

# CS 100: Search, Google, Algorithms, Heuristics

Chris Kauffman

Week 8-1

# Logistics

## Feedback

- ▶ Advisory Grades to be posted by Thursday
- ▶ Feedback from you today

## Today

- ▶ Searching for things
- ▶ Google

## HW 4

Due Thursday

## HW 5

Will post later this week

## Reading

"Pattern": Chapters 5-6

# I couldn't Remember

- ▶ Wanted a particular album from my stack of CDs
- ▶ Shelf is roughly alphabetical by artist
- ▶ Could not remember the name of the artist
- ▶ Did remember the album was something like "Now He Sings, Now He Sleeps"
- ▶ Genre: jazz, piano trio
- ▶ How do I find the CD?
- ▶ Discuss with a neighbor



Source

# Searches

## Linear Search

- ▶ Start at the beginning
- ▶ Examine, each item sequentially against the search query
- ▶ If the query matches the item, report where it is located and quit
- ▶ When all items have been scanned report that nothing matched

## Question

- ▶ Are these **algorithms** or **heuristics**?
- ▶ What's an algorithm? What's a heuristic?
- ▶ What's the difference?

## Alternative: "Indexed" Search

- ▶ Know a better starting point
  - ▶ The artist's name starts with 'C'
  - ▶ CDs organized by name
- ▶ Search items sequentially at that starting point
- ▶ When it's obvious item is not present bail out
  - ▶ At 'D' artists, CD must be in the car

# Chick



- ▶ Google query "now he sings now he sleeps jazz piano"
- ▶ Results: *Now He Sings, Now He Sobs* by Chick Corea
- ▶ Chick = 20 Grammy wins, 59 nominations
- ▶ He's kind of a big deal

# Music and Programming

- ▶ Have a lot in common
- ▶ Both written down in a coded form
- ▶ A language that is opaque to the uninitiated
- ▶ Both involve conditionals, repetition, jumps
- ▶ Static and dynamic behavior
- ▶ The code is not the action
- ▶ Surprising action comes from simple code
- ▶ Encoding dynamic as static is difficult

216

*\* NOW HE SINGS;  
NOW HE SOBS*

CHICK COREA

In One

LAST X ONLY

\* From Chick's Lead Sheet

## In-Class: How Does Google Search Work?

- ▶ How did Google know about Chick?
- ▶ How does Google search work?
- ▶ Use google search to help you

Define each of these as related to Google Search

1. Crawling
2. Indexes / Indices
3. Page Rank / Popularity
4. Freshness
5. Credibility / Quality
6. Personalization
7. SPAM removal...

Work in Groups of 2-4

Hand in one sheet explaining at least 1-3 with names, NetIDs of all group members

## Gaming the System

*During the 2004 U.S. presidential election, Google bombs were used to further various political agendas. Two of the first were the "miserable failure" Google bomb linked to George W. Bush's White House biography and the "waffles" Google bomb linked to John Kerry's website. Supporters and detractors of the candidates were able to manipulate search engine rankings so that searches for "miserable failure" and "waffles" returned links to the targets' sites.*

*Political Google bombs in the 2004 U.S. Presidential election*

Now that you understand some basics of how google evaluates "popularity" of a page, how would you bomb your personal nemesis?

("irritating professor" → <https://cs.gmu.edu/~kauffman/>)



# Important Lessons

## Search Engine Optimization

- ▶ Gaming search engines is a big business
- ▶ SEO: Search Engine Optimization
- ▶ Try to make sure that your page shows up high in rankings
- ▶ Tricks to use for this that you've learned?

## Algorithms Encode Human Values

- ▶ Algorithms will affect all you for your entire lives
- ▶ Once created, algorithms act completely objectively, no bias
- ▶ Humans always specify algorithms in the first place and frequently encode values and biases into them

# Algorithm or Heuristic?

From *The Pattern on the Stone*

**Algorithm** A set of well-defined steps to go from input to answer

**Heuristic** A set of steps that may produce a correct answer for some inputs but may also fail

Is Google search an algorithm or a heuristic?