

# Use Generalized Representations, But Do Not Forget Surface Features!

**Nafise Sadat Moosavi and Michael Strube**

**Heidelberg Institute for Theoretical Studies**

**Heidelberg, Germany**

# Coreference Resolution



Freddie Mac stock is down another 6% today.  
It has lost more than 20% since Monday.  
Monday, of course, is when the company fired  
its president for not cooperating with efforts  
to restate earnings for the last few years.  
Freddie Mac's CEO forced out the same day.



# Mention Space Pruning



“The world’s fifth **Disney** park will soon open to the public here. The most important thing about **Disney** is that **it** is a global brand.”

fifth

**Disney**

The world ’s fifth Disney park

The world ’s

the public

**Disney**

The most important thing about Disney

**it**

a global brand

# Mention Types



Non-referential mentions

Coreferent mentions

Discourse-old mentions

# Non-Referential Mentions



“It is raining.”

“Vatican City is actually a country.”

# Coreferent Mentions



“The world’s fifth **Disney** park will soon open to the public here. The most important thing about **Disney** is that **it** is a global brand.”

Mentions that appear in a coreference chain

# Discourse-old Mentions



“The world’s fifth **Disney** park will soon open to the public here. The most important thing about **Disney** is that **it** is a global brand.”

## Discourse-new

Refers to an entity that has not occurred before

## Discourse-old

Refers to an entity that is already evoked

# Discourse-old Mentions



Discourse-new

Discourse-old

Anaphoricity determination



# Anaphoricity Determination



Neural models

SVM

# Neural Models

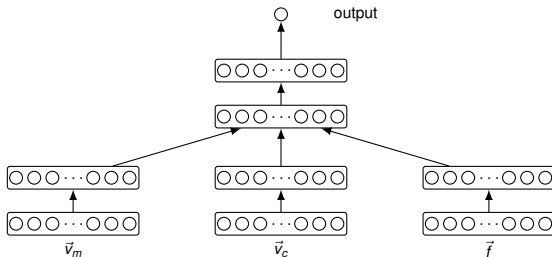


deep-coref anaphoricity determination by Clark and Manning (2016)

$\vec{v}_m$ : embeddings of head, first, last, two previous/following words

$\vec{v}_c$ : averaged embeddings of 5 previous/following words, all mention words, sentence words, document words

$\vec{f}$ : type, length, position and nested



# Neural Models



## Bidirectional LSTM

### Generalized mention representation

“The world’s fifth **Disney** park will soon open to the public here. The most important thing about Disney is that it is a global brand.”

the world’s fifth **MENTION** park will soon open  
to the public here . the **## disney ## disney**

# Neural Models



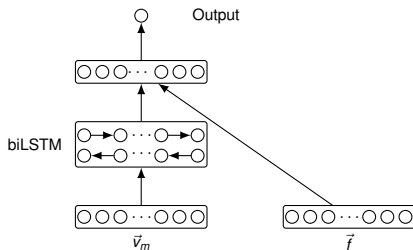
## Bidirectional LSTM

Generalized mention representation

type, string match, previous string match

head match, previous head match

tokens contained, previous tokens contained, nested



# Neural Models



	Non-Anaphoric			Anaphoric		
	R	P	F1	R	P	F1
CoNLL English test set						
joint	-	-	-	81.81	77.18	79.43
LSTM	90.71	92.64	91.66	85.00	81.48	83.20
LSTM—surface	90.51	87.31	88.88	72.64	78.64	75.52

# Simple SVM Model



SVM achieves state-of-the-art results for coreferent mention detection [Moosavi and Strube, NAACL 2016]

lemmas and POS tags of mention words, 2 previous/following words

mention string, length, type

string match, head match, tokens contained

previous string match, head match, and tokens contained

# Simple SVM Model



	Non-Anaphoric			Anaphoric		
	R	P	F1	R	P	F1
CoNLL English test set						
joint	-	-	-	81.81	77.18	79.43
LSTM	90.71	92.64	91.66	85.00	81.48	83.20
SVM	92.42	92.61	92.51	84.66	84.30	84.48

# Simple SVM Model



	Non-Anaphoric			Anaphoric		
	R	P	F1	R	P	F1
CoNLL English test set						
joint	-	-	-	81.81	77.18	79.43
LSTM	90.71	92.64	91.66	85.00	81.48	83.20
SVM	92.42	92.61	92.51	84.66	84.30	84.48

How do they perform on a new domain?



# Out-of-Domain Evaluation



	Non-Anaphoric			Anaphoric		
	R	P	F1	R	P	F1
	WikiCoref					
LSTM	93.25	92.78	93.01	79.41	80.57	79.99
SVM	93.83	93.05	93.43	80.11	82.07	81.08

# Analysis



Anaphoric						
	R	P	F1	R	P	F1
	<b>Proper names</b>			<b>Common nouns</b>		
LSTM	79.49	82.31	80.88	62.96	65.04	63.99
LSTM—surface	47.60	70.09	56.69	46.30	57.75	51.40
SVM	83.80	85.71	84.74	52.46	71.98	60.69
	<b>Pronouns</b>			Other		
LSTM	94.67	85.60	89.91	29.11	63.88	40.00
LSTM—surface	92.67	86.01	89.22	10.13	34.78	15.69
SVM	95.59	86.29	90.71	32.91	76.47	46.02

**Proper names:** SVM with surface features

**Common nouns:** LSTM

**Pronouns:** not much affected by surface features

# Take-home Message



- exact match
- head match
- tokens contained
- head contained
- alias
- gender
- number
- refined mention type
- lexical features
- combinatorial features

- **word embeddings**

- exact match - head contained



# Take-home Message



- ✓ Replacing exact lexical forms with word embeddings
- × Disregarding traditional features including surface features

# Take-home Message



- ✓ Replacing exact lexical forms with word embeddings

“Lexical Features in Coreference Resolution: To be Used With Caution”

- × Disregarding traditional features including surface features

# Thank You!



Don't throw the baby out with the bathwater!

