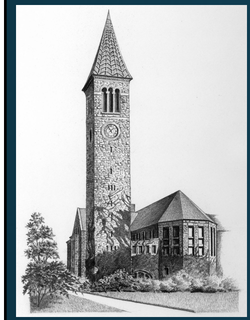


Claire Cardie
Cornell University
Departments of Computer Science
and Information Science

Information Extraction :
from opinions to arguments (to persuasion)

Plan for the talk

- Two real-world structured prediction tasks in NLP

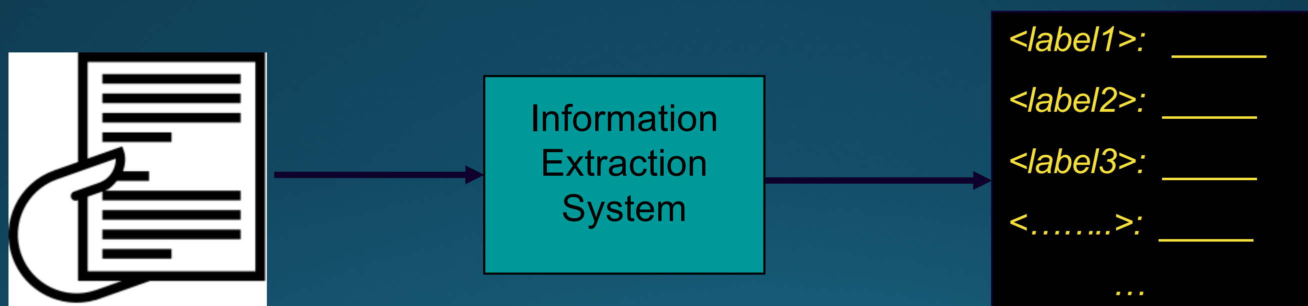
Opinion extraction
1. Formulation as structured prediction
2. ML methods employed (covered some this morning)
3. Performance results

Argument extraction
1. Formulation as structured prediction

- Our current research on persuasion

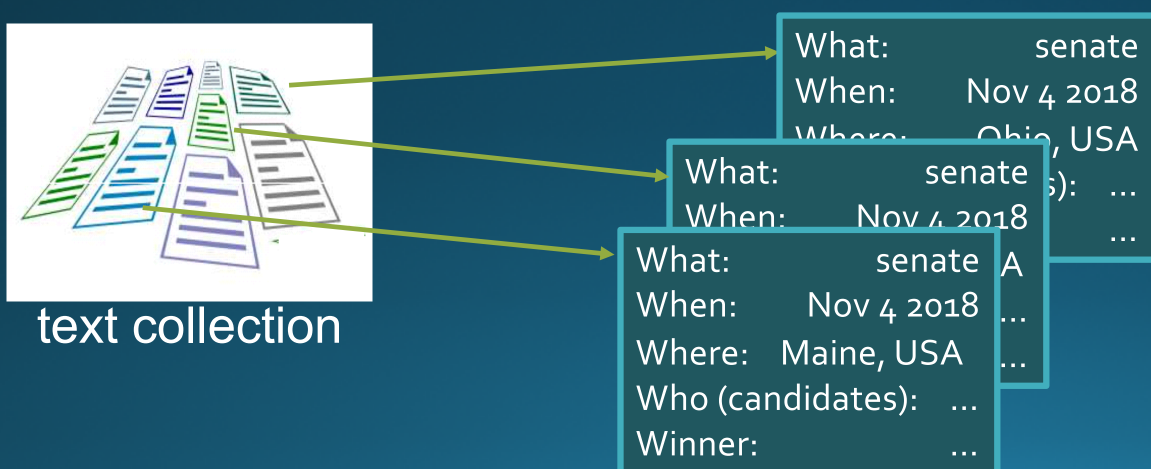
Information extraction

- Unstructured text → structured representation



Information extraction

- Usually domain-specific focus, usually fact- or event-oriented



Create database from text

Opinion extraction

- Sentence-level task (typically)

