nlogn)
- Как без доп. памяти?

Построение кучи по массиву

- Последовательный insert всех элементов
  - O(nlogn)
- Можно ли быстрее
  - Идем с конца и просеиваем вниз
  - Будет O(n)
  - Почему?

Все!