Natural Language Processing

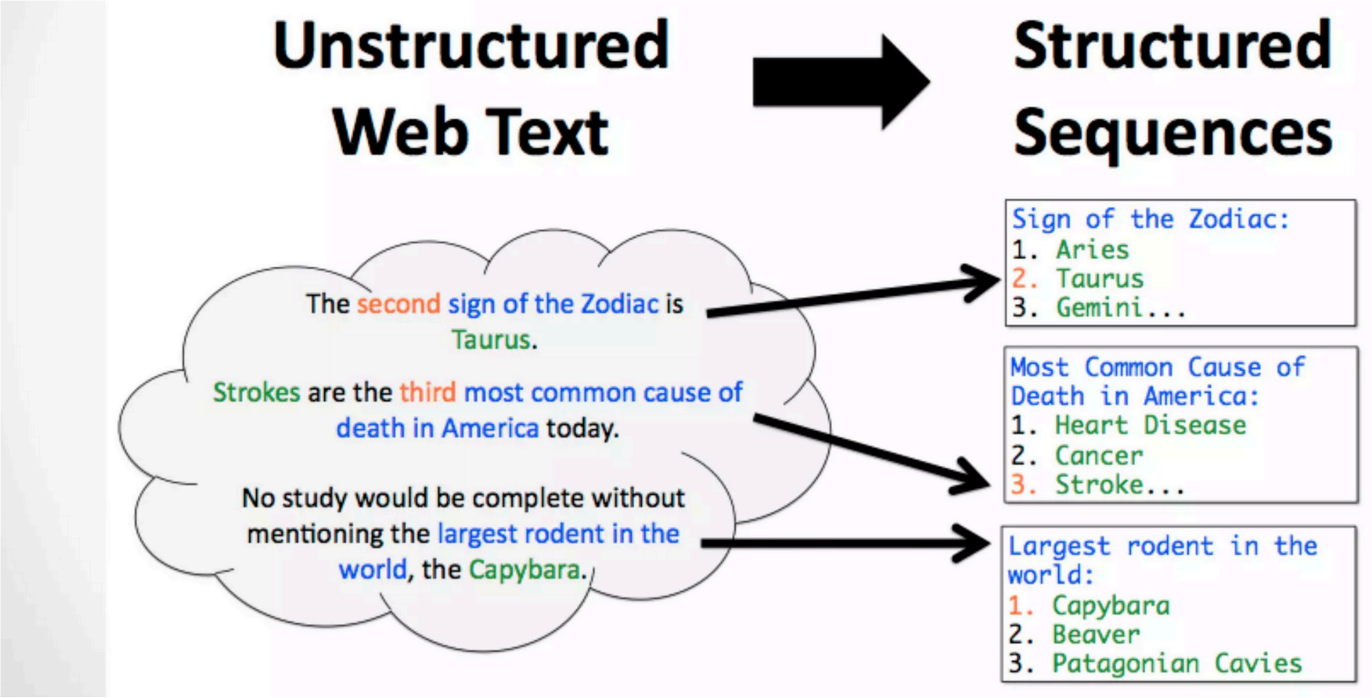
Info 159/259
Lecture 18: Information Extraction (April 1, 2024)

Many slides & instruction ideas borrowed from:
David Bamman, Dan Jurafsky & Ruben Izquirdo

Logistics

- HW4, Exam2 & HW5 are being graded.
- AP2: due next Tuesday (April 9)
- 259 Mid-project report due next Thursday (April 11)
- Quiz 8 will be out this Friday afternoon (due Monday night).

Information extraction



IE Presentation by Ruben Izquirdo

Investigating(SEC, Tesla)

SEC Probing Tesla CEO Musk's Tweets: Reports

By Deborah DSouza | Updated August 9, 2018 — 4:47 AM EDT

On Tuesday, Tesla Inc. (TSLA) CEO Elon Musk made the dramatic announcement that he was considering taking Tesla private for $420 a share on Twitter. In an email sent to Tesla employees posted on the company's official blog, Musk explained that he is mulling taking the firm private to protect it from short sellers and wild swings in stock prices. However, the email didn't provide any details regarding financing. (See also: What if Tesla Goes Private?)

Withdraws-from(Sanders, US Presidential Race)

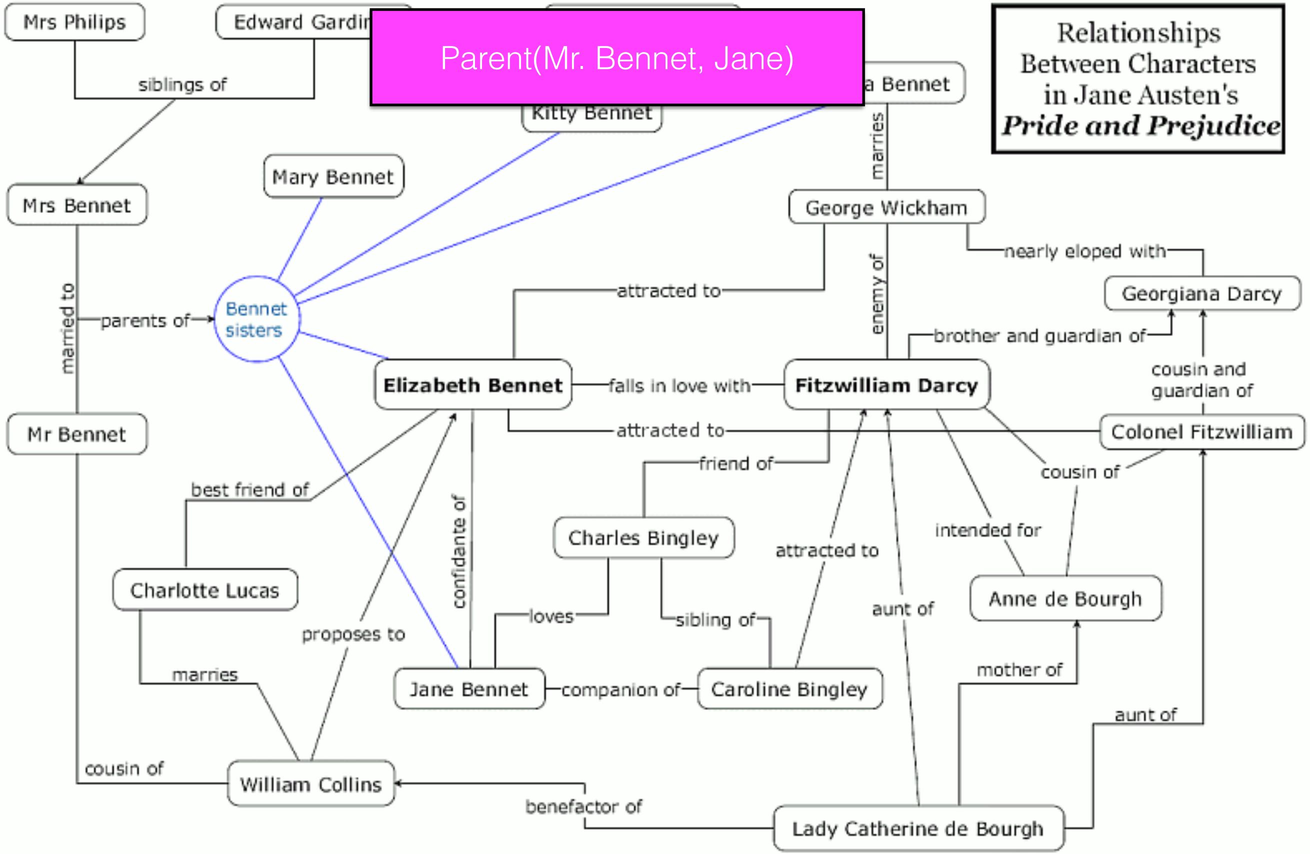
Bernie Sanders Drops Out of 2020 Democratic Race for President

By Sydney Ember

April 8, 2020

4497

Senator Bernie Sanders of Vermont ended his presidential candidacy on Wednesday, concluding a quest that elevated him as a standard-bearer of American liberalism and clearing the way for a general election between the presumptive Democratic nominee, Joseph R. Biden Jr., and President Trump at a time of national crisis.



https://en.wikipedia.org/wiki/Pride_and_Prejudice

Information extraction

- Named entity recognition
- Entity linking
- Relation extraction

Named entity recognition

[tim cook]$_{\text{PER}}$ is the ceo of [apple]$_{\text{ORG}}$

- Identifying spans of text that correspond to typed entities

Named entity recognition

Type	Tag	Sample Categories	Example sentences
People	PER	people, characters	**Turing** is a giant of computer science.
Organization	ORG	companies, sports teams	The **IPCC** warned about the cyclone.
Location	LOC	regions, mountains, seas	The **Mt. Sanitas** loop is in **Sunshine Canyon**.
Geo-Political Entity	GPE	countries, states, provinces	**Palo Alto** is raising the fees for parking.
Facility	FAC	bridges, buildings, airports	Consider the **Golden Gate Bridge**.
Vehicles	VEH	planes, trains, automobiles	It was a classic **Ford Falcon**.

