

Introduction

Lecture 1: Sept 11, 2013

CS886-2: Natural Language Understanding
University of Waterloo

CS886-2 Lecture Slides (c) 2013 P. Poupart

1

Logistics

- Web: cs.uwaterloo.ca/~ppoupart/teaching/cs886-fall13/cs886-fall13.html
- Term: Fall 2013
- Lectures: Wed & Fri 9-10:20 (DC2568)
- Prerequisite: undergrad course in Statistics
- Instructor: Pascal Poupart
 - ppoupart@uwaterloo.ca (DC2514)
 - Interests: Machine Learning, Natural Language Processing, Artificial Intelligence and Health Informatics

CS886-2 Lecture Slides (c) 2013 P. Poupart

2

Natural Language Understanding



Siri. Beta
Your wish is
its command.



 **WolframAlpha**™ computational...
knowledge engine

What is Stephen Wolfram up to now?

CS886-2 Lecture Slides (c) 2013 P. Poupart

3

Why Natural Language Understanding instead of Natural Language Processing?

- **Processing** indicates that natural text or speech will be transformed by some algorithm to
 - Ease search, information retrieval, classification
 - Infer author/speaker, sentiment, topic
 - Derive document or corpus level statistics
- **Understanding** emphasizes algorithms that focus on
 - The extraction of meaning/semantics of words/sentences

CS886-2 Lecture Slides (c) 2013 P. Poupart

4

Topics

- Semantic parsing and representation
- Latent semantic analysis
- Topic modeling
- Entity and relation extraction
- Co-reference resolution
- **Applications:** question answering, information retrieval, sentiment analysis, dialog management

CS886-2 Lecture Slides (c) 2013 P. Poupart

5

Who can take this course?

- **Everyone!**
- English is not my native language, can I take this course?
 - Yes, English is not my native language either
 - **Proficiency in English will not help you in this course**
- What will help me in this course?
 - **Statistics and machine learning**
- I haven't taken any course in Machine Learning, can I take this course?
 - Yes, **I will teach what you need to know about machine learning**

CS886-2 Lecture Slides (c) 2013 P. Poupart

6

Consider the following article

iPhone 5S, fingerprint scanner and everything else Apple announced at its event Tuesday

As expected, Apple lifted the curtain on a new version of its iPhone on Tuesday. Two of them, actually. Phil Schiller, Apple's senior VP of worldwide marketing, debuted the 5S during an event at the tech giant's headquarters in Cupertino, California, calling it the "most forward-thinking phone anyone's ever made." The company also announced the 5C, a new iPhone that costs less than the 5S. Apple unveiled a bunch of other stuff, too...

CS886-2 Lecture Slides (c) 2013 P. Poupart

7

Tasks that require some understanding

- **Topic classification:**
 - Is this a financial, technology, sport or political news?
- **Text summarization:**
 - Summarize the paragraph in one sentence
- **Question answering:**
 - Where was the iPhone 5S unveiled?
 - Which new iPhone is cheaper?

CS886-2 Lecture Slides (c) 2013 P. Poupart

8

Bag of words

- Ignore word order
- Treat each word as a feature
- Example:
 - address U.S. chemical Obama said a television threat in to precede country Barack his on night Tuesday not responding that to a in Syria would be a to U.S. the, to assure sought Americans strike President that diplomatic but efforts attack would any possible

CS886-2 Lecture Slides (c) 2013 P. Poupart

9

Suitability of bag of words

- What is the topic?
- Can you summarize the paragraph in one sentence?
- Can you answer questions such as:
 - What would precede a strike?

