Computer Vision, Speech Communication & Signal Processing Group,
National Technical University of Athens, Greece (NTUA)
Robotic Perception and Interaction Unit,
Athena Research and Innovation Center (Athena RIC)

Part 4.
Text Processing and Saliency

Alexandros Potamianos and Elias Iosif

Tutorial at IEEE International Conference on Acoustics, Speech and Signal Processing 2017,
New Orleans, USA, March 5, 2017


services
prices
payment
links
Coin
specie
century
Bank
systems
Theory
country
value
note
articles
theory
held
deferred
Juno
reserve
January
example
economies
policy
often
account
Fiat
store
stable
may
trade
debts
time
part
consists
will
link
Market
term
record
dollar
medium
silver
early
deposits
using
notes
one
Business
fixed
portal
U.S
Free
plus
system
bank
article
thus
rate
coin
means
types
legal
amount
form
process
face
also
First
York
easily
item
John
Paper
banks
role
Law
unit
taxes
terms
New
cash
gold
world
Polo
metals
Gold
goods
credit
market
page
Credit
Song
David
central
copper

Saliency Models for Text

Did you see/think *money*?

Semantic Priming

- Semantic priming
  - The presence of a word (prime) facilitates the cognitive processing of another word e.g., bank-money
- Semantic priming as explanation of false memories
  - Remembering events never happened (or remember differently)
  - Experiment: remembering words never presented in lists, e.g.,

chair and *sleep*

table	seat	legs	…	desk	sofa	wood	*chair*
bed	rest	dream	…	snooze	nap	snore	*sleep*

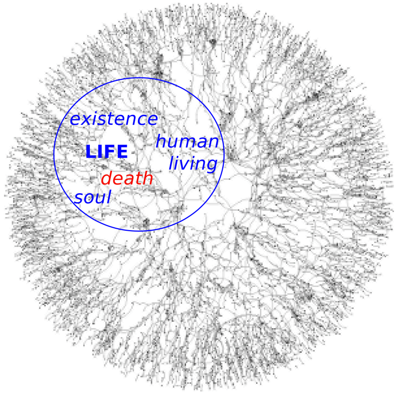
- Affective priming: emotional analogue of semantic priming
  - E.g., fusion of semantic and affective spaces

[Collins and Loftus, *A Spreading-Activation Theory of Semantic Processing,* Psychological Review,1975]
[Roediger and McDermott, *Creating False Memories: Remembering Words not Presented in Lists*, Journal of experimental psychology: Learning, Memory, and Cognition, 1995]

Semantic & Affective Priming: Example

When semantic semantic priming only is not enough

- Consider the semantic sub-space activated for “life”



- Antonym (“death”) is also activated
  - Antonymy embodies both semantic proximity and distance
  - Easily recognized by humans
  - Lexical models fail – need to also consider affective info

[Iosif and Potamianos, *Feeling is Understanding: From Affective to Semantic Spaces,* IWCS, 2015]

Saliency Models for Text

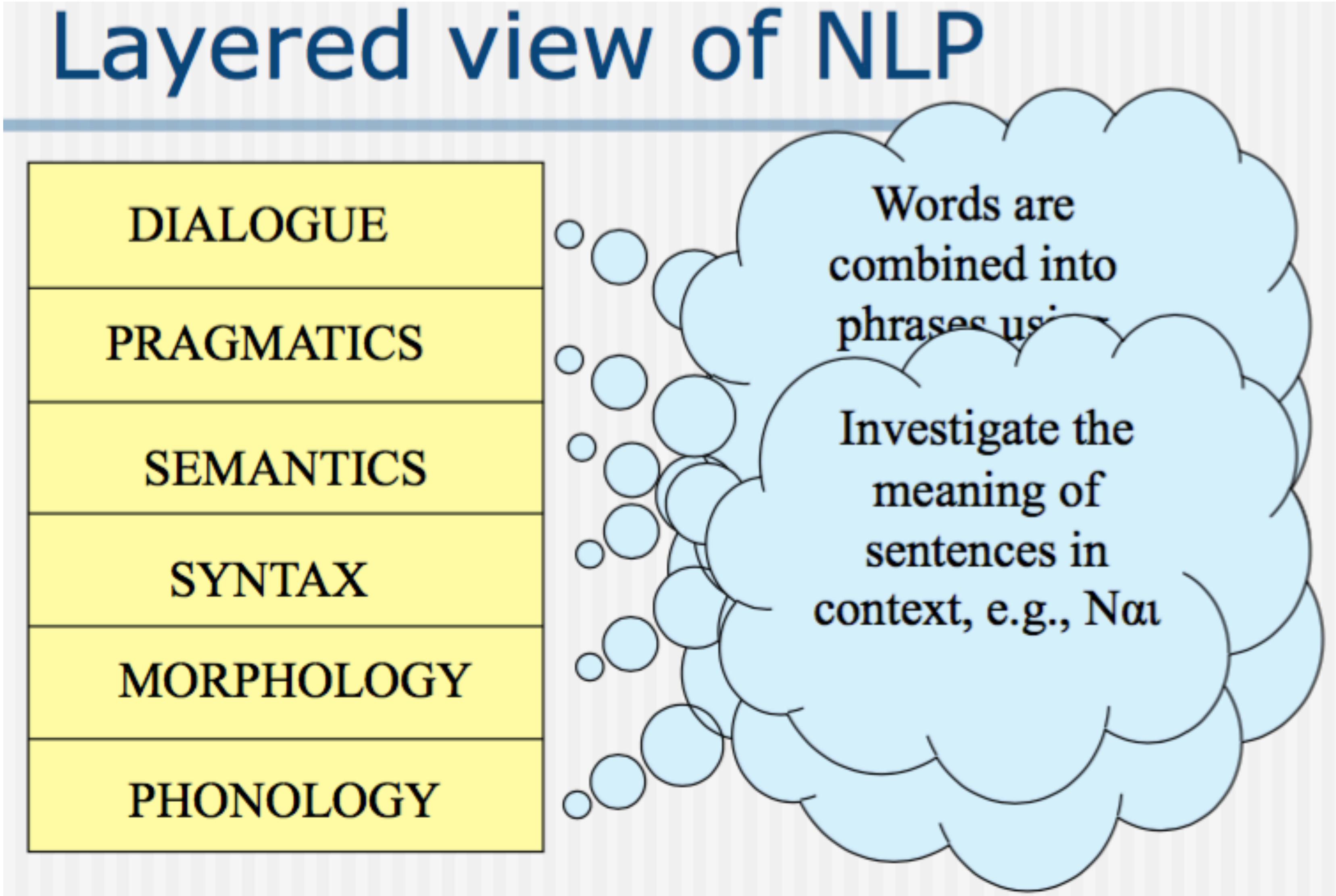
- Application of saliency models: less-investigated area for text
- Text captures attention in visual scenes
  - E.g., in free viewing and search tasks in images
- Linguistic info used for detecting salient words in speech
  - Saliency as intonational emphasis
  - E.g., part-of-speech, freq.-based