Figure 17.1 A list of generic named entity types with the kinds of entities they refer to.

ACE NER categories (+weapon)

Named entity recognition

- GENIA corpus of MEDLINE abstracts (biomedical)

We have shown that [interleukin-1]PROTEIN ([IL-1]PROTEIN) and [IL-2]PROTEIN control [IL-2 receptor alpha (IL-2R alpha) gene]DNA transcription in [CD4-CD8- murine T lymphocyte precursors]CELL LINE

protein
cell line
cell type
DNA
RNA

http://www.aclweb.org/anthology/W04-1213

BIO notation

B-PERS I-PERS O O O O B-ORG

tim cook is the ceo of apple

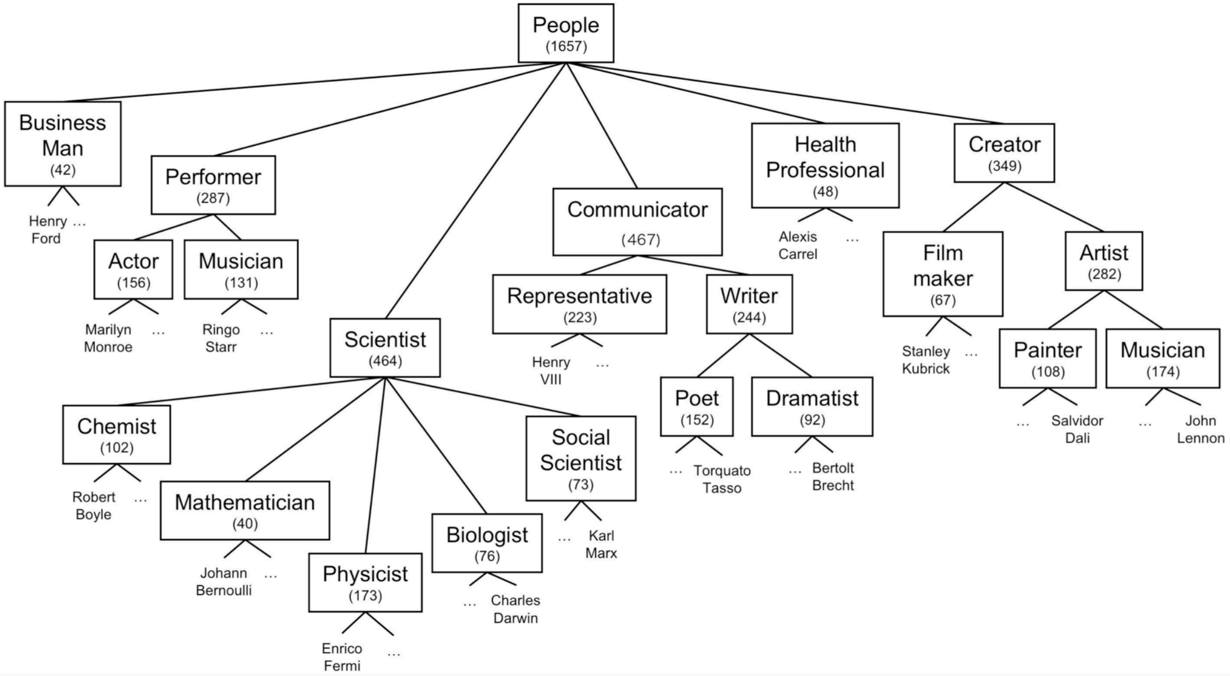
- **B**eginning of entity
- **I**nside entity
- **O**utside entity

[tim cook]$_{\text{PER}}$ is the ceo of [apple]$_{\text{ORG}}$

Named entity recognition

			B-PER	B-PER					
After	he	saw	Harry	Tom	went	to	the	store	

Fine-grained NER



Giuliano and Gliozzo (2008)

Fine-grained NER

WordNet Search - 3.1

- WordNet home page - Glossary - Help

Word to search for: Bertolt Brecht Search WordNet

Display Options: (Select option to change) Change

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations
Display options for sense: (gloss) "an example sentence"

Noun

- S: (n) Brecht, **Bertolt Brecht** (German dramatist and poet who developed a style of epic theater (1898-1956))
  - ***instance***
    - S: (n) dramatist, playwright (someone who writes plays)
    - S: (n) poet (a writer of poems (the term is usually reserved for writers of good poetry))

Entity recognition

Person	… named after [the daughter of a Mattel co-founder] …
Organization	[The Russian navy] said the submarine was equipped with 24 missiles
Location	Fresh snow across [the upper Midwest] on Monday, closing schools
GPE	The [Russian] navy said the submarine was equipped with 24 missiles
Facility	Fresh snow across the upper Midwest on Monday, closing [schools]
Vehicle	The Russian navy said [the submarine] was equipped with 24 missiles
Weapon	The Russian navy said the submarine was equipped with [24 missiles]

ACE entity categories
https://www.ldc.upenn.edu/sites/www.ldc.upenn.edu/files/english-entities-guidelines-v6.6.pdf

Named entity recognition

- Most named entity recognition datasets have flat structure (i.e., non-hierarchical labels).

✔ [The University of California]$_{\text{ORG}}$

✖ [The University of [California]$_{\text{GPE}}$]$_{\text{ORG}}$

- Mostly fine for named entities, but more problematic for general entities:

[[John]$_{\text{PER}}$'s mother]$_{\text{PER}}$ said …

Nested NER

named	after	the	daughter	of	a	Mattel	co-founder
						B-ORG	
					B-PER	I-PER	I-PER
		B-PER	I-PER	I-PER	I-PER	I-PER	I-PER

Sequence labeling

$$x = \{x_1, \ldots, x_n\}$$

$$y = \{y_1, \ldots, y_n\}$$

- For a set of inputs x with n sequential time steps, one corresponding label y_i for each x_i
- Model correlations in the labels y.

Sequence labeling (methods)

- Hidden Markov Model (HMM)
- Maximum Entropy Markov Model (MEMM)
- Conditional Random Field (CRF)
    - Another discriminative model with better handling of label dependency
- Decoding (finding the optimal label sequence): Viterbi

Sequence labeling

- Feature-based models (MEMM, CRF)

identity of w_i, identity of neighboring words
embeddings for w_i, embeddings for neighboring words
part of speech of w_i, part of speech of neighboring words
base-phrase syntactic chunk label of w_i and neighboring words
presence of w_i in a **gazetteer**
w_i contains a particular prefix (from all prefixes of length ≤ 4)
w_i contains a particular suffix (from all suffixes of length ≤ 4)
w_i is all upper case
word shape of w_i, word shape of neighboring words
short word shape of w_i, short word shape of neighboring words
presence of hyphen

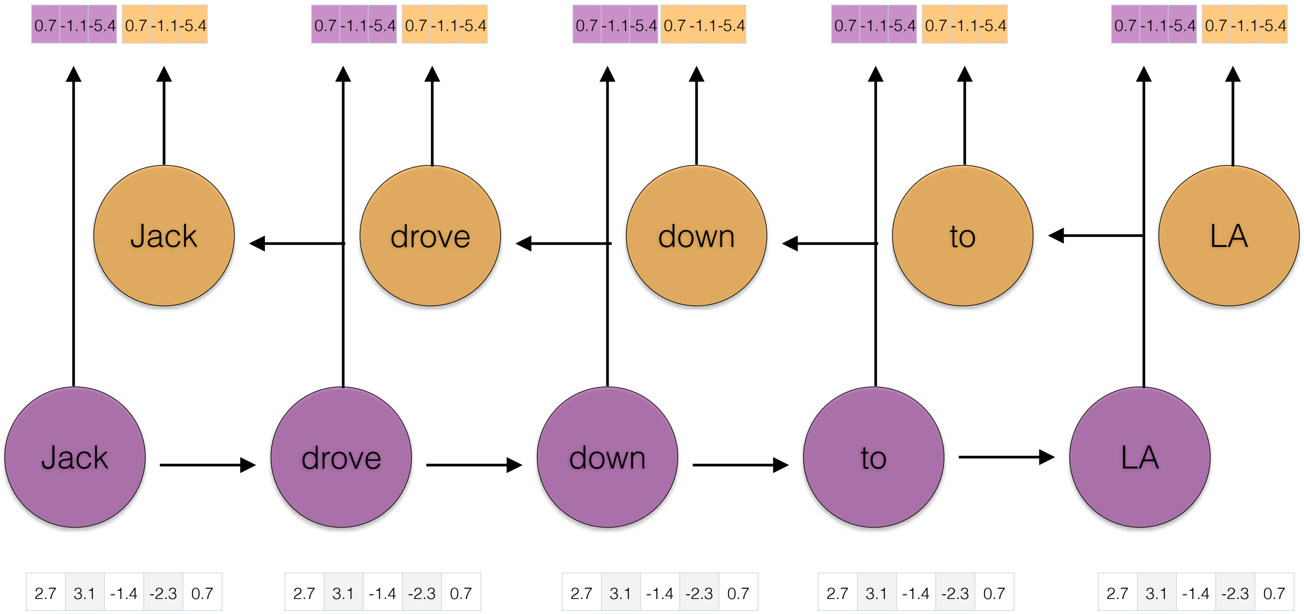
Figure 17.5 Typical features for a feature-based NER system.

