What's Hot in Human Language Technology: Highlights from NAACL HLT 2015

Joyce Y. Chai, *Michigan State University*

Anoop Sarkar, *Simon Fraser University*

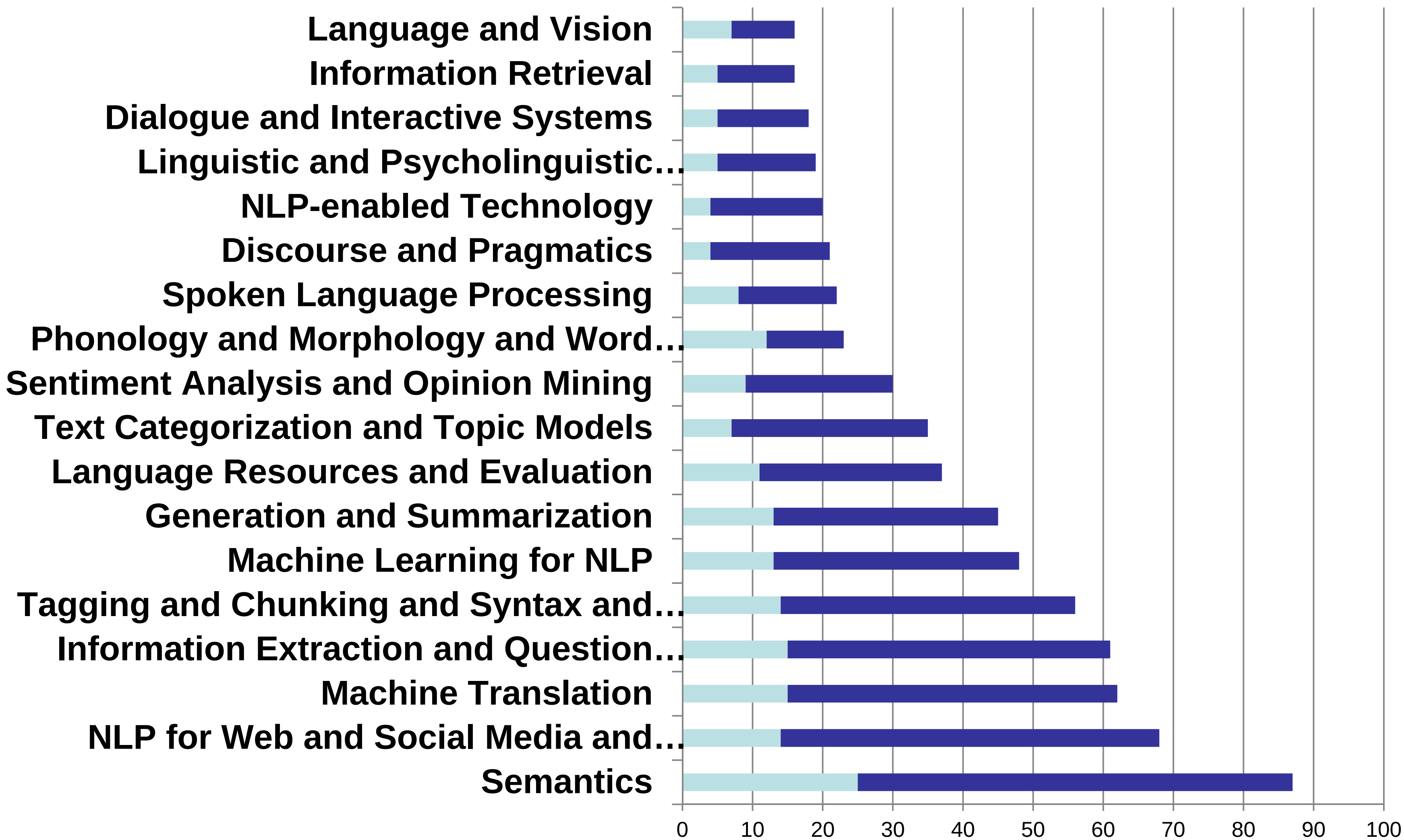
Rada Mihalcea, *University of Michigan*

NAACL HLT 2015 May 31 – June 5 Denver, Colorado, USA

A Quick Glance

- 14th Conference of the North American chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT)
- New records in NAACL history
  - Received 714 submissions (overall 26% acceptance rate)
  - Closed to 1000 registered participants
- Eighteen topic areas

Topic Areas


Language and Vision
Information Retrieval
Dialogue and Interactive Systems
Linguistic and Psycholinguistic…
NLP-enabled Technology
Discourse and Pragmatics
Spoken Language Processing
Phonology and Morphology and Word…
Sentiment Analysis and Opinion Mining
Text Categorization and Topic Models
Language Resources and Evaluation
Generation and Summarization
Machine Learning for NLP
Tagging and Chunking and Syntax and…
Information Extraction and Question…
Machine Translation
NLP for Web and Social Media and…
Semantics
0
10
20
30
40
50
60
70
80
90
100

EA1: Language and Vision

- Invited speaker: Fei-Fei Li, Stanford University
  – Title: *A Quest for Visual Intelligence in Computers*
- Main problems:
  – Alignment between language instructions and video segments (Malmuaud et al., 2015; Naim et al., 2015)
  – Video/Image to text generation (Ortiz et al., 2015; Venugopalan et al., 2015).