[Cerf et al., *Faces and Text Attract Gaze Independent of the Task: Experimental Data and Computer Model,* Journal of vision, 2009]

[Hirschberg, *Pitch Accent in Context Predicting Intonational Prominence from Text*, Artificial Intelligence, 1993]

[Brenier et al., *The Detection of Emphatic Words Using Acoustic and Lexical Features*, Interspeech, 2005]

Natural Language Proc.: Layers and Attention



Natural Language Proc.: Layers and Attention

Layers:

- Phonetics: salient words are emphasized
- Morphology: identify core components of words
- Lexical: words into part-of-speech classes
- Syntax: structurally relate words
- Semantics: identify relevant word senses
- Pragmatics: ground to situational context
- Dialogue/Discourse: identify salient spots in large linguistic units

Saliency: Definition

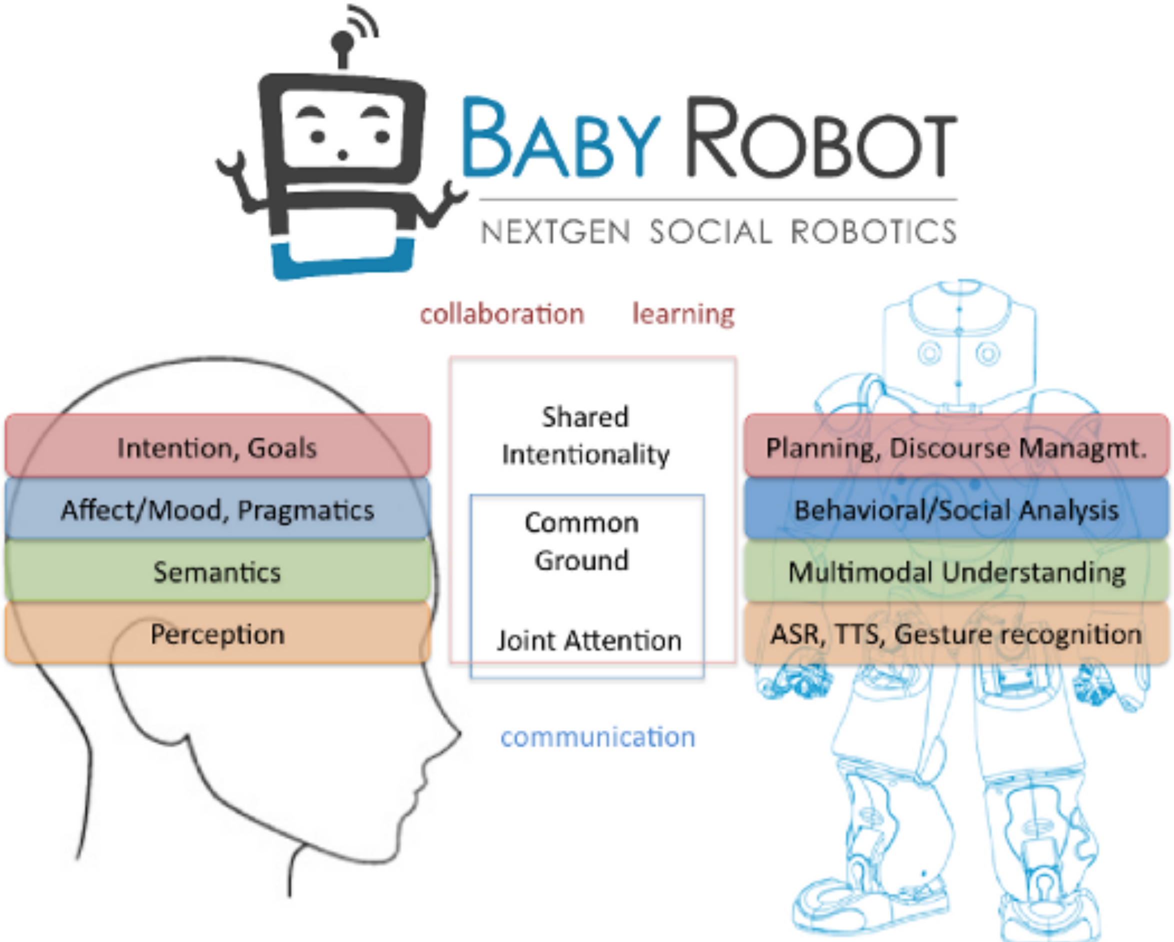
- Saliency: refers to the properties of an entity
- Salient entity (aka target): distinguished from other entities
  - Distinguishment within context (e.g., sentence, dialogue, etc.)
- Saliency detection: fundamental attentional mechanism
  - Facilitation of learning and survival
  - Perceptual & cognitive resources focused on “important” info
- Saliency detection via contrasting
  - Physical properties, e.g., color, intensity, size, orientation, etc.
- Also, other factors can contribute to saliency
  - Emotional, motivational, cognitive