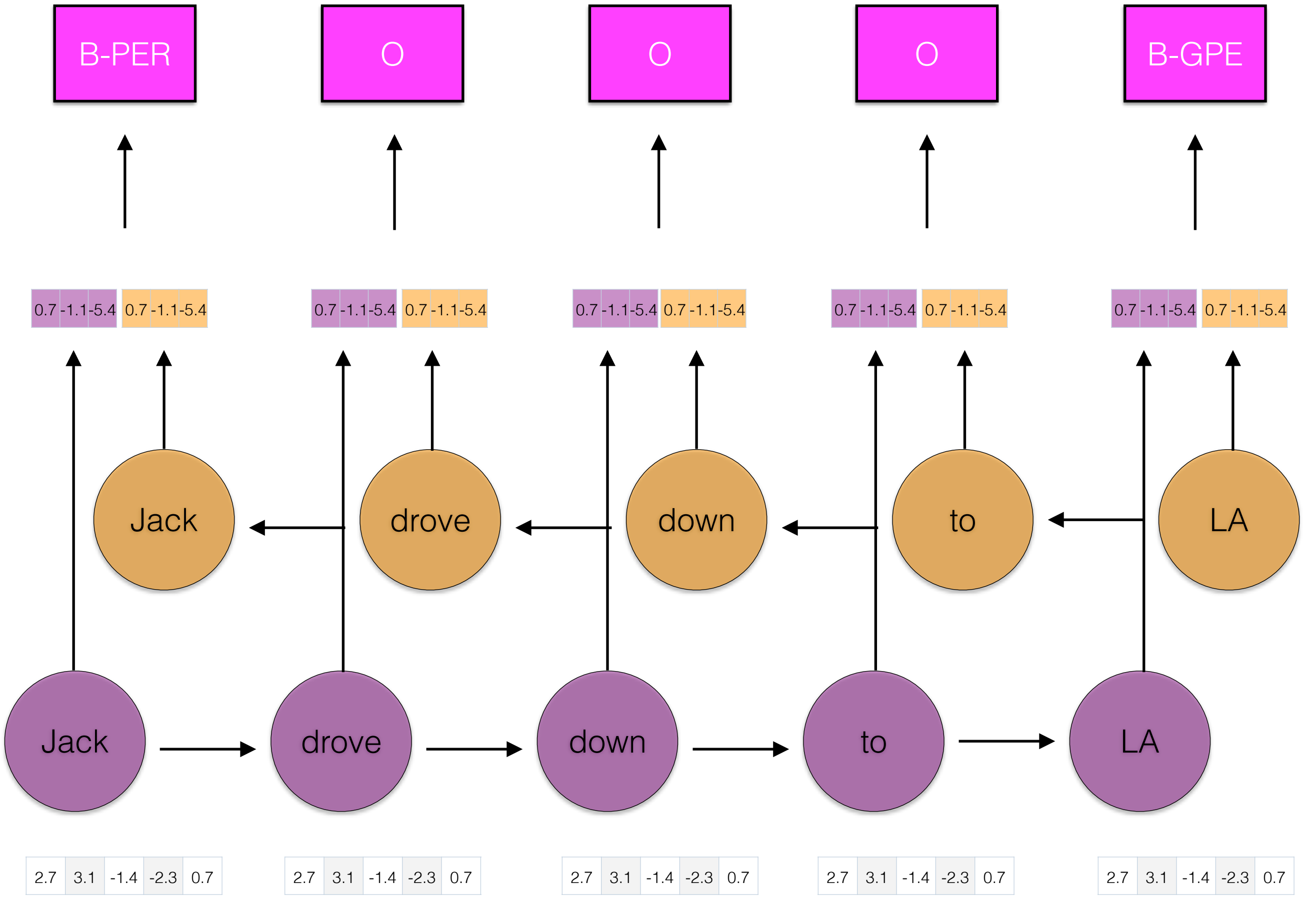
Gazetteers

- List of place names; more generally, list of names of some typed category
- GeoNames (GEO), US SSN (PER), Getty Thesaurus of Geographic Placenames, Getty Thesaurus of Art and Architecture

Cliff
Bun Cranncha
Dromore West
Dromore
Youghal Harbour
Youghal Bay
Youghal
Eochaill
Yellow River
Yellow Furze
Woodville
Wood View
Woodtown House
Woodstown
Woodstock House
Woodsgift House
Woodrooff House
Woodpark
Woodmount
Wood Lodge
Woodlawn Station
Woodlawn
Woodlands Station
Woodhouse
Wood Hill
Woodfort
Woodford River
Woodford
Woodfield House
Woodenbridge Junction Station
Woodenbridge
Woodbrook House
Woodbrook
Woodbine Hill
Wingfield House
Windy Harbour
Windy Gap

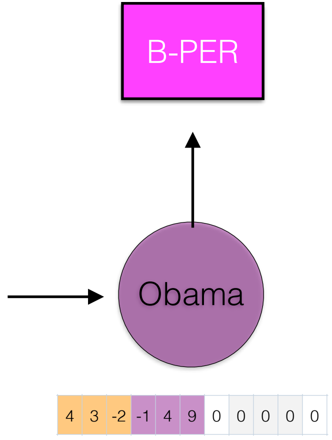
Bidirectional RNN

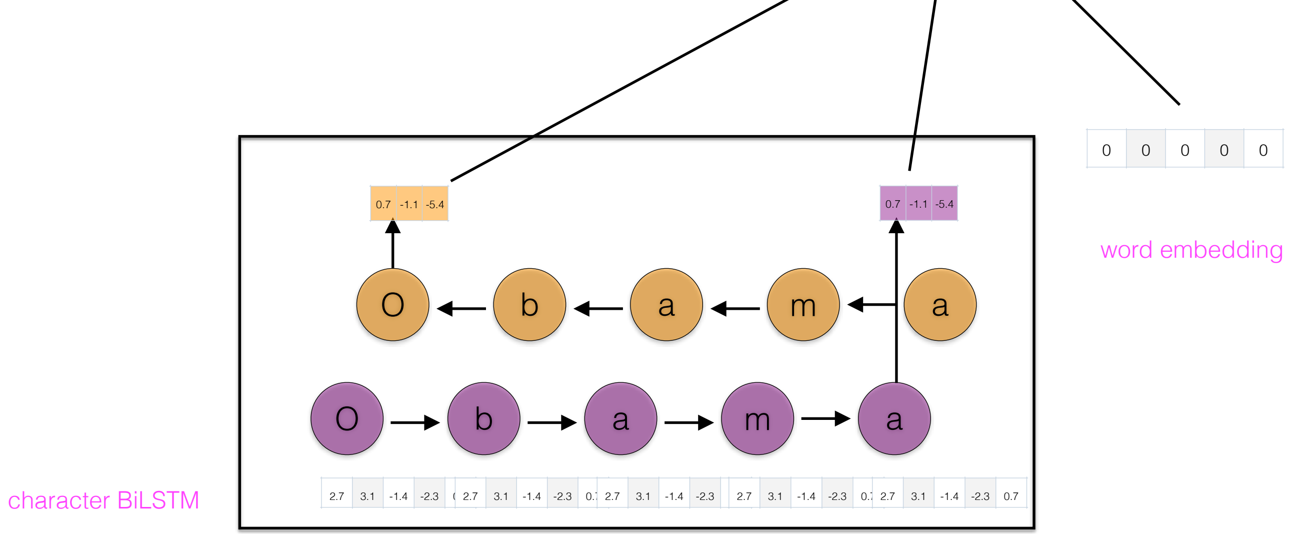



B-PER
O
O
O
B-GPE
0.7 -1.1 -5.4 0.7 -1.1 -5.4
Jack
drove
down
to
LA
Jack
drove
down
to
LA
2.7 3.1 -1.4 -2.3 0.7

BiLSTM for each word; concatenate final state of forward LSTM, backward LSTM, and word embedding as representation for a word.