The White House press corps launched a bitter attack on Trump…

text collection

Opinion holder:	“White House press corps”
Target:	“Trump”
Polarity:	negative
Intensity:	high
Date:	…

Opinion extraction

text collection

Opinion holder:	“White House press corps”
Target:	“Trump”
Polarity:	negative
Intensity:	high
Date:	…

Opinion extraction

President Trump **hopes** to build a wall along the U.S.-Mexico border.

text collection

Opinion holder:	"President Trump"
Target:	"build a wall along…"
Polarity:	positive
Intensity:	low

O	
Target:	"Trump"
Polarity:	negative
Intensity:	high
Date:	…

Opinion extraction

Broccoli is not one of Bush's **favorite** foods.

text collection

Opinion holder: "Bush"
Target: "broccoli"
Polarity: negative
Intensity: medium

Opini
Target: "build a wall along…"
Polarity: positive
Intensity: low

O
Target:
Polarity: negative
Intensity: high
Date: …

Opinion extraction (and IE generally)

- Connections to
  - Relation extraction
  - Event extraction
  - Slot-filling

ML perspective: structured prediction

General approach: sequence tagging

Plan for the talk

- Two real-world structured prediction tasks in NLP

Opinion extraction

1. **Formulation as structured prediction**
2. **ML methods employed (covered some this morning)**
3. Performance results

Argument extraction

1. Formulation as structured prediction

- Our current research on persuasion

Sequence tagging

Lose the information that there were two opinion expressions.

Polarity: negative
Source : “environmentalists”
Target: “the proposal”

Target: The proposal

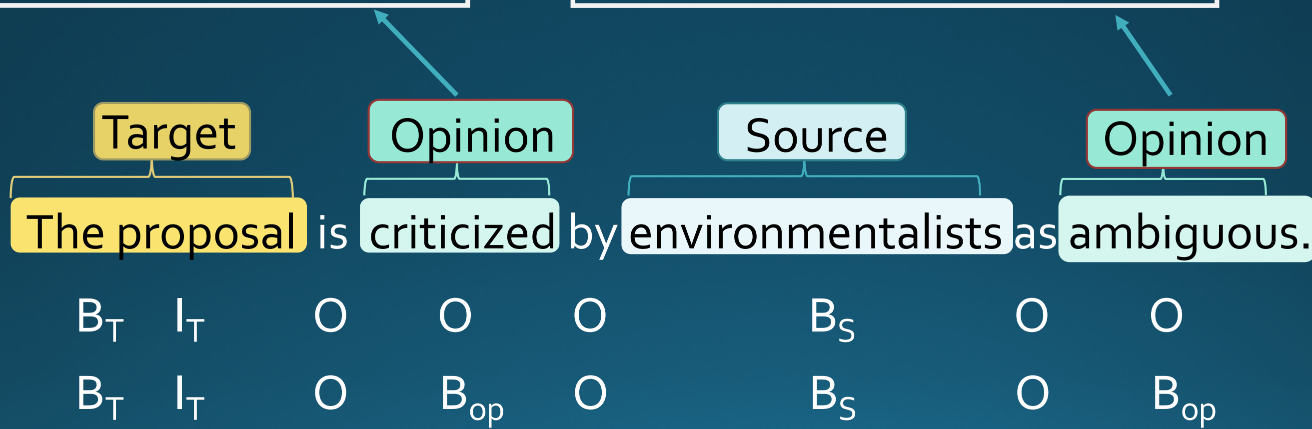
Source: environmentalists

The	proposal	is	criticized	by	environmentalists	as	ambiguous.
T	T	--	--	--	S	--	--
T_{neg}	T_{neg}	--	--	--	S_{neg}	--	--
B_T	I_T	O	O	O	B_S	O	O

Sequence tagging

OpExpr:	“criticized”
Polarity:	negative
Source :	“environmentalists”
Target:	“the proposal”

OpExpr:	“ambiguous”
Polarity:	negative
Source :	“environmentalists”
Target:	“the proposal”



Sequence tagging with complications…

Polarity: negative
Source : “environmentalists”
Target: “the proposal”

Polarity: negative
Source: “Trump”
Target: “environmentalists”

The proposal is **criticized** by environmentalists that are Trump’s **enemies**.