Discriminative Unsupervised Alignment of Natural Language Instructions with Corresponding Video Segments

Iftekhar Naim, Young Chol Song, Qiguang Liu, Liang Huang, Henry Kautz, Jiebo Luo, and Daniel Gildea

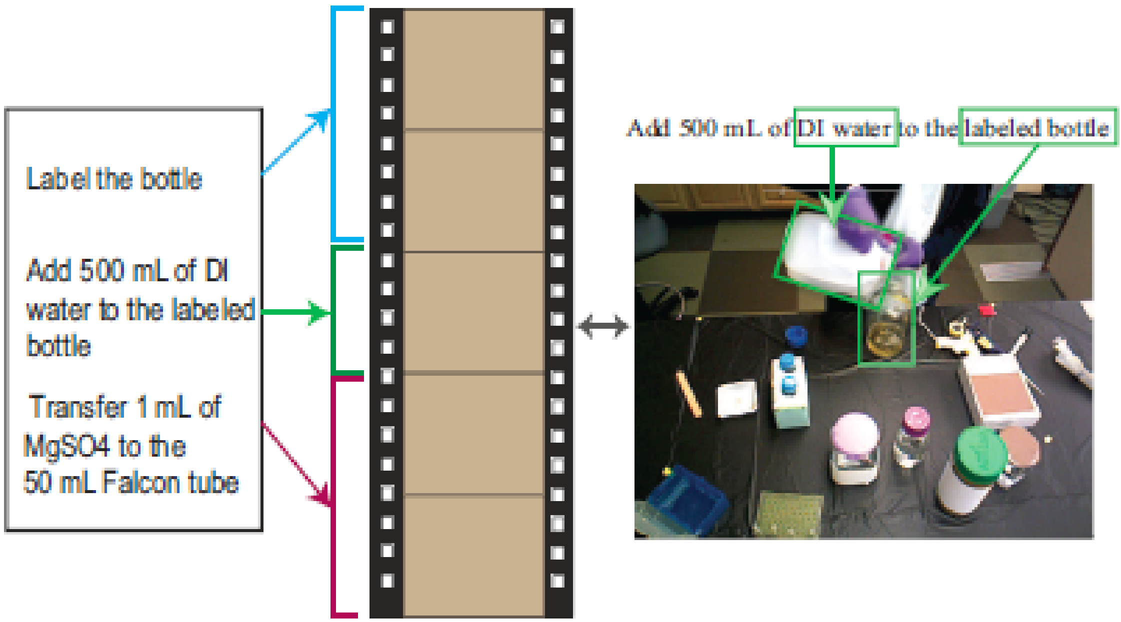
NAACL, 2015

UNIVERSITY of ROCHESTER

Credit: I. Naim

Unsupervised Alignment

- *Unsupervised Grounded Language Learning*
  - No manual annotation
  - Intuition: *Objects and actions in videos usually appear in the same sequential order as their text mentions.*



Latent alignment: (1) Sentences to Video and (2) Nouns to Blobs

Credit: I. Naim

Unsupervised Alignment Models

- Generative Alignment (Naim et al. AAAI [2014])
  - Hidden Markov Model (HMM) + IBM Model 1
- Discriminative Alignment
  - Latent CRF, Latent Perceptron, and Latent SVM

System	Alignment Accuracy (%)	
	Manual Tracking	Auto Tracking
HMM	75.50	63.88
Latent HMM	75.58	64.04
LSP-H	80.41	65.27
LSSVM-H	80.41	65.27
Latent CRF	**85.09**	**65.59**

Credit: I. Naim

Translating Videos to Natural Language Using Deep Recurrent Neural Networks

S. Venugopalan, H. Xu, M. Rohrbach, J. Donahue, R. Mooney, K. Saenko

A monkey pulls a dog's tail and is chased by the dog.