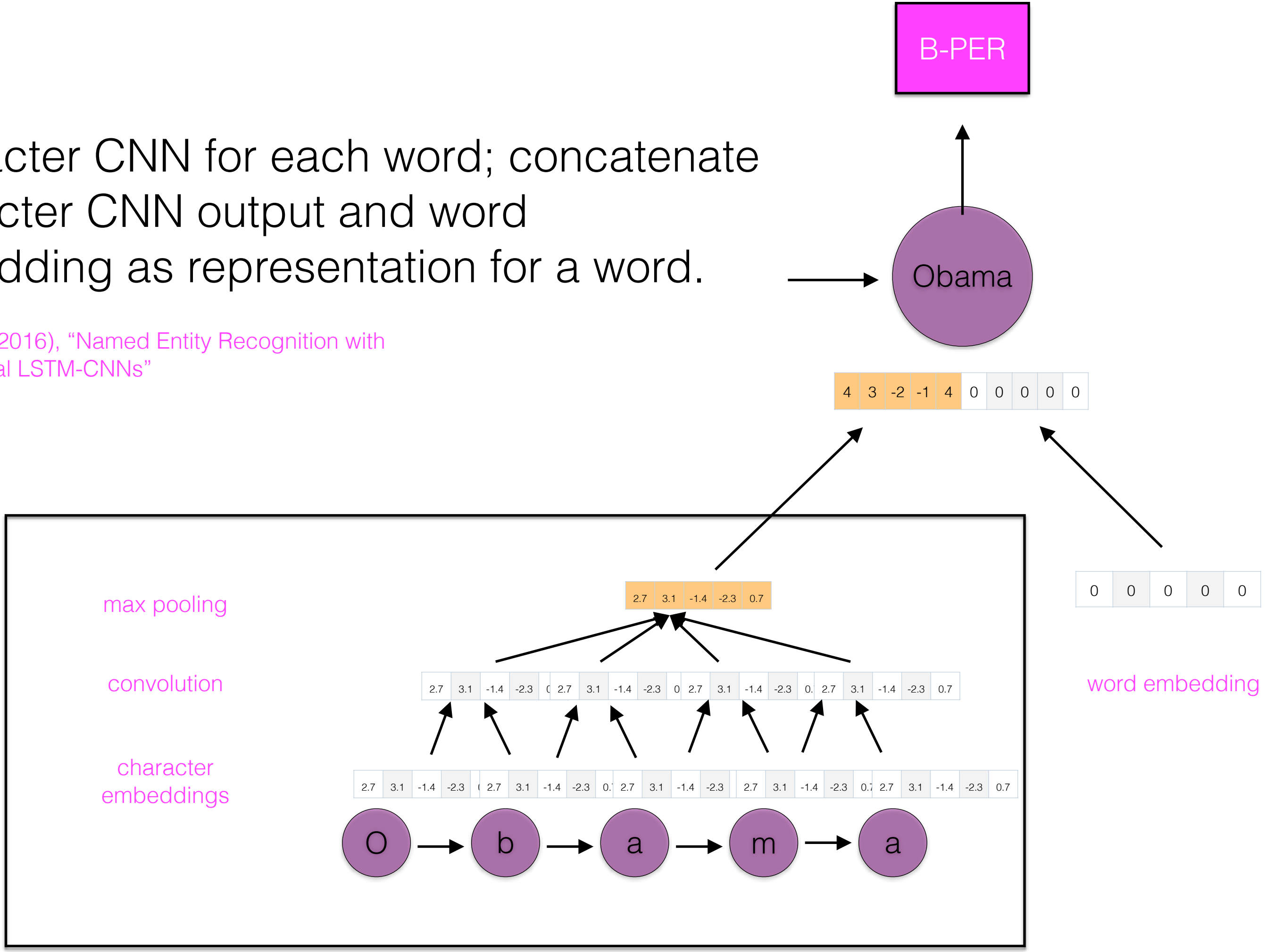
Lample et al. (2016), "Neural Architectures for Named Entity Recognition"

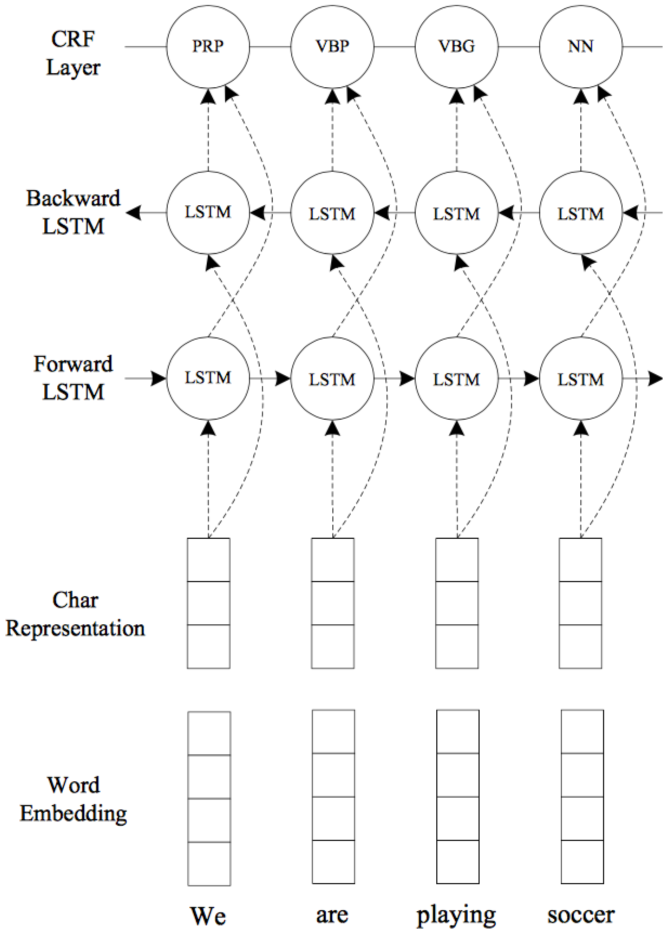




Character CNN for each word; concatenate character CNN output and word embedding as representation for a word.

Chu et al. (2016), “Named Entity Recognition with Bidirectional LSTM-CNNs”





Ma and Hovy (2016), "End-to-end Sequence Labeling via Bi-directional LSTM-CNNs-CRF"

Some transitions (e.g. `O I-PER`) are impossible should never be predicted in NER (or any task that uses BIO notation)

	POS		NER					
	Dev	Test	Dev			Test		
Model	Acc.	Acc.	Prec.	Recall	F1	Prec.	Recall	F1
BRNN	96.56	96.76	92.04	89.13	90.56	87.05	83.88	85.44
BLSTM	96.88	96.93	92.31	90.85	91.57	87.77	86.23	87.00
BLSTM-CNN	97.34	97.33	92.52	93.64	93.07	88.53	90.21	89.36
BRNN-CNN-CRF	97.46	97.55	94.85	94.63	94.74	91.35	91.06	91.21

Ma and Hovy (2016), "End-to-end Sequence Labeling via Bi-directional LSTM-CNNs-CRF"

Evaluation

- We evaluate NER with precision/recall/F1 over typed chunks.

Evaluation

	1	2	3	4	5	6	7
	tim	cook	is	the	CEO	of	Apple
gold	B-PER	I-PER	O	O	O	O	B-ORG
system	B-PER	O	O	O	B-PER	O	B-ORG

<start, end, type>

Precision	1/3
Recall	1/2

gold

<1,2,PER>
<7,7,ORG>

system

<1,1,PER>
<5,5,PER>
<7,7,ORG>

Entity linking

Michael	Jordan	can	dunk	from	the	free	throw	line
B-PER	I-PER							

Entity linking

- Task: Given a database of candidate referents, identify the correct referent for a mention in context.

Text	True wikipedia page
Hornets owner **Michael Jordan** thinks having one or two "superteams" is a detriment to the NBA because the other 28 teams "are going to be garbage."	`wiki/Michael_Jordan`
In 2001, **Michael Jordan** and others resigned from the Editorial Board of *Machine Learning*.	`wiki/Michael_I._Jordan`
The stars are aligning for leading man **Michael Jordan**, who just signed on for a new film, according to Variety.	`wiki/Michael_B._Jordan`
Michael Jordan played in 1,072 regular-season games in his 15-season career	`wiki/Michael_Jordan`

Michael Jordan (disambiguation)

From Wikipedia, the free encyclopedia

Michael Jordan (born 1963) is an American basketball player.

