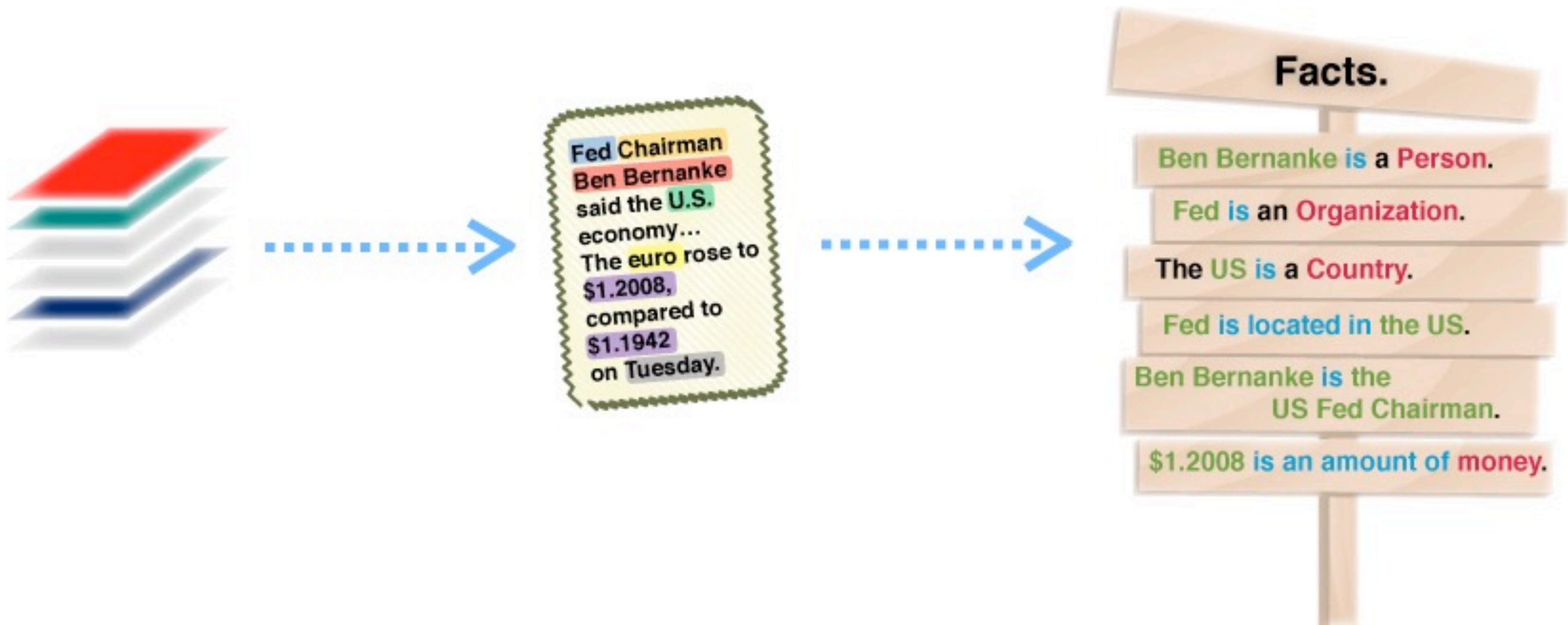


NLP Special Topics

Ruihong Huang
Texas A&M University

Information Extraction





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.

Person
Date
Location
**Organi-
zation**

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.
- Example: B. Gates married Melinda French on Jan. 1, 1994. => Spouse (B. Gates, Melinda French)



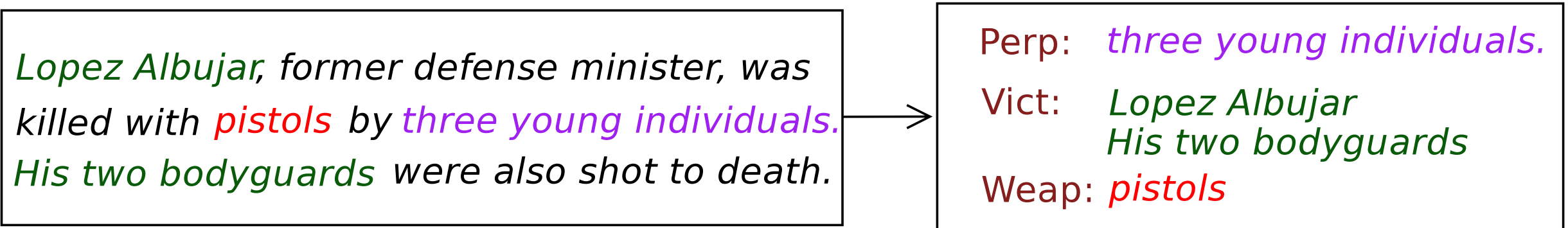
Why Relation Extraction?

- 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.”?
(acted-in ?x “E.T.”)(is-a ?y actor)(granddaughter-of ?x ?y)
- But which relations should we extract?

Event Extraction

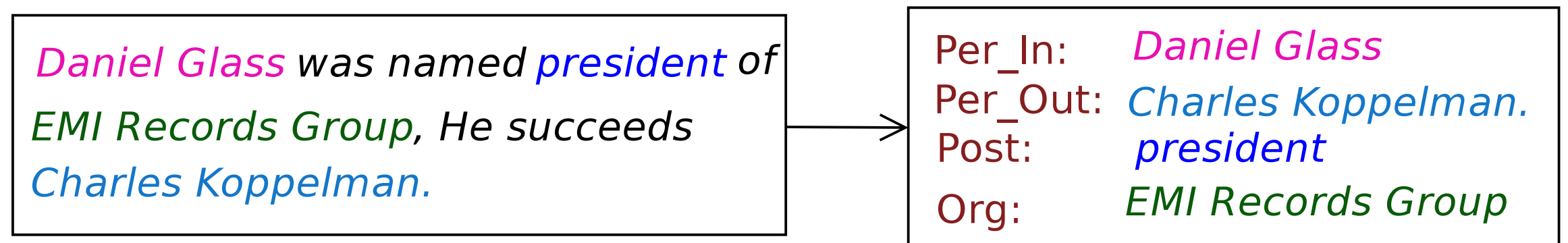
Event Roles in Terrorism Domain:

- Perpetrators, victims, targets, weapons



Event Roles in Management Succession Domain:

- People and companies in corporate management changes.



To identify civil unrest events...



- Rally?



- Strike?



Coreference Resolution

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

S3 The NF-kappa B p65^{T7} subunit provides the transactivation activity in this complex^{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 p50^{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^{T9}, a series of human NF-kappa B p65^{T10 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^{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? : `answer(x1, count(x2, (city(x2), major(x2), loc(x2, x3), next to(x3, x4), state(x3), equal(x4, stateid(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.