Top-down vs. Bottom-up Attention

- Top-down perspective: knowledge-driven
  - A-priori knowledge about the target, e.g., its (anticipated) location
- Bottom-up perspective: stimuli-driven
  - Detection of the target based on sensory saliency
- Overlap between those perspectives
- Synergy between top-down & bottom-up attention
  - Hard to recognize entities in a scene & understand their relations
  - Selective attention: optimization of attentional performance
- Cognitive load theory: 2 mechanisms of selective attention
  - Perceptual: perceive or ignore stimuli
  - Cognitive: process stimuli

[Sarter et al., *The Cognitive Neuroscience of Sustained Attention: Where Top-Down Meets Bottom-Up*, Brain Research Reviews, 2001]

Attention and Discourse Processing



BabyRobot project: www.babyrobot.eu

Attention and Discourse Processing

- Discourse: piece of language behavior
  - Typically, involves multiple utterances and participants
  - Produced by: speakers or writers
  - Consumed by: hearers or readers
- Constituents of a model of discourse
  1. Linguistic structure: arrangement of words/phrases into utterances
  2. Intentional structure: intentions of participants in discourse segments
  3. Attentional structure: information about word and their relations, as well as saliency in discourse segments

[Grosz and Sidner, *Attention, Intentions, and the Structure of Discourse,* Computational Linguistics, 1986]

Attention and Discourse Processing

Discourse model as a composite of interacting constituents for:

- Assessment of the coherence of utterances
  - Fit of an utterance wrt rest utterances
  - Why it was said
  - Its meaning
- Formulation of a basis of anticipations
  - Facilitates the accommodation of new utterances
- The attentional structure has an additional role:
  - Creation of the means for exploiting the lexical information in the linguistic and intentional structures during the generation and interpretation of individual utterances

[Grosz and Sidner, *Attention, Intentions, and the Structure of Discourse,* Computational Linguistics, 1986]

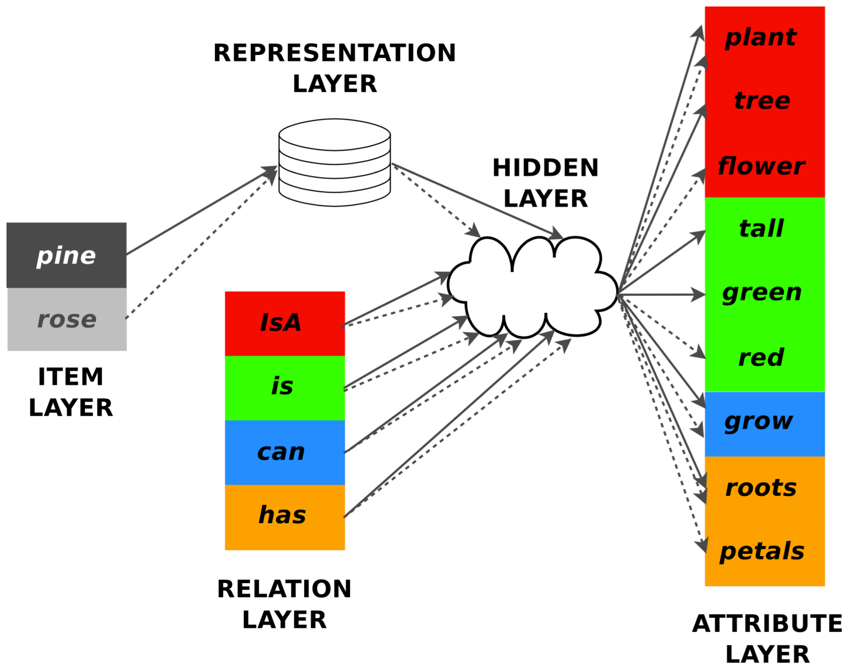
Attention and Discourse Processing

Attention structure:

