Systems & Applications: Introduction

Ling 573
NLP Systems and Applications
March 31, 2015

Roadmap

- Motivation
- 573 Structure
- Summarization
- Shared Tasks

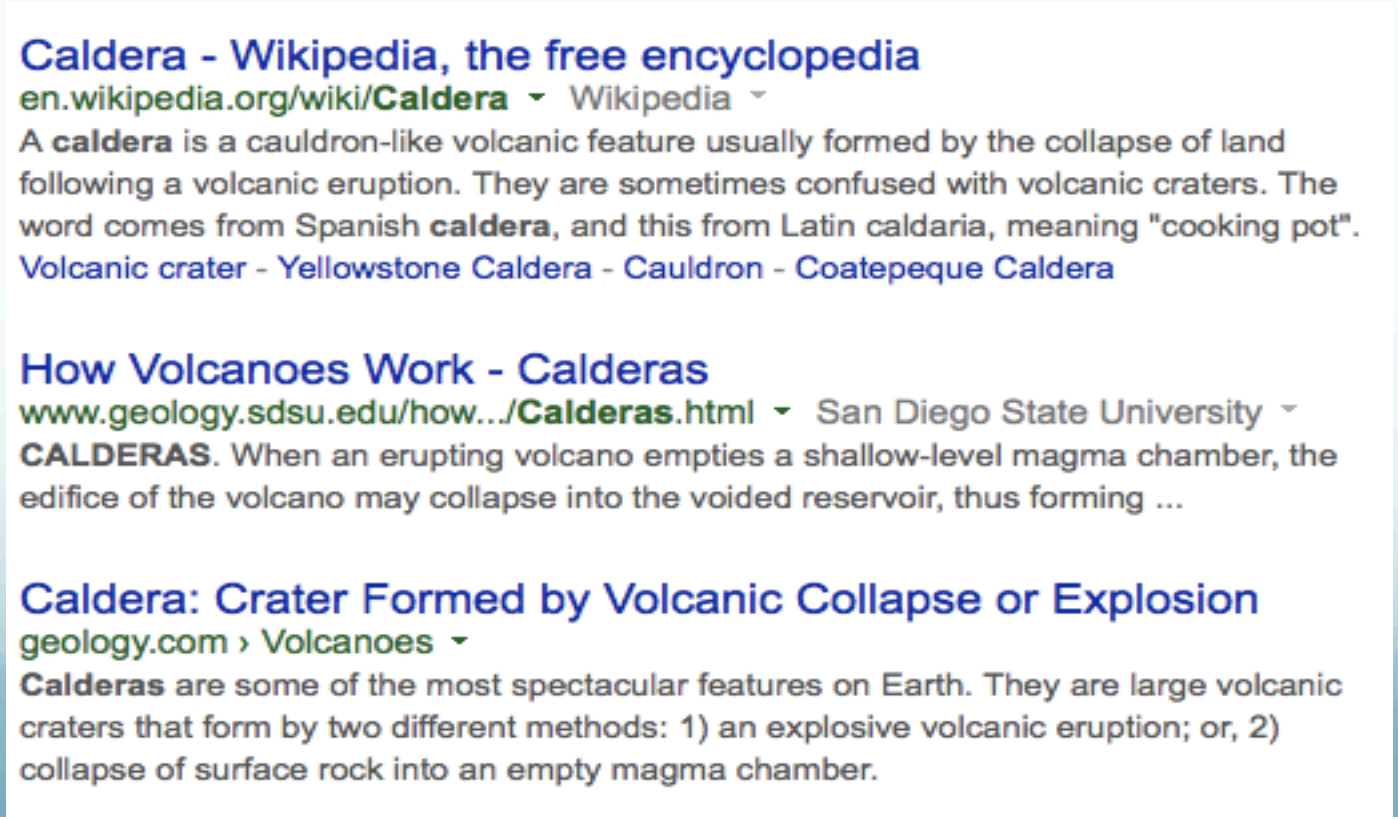
Motivation

- Information retrieval is very powerful
  - Search engines index and search enormous doc sets
  - Retrieve billions of documents in tenths of seconds

- But still limited!
  - Technically – keyword search (mostly)
  - Conceptually
    - User seeks information
      - Sometimes a web site or document
      - Sometimes the answer to a question
      - But, often a summary of document or document set

Why Summarization?

- Even web search relies on simple summarization
  - Snippets!
    - Provide thumbnail summary of ranked document



Caldera - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/**Caldera** ▾ Wikipedia ▾
A **caldera** is a cauldron-like volcanic feature usually formed by the collapse of land following a volcanic eruption. They are sometimes confused with volcanic craters. The word comes from Spanish **caldera**, and this from Latin caldaria, meaning "cooking pot".
Volcanic crater - Yellowstone Caldera - Cauldron - Coatepeque Caldera

How Volcanoes Work - Calderas
www.geology.sdsu.edu/how.../**Calderas**.html ▾ San Diego State University ▾
CALDERAS. When an erupting volcano empties a shallow-level magma chamber, the edifice of the volcano may collapse into the voided reservoir, thus forming ...

Caldera: Crater Formed by Volcanic Collapse or Explosion
geology.com › Volcanoes ▾
Calderas are some of the most spectacular features on Earth. They are large volcanic craters that form by two different methods: 1) an explosive volcanic eruption; or, 2) collapse of surface rock into an empty magma chamber.

