COMPUTER SCIENCE 51 | DECEMBER 3, 2008

COMPUTER SCIENCE 51
Spring 2009

Prof. Greg Morrisett
Prof. Ramin Zabih

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

CS51 Perspectives

- ***From the right perspective, many problems that look hard become easy***
- CS51 will let you do amazing things
  - Replace 10 pages of horrible code with 1 line of good code
    - Code that can be understood, and easily modified when requirements evolve
  - Solve problems in seconds where the obvious solution would take millennia
    - Recognize some easy-looking problems that are actually hard

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Programs and Efficiency

- The right perspective can vastly simplify programming tasks
  - Goal: simple, readable, extensible code
  - Can even prove correctness!
- The right perspective can vastly improve efficiency
  - Goal: fast code where obvious way is slow
  - Or, identify a nearby problem that can be solved fast (as an alternative)

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Perspectives on programming

- There are thousands of general purpose programming languages:
  - Java, C, Scheme, Python, Mercury, Erlang,...
  - SQL, regexp, XML, CUDA, Google, Elisp, ...
  - Key: understand critical concepts so you can pick up and apply the right tool for the job.
- The right language helps formulate solutions
  - Try writing multiplication for Roman Numerals
  - Try doing optimistic concurrency for C++ code

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Proofs and code

- One way to get the right code is to construct a proof first and extract code from the proof
  - Proofs and functional code coincide deeply
    - e.g., proof by induction ==> code by recursion
    - e.g., proof by cases ==> pattern matching
    - e.g., proof by lemma ==> function application
- Works both ways:
  - Got functional model of code?
    - Extract proof sketch
  - Got (constructive) proof of algorithm?
    - Extract code sketch

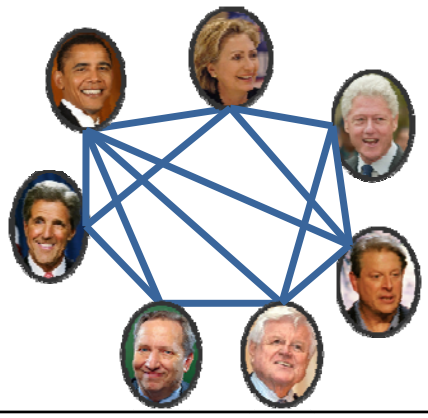
COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Perspectives on efficiency

- There is usually an obvious way to solve a computational problem
  - Example: CS50 spell-checker dictionary
- What if this is way too slow?
  - Sometimes we can do vastly better
  - Or we can prove we can't
    - Perhaps we can modify the problem?
- CS51 will cover all these possibilities
- Let's preview a few sample problems

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Some Facebook friends



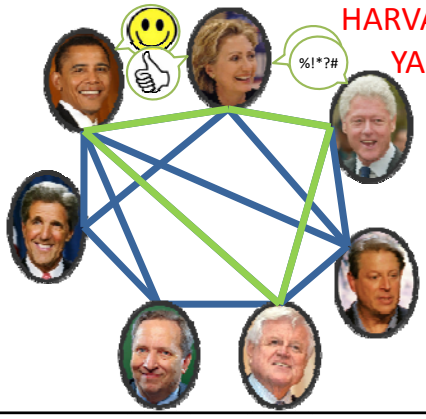
COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Barack shares good news (1)

- Conversation Problem: Each person has a conversation with each friend

  Can we find a way for the good news to spread?

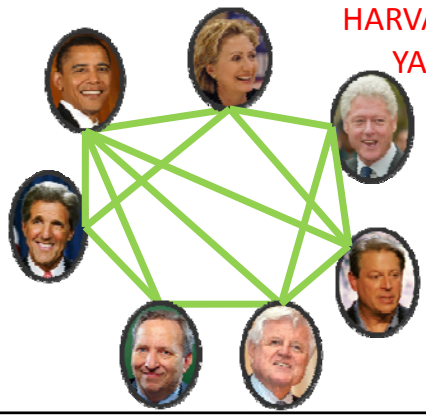
Conversation Problem



HARVARD 10,
YALE 0

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Can the good news spread such that each pair of friends has one conversation?



HARVARD 10,
YALE 0

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

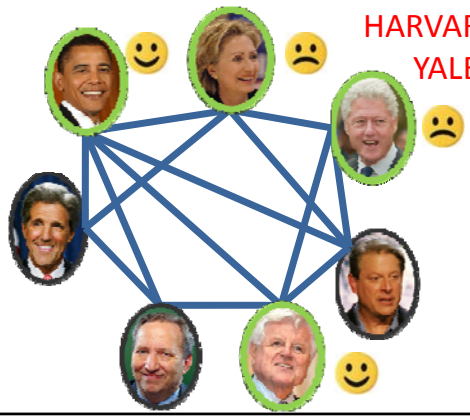
Barack shares good news (2)

- Notification Problem: Each person hears the news once through a friend

Can we find a way for the good news to spread?

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Notification Problem



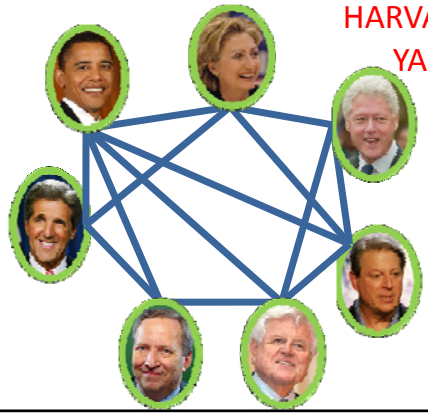
HARVARD 10,
YALE 0

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Can the good news spread so that each person learns it once?