- An abstraction of participants' focus (center of attention)
  - Role: summarization of information from previous utterances required for subsequent processing
- Can be regarded as a stack of focus spaces
  - A focus space is associated with a discourse segment
  - A focus space contains the salient entities of the respective segment
- Evolves with the unfolding of the discourse
  - Additions and deletions of focus spaces
  - Those operations are determined by the intentions that signify the initiations of new discourse segments

[Grosz and Sidner, *Attention, Intentions, and the Structure of Discourse,* Computational Linguistics, 1986]

Models: Semantic Cognition



- Representation: based on semantic attributes
  - Similarity: common vs. distinctive attributes; distributed representation in neural nets

[Tversky, *Features of Similarity*, Psychological review, 1977]

[Rogers and McClelland, *Semantic Cognition: A Parallel Distributed Processing Approach*, MIT press, 2004]

Models: Distributional Representation of Meaning

How do we represent the meaning of a word?

- Possible approaches: use of resources, e.g., WordNet
  - Disadvantages: manual effort, words as atomic symbols, etc
- Distributional hypothesis of meaning
  - The meaning of a word *w* can be represented by its neighbors

    "You shall know a word by the company it keeps" - Firth
  - Neighbors of *w*: words that co-occur with *w* in linguistic context

Toy corpus

> *"Cars are motor vehicles with four wheels; usually propelled by an internal combustion engine.*
> *A tree is a tall perennial woody plant having a main trunk and branches forming a distinct elevated crown.*
> *...*
> *They built a large plant to manufacture a special type of engine for cars.*
> *He reads his newspaper at breakfast."*

Models: Distributional Representation of Meaning

Example of word-context matrix (aka Vector Space Model-VSP)

Target words	Contextual neighbors and co-occurrence counts for targets-neighbors								
	breakfast	*cars*	*crown*	*large*	*motor*	...	*tall*	*trunk*	*vehicles*
engine	0	11	0	0	12	…	0	0	9
newspaper	5	0	0	0	0	…	0	0	0
plant	0	6	1	8	0	…	2	2	0
tree	0	0	1	1	0	…	4	3	0

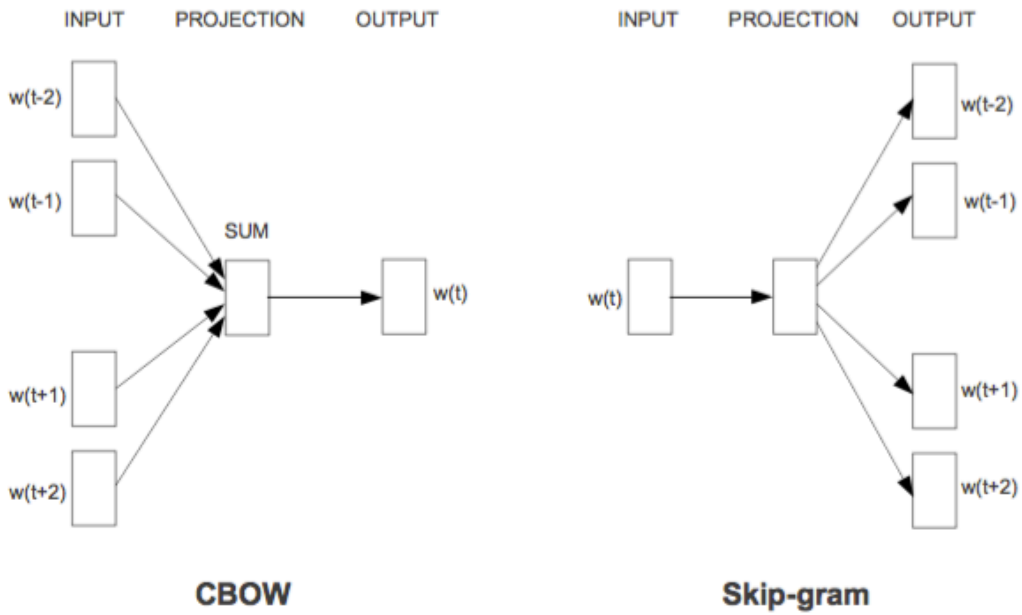
- Basic parameters of VSP
  - Corpus pre-processing (e.g., tokenization, lemmatization, etc.)
  - Size of context window (typically, 1-5)
  - Weighting of contextual neighbors (e.g., freq.-based, mutual info.)
  - Dimensionality reduction (e.g., Singular Value Decomposition)

Models: Distributional Representation of Meaning

- Limitations of traditional VSP
  - Increases wrt. vocabulary size
  - High dimensional – storage issues
  - Sparsity issues (especially for rare words)
- Solution: salient info in low-dimensional dense vectors
  - Typically, 100-500 dimensions
- Recently: word embeddings based on neural networks
  - Originally from the field of statistical language modeling
  - Application to Distributional Semantic Models
  - Examples: word2vec and GloVe

[Bengio et al., *A Neural Probabilistic Language Model*, Journal of Mach. Learning Research, 2003]
[Mikolov et al., *Efficient Estimation of Word Representations in Vector Space*, In Proc. ICLR, 2013]
[Pennington et al., *GloVe: Global Vectors for Word Representation*, In Proc. EMNLP, 2014]

