Introduction to Linguistics
165: Computational Linguistics

Roger Levy
23 September 2010
What is computational linguistics?

• Formal tools for computing with natural language structure: lexicon, phonology, morphology, syntax
• Automatically classifying documents (Spam detection)
• Automated speech recognition by computers
• Automated spelling and grammar correction
• Information retrieval (web search!) and extraction
• Machine translation
• Question answering (Ask.com)
• Automated document summarization
Why study computational linguistics?

• Financial payoff!

...if someone told me I had to make a million bucks in one year...I'd grab a copy of [a computational linguistics textbook] and start a web text-processing company.

-- Peter Norvig, director of research, Google
Why study computational linguistics?

• Scientific interest!

I have emphasized in my writings on this subject that the challenge to psychological theory made by linguists to provide an adequate theory of language learning may well be regarded as the most significant intellectual challenge to theoretical psychology in this century. At the present time numerous difficult problems of providing a completely adequate scientific theory of language learning and language performance are enough to make even the most optimistic theorist uneasy. In very developed areas of science or mathematics, it is familiar to find the statement made that certain kinds of problems are simply far beyond the resources currently available but that certain more restricted problems are amenable to serious attack and likely solution.

-- Patrick Suppes, Lucie Stern Professor of Philosophy, Emeritus
Why study computational linguistics?

• Fun!

[At least, I hope to convince you of that over the next eleven weeks!]
Examples you can try at home

• Machine translation. It works pretty well sometimes!

```
How many teaspoons of sugar should I put in this cake mix?
```

```
Cuántas cucharillas de azúcar debo poner en esta mezcla para pastel?
```

```
How many teaspoons of sugar I put in the cake mix?
```

```
Cuántas cucharillas de azúcar debo poner en esta mezcla para pastel?
```

Examples you can try at home

• ...and sometimes it doesn’t work nearly as well.

`Google translate`

From: English  ➔  To: Vietnamese  ➔  Translate

Whichever restaurant he decides to eat at, I'm going to another.

Nhà hàng nào anh quyết định ăn ở, tôi đi đến khác.

`Listen`

`Google translate`

From: Vietnamese  ➔  To: English  ➔  Translate

Nhà hàng nào anh quyết định ăn ở, tôi đi đến khác.

Restaurants in what he decides to eat, I go to the other.

`Listen`
What we’ll do today, and in the course

- This is a hands-on course designed to get you familiar with using computers to process human language
- You’ll learn how to do this with some computational tools, including the programming language Python
- Class time will be a mix of lectures and practicum sessions
- You’ll have both pen-and-paper and simple computer-programming assignments
- All the computational tools you’ll need are on instructional servers (ieng6.ucsd.edu)