Michael or **Mike Jordan** may also refer to:

People [edit]

Sports [edit]

- Michael Jordan (footballer) (born 1986), English goalkeeper
- Mike Jordan (racing driver) (born 1958), English racing driver
- Mike Jordan (baseball, born 1863) (1863–1940), baseball player
- Mike Jordan (cornerback) (born 1992), American football cornerback
- Michael-Hakim Jordan (born 1977), American professional basketball player
- Michal Jordán (born 1990), Czech ice hockey player

Other people [edit]

- Michael B. Jordan (born 1987), American actor
- Michael Jordan (insolvency baron) (born 1931), English businessman
- Michael Jordan (Irish politician), Irish Farmers' Party TD from Wexford, 1927–1932
- Michael I. Jordan (born 1956), American researcher in machine learning and artificial intelligence
- Michael H. Jordan (1936–2010), American executive for CBS, PepsiCo, Westinghouse
- Michael Jordan (mycologist), English mycologist

Learning to rank

- Entity linking is often cast as a learning to rank problem: given a mention x, some set of candidate entities $\mathcal{Y}$(x) for that mention, and context c, select the highest scoring entity from that set.

$$\hat{y} = \underset{y \in \mathcal{Y}(x)}{\arg\max}\ \Psi(y, x, c)$$

Some scoring function over the mention x, candidate y, and context c

Eisenstein 2018

Learning to rank

- We learn the parameters of the scoring function by minimizing the ranking loss

$$\ell(\hat{y}, y, x, c) = \max\left(0, \Psi(\hat{y}, x, c) - \Psi(y, x, c) + \ldots\right)$$

Eisenstein 2018

Learning to rank

Some scoring function over the mention x, candidate y, and context c

$$\Psi(y, x, c)$$

feature = f(x,y,c)
string similarity between x and y
popularity of y
NER type(x) = type(y)
cosine similarity between c and Wikipedia page for y

$$\Psi(y, x, c) = f(x, y, c)^{\top}\beta$$

Neural learning to rank

Parameters measuring the compatibility of the candidate and mention

Parameters measuring the compatibility of the candidate and context

$$\Psi(y, x, c) = v_y^\top \Theta^{(x,y)} x + v_y^\top \Theta^{(y,c)} c$$

Embedding for candidate

Embedding for mention

Embedding for context

Learning to rank

- We learn the parameters of the scoring function by minimizing the ranking loss; take the derivative of the loss and backprop using SGD.

$$\ell(\hat{y}, y, x, c) = \max\left(0, \Psi(\hat{y}, x, c) - \Psi(y, x, c) + 1\right)$$

Eisenstein 2018

Entity linking

- Task: Given a database of candidate referents, identify the correct referent for a mention in context.

Text	True wikipedia page
Hornets owner **Michael Jordan** thinks having one or two "superteams" is a detriment to the NBA because the other 28 teams "are going to be garbage."	`wiki/Michael_Jordan`
In 2001, **Michael Jordan** and others resigned from the Editorial Board of *Machine Learning*.	`wiki/Michael_I._Jordan`
The stars are aligning for leading man **Michael Jordan**, who just signed on for a new film, according to Variety.	`wiki/Michael_B._Jordan`
Michael Jordan played in 1,072 regular-season games in his 15-season career	`wiki/Michael_Jordan`

Relation extraction

The Big Sleep is a 1946 film noir directed by Howard Hawks,[2][3] the first film version of Raymond Chandler's 1939 novel of the same name. The film stars Humphrey Bogart as private detective Philip Marlowe and Lauren Bacall as Vivian Rutledge in a story about the "process of a criminal investigation, not its results."[4] William Faulkner, Leigh Brackett and Jules Furthman co-wrote the screenplay.

subject	*predicate*	*object*
The Big Sleep	directed_by	Howard Hawks
The Big Sleep	stars	Humphrey Bogart
The Big Sleep	stars	Lauren Bacall
The Big Sleep	screenplay_by	William Faulkner
The Big Sleep	screenplay_by	Leigh Brackett
The Big Sleep	screenplay_by	Jules Furthman

Relation extraction

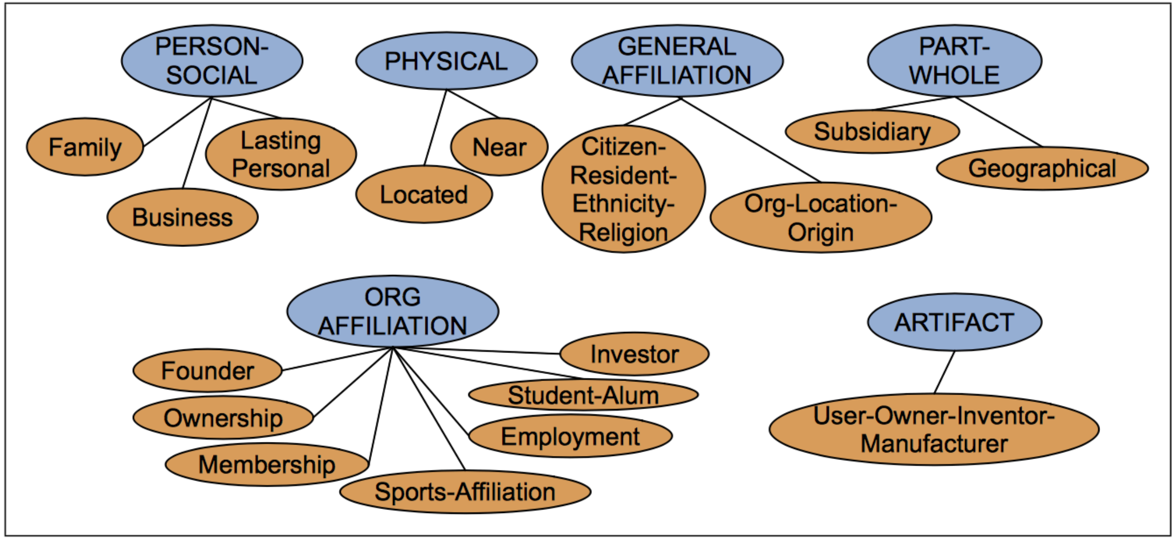


Figure 17.9 The 17 relations used in the ACE relation extraction task.

ACE relations, SLP3

Relation extraction

Entity	Relation	Entity
Injury	disrupts	Physiological Function
Bodily Location	location-of	Biologic Function
Anatomical Structure	part-of	Organism
Pharmacologic Substance	causes	Pathological Function
Pharmacologic Substance	treats	Pathologic Function

Unified Medical Language System (UMLS), SLP3

Wikipedia Infoboxes

The Big Sleep is a 1946 film noir directed by Howard Hawks,[2][3] the first film version of Raymond Chandler's 1939 novel of the same name. The film stars Humphrey Bogart as private detective Philip Marlowe and Lauren Bacall as Vivian Rutledge in a story about the "process of a criminal investigation, not its results."[4] William Faulkner, Leigh Brackett and Jules Furthman co-wrote the screenplay.

A remake starring Robert Mitchum as Philip Marlowe was released in 1978. This was the second film in three years featuring Mitchum as Marlowe. The remake was arguably more faithful to the novel, possibly due to fewer restrictions in 1978 on what could be portrayed on screen, however, it was far less successful than the original 1946 version. In 1997, the U.S. Library of Congress deemed the film "culturally, historically, or aesthetically significant," and added it to the National Film Registry.

The Big Sleep

Theatrical release lobby card

Directed by	Howard Hawks
Produced by	Howard Hawks
Screenplay by	William Faulkner Leigh Brackett Jules Furthman
Based on	*The Big Sleep* by Raymond Chandler
Starring	Humphrey Bogart Lauren Bacall
Music by	Max Steiner
Cinematography	Sidney Hickox
Edited by	Christian Nyby
Distributed by	Warner Bros.
Release date	August 23, 1946 (United States)
Running time	114 minutes (released cut) 116 minutes (re-released original cut)

Regular expressions

- Regular expressions are precise ways of extracting high-precision relations
    - "NP_1 is a film directed by NP_2" → directed_by(NP_1, NP_2)
    - "NP_1 was the director of NP_2"→ directed_by(NP_2, NP_1)

Hearst patterns

pattern	*sentence*
NP {, NP}* {,} (and|or) other NP_H	temples, treasuries, and other important civic buildings
NP_H such as {NP,}* {(or|and)} NP	red algae such as Gelidium
such NP_H as {NP,}* {(or|and)} NP	such authors as Herrick, Goldsmith, and Shakespeare
NP_H {,} including {NP,}* {(or|and)} NP	common-law countries, including Canada and England
NP_H {,} especially {NP}* {(or|and)} NP	European countries, especially France, England, and Spain

Hearst 1992; SLP3

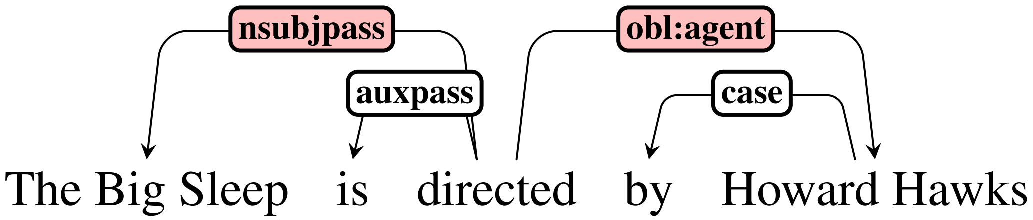
Supervised relation extraction

[The Big Sleep]$_{m1}$ is a 1946 film noir directed by [Howard Hawks]$_{m2}$, the first film version of Raymond Chandler's 1939 novel of the same name.

feature(m1, m2)
headwords of m1, m2
bag of words in m1, m2
bag of words between m1, m2
named entity types of m1, m2
syntactic path between m1, m2

Supervised relation extraction

[The Big Sleep]$_{m1}$ is a 1946 film noir directed by [Howard Hawks]$_{m2}$, the first film version of Raymond Chandler's 1939 novel of the same name.