B_S ?
B_T ?

Sequence tagging with complications…

- Generally viewed as **two tasks**
  - Opinion entity identification
    - Opinion expression
    - Source
    - Target
  - Relation detection among entities
    - For each opinion expression
      - <opinion expression> IS-FROM <source>
      - <opinion expression> IS-ABOUT <target>

ML Pipeline

- Extract candidate entities: **sequence tagging**

Target			Opinion		Source		Opinion
The	proposal	is	criticized	by	environmentalists	as	ambiguous.
B_T	I_T	O	B_{Op}	O	B_S	O	B_{Op}

ML Pipeline

- Sequence tagging

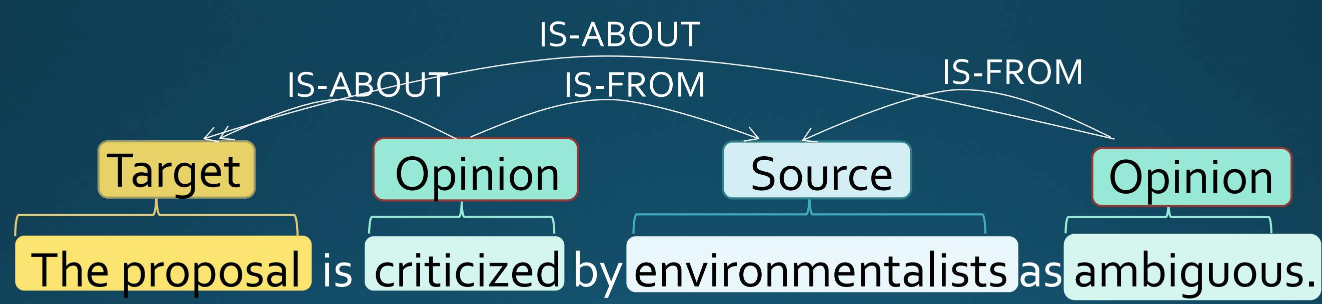
  \+ **classification** (relation identification)

$\overbrace{\text{The proposal}}^{\text{Target}}$ is $\overbrace{\text{criticized}}^{\text{Opinion}}$ by $\overbrace{\text{environmentalists}}^{\text{Source}}$ as $\overbrace{\text{ambiguous.}}^{\text{Opinion}}$

ML Pipeline

- Sequence tagging

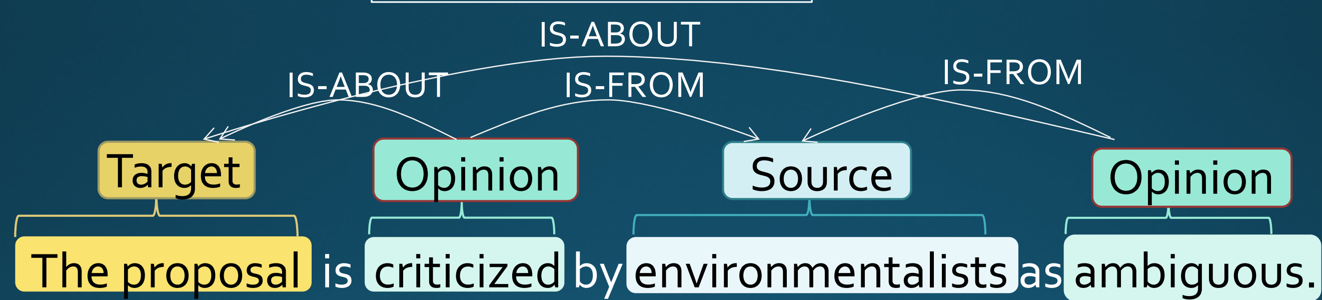
  + **classification** (relation identification)