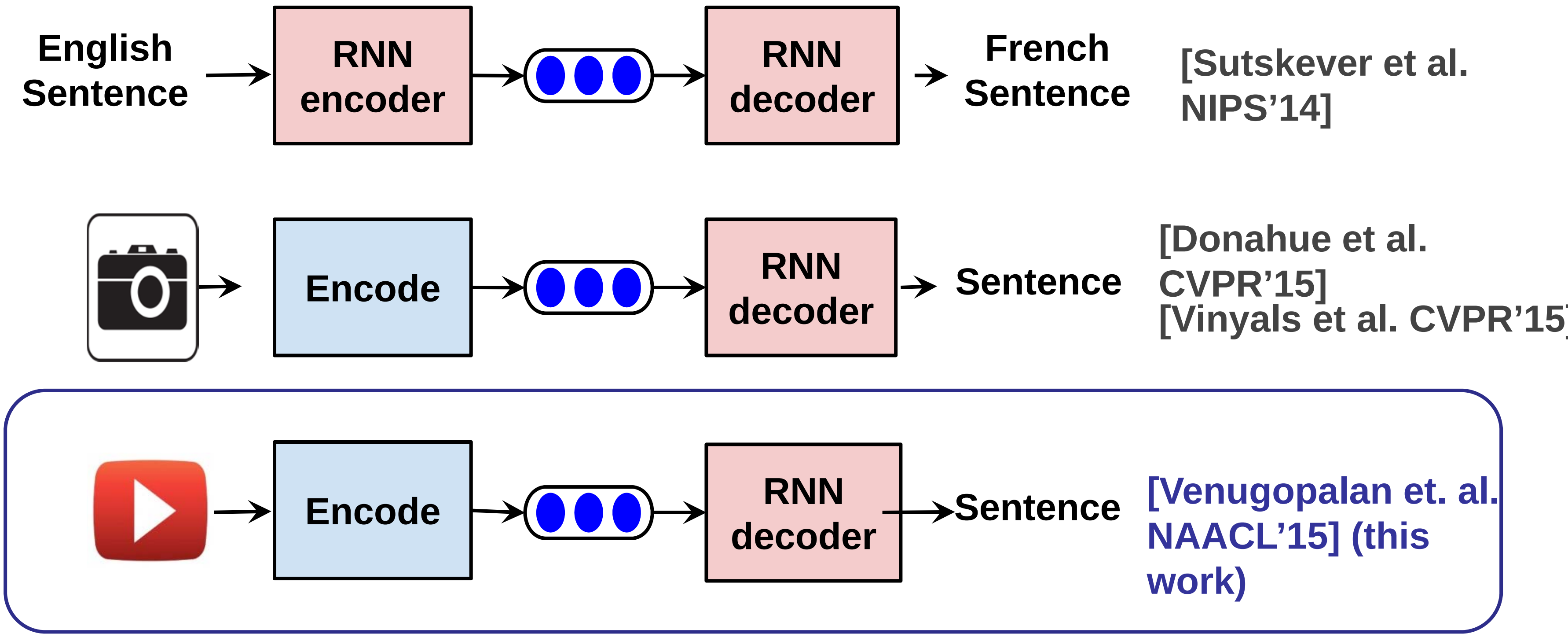
Motivation: Given a video clip generate a sentence to describe the event.

Translating Videos to Natural Language using Deep RNNs (NAACL-HLT 15)
S. Venugopalan, H. Xu, M. Rohrbach, J. Donahue, R. Mooney, K. Saenko

http://bit.ly/TranslatingVideos

Credit: S. Venugopalan

Decode a visual vector to a sequence using Recurrent Neural Networks (RNNs)



Key Insight:

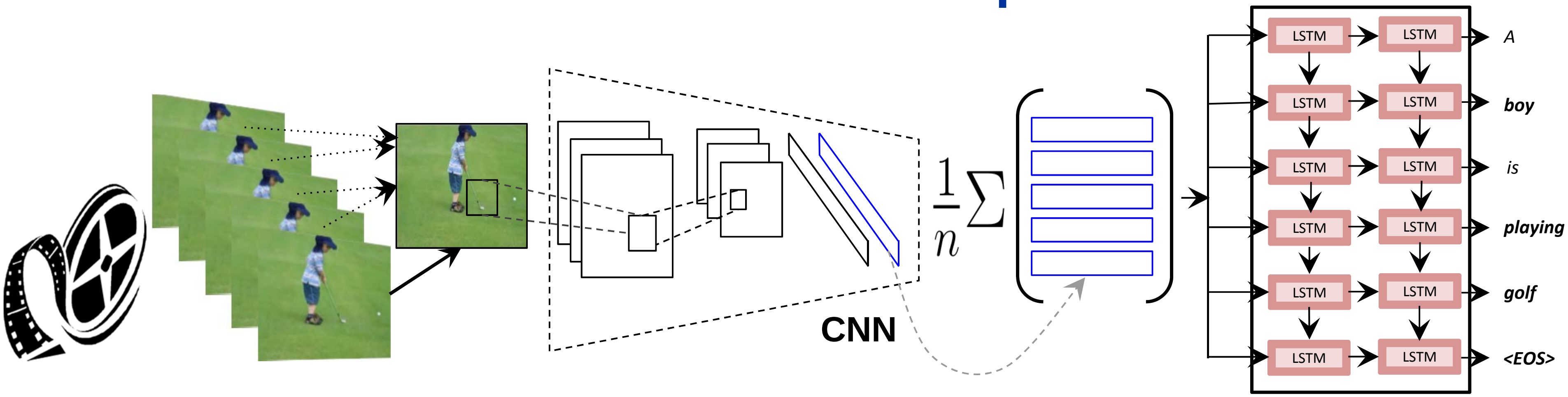
Generate feature representation of the video and "decode" it to a sentence