Models: Distributional Representation of Meaning



- word2vec: instead of counting word co-occurrences
  - CBOW: predicts current word *w(t)* based on local context
  - Skip-gram: predicts local context based on *w(t)*
- GloVe: can be regarded as a global skip-gram model

[Mikolov et al., *Efficient Estimation of Word Representations in Vector Space*, ICLR, 2013]
[Pennington et al., *Glove: Global Vectors for Word Representation*, EMNLP, 2014]

Models: Distributional Representation of Meaning

- Related sub-tasks of lexical semantics (not exhaustive list):
  - Similarity computation (e.g., “gem-jewel” vs. “gem-apple”)
  - Analogy (“Greece:Athens” vs. “Italy:Rome”)
  - Concept categorization (e.g., “cat” IsA “mammal”)
  - Verb selectional preferences (e.g., “eat an apple” vs. “eat a car”)
  - Relation classification (e.g., “wealth-happiness” CauseEffect)
  - Paraphrasing, summarization
  - Affective analysis of text (e.g., positively valenced words)
- The notion of saliency exists in the aforementioned sub-tasks
  - For a given word (or relation between words) the lexico-semantic space is filtered and only the lexically/semantically relevant sub-spaces are activated

Models: Word Graphs-Introduction

- Basic idea: graphs (networks) as mental representation for language units and their relationships
  - Originates with early work in psychology
  - Cognitive sciences
  - Various applications in NLP
- Cognitive perspective: model of semantic memory

  "The memory that a person calls upon in his everyday language behavior" - Quilian
- Various types of networks, e.g.,
  - Word co-occurrence networks
  - Syntactic dependency networks; semantic networks

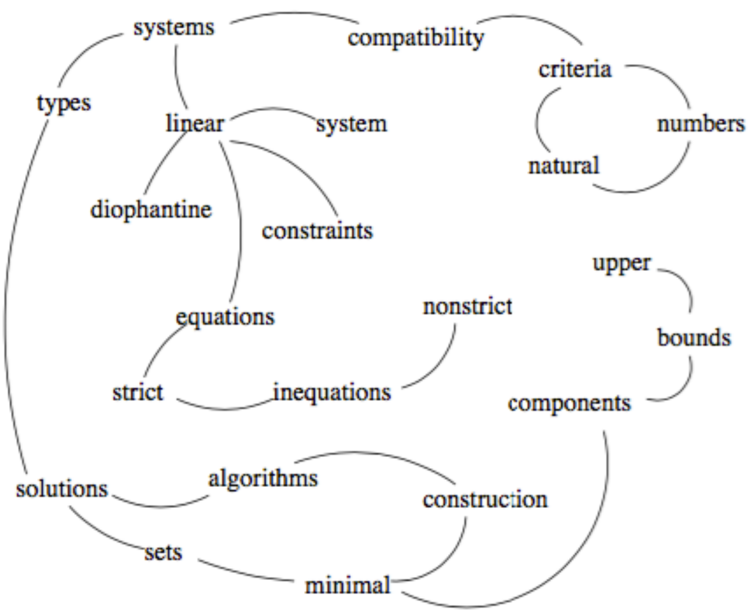
[Freud, Psychopathology of Everyday Life, Payot, 1901]

[Quilian, *Semanic Memory,* In M. Minsky (ed.) Semantic Information Processing, 1968]

[Mihalcea and Radev, *Graph-based Natural Language Processing and Information Retrieval,* Cambridge University Press, 2011]

Models: Word Co-occurrences

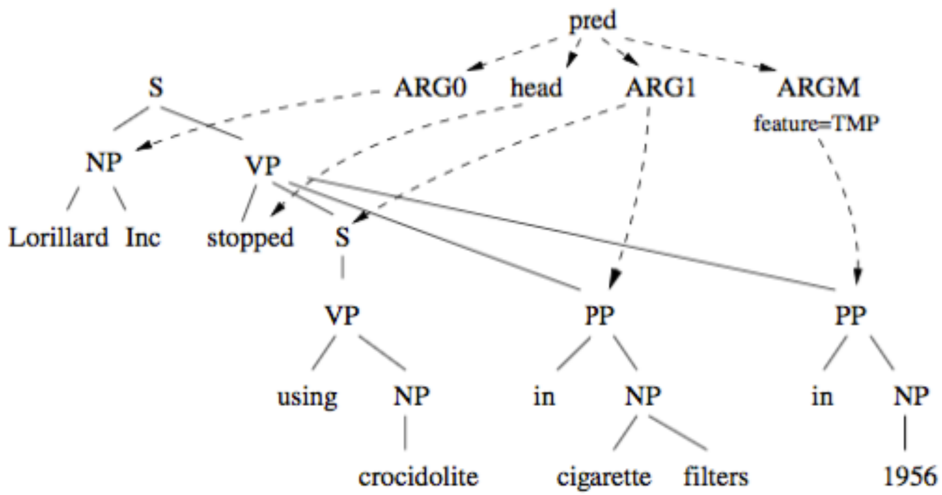
- Extracted from the abstract of a scientific article



  - Enables keyword extraction

[Mihalcea and Tarau, *TextRank: Bringing Order into Texts*, ACL, 2004]