CS886-2 Lecture Slides (c) 2013 P. Poupart

10

Original article

- U.S. President Barack Obama said in a television address to his country on Tuesday night that not responding to a chemical attack in Syria would be a threat to the U.S., but sought to assure Americans that diplomatic efforts would precede any possible strike

Difficulty of Understanding

Entities and Relations

- The essence of a document can often be captured by the entities and relations that are mentioned
- **Entity**: object, person, organization, date, etc.
 - Most things denoted by a noun or pronoun
- **Relation**: property that links one or several entities
 - Most things denoted by an adjective, verb or adverb

CS886-2 Lecture Slides (c) 2013 P. Poupart

13

Circle the entities

iPhone 5S, fingerprint scanner and everything else Apple announced at its event Tuesday

As expected, Apple lifted the curtain on a new version of its iPhone on Tuesday. Two of them, actually. Phil Schiller, Apple's senior VP of worldwide marketing, debuted the 5S during an event at the tech giant's headquarters in Cupertino, California, calling it the "most forward-thinking phone anyone's ever made." The company also announced the 5C, a new iPhone that costs less than the 5S. Apple unveiled a bunch of other stuff, too...

CS886-2 Lecture Slides (c) 2013 P. Poupart

14

Named Entities

- Definition: **subset of entities referred by a “rigid designator”**
- Rigid designator: expression that always refers to the same thing in all possible worlds
- Categories of named entities:
 - Location, organization, person, object, time, etc.

CS886-2 Lecture Slides (c) 2013 P. Poupart

15

Circle the named entities

**iPhone 5S, fingerprint scanner and everything else
Apple announced at its event Tuesday**

As expected, Apple lifted the curtain on a new version of its iPhone on Tuesday. Two of them, actually. Phil Schiller, Apple’s senior VP of worldwide marketing, debuted the 5S during an event at the tech giant’s headquarters in Cupertino, California, calling it the “most forward-thinking phone anyone’s ever made.” The company also announced the 5C, a new iPhone that costs less than the 5S. Apple unveiled a bunch of other stuff, too...

CS886-2 Lecture Slides (c) 2013 P. Poupart

16

Co-reference resolution

- Problem of recognizing which entities denote the same thing
- Task: partition entities into groups that denote the same thing

CS886-2 Lecture Slides (c) 2013 P. Poupart

17

Circle co-references of Apple and put co-references of 5S in boxes

iPhone 5S, fingerprint scanner and everything else Apple announced at its event Tuesday

As expected, Apple lifted the curtain on a new version of its iPhone on Tuesday. Two of them, actually. Phil Schiller, Apple's senior VP of worldwide marketing, debuted the 5S during an event at the tech giant's headquarters in Cupertino, California, calling it the "most forward-thinking phone anyone's ever made." The company also announced the 5C, a new iPhone that costs less than the 5S. Apple unveiled a bunch of other stuff, too...

CS886-2 Lecture Slides (c) 2013 P. Poupart

18

Circle relations and link entities

iPhone 5S, fingerprint scanner and everything else Apple announced at its event Tuesday

As expected, Apple lifted the curtain on a new version of its iPhone on Tuesday. Two of them, actually. Phil Schiller, Apple's senior VP of worldwide marketing, debuted the 5S during an event at the tech giant's headquarters in Cupertino, California, calling it the "most forward-thinking phone anyone's ever made." The company also announced the 5C, a new iPhone that costs less than the 5S. Apple unveiled a bunch of other stuff, too...

CS886-2 Lecture Slides (c) 2013 P. Poupart

19

Question Answering

- Possible approach:
 - Extract entities and relations from document and question
 - Find answer by co-reference resolution
- Example: Where was the new iPhone launched?

CS886-2 Lecture Slides (c) 2013 P. Poupart

20

Semantic Representations

- Traditional representations:
 - First order logic, description logic, temporal logic
 - Semantic network
- Problem:
 - Difficult to represent uncertainty due to ambiguities

CS886-2 Lecture Slides (c) 2013 P. Poupart

21

Probabilistic Representations

- Idea:
 - Use random variables and distributions
- Probabilistic graphical models
 - Bayesian networks, Markov networks
- Question answering by probabilistic inference

CS886-2 Lecture Slides (c) 2013 P. Poupart

22

Probabilistic co-references in iPhone article