ML Pipeline

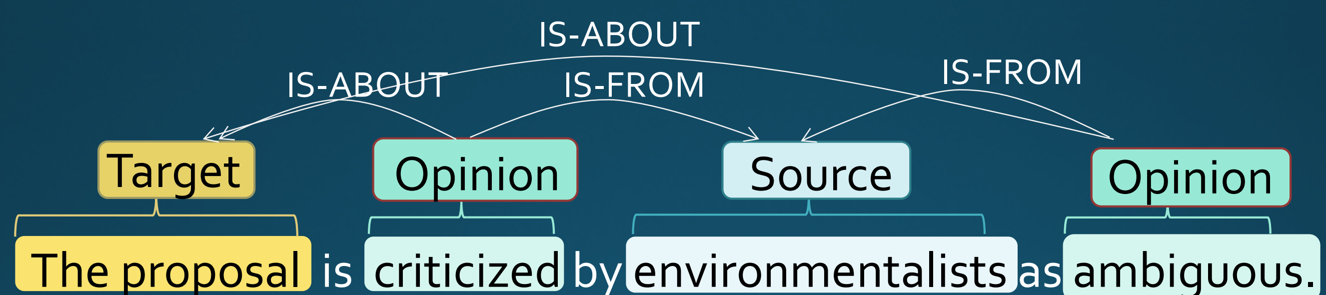
- Sequence tagging e.g., CRFs [Lafferty et al., 2001]

  \+ classification e.g., SVMs, MaxEnt

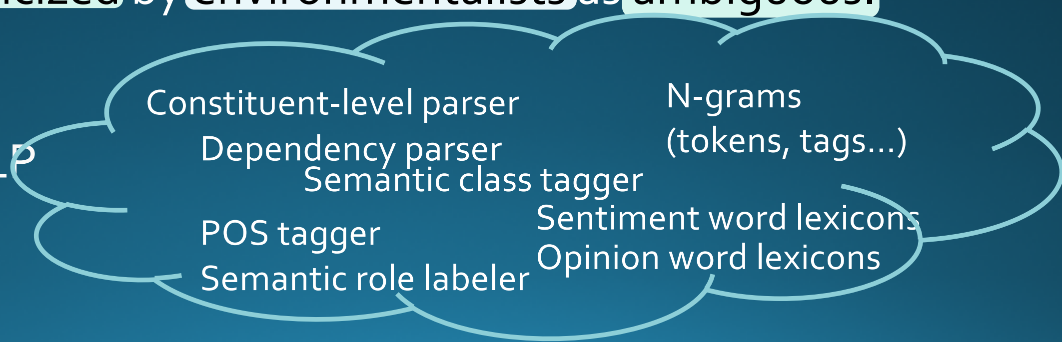


ML Pipeline

- Sequence tagging
  + classification



Hand-crafted features
Drawn from output of NLP components



Well-known problems

- Error propagation
  - Errors made in entity extraction limit performance of relation classification
- Relation extraction cannot influence entity candidate generation

Mitigated, in part, by ML methods for:

1. Joint inference (E.g. ILP-based, AD^3)
2. Joint learning
   - End-to-end neural methods

Joint inference models (e.g. Roth & Yih, 2004)

- Allow modeling of global constraints on the output structure
- Simple models are learned separately
  - Top-k results are used
- Incorporation of **task-specific constraints** can bias (re-rank or remove) decisions made by simpler models
- Constraints employed (only) at the decision time
- Can be solved for using, e.g., Integer Linear Programming (ILP), AD^3 (Martins & Smith)

For opinion extraction: See Choi & Cardie (2006); Yang & Cardie, 2013

Limitations

- Entities and relations are still learned separately
  - Relation information cannot influence the entity extraction
- Linear constraints