HARVARD 10, YALE 0



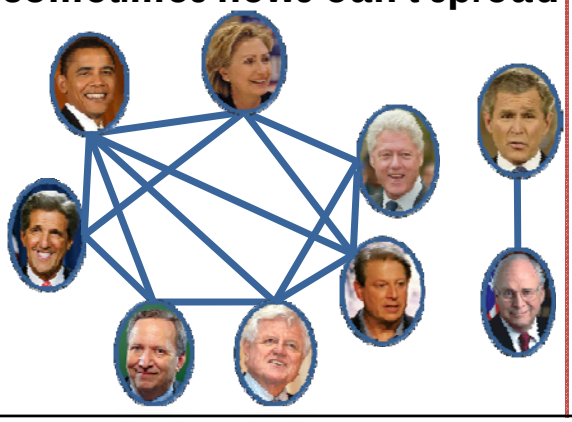
COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Start spreading the news

- There is an easy way to solve both these problems
  - Note: I said "easy", not "smart"
- Exhaustive search: try every possibility
- This could take a ***really*** long time
  - Especially with lots of popular people!
- Greedy algorithm: pick a neighbor
  - Notify them, or have a conversation
  - Might be clever about which neighbor

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Sometimes news can't spread



COMPUTER SCIENCE 51 | DECEMBER 3, 2008

These problems look similar

- Spreading the news on Facebook
  - Between friends, avoiding repetition
- One of these problems can be solved pretty easily in O(n) time
  - Greedy method, with cleverness
- For the other problem, no one can do much better than exhaustive search
  - If you can do better, you will be the most famous computer scientist in the world
  - You will also win some fabulous prizes

Fabulous prizes

+

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

The world's toughest job?

[J. Allard, Corporate Vice President, Microsoft]

- J.'s job is to make the Zune popular
- What, you may ask, is a Zune?
  - It's kind of like an iPod
  - But with 2 important features!
  - Feature 1: can't use iTunes

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Feature 2: made by Microsoft

Help J. make Zunes popular

- J. has a limited advertising budget, but he can scrape up 25 free Zunes
  - He wants to give them to Facebook users
    - Hopefully they like it & tell their friends
- Only the first time someone hears about the Zune matters
  - Having another friend with a free Zune doesn't make you any more interested
    - Somewhat like immunization

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

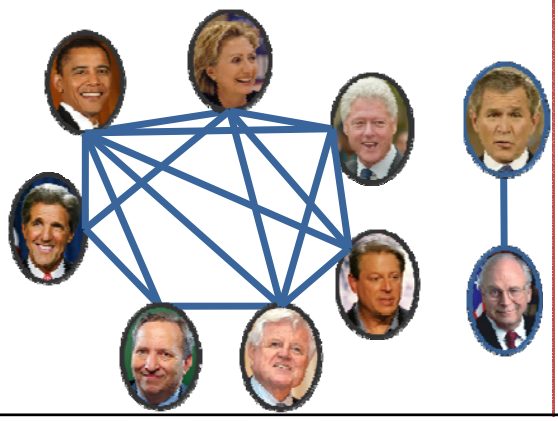
COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Two options for J.

- Exhaustive search:
  - Try every group of 25 people on Facebook
  - Guaranteed to find the best group!
    - I.e., reach the most people
  - Hope the universe doesn't end first…
- Greedy algorithm:
  - Give 1st to Barack (has the most friends)
  - Give 2nd to Hillary (next most friends)
- Does this work? Can we do better?

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Greedy Algorithm



COMPUTER SCIENCE 51 | DECEMBER 3, 2008

What's wrong with this idea?

- Barack & Hillary's friends overlap a lot
  - In this little example, she adds nothing
    - Please don't tell her we said this…
- We want popular people, but also ones with distinct friends

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Exhaustive solution

- Suppose Facebook had just 10M users
  - How many of groups 25 would there be?
- Answer: over 10^{149}
  - Age of universe is less than 10^{18} seconds
    - If you could try 10^{100} groups in a second, you wouldn't be nearly done
  - There are about 10^{80} elementary particles
- Can you beat exhaustive search?
  - If so, you will win some fabulous prizes!

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

A clever algorithm?

- Give Zune to Barack
- Forget about him and all his friends
- Give it to whoever has the most friends
- Repeat until Zune budget exhausted

- Suppose that exhaustive search will (slowly!) find a way to reach 1M people
  - This simple algorithm is **guaranteed** to quickly reach over 632,120 people

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Netflix/Walmart problem

Basket #1:

Basket #2:

Basket #3:

Basket #4:

- What can we conclude?
  - Beer is very popular
  - Beer & diapers popular, but slightly less

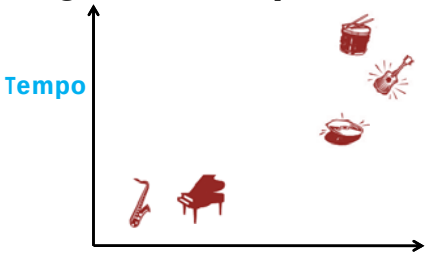
COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Legal gibberish problem

- Which building is the source?

COMPUTER SCIENCE 51 | DECEMBER 3, 2008

Algorithms to partition data



- This problem has many formulations
  - With right perspective, can solve it fast