[The Big Sleep]$_{m1}$ ←*nsubjpass* directed→*obl:agent* [Howard Hawks]$_{m2}$,

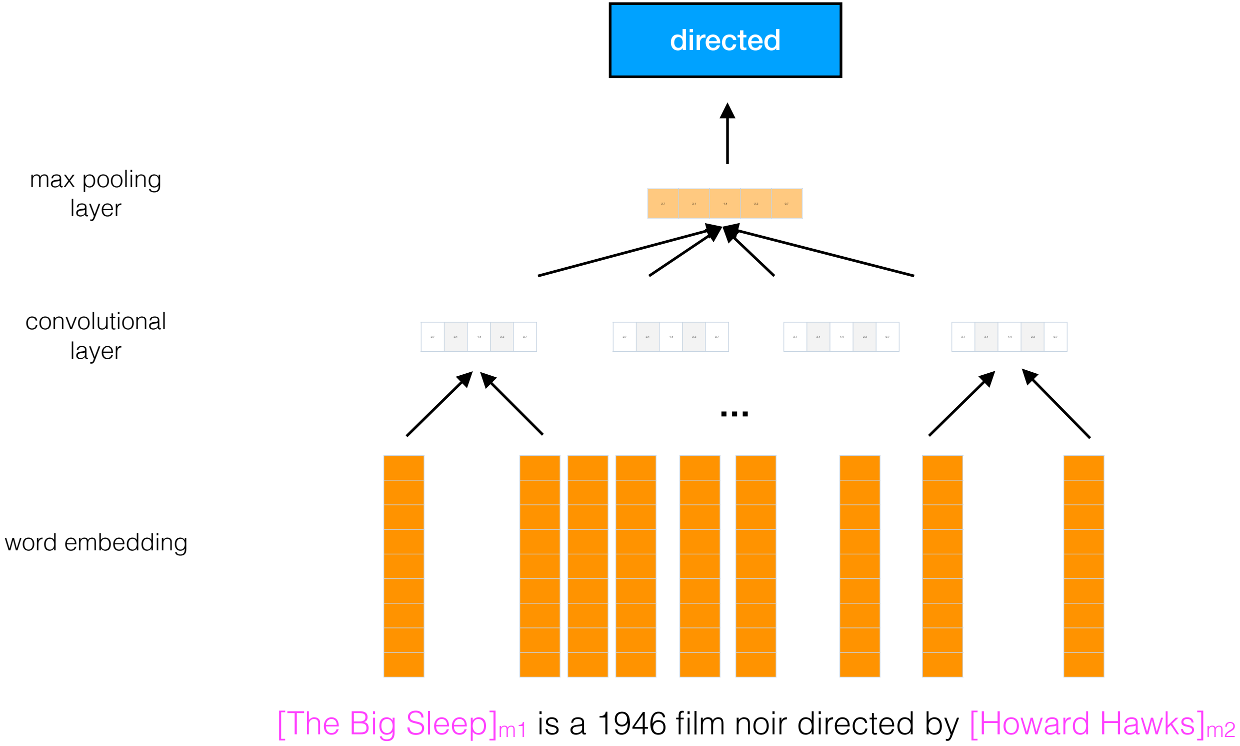
m1←*nsubjpass* ← directed→*obl:agent* → m2

Supervised relation extraction

```
function FINDRELATIONS(words) returns relations

    relations ← nil
    entities ← FINDENTITIES(words)
    forall  entity pairs ⟨e1, e2⟩ in entities do
        if RELATED?(e1, e2)
            relations ← relations+CLASSIFYRELATION(e1, e2)
```

Figure 17.13 Finding and classifying the relations among entities in a text.



directed
max pooling layer
convolutional layer
...
word embedding
[The Big Sleep]m1 is a 1946 film noir directed by [Howard Hawks]m2

Neural RE

- To encode relative position to entities, we'll add positional embeddings to our representation of each word — the distance from each word w in the sentence to m1 and m2

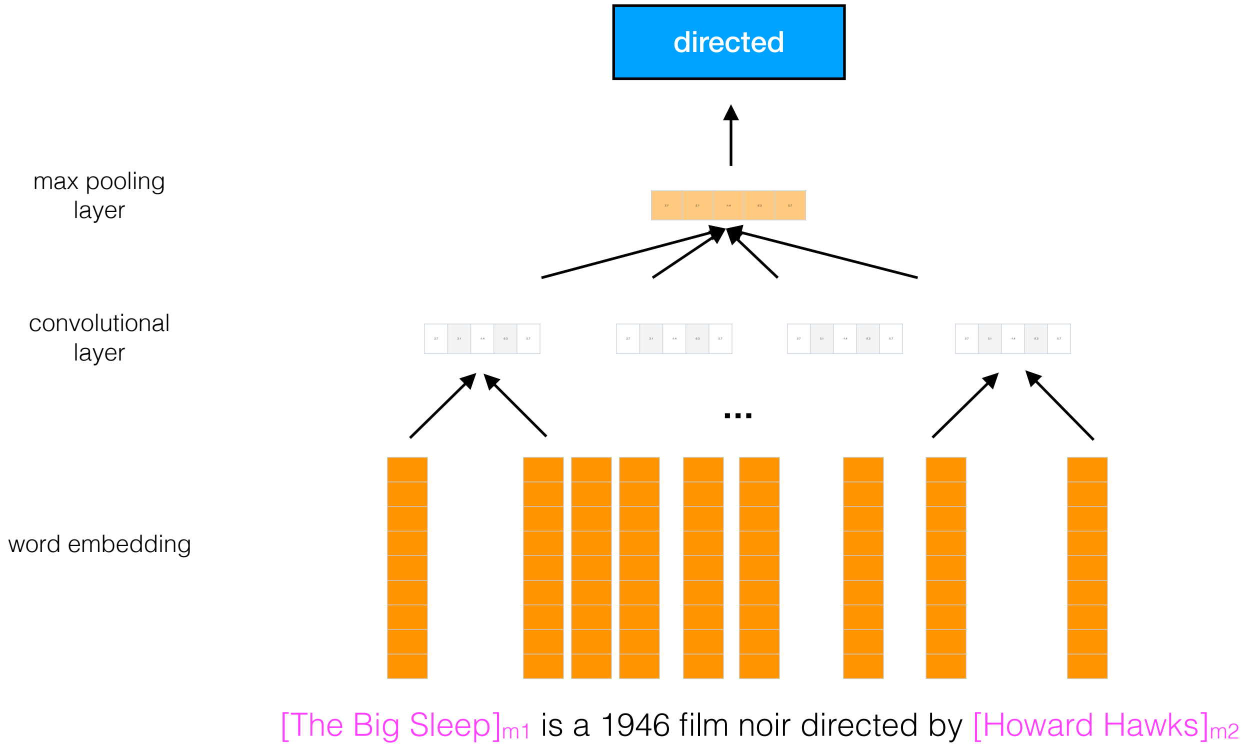
dist from m1	0	1	2	3	4	5	6	7	8
dist from m2	-8	-7	-6	-5	-4	-3	-2	-1	0
	[The Big Sleep]	is	a	1946	film	noir	directed	by	[Howard Hawks]

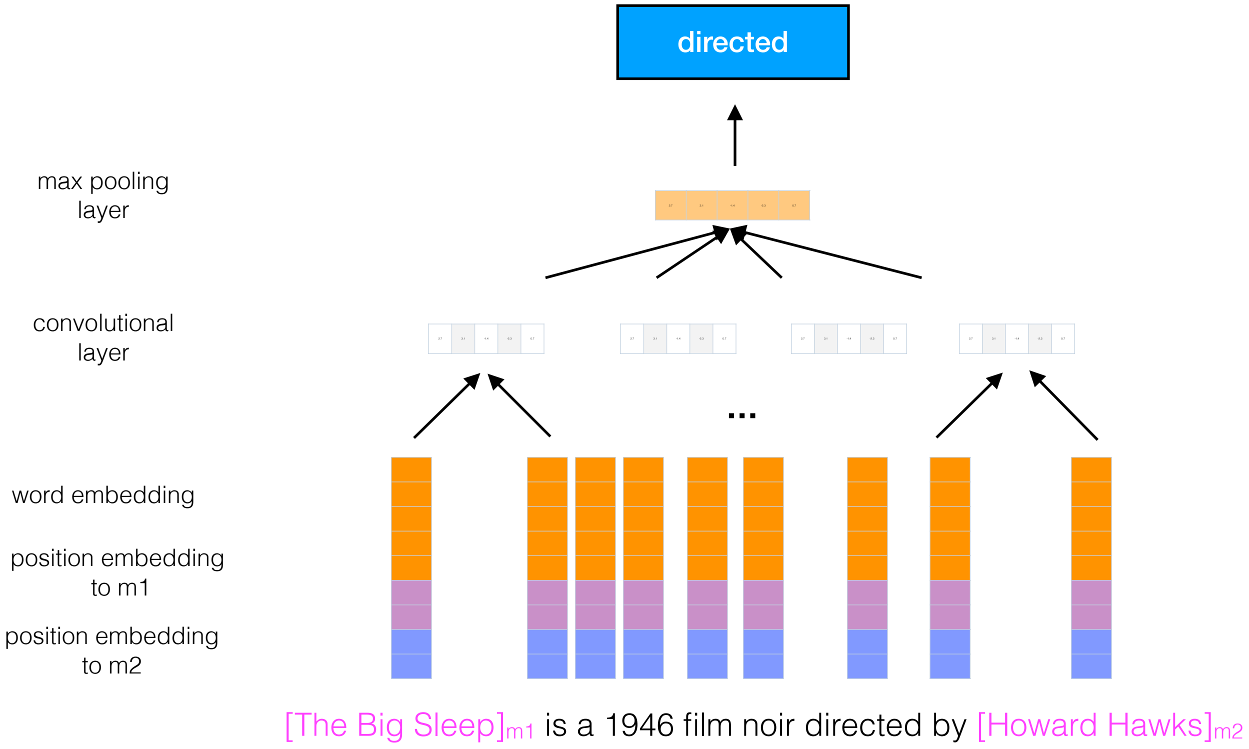
- 0 here uniquely identifies the head and tail of the relation; other position indicate how close the word is (maybe closer words matter more)