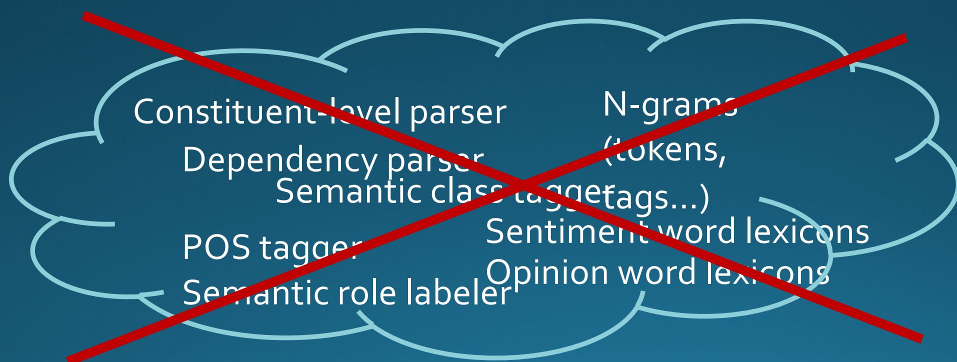
Joint extraction of entities and relations

- Entities and relations are learned jointly
- Disadvantage
  - Heavily feature-engineered
    - E.g. Li and Ji, 2014; Miwa and Sasaki, 2014

For opinion extraction: See Yang & Cardie, 2013

End-to-end neural network approaches

- Joint extraction of entities and relations
  - **Without** NLP components, **without** feature engineering, **without** manually procured lexicons
- Comparable (and sometimes better) performance than feature-based approaches