Why Summarization?

- Complex questions go beyond factoids, infoboxes
  - Require explanations, analysis
    - E.g. In children with an acute febrile illness, what is the efficacy of single-medication therapy with acetaminophen or ibuprofen in reducing fever?
    - Highest search hit is manually created summary site
      - Umich medical
        - Vs 5 articles cited in creating

Ibuprofen is More Likely to Normalize Temperature than Acetaminophen, Though Both are Safe and Effective Antipyretics for Short-Term Use in Children

Question

- In children with an acute febrile illness, what is the efficacy of single-medication therapy with acetaminophen or ibuprofen compared with combination therapy combining the two medications in reducing fever while avoiding adverse effects?

Clinical Bottom Lines

1. Both acetaminophen and ibuprofen are effective antipyretics and are well-tolerated in short-term use in febrile children.
2. Ibuprofen is more effective at achieving temperature normalization than acetaminophen, though both effectively lower temperatures >1.5 C in most patients with standard dosing.
3. There is no data currently available comparing the efficacy and tolerability alternating regimens with ibuprofen and acetaminophen to single-drug regimens.[1]

Summary of Key Evidence

1. 628 children aged 6 months to 6 years with initial temperature >38.5C were randomized to receive ibuprofen, acetaminophen, or dipyrone (banned in the US) in a 1:1:1 ratio. The study was double-blinded and multinational. There was no placebo arm.

Why Summarization?

- Complex questions go beyond factoids, infoboxes
  - Require explanations, analysis
    - E.g. In children with an acute febrile illness, what is the efficacy of single-medication therapy with acetaminophen or ibuprofen in reducing fever?
    - Summ: Ibuprofen provided greater temperature decrement and longer duration of antipyresis than acetaminophen when the two drugs were administered inapproximately equal doses. (PubMedID: 1621668)

Why Summarization?

- Huge scale, explosive growth in online content
  - 2-4K articles in PubMed daily, 41.7M articles/mo on WordPress alone (2014)
  - How can we manage it?
    - Lots of aggregation sites
      - Effective summarization rarer
- Recordings of meetings, classes, MOOCs
  - Slow to access linearly, awkward to jump around
  - Structured summary can be useful
    - Outline of: how-tos, to-dos,

Perspectives on Summarization

- DUC, TAC (2001-...):
  - Single-, multi-document summarization
    - Readable concise summaries
    - Largely news-oriented
      - Later blogs, etc; also query-focused
- Text simplification:
  - Compress, simplify text for enhanced readability
    - Application to CALL, reading levels (e.g. Simple Wikipedia), assistive technology
      - Also aims to support greater automation

Natural Language Processing and Summarization

- Rich testbed for NLP techniques:
  - Information retrieval
  - Named Entity Recognition
  - Word, sentence segmentation
  - Information extraction
  - Parsing
  - Semantics, etc..
  - Discourse relations
  - Co-reference
  - Generation
  - Paraphrasing
- Deep/shallow techniques; machine learning

573 Structure

- Implementation:
  - Create a summarization system
    - Extend existing software components
    - Develop, evaluate on standard data set
- Presentation:
  - Write a technical report
  - Present plan, system, results in class
  - Give/receive feedback

Implementation: Deliverables

- Complex system:
  - Break into (relatively) manageable components
  - Incremental progress, deadlines
- Key components:
  - D1: Setup
  - D2: Baseline system, Content selection
  - D3: Content selection, Information ordering
  - D4: : Content selection, Information ordering, Surface realization, final results
- Deadlines:
  - Little slack in schedule; please keep to time
  - Timing: ~12 hours week; sometimes higher

Presentation

- Technical report:
  - Follow organization for scientific paper
    - Formatting and Content

- Presentations:
  - 10-15 minute oral presentation for deliverables
  - Explain goals, methodology, success, issues
  - Critique each others' work
  - Attend **ALL** presentations

Working in Teams

- Why teams?
  - Too much work for a single person
  - Representative of professional environment
- Team organization:
  - Form groups of 3 (possibly 2) people
  - Arrange coordination
  - Distribute work equitably
    - All team members receive the same grade
      - End-of-course evaluation

First Task

- Form teams:
  - Email Glenn gslayden@uw.edu with the team list

Resources

- Readings:
  - Current research papers in summarization
  - Jurafsky & Martin/Manning & Schutze text
    - Background, reference, refresher

- Software:
  - Build on existing system components, toolkits
    - NLP, machine learning, etc
    - Corpora, etc

Resources: Patas

- System should run on patas
  - Existing infrastructure
    - Software systems
    - Corpora
    - Repositories

Shared Task Evaluations

- Goals:
  - Lofty:
    - Focus research community on key challenges
      - ‘Grand challenges’
    - Support the creation of large-scale community resources
      - Corpora: News, Recordings, Video
      - Annotation: Expert questions, labeled answers,..
    - Develop methodologies to evaluate state-of-the-art
      - Retrieval, Machine Translation, etc
    - Facilitate technology/knowledge transfer b/t industry/acad.

Shared Task Evaluation

- Goals:
  - Pragmatic:
    - Head-to-head comparison of systems/techniques
      - Same data, same task, same conditions, same timing
    - Centralizes funding, effort
    - Requires disclosure of techniques in exchange for data
  - Base:
    - Bragging rights
    - Government research funding decisions