Translating Videos to Natural Language using Deep RNNs (NAACL-HLT 15)
S. Venugopalan, H. Xu, M. Rohrbach, J. Donahue, R. Mooney, K. Saenko

http://bit.ly/TranslatingVideos

Credit: S. Venugopalan

CNN+LSTM network for video description.



Train on image-caption datasets and tune on videos.

Translating Videos to Natural Language using Deep RNNs (NAACL-HLT 15)
S. Venugopalan, H. Xu, M. Rohrbach, J. Donahue, R. Mooney, K. Saenko

http://bit.ly/TranslatingVideos

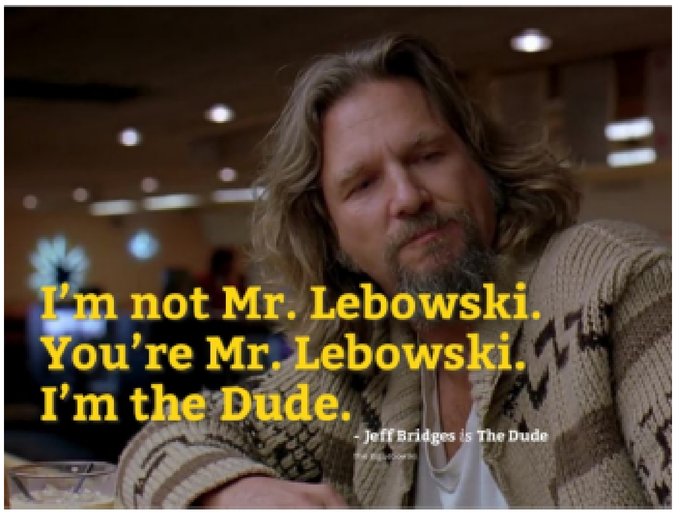
Credit: S. Venugopalan

EA2: NLP for Web, Social Media and Social Sciences

- Invited speaker: Lillian Lee, Cornell University
  - Title: *"Big data pragmatics!", or "Putting the ACL in Computational Social Science", or, if you think these title alternatives could turn people on, turn people off, or otherwise have an effect, this talk might be for you.*
- Some Topics:
  - Framing (Baumer et al., 2015; Guerini et al., 2015)
  - Social structure and dynamics (Krishnan and Eisenstein 2015).

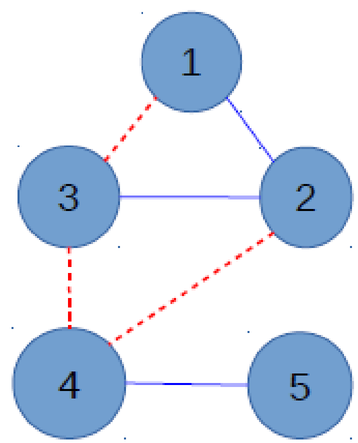
Research question: how does *language* reflect and define *social relationships*?

Krishnan & Eisenstein
NAACL 2015



Address terms like "Dude" and "Mr." play an important role in setting the nature of social relationships.

Setting: network structure and *address terms* are observed.



y_{12} = `formal`

x_{12} = {'*Mr Lebowski*':3, '*sir*':1}

x_{21} = {'*sir*': 1}

Edge labels are *latent variables*.

Credit: J. Eisenstein

Probabilistic model

Krishnan & Eisenstein NAACL 2015

Latent edge labels

Observed address terms

Network structure

$$P(y, x \mid G; \Theta, \beta, \eta) = P(x \mid y; \Theta)\color{red}{P(y \mid G; \beta, \eta)}$$

The **likelihood** explains how social factors determine linguistic features.

This term factors over the dyads in the social network.

The **prior** explains which network label configurations are likely.

This term factors over the **triads** in the network.

All parameters are estimated using **variational inference.**
Noise-contrastive estimation lets us avoid summing over all label configurations.

