NLP
Introduction to NLP

*Evaluation of QA*
The TREC Q&A evaluation

• Run by NIST
  – Voorhees and Tice 1999, 2000
  – 2GB of input
  – AQUAINT corpus
  – 200 questions in 1999
  – 693 questions in 2000

• Essentially fact extraction
  – Who was Lincoln’s secretary of state?
  – What does the Peugeot company manufacture?

• Assumptions
  – Questions are based on text
  – Answers are assumed to be present
  – Systems return five passages of 50 or 250 bytes
  – After 2002, only a single passage of 50 bytes + confidence score + NIL
  – No inference needed
TREC 1999

- What date in 1989 did East Germany open the Berlin Wall?
  - Nov 9

- Who was Johnny Mathis' high school track coach?
  - Lou Vasquez

- What is the shape of a porpoises' tooth?
  - spade-shaped

- What is the number of buffaloes thought to have been living in North America when Columbus landed in 1492?
  - 60 million

- The Faroes are a part of what northern European country?
  - Denmark

- The symptoms of Parkinson's disease are linked to the demise of cells in what area of the brain?
  - substantia nigra

- What was the date of the Chelsea flower show in 1992?
  - May 18

- What hotel was used for a setting of the Agatha Christie novel, "And Then There Were None"?
  - Burgh Island Hotel

- What year was the Magna Carta signed?
  - 1215

- Who was Lincoln's Secretary of State?
  - William Seward

- How long does it take to fly from Paris to New York in a Concorde?
  - 3 1/2 hours
• Who is the author of the book, "The Iron Lady: A Biography of Margaret Thatcher"?
• What was the monetary value of the Nobel Peace Prize in 1989?
• What does the Peugeot company manufacture?
• How much did Mercury spend on advertising in 1993?
• What is the name of the managing director of Apricot Computer?
• Why did David Koresh ask the FBI for a word processor?
• What debts did Qintex group leave?
• What is the name of the rare neurological disease with symptoms such as: involuntary movements (tics), swearing, and incoherent vocalizations (grunts, shouts, etc.)?
• How far is Yaroslavl from Moscow?
• Name the designer of the shoe that spawned millions of plastic imitations, known as "jellies".
• Who was President Cleveland's wife?
• How much did Manchester United spend on players in 1993?
• How much could you rent a Volkswagen bug for in 1966?
• What country is the biggest producer of tungsten?
• When was London's Docklands Light Railway constructed?
• What two US biochemists won the Nobel Prize in medicine in 1992?
• How long did the Charles Manson murder trial last?
• Who was the first Taiwanese President?
• Who was the leader of the Branch Davidian Cult confronted by the FBI in Waco, Texas in 1993?
• Where is Inoco based?
• Who was the first American in space?
• When did the Jurassic Period end?
• When did Spain and Korea start ambassadorial relations?
• When did Nixon visit China?
• Who was the lead actress in the movie "Sleepless in Seattle"?
• What is the name of the "female" counterpart to El Nino, which results in cooling temperatures and very dry weather?
• What is the brightest star visible from Earth?
• What are the Valdez Principles?
Evaluation

• MRR
  – Mean Reciprocal Rank
  – Introduced by TREC in 1999

\[
MRR = \frac{1}{n} \left( \sum_{i=1}^{n} \frac{1}{\text{rank}_i} \right)
\]

• Example
  – What is the capital of Canada?
  – Correct answer ranks = 2, 5 (highest rank = 2)
  – MRR = 1/2 = 0.5
  – TRR = 1/2 + 1/5 = 0.7 = total reciprocal rank

• Later years
  – Confidence-weighted score
System Rankings

- **1999**
  - Cymphony, SMY, AT&T, IBM, XRCE, Um
d
- **2000**
  - SMU, ISI, Waterloo, IBM, CUNY–QC
- **2001**
  - InsightSoft, LCC, Oracle, ISI
- **2002**
  - LCC, InsightSoft, NUS
- **2003**
  - LCC, NUS, LexiClone, ISI
- **2004**
  - LCC, NUS, Uwales, IBM
- **2005**
  - LCC, NUS, IBM, Albany
Other Types of Questions

• **Definitional**
  – What is a boll weevil?

• **List**
  – Which states signed the US Declaration of Independence?

• **Crosslingual**
  – E.g., questions in Spanish, documents in English

• **Series**
  – Next slide
TREC Series Questions

- What are prions made of?
- Who discovered prions?
- What researchers have worked with prions?

- Who is the lead singer/musician in Nirvana?
- Who are the band members?
- When was the band formed?
- What is their biggest hit?
- What are their albums?
- What style of music do they play?

- What industry is Rohm and Haas in?
- Where is the company located?
- What is their annual revenue?
- How many employees does it have?

- What kind of insect is a boll weevil?
- What type of plant does it damage?
- What states have had problems with boll weevils?
NLP