

PfTWAFB

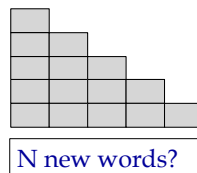
- **Understanding algorithmic trade-offs**
 - When is simple good enough, when is slow good enough
 - How do we have discussions about fast and slow?
- **Understanding structuring data trade-offs**
 - Very much related to algorithmic trade-offs
 - Python supplies built-ins: list, set, dictionary
- **Toward understanding dictionaries**
 - Versatile, useful, cross-language idiom/tool/concept
 - Facilitates efficient solutions to many problems

Comparing Algorithms

- **Searching a list of N elements for a specific value**
 - Worst case is
 - What if we search a list of size $1+2+3+...+N$?
- **Doing binary search (guess a number), sorted list**
 - Worst case is ...
 - What if we search a list of size $1+2+3+...+N$
- **Search using a dictionary: magically find location**
 - Worst case is ...
 - What if we search a list of size $1+2+3+...+N$

See DictionaryTimings.py

```
def linear(words):  
    data = []  
    for w in words:  
        found = False  
        for elt in data:  
            if elt[0] == w:  
                elt[1] += 1  
                found = True  
        if not found:  
            data.append([w, 1])  
    return data
```



Analysis: Algorithms and Data Structures

- **We need a vocabulary to discuss performance**
 - Reason about alternative algorithms and implementations
 - It's faster! It's more elegant! It's safer! It's less filling!
- **Need empirical tests, analytical/mathematical tools**
 - Given two methods, which is better? Run them to check.
 - 30 seconds vs. 3 seconds, easy! 5 hours vs. 2 minutes, harder
 - What if it takes two weeks to implement the methods?
 - Use mathematics to analyze the *algorithm*,
 - The implementation is another matter, cache, compiler optimizations, OS, memory,...

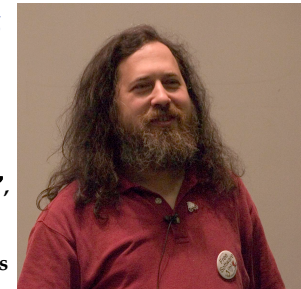
Running times @ 10^9 instructions/sec

N	$O(\log N)$	$O(N)$	$O(N \log N)$	$O(N^2)$
10^2	0.0	0.0	0.0	0.00001
10^3	0.0	0.0000001	0.00001	0.001
10^6	0.0	0.001	0.02	16.7 min
10^9	0.0	1.0	29.9	31.7 years
10^{12}	9.9 secs	16.7 min	11.07 hr	31.7 million years

This is a real focus in Compsci 201

Richard Stallman (b.1953, Hopper '90)

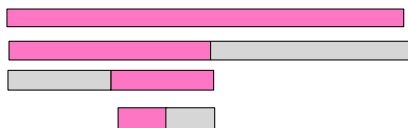
- Transformed programming
 - > Free Software Foundation
- "World's Best Programmer"
 - > Gnu/Linux: g++, emacs
 - > Believes all software should be free, but like "free speech", not "free beer"
 - > Won MacArthur award for his efforts and contributions



If you want to accomplish something in the world, idealism is not enough-- you need to choose a method that works to achieve the goal. In other words, you need to be pragmatic.

Binary Search

- Before the first guess, there are 1024 numbers



How many times can we divide list in half?
 $\log_2(N)$ for N element list, why?
 What must be true to use binary search?
 How is this done in Python?

See DictionaryTimings.py

```
def binary(words):
    data = []
    for w in words:
        elt = [w, 1]
        index = bisect.bisect_left(data, elt)
        if index == len(data):
            data.append(elt)
        elif data[index][0] != w:
            data.insert(index, elt)
        else:
            data[index][1] += 1
    return data
```