Neural RE

Each position then has an embedding

-4	2	-0.5	1.1	0.3	0.4	-0.5
-3	-1.4	0.4	-0.2	-0.9	0.5	0.9
-2	-1.1	-0.2	-0.5	0.2	-0.8	0
-1	0.7	-0.3	1.5	-0.3	-0.4	0.1
0	-0.8	1.2	1	-0.7	-1	-0.4
1	0	0.3	-0.3	-0.9	0.2	1.4
2	0.8	0.8	-0.4	-1.4	1.2	-0.9
3	1.6	0.4	-1.1	0.7	0.1	1.6
4	1.2	-0.2	1.3	-0.4	0.3	-1.0


directed
max pooling layer
convolutional layer
...
word embedding
[The Big Sleep]m1 is a 1946 film noir directed by [Howard Hawks]m2



directed
max pooling layer
convolutional layer
...
word embedding
position embedding to m1
position embedding to m2
[The Big Sleep]m1 is a 1946 film noir directed by [Howard Hawks]m2

Distant supervision

- It’s uncommon to have labeled data in the form of <sentence, relation> pairs

sentence	*relations*
[The Big Sleep]$_{m1}$ is a 1946 film noir directed by [Howard Hawks]$_{m2}$, the first film version of Raymond Chandler's 1939 novel of the same name.	directed_by(The Big Sleep, Howard Hawks)

Distant supervision

- More common to have knowledge base data about entities and their relations that's separate from text.
- We know the text likely expresses the relations somewhere, but not exactly where.

Wikipedia Infoboxes

The Big Sleep is a 1946 film noir directed by Howard Hawks,[2][3] the first film version of Raymond Chandler's 1939 novel of the same name. The film stars Humphrey Bogart as private detective Philip Marlowe and Lauren Bacall as Vivian Rutledge in a story about the "process of a criminal investigation, not its results."[4] William Faulkner, Leigh Brackett and Jules Furthman co-wrote the screenplay.

A remake starring Robert Mitchum as Philip Marlowe was released in 1978. This was the second film in three years featuring Mitchum as Marlowe. The remake was arguably more faithful to the novel, possibly due to fewer restrictions in 1978 on what could be portrayed on screen, however, it was far less successful than the original 1946 version. In 1997, the U.S. Library of Congress deemed the film "culturally, historically, or aesthetically significant," and added it to the National Film Registry.

The Big Sleep

Theatrical release lobby card

Directed by	Howard Hawks
Produced by	Howard Hawks
Screenplay by	William Faulkner Leigh Brackett Jules Furthman
Based on	*The Big Sleep* by Raymond Chandler
Starring	Humphrey Bogart Lauren Bacall
Music by	Max Steiner
Cinematography	Sidney Hickox
Edited by	Christian Nyby
Distributed by	Warner Bros.
Release date	August 23, 1946 (United States)
Running time	114 minutes (released cut) 116 minutes (re-released original cut)

Relation name	Size	Example
/people/person/nationality	281,107	John Dugard, South Africa
/location/location/contains	253,223	Belgium, Nijlen
/people/person/profession	208,888	Dusa McDuff, Mathematician
/people/person/place_of_birth	105,799	Edwin Hubble, Marshfield
/dining/restaurant/cuisine	86,213	MacAyo's Mexican Kitchen, Mexican
/business/business_chain/location	66,529	Apple Inc., Apple Inc., South Park, NC
/biology/organism_classification_rank	42,806	Scorpaeniformes, Order
/film/film/genre	40,658	Where the Sidewalk Ends, Film noir
/film/film/language	31,103	Enter the Phoenix, Cantonese
/biology/organism_higher_classification	30,052	Calopteryx, Calopterygidae
/film/film/country	27,217	Turtle Diary, United States
/film/writer/film	23,856	Irving Shulman, Rebel Without a Cause
/film/director/film	23,539	Michael Mann, Collateral
/film/producer/film	22,079	Diane Eskenazi, Aladdin
/people/deceased_person/place_of_death	18,814	John W. Kern, Asheville
/music/artist/origin	18,619	The Octopus Project, Austin
/people/person/religion	17,582	Joseph Chartrand, Catholicism
/book/author/works_written	17,278	Paul Auster, Travels in the Scriptorium
/soccer/football_position/players	17,244	Midfielder, Chen Tao
/people/deceased_person/cause_of_death	16,709	Richard Daintree, Tuberculosis
/book/book/genre	16,431	Pony Soldiers, Science fiction
/film/film/music	14,070	Stavisky, Stephen Sondheim
/business/company/industry	13,805	ATS Medical, Health care

Table 2: The 23 largest Freebase relations we use, with their size and an instance of each relation.

Mintz et al. 2009

Distant supervision

mayor(Maynard Jackson, Atlanta)

Elected mayor of Atlanta in 1973, Maynard Jackson…
Atlanta's airport will be renamed to honor Maynard Jackson, the city's first Black mayor
Born in Dallas, Texas in 1938, Maynard Holbrook Jackson, Jr. moved to Atlanta when he was 8.

mayor(Fiorello LaGuardia, New York)

Fiorello LaGuardia was Mayor of New York for three terms...
Fiorello LaGuardia, then serving on the New York City Board of Aldermen...

Eisenstein 2018

Distant supervision

- For feature-based models, we can represent the tuple <m1, m2> by aggregating together the representations from all the sentences they appear in

Distant supervision

[The Big Sleep]$_{m1}$ is a 1946 film noir directed by [Howard Hawks]$_{m2}$, the first film version of Raymond Chandler's 1939 novel of the same name.

[Howard Hawks]$_{m2}$ directed the [The Big Sleep]$_{m1}$

feature(m1, m2)	value (e.g., normalized over all sentences)
“directed” between m1, m2	0.37
“by” between m1, m2	0.42
m1←*nsubjpass* ← directed→*obl:agent* → m2	0.13
m2←*nsubj* ← directed→*obj* → m2	0.08

Distant supervision

- Discovering Hearst patterns from distant supervision using WordNet (Snow et al. 2005)

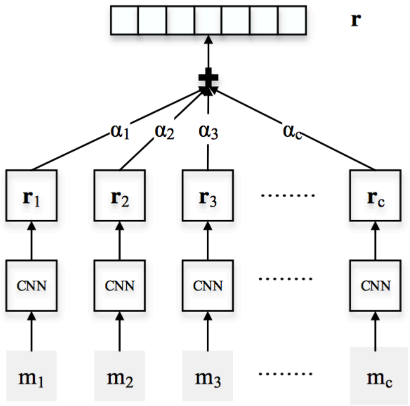
pattern	*sentence*
NP_H like NP	Many hormones like leptin...
NP_H called NP	a markup language called XHTML
NP is a NP_H	Ruby is a programming language...
NP, a NP_H	IBM, a company with a long...