Credit: J. Eisenstein

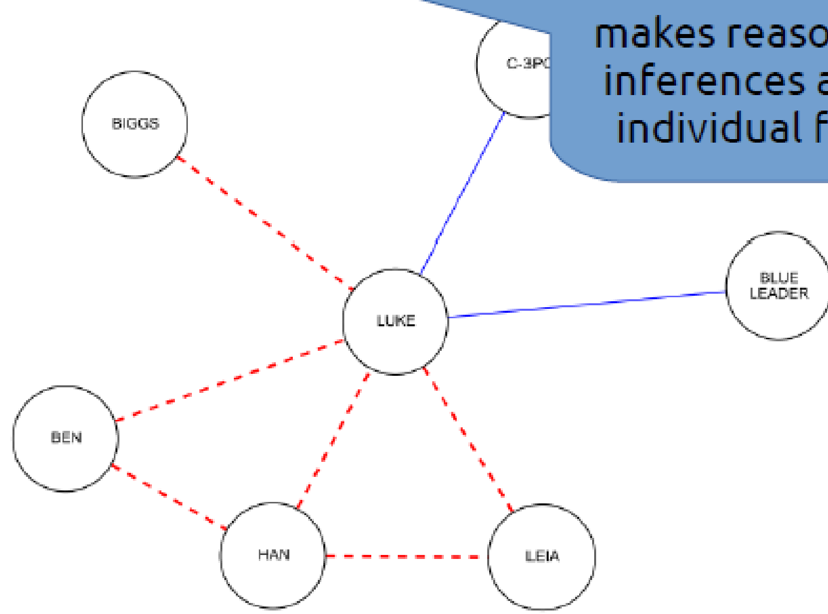
Movie script evaluation

Krishnan & Eisenstein
NAACL 2015

V-cluster	T-cluster
sir	FirstName
mr+LastName	man
mr+FirstName	baby
mr	honey
miss+LastName	darling
son	sweetheart
mister+FirstName	buddy
mrs	sweetie

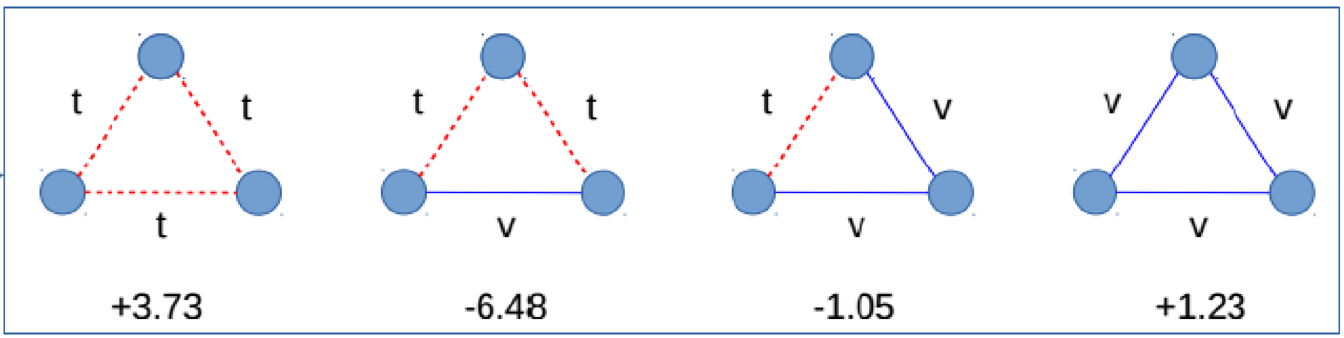
1. Address terms naturally cluster by formality