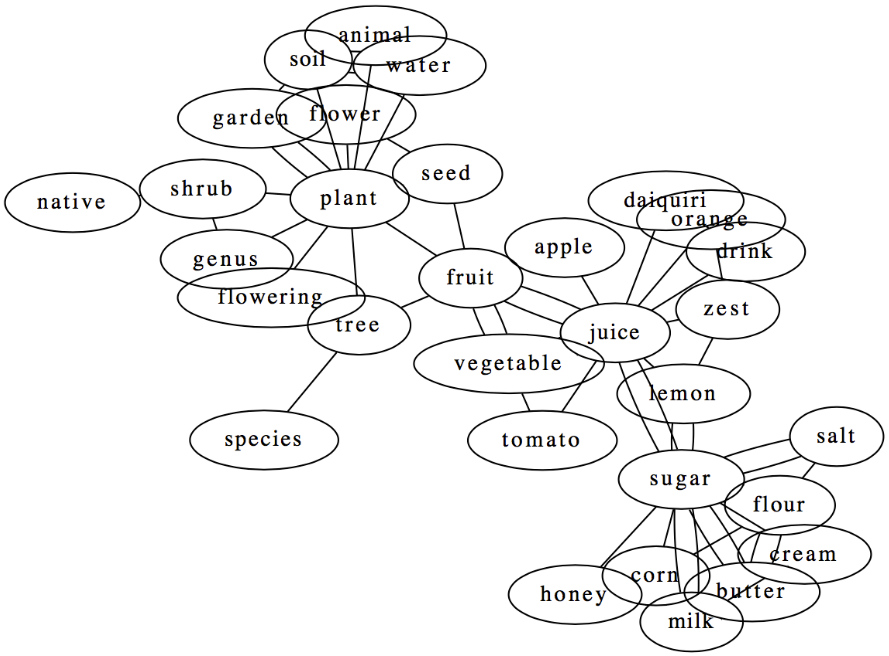
Models: Syntactic Dependencies



- Solid edges: based on surface syntactic structure
- Dashed edges: based on verb "stopped" and its arguments

[Jijkoun and De Rijke, *Learning to transform linguistic graphs*, HLT-NAACL, 2007]

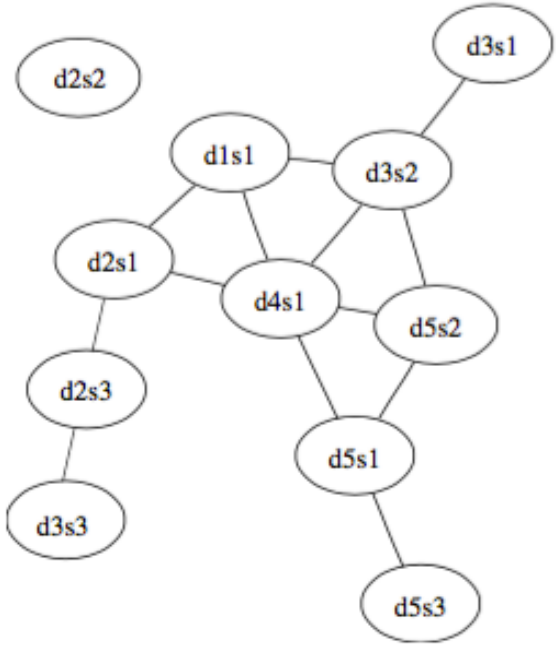
Models: Semantic Similarity Graphs



- Edges: semantic sim. between nodes (subj. to thresholding)
  - Similarity computation via distributional semantic models
  - Enables the discovery of semantic cliques

[Athanasopoulou et al., *Low-Dimensional Manifold Distributional Semantic Models*, COLING, 2014]