Multiple Instance Learning

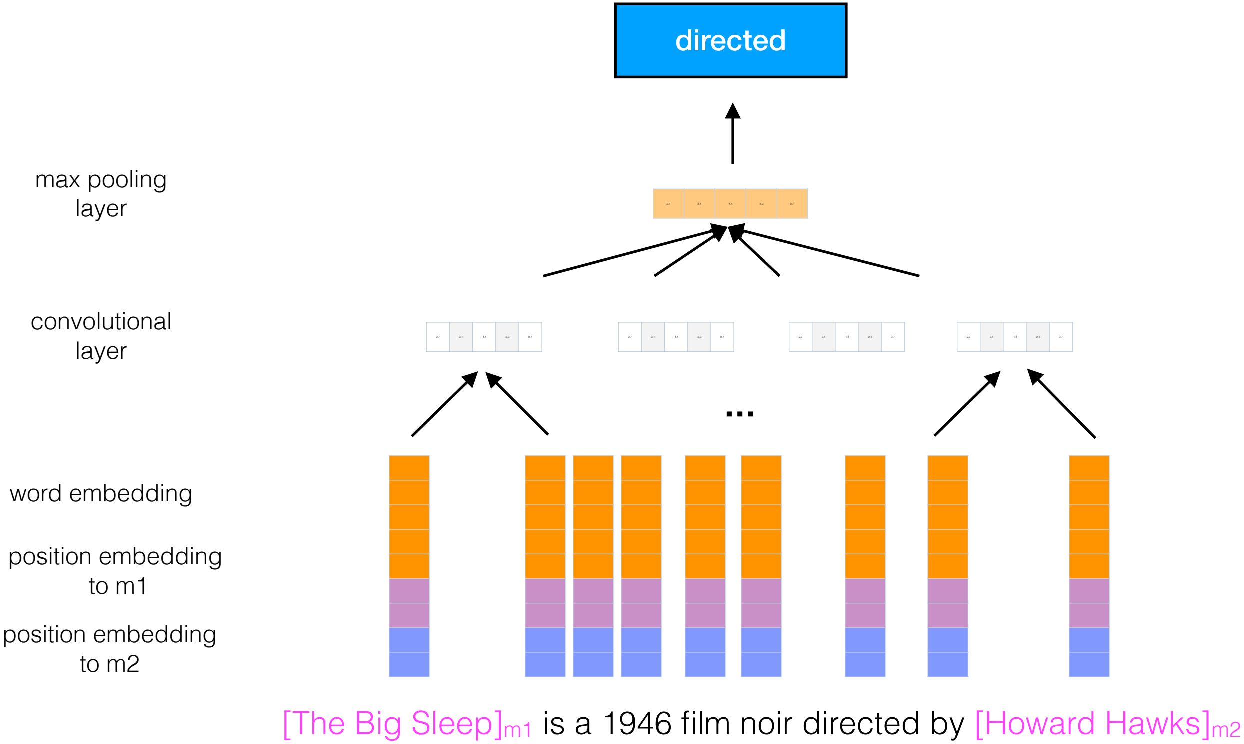
- Labels are assigned to a set of sentences, each containing the pair of entities m1 and m2; not all of those sentences express the relation between m1 and m2.

Attention

- Let's incorporate structure (and parameters) into a network that captures which sentences in the input we should be attending to (and which we can ignore).

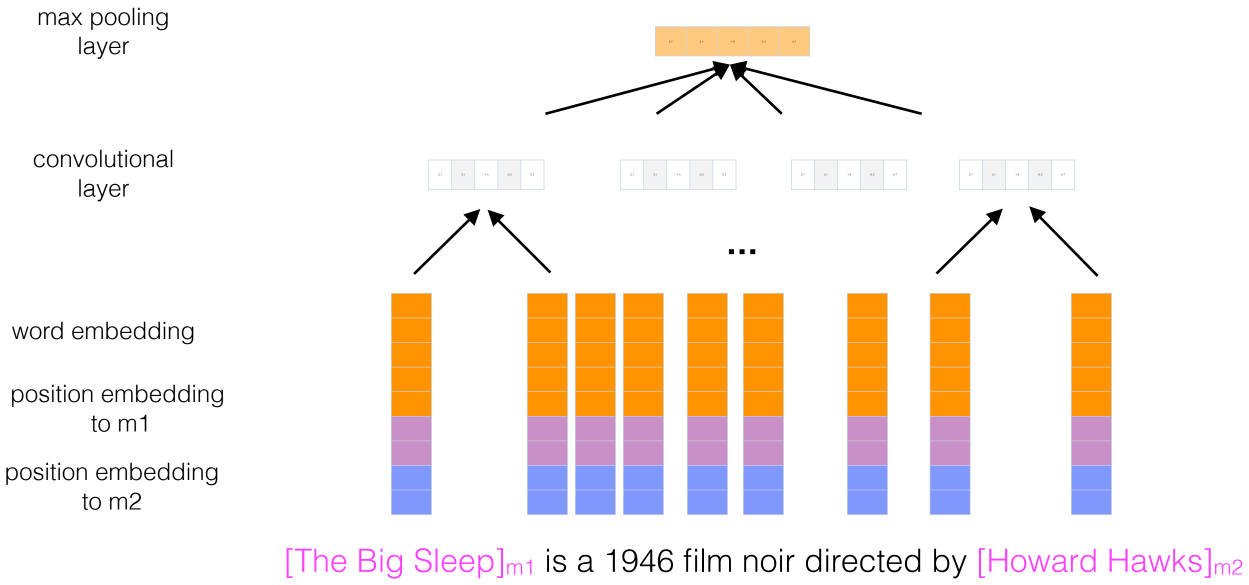


Lin et al (2016), "Neural Relation Extraction with Selective Attention over Instances" (ACL)

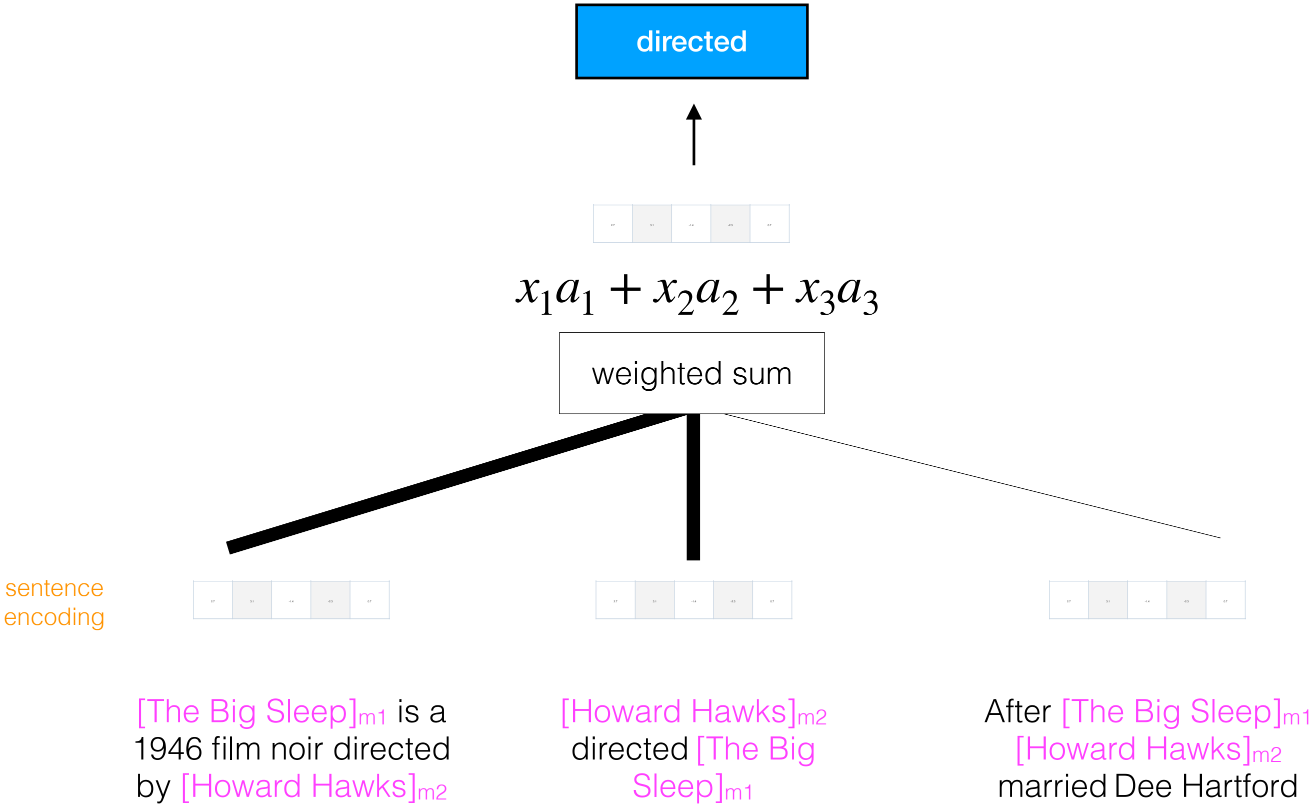


Lin et al (2016), “Neural Relation Extraction with Selective Attention over Instances” (ACL)

Now we just have an encoding of a sentence



Lin et al (2016), “Neural Relation Extraction with Selective Attention over Instances” (ACL)



directed
$x_1a_1 + x_2a_2 + x_3a_3$
weighted sum
sentence encoding
[The Big Sleep]m1 is a 1946 film noir directed by [Howard Hawks]m2
[Howard Hawks]m2 directed [The Big Sleep]m1
After [The Big Sleep]m1 [Howard Hawks]m2 married Dee Hartford

Information Extraction

- Named entity recognition
- Entity linking
- Relation extraction
- Template filling
- Event detection
- Event coreference
- …