Star Wars



2. The model makes reasonable inferences about individual films.

3. We can estimate the stability of signed triads.



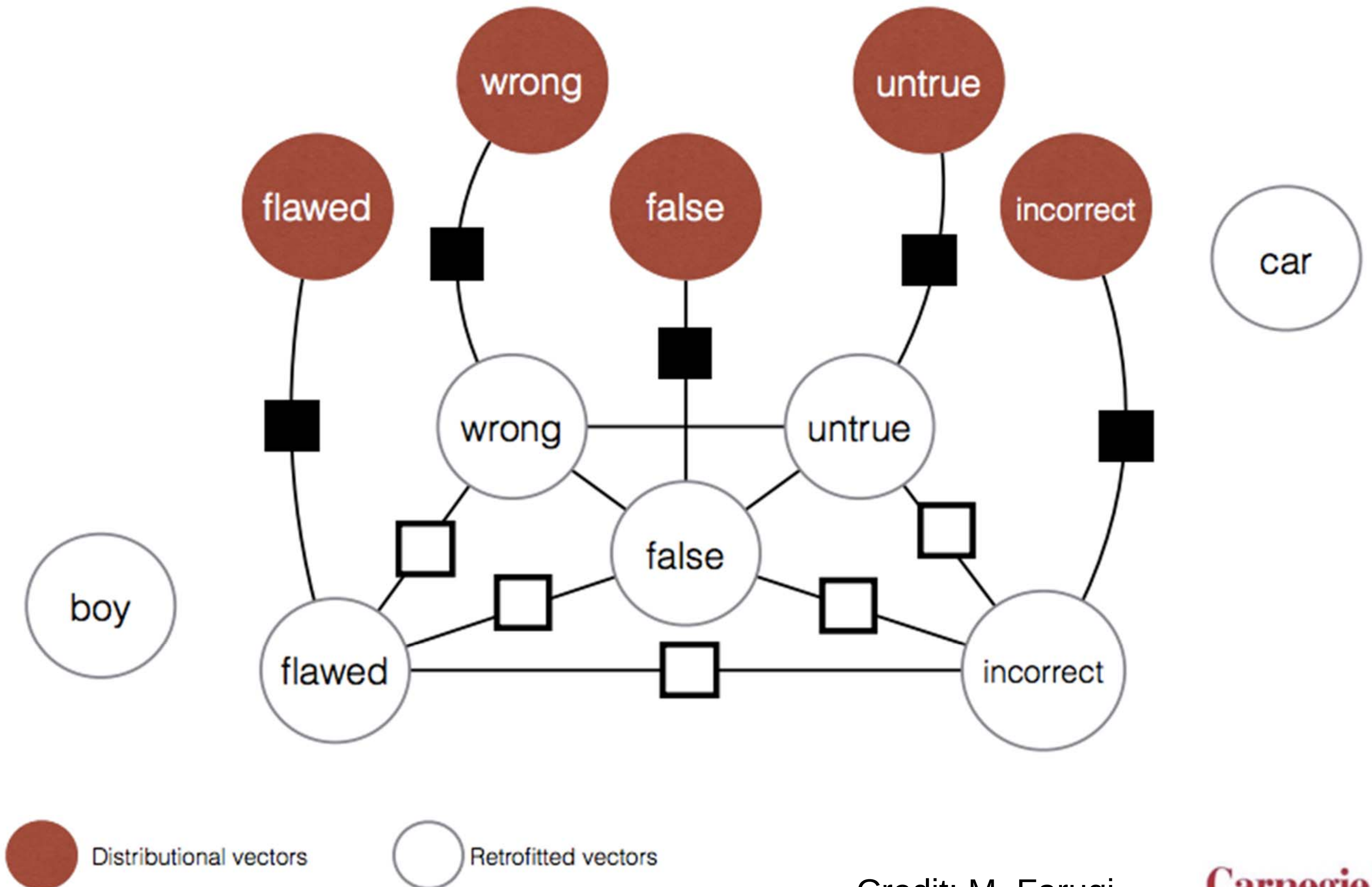
Trends in Methodology

Distributional Semantic Models

- ***Distributional Hypothesis***: words that are used in the same context tend to have similar meanings
- Vector space representation for words – word vectors
- Semantic compositionality (Fyshe et al. 2015), verbs with multiple senses (Greenberg et al. 2015), incorporating visual features (Lazaridou et al. 2015), incorporating semantic resources (Faruqui et al., 2015)

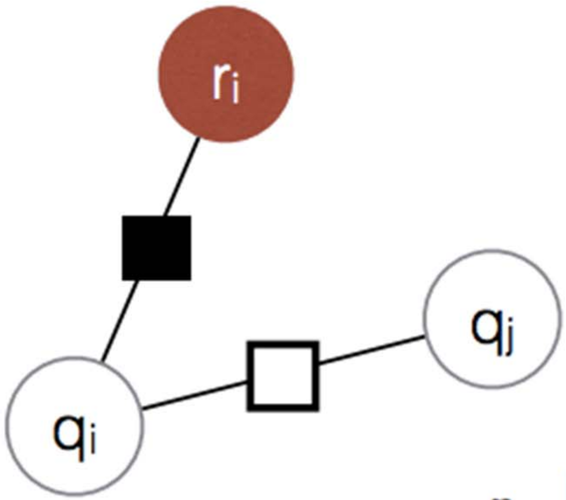
M. Faruqi, J. Dodge, S. Jauhar, C. Dyer, E. Hovy, and N. Smith, *Retrofitting Word Vectors to Semantic Lexicons,* NAACL 2015

Retrofitting



Credit: M. Faruqi

Retrofitting



$$\alpha_i \|q_i - \hat{q}_i\|^2$$

$$\beta_{ij} \|q_i - q_j\|^2$$

$$\Psi(Q) = \sum_{i=1}^{n} \left[\alpha_i \|q_i - r_i\|^2 + \sum_{(i,j)\in E} \beta_{ij} \|q_i - q_j\|^2 \right]$$

