NLP Special Topics

Ruihong Huang
Texas A&M University
Information Extraction

Fed Chairman Ben Bernanke said the U.S. economy... The euro rose to $1.2008, compared to $1.1942 on Tuesday.

Facts.
- Ben Bernanke is a Person.
- Fed is an Organization.
- The US is a Country.
- Fed is located in the US.
- Ben Bernanke is the US Fed Chairman.
- $1.2008 is an amount of money.
Christopher Manning

Named Entity Recognition (NER)

- A very important sub-task: find and classify names in text, for example:

- The decision by the independent MP Andrew Wilkie to withdraw his support for the minority Labor government sounded dramatic but it should not further threaten its stability. When, after the 2010 election, Wilkie, Rob Oakeshott, Tony Windsor and the Greens agreed to support Labor, they gave just two guarantees: confidence and supply.
One Specific Domain

[A 14yo doxy] **ANIMAL** owned by [a reputable breeder] **HUMAN** is being treated for [IBD] **DISEASE** with [pred] **DRUG**. But [she] **ANIMAL** has not been improving.
Relation Extraction

• Goal: detecting semantic relations between two entities.
• Relations are domain dependent.
Why Relation Extraction?

- Create new structured knowledge bases, useful for any app
- Augment current knowledge bases
  - Adding words to WordNet thesaurus, facts to FreeBase or DBPedia
- Support question answering
  - The granddaughter of which actor starred in the movie “E.T.”?
    \[(\text{acted-in } ?x \text{ “E.T.”})(\text{is-a } ?y \text{ actor})(\text{granddaughter-of } ?x ?y)\]
- But which relations should we extract?
Event Extraction

Event Roles in Terrorism Domain:

- Perpetrators, victims, targets, weapons

Lopez Albajar, former defense minister, was killed with pistols by three young individuals. His two bodyguards were also shot to death.

Perp: three young individuals.
Vict: Lopez Albajar
     His two bodyguards
Weap: pistols

Event Roles in Management Succession Domain:

- People and companies in corporate management changes.

Daniel Glass was named president of EMI Records Group. He succeeds Charles Koppelman.

Per_In: Daniel Glass
Per_Out: Charles Koppelman.
Post: president
Org: EMI Records Group
To identify civil unrest events...
• Rally?

• Strike?
Coreference Resolution

S2 The active nuclear form of the NF-kappa B transcription factor complex\(^ \text{<T27>} \) is composed of two DNA binding subunits, NF-kappa B\(^ >\text{p65}<\text{T4} \) and NF-kappa B\(^ >\text{p50}<\text{T5}\text{T28}\), both of which\(^ >\text{share}<\text{T29}\) share extensive N-terminal sequence homology with the \( v-rel \text{T6} \) oncogene product.

S3 The NF-kappa B\(^ >\text{p65}<\text{T7} \) subunit provides the transactivation activity in this complex\(^ >\text{<T30}>\) and serves as an intracellular receptor for a cytoplasmic inhibitor of NF-kappa B, termed I kappa B.

S4 In contrast, NF-kappa B\(^ >\text{p50}<\text{T8} \) alone fails to stimulate kappa B-directed transcription, and based on prior in vitro studies, is not directly regulated by I kappa B.

S5 To investigate the molecular basis for the critical regulatory interaction between NF-kappa B and I kappa B/MAD-3\(^ >\text{T2}\), a series of human NF-kappa B\(^ >\text{p65}<\text{T10}\text{T31} \) mutants was identified that functionally segregated DNA binding, I kappa B-mediated inhibition, and I kappa B-induced nuclear exclusion of this transcription factor\(^ >\text{T32}\).
Why sentiment analysis?

- **Movie**: is this review positive or negative?
- **Products**: what do people think about the new iPhone?
- **Public sentiment**: how is consumer confidence? Is despair increasing?
- **Politics**: what do people think about this candidate or issue?
- **Prediction**: predict election outcomes or market trends from sentiment
Scherer Typology of Affective States

- **Emotion**: brief organically synchronized ... evaluation of a major event
  - angry, sad, joyful, fearful, ashamed, proud, elated
- **Mood**: diffuse non-caused low-intensity long-duration change in subjective feeling
  - cheerful, gloomy, irritable, listless, depressed, buoyant
- **Interpersonal stances**: affective stance toward another person in a specific interaction
  - friendly, flirtatious, distant, cold, warm, supportive, contemptuous
- **Attitudes**: enduring, affectively colored beliefs, dispositions towards objects or persons
  - liking, loving, hating, valuing, desiring
- **Personality traits**: stable personality dispositions and typical behavior tendencies
  - nervous, anxious, reckless, morose, hostile, jealous
Semantic Role Labeling

- Shallow semantic parsing
- Goal: detecting semantic arguments (semantic roles) of a predicate in a sentence.
- a higher-level abstraction
- Example: Mary (agent) sold the book (theme) to John (recipient).
- Example: The book (theme) was sold by Mary (agent) to John (recipient).
Semantic Parsing

- Goal: mapping a natural-language sentence into a formal representation of its meaning, e.g., logical formulas.

- How many major cities are in states bordering Texas? : \( \text{answer}(x1, \text{count}(x2, (\text{city}(x2), \text{major}(x2), \text{loc}(x2, x3), \text{next to}(x3, x4), \text{state}(x3), \text{equal}(x4, \text{stateid}(\text{texas}))))) \)
• Please email me at huangrh@cse.tamu.edu, your preferences of papers to present.

• Please give 3 options, ranked #1, #2, #3.

• Please indicate if you’re willing to present two papers.