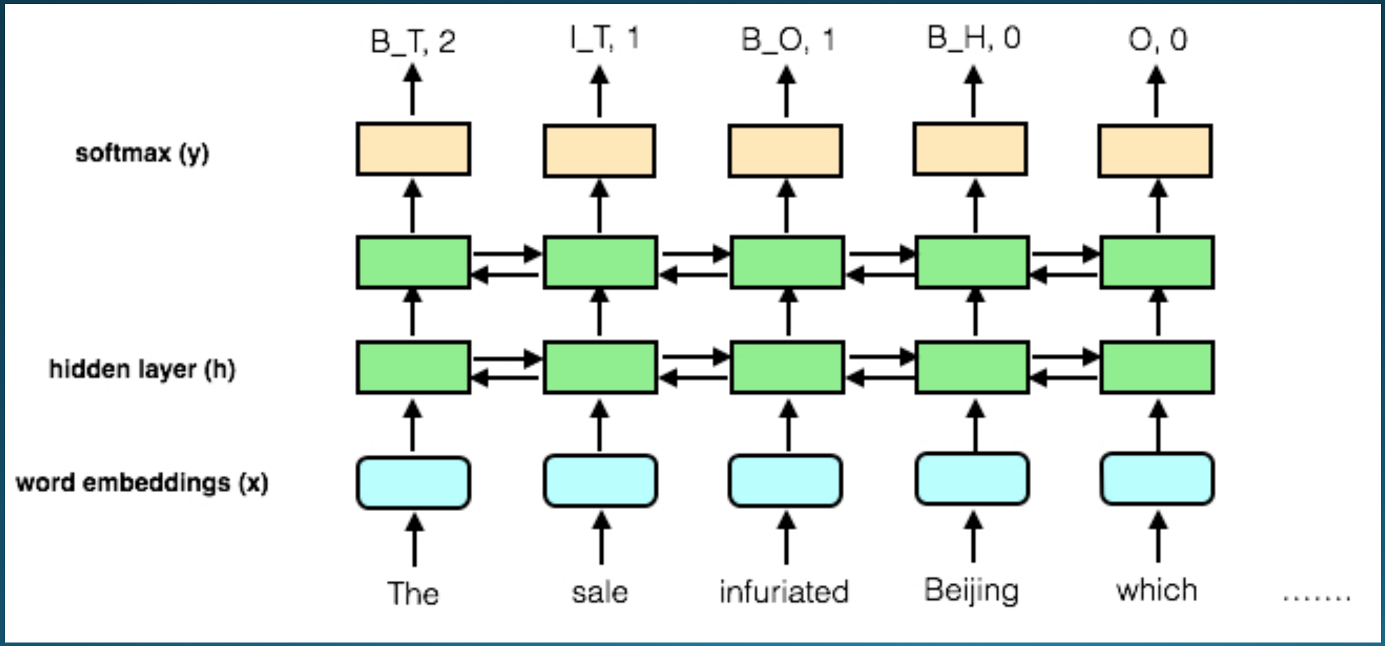
Opinion extraction performance

- NOT measured at the token-level !!!!!
- Measured at the entity and relation level
  - Recall, precision, F-measure
- Data set
  - MPQA
  - ~500 documents with fine-grained opinion extraction information
  - 10's of thousands of opinions

Use a multi-layer bi-directional LSTM

[Katiyar &Cardie, ACL 2016]

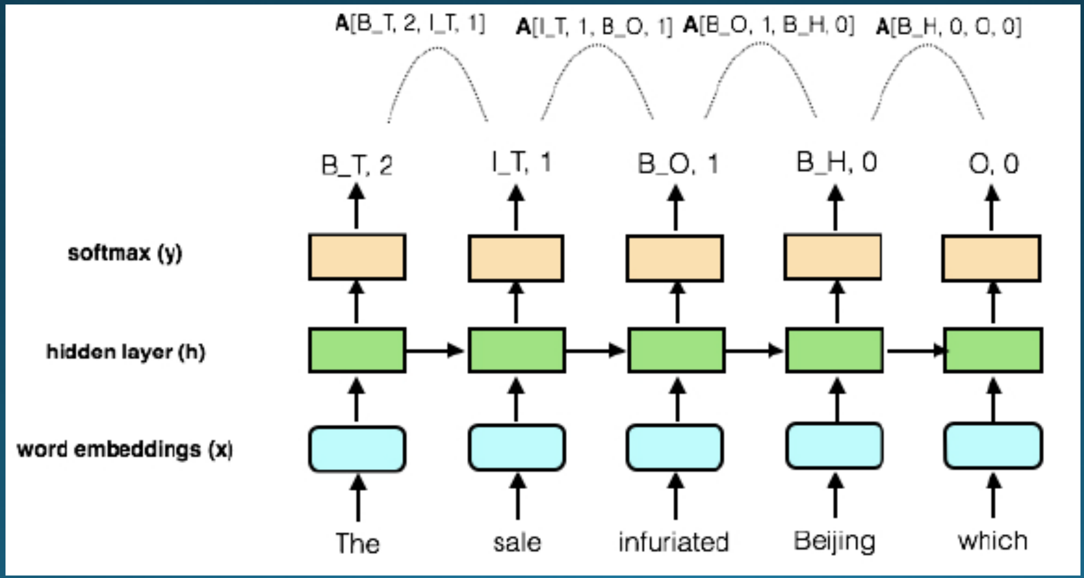
- Not competitive with best CRF+ILP joint inference approach
  - Yang & Cardie (ACL 2013)



Add sentence+relation-level likelihood

[Katiyar &Cardie, ACL 2016]