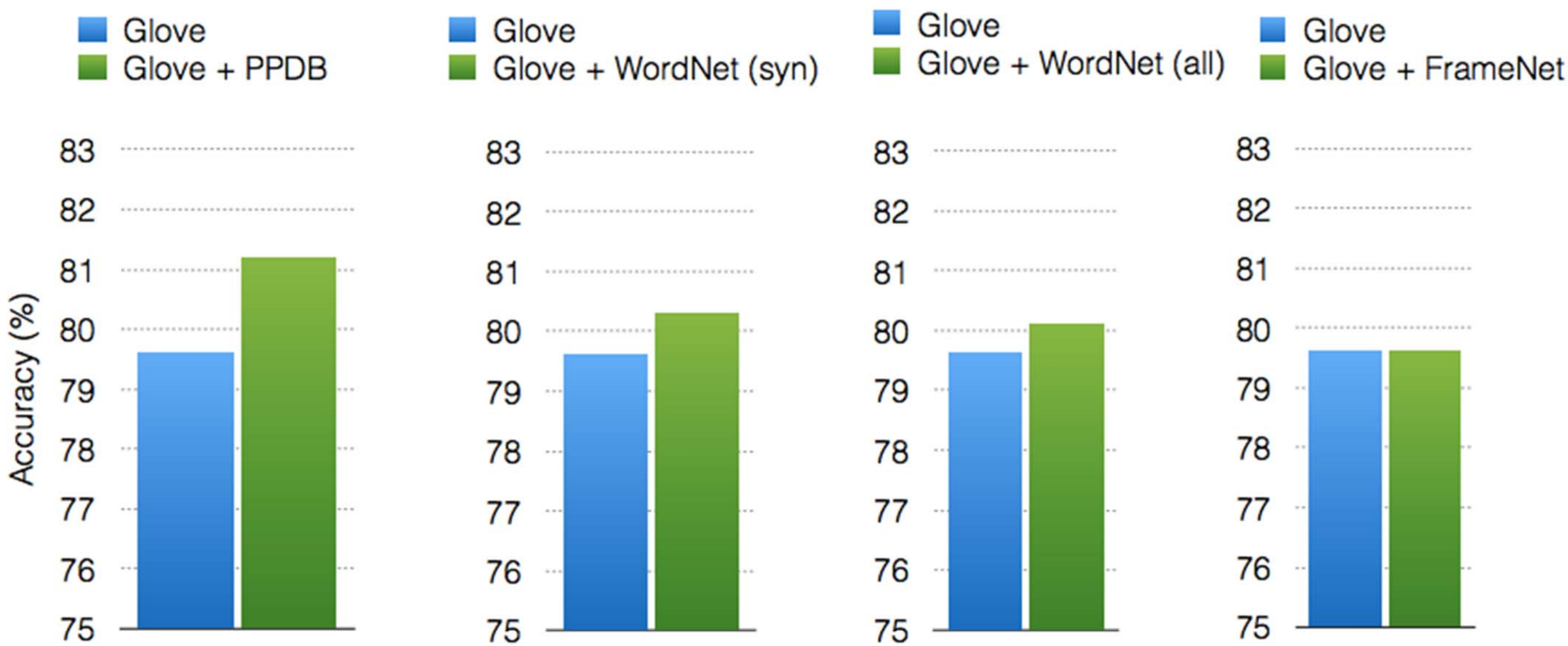
Iterative Updates: $$q_i = \frac{\sum_{j:(i,j)\in E} \beta_{ij} q_j + \alpha_i r_i}{\sum_{j:(i,j)\in E} \beta_{ij} + \alpha_i}$$

Distributional vectors   Retrofitted vectors

Credit: M. Faruqi

Carnegie Mellon

Sentiment Analysis



Credit: M. Faruqi

Carnegie Mellon

Google research

Unsupervised Morphology Induction using Word Embeddings

Radu Soricut, Franz J. Och*
NAACL 2015

*now at Human Longevity Inc.

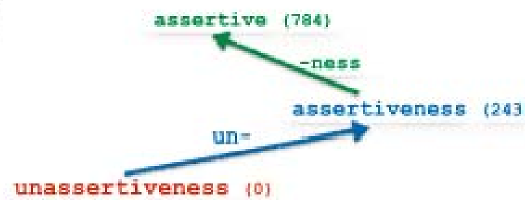
Credit: R. Soricut

Unsupervised Morphology Induction

Google

Q: What do we want?

A: We want *morphology-based transformations* that can accurately analyze words (even ones unseen at training time)



prefix:un:ε

unabated unable unabridged…
unaware unbalance unbeaten…
undoing undone undoubted…
untrusted untrustworthy…