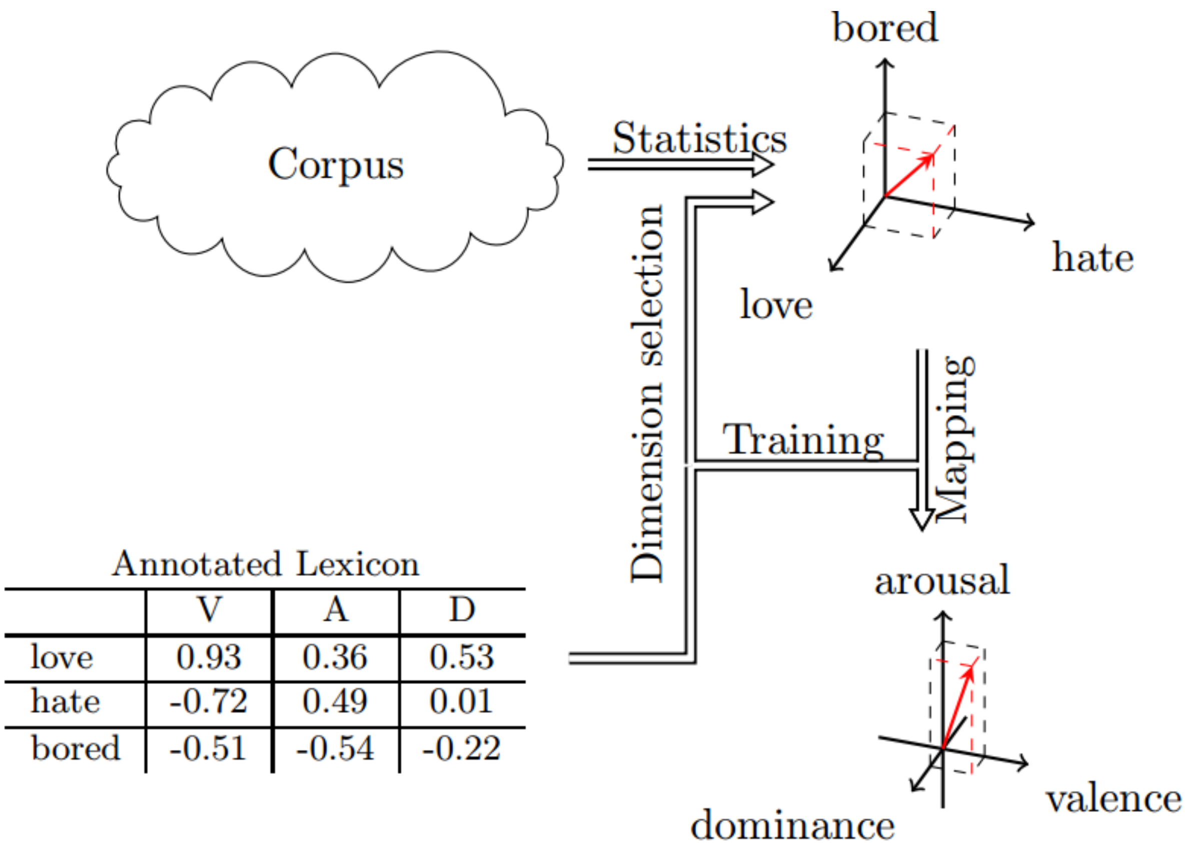
Models: Multi-Document Summarization



- Nodes: sentences (*d*<u>s*</u>*) from different documents (*<u>d*</u>s**)
- Edges: similarity between sentences (subj. to thresholding)

[Erkan and Radev, *LexRank: Graph-based Lexical Centrality as Salience in Text Summarization*, JAIR, 2004]

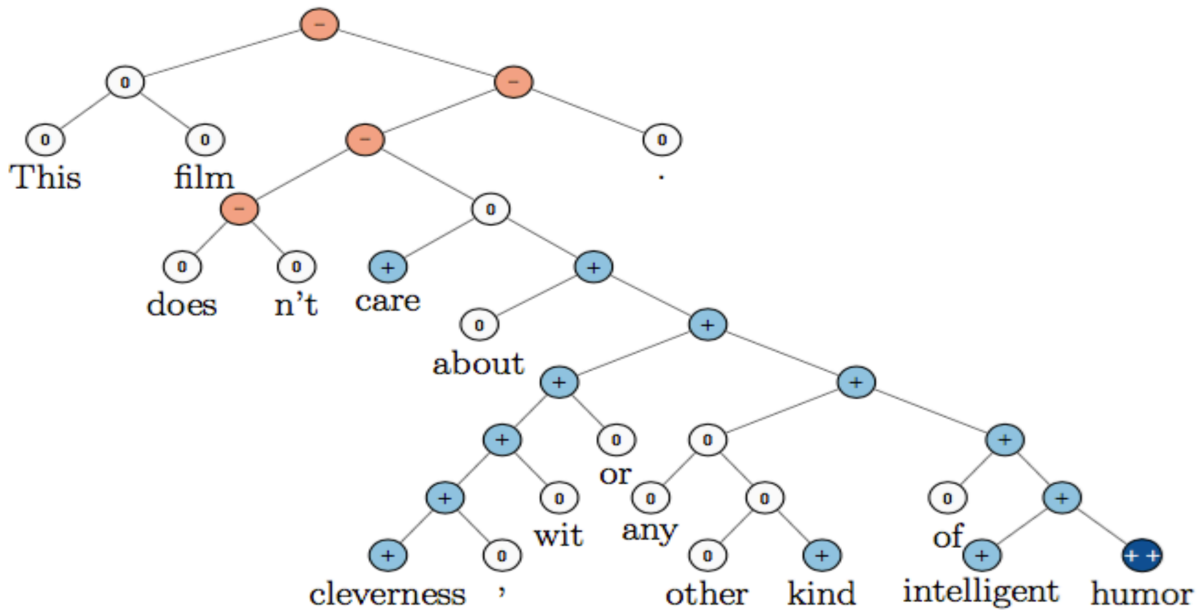
Models: Word-level Semantic-Affective Mapping



	V	A	D
love	0.93	0.36	0.53
hate	-0.72	0.49	0.01
bored	-0.51	-0.54	-0.22

[Malandrakis et al., *Distributional Semantic Models for Affective Text Analysis*, IEEE TASLP, 2013]

Models: Sentence-Level Sentiment Analysis



- Sentimental polarity of sentences
  - Based on parse trees
  - Use of DNN for modeling compositional effects

[Socher et al., *Recursive Deep Models for Semantic Compositionality over a Sentiment Treebank*, EMNLP, 2013]

Models: Story Analysis

Example: analysis of children’s tales (here: “Hans in Luck”)