- Incorporate dependencies between consecutive labels
  - Via CRF at top layer



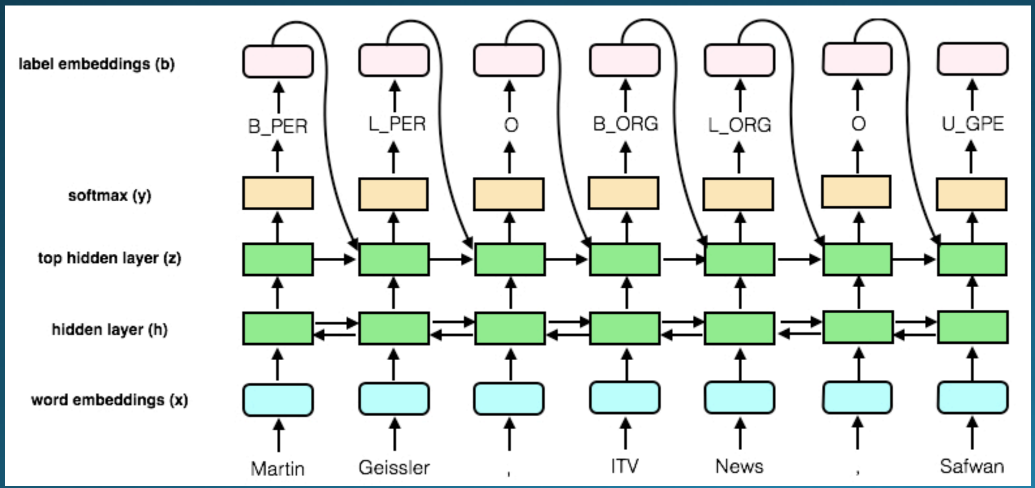
Results

- As good as CRF+ILP for IS-ABOUT
- Within 1-3% F-score for opinion entity extraction + IS-FROM

S1 :	[Australia's involvement in Kyoto]$_{T_1}$ [has been in doubt]$_{O_1}$ ever since [the US President, George Bush]$_{H_2}$, [announced]$_{O_2}$ last year that [ratifying the protocol]$_{T_2}$ would hurt the US economy.
CRF+ILP	Australia's involvement in Kyoto [has been in doubt]$_{O_1}$ ever since the US President, George Bush, announced last year that [ratifying the protocol]$_{T_1}$ would hurt the US economy.
LSTM	[Australia's involvement in Kyoto]$_{T}$ [has]$_{O}$ been in doubt ever since the US [President]$_{H}$, [George Bush]$_{H}$, announced last year that ratifying the protocol would hurt the US economy.
SLL	[Australia's involvement in Kyoto]$_{T}$ [has been in doubt]$_{O}$ ever since the US President, George Bush, announced last year that ratifying the protocol would hurt the US economy.
SLL+RLL	[Australia's involvement in Kyoto]$_{T}$ [has been in doubt]$_{O}$ ever since the US President, [George Bush]$_{H_2}$, [announced]$_{O_2}$ last year that [ratifying the protocol]$_{T_2}$ would hurt the US economy.

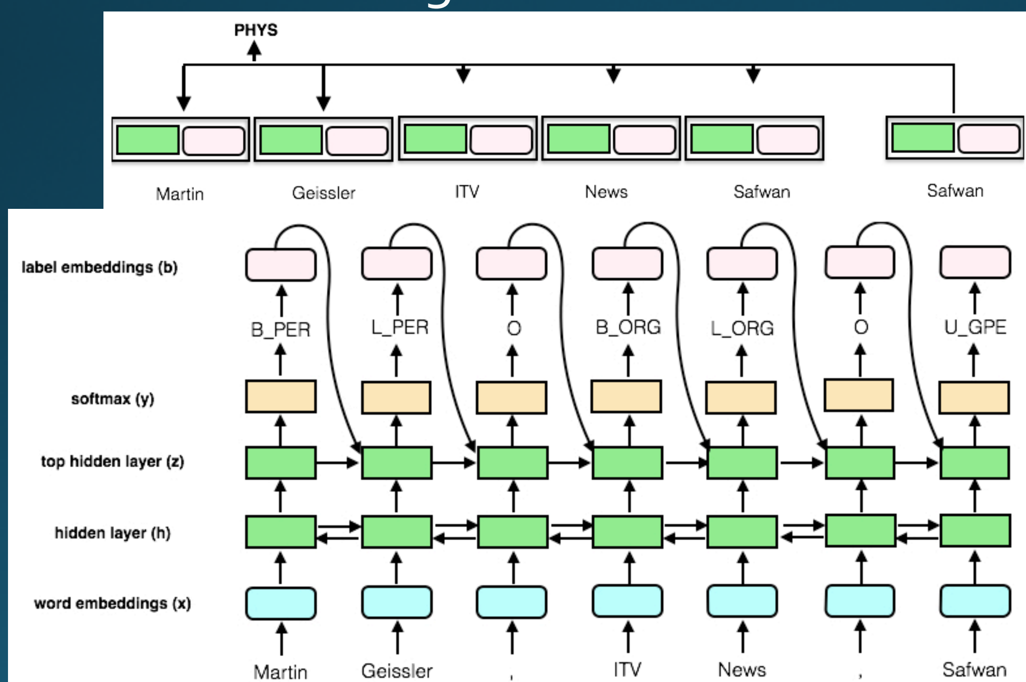
Katiyar & Cardie (ACL 2017)