rank(unaware → aware) = 0
rank(undone → done) = 129

↑un-

clear – unclear
delivered – undelivered
truthful – untruthful

rank(undone + ↑un- → done) = 4

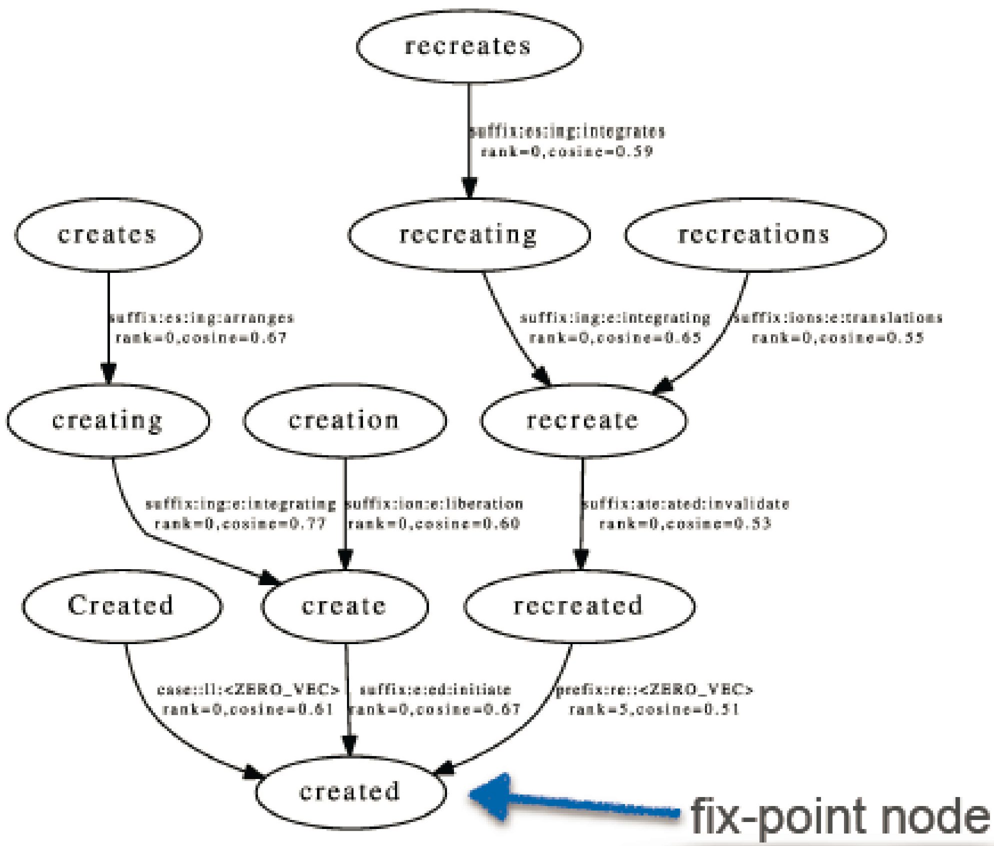
Query against embedding space: *morphology shifts meaning consistently*

Credit: R. Soricut

Unsupervised Morphology Induction: Algorithm

Output: labeled, weighted, acyclic, directed graph D^{V}_{Morph}

- words are nodes, morphological mappings are weighted edges



Credit: R. Soricut

Unsupervised Morphology Induction: Evaluation

Google

Evaluation on similarity datasets (RG-DE, RW-EN)

Language	Train Set	|Tokens|	|V|	$\|G^V_{Morph}\|$	$\|D^V_{Morph}\|$
EN	Wiki-EN	1.1b	1.2m	780k	75,823
DE	WMT-DE	1.2b	2.9m	3.7m	169,017
EN	News-EN	120b	1.0m	2.9m	98,268
DE	News-DE	20b	1.8m	6.7m	351,980

impossibilities	unattainableness	8.8
deregulating	liberation	8.0
baseness	unworthiness	4.0
transmigrating	born	1.1

RW-EN Testset

System	|Unembedded| Wiki-EN	|Unembedded| News-EN	Spearman ρ Wiki-EN	Spearman ρ News-EN
SkipGram	80	177	35.8	44.7
SkipGram+Morph	1	0	41.8	52.0

+9 +7

Edelstein	Juwel	3.8
Autogramm	Unterschrift	3.5
Irrenhaus	Friedhof	0.3
Kraftfahrzeug	Magier	0.0

RG-DE Testset

System	|Unembedded| WMT-DE	|Unembedded| News-DE	Spearman ρ WMT-DE	Spearman ρ News-DE
SkipGram	0	20	62.4	62.1
SkipGram+Morph	0	0	64.1	69.1

+0 +7

Credit: R. Soricut

Trends in Methodology

Deep Neural Network

- Document categorization (Johnson and Zhang 2015)
- Word sense disambiguation (Taghipour and Ng 2015)
- Paraphrase identification (Yin and Schutze 2015; Zhao et al. 2015)
- Response generation (Sordoni et al. 2015)
- Machine Translation (Xing et al. 2015; de Gispert et al. 2015)
- Video-to-text generation (Venugopalan et al. 2015)

Summary

- NLP has grown into one of the most exciting and diverse community
- Conference website:
  - http://naacl.org/naacl-hlt-2015.
- ACL Anthology: an open-access repository
  - http://aclweb.org/anthology/
  - Journal, conference, and workshop papers since 1965