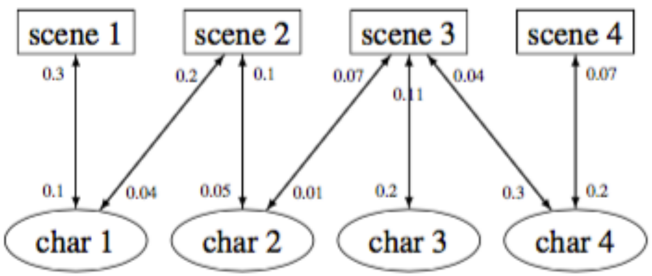
butcher (Gender:M) (Age:A)	36	`` What is the matter with you , my man ? '' said the butcher , as he helped him up .
	37	Hans told him what had happened , how he was dry , and wanted to milk his cow , but found the cow was dry too .
cow (Gender:F) (Age:A)	38	Then the butcher gave him a flask of ale , saying , `` There , drink and refresh yourself ; your cow will give you no milk : do n't you see she is an old beast , good for nothing but the slaughter-house ? ''
Hans (Gender:M) (Age:A)	39 (NEG)	`` Alas , alas ! '' said Hans , `` who would have thought it ? What a shame to take my horse , and give me only a dry cow ! If I kill her , what will she be good for ? I hate cow-beef ; it is not tender enough for me . If it were a pig now?like that fat gentleman you are driving along at his ease?one could do something with it ; it would at any rate make sausages . ''
butcher (Gender:M) (Age:A)	40 (POS)	`` Well , '' said the butcher , `` I do n't like to say no , when one is asked to do a kind , neighbourly thing . To please you I will change , and give you my fine fat pig for the cow . ''

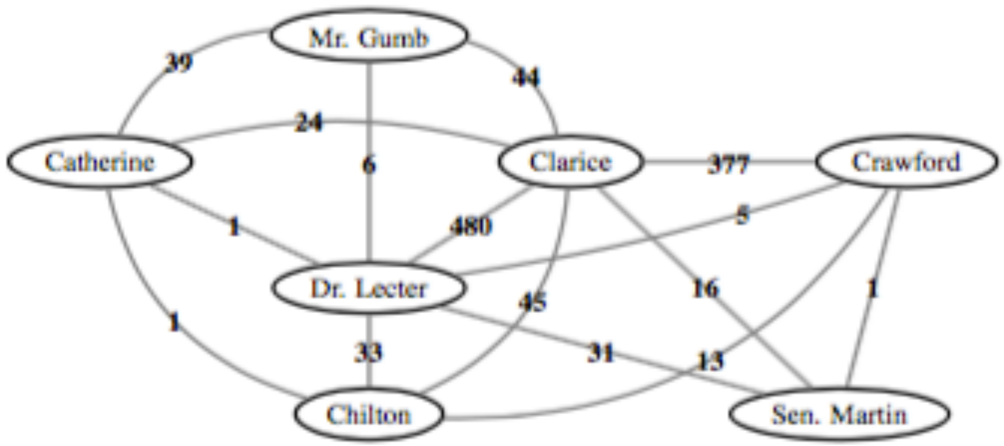
- Identification of:
  - Story characters and speakers; attribution of utterances to speakers
  - Speakers’ gender and age
  - Emotional utterances (positive – neutral - negative)

[Iosif and Mishra, *From Speaker Identification to Affective Analysis*, EACL, 2014]

Models: Movie Script Summarization

Example from “Salience of the lambs”





- Important scenes involve main characters
- Summarization: computation of the optimal chain of important scenes

[Gorinski and Lapata, *Movie Script Summarization as Graph-based Scene Extraction*, HLT-NAACL, 2015]

Text-based Features: Overview

- Various types of info extracted from the layered NLP model
- Two basic computational tasks (often application-specific)
  - Identify important entities and score their saliency
  - Identify relationships between entities and link them
- Examples of entities
  - Topic-specific words, nouns, pronouns, named entities, sentimental words
  - Tools: text analytics, PoS tagging, named entity recognition, syntactic parsing, co-reference resolution, affective lexica, etc.
- Examples of cues indicating relationships between entities
  - Proximity in discourse, actors in semantic relations
  - Tools: discourse analysis, semantic role label., (+ heuristics), etc.

Text-based Features: Doc Summarization

<u>Word importance</u>

- Word frequency and probability: freq. as importance indicator
  - Consider document length: use word probability instead of absolute freq.
- Term Freq. – Inverse Doc. Freq.

$$TF\text{–}IDF(w) = c(w) \log \frac{D}{d(w)}$$

  - *c(w)*: frequency of word *w*
  - *d(w)*: num. of docs in which *w* occurs; *D*: num. of docs in collection
- Topic signatures
  - Words being frequent in text *I* but rare wrt background corpus *B*
  - A word *w* is considered as a topic signature if *P(w|I) > P(w|B)*

[Nenkova and McKeown, *Automatic Summarization*, Foundations and Trends in Information Retrieval, 2011]

Text-based Features: Doc Summarization

<u>Sentence importance</u>

- Proportion of topic signatures in sentence
- Centroid-based summarization
  - Documents are represented by a sentence-level centroid
  - Sentence importance: based on the distance from the centroid
- Graph-based: sentences represented as nodes
  - Centrality-based metrics
- Machine learning based
  - Features