- Entity extraction using sequence labeling



Katiyar & Cardie (ACL 2017)

- Entity extraction using sequence labeling
- Relation extraction using attention



Plan for the talk

- Two real-world structured prediction tasks in NLP
  - Opinion extraction
  - **Argument extraction**
- **Our current research on persuasion**

Argumentation Mining

Want to understand not only

WHAT people are thinking (i.e., opinions),

but **WHY** they are thinking it

Ultimately

- distinguish good vs. bad arguments
- understand what makes an argument persuasive

Expose the reasoning behind an opinion

- Argument parsing

[1]There should be a full ban of peanut products on all airlines,[2]because peanut allergy could have terrible effects.[3]Peanut reactions can be life threatening.[4]Restricting to certain flights is not enough,[5]as residue from previous flights can remain on the seats.[6]Recently we flew across the country[7]and I find left over peanuts in our seats!

[Joonsuk Park, Cornell PhD thesis, 2016]
[Niculae, Park & Cardie, ACL 2017]

Structured prediction ...at the **discourse** level

1. Proposition classification
2. Identification of support relations

[1]There should be a full ban of peanut products on all airlines,[2]because peanut allergy could have terrible effects.[3]Peanut reactions can be life threatening.[4]Restricting to certain flights is not enough,[5]as residue from previous flights can remain on the seats.[6]Recently we flew across the country[7]and I find left over peanuts in our seats!

Value

See Niculae, Park & Cardie (ACL 2017)

- Joint learning approach
- **Based on factor graph construction**
  - Heavily feature engineered OR not
  - Allows arbitrary task-specific constraints

Neural nets do *not* perform the best.

What makes a convincing argument?

- **Previous work in NLP** identified linguistic features important for discriminating persuasive language from non-persuasive language.

[Tan et al., 2016]

[Zhang et al., 2016]

[Potash and Rumshisky et al., 2017]

Findings using on-line debates

- Winners
  - actively pursue opponents' points rather than promoting their own ideas
  - have longer argument – more words, more sentences, more paragraphs
  - use calmer language
  - use more 1st person singular pronouns (self affirmation)
  - use fewer person plurals (distancing from presented view)
  - …

Our current work…

- Are logical/well-formed arguments more persuasive than less logical arguments?
  - Initial results: NO

Argument Parsing

- Are logical/well-formed arguments judged as higher quality than less logical arguments?
  - Initial results: MAYBE
- What aspects of the argument structure are most associated with persuasiveness / quality?
  - Initial results: more support links = more persuasive and higher quality

But…

What makes a persuasive argument depends on:

who is voting/reading/listening

their prior beliefs on the topic of the argument

More important than the language used

Affect the ranking of linguistic features that predictive persuasion

See work of PhD student Esin Durmus

NAACL 2018, PEOPLES@NAACL 2018, WWW 2019,

ACL 2019, ArgMining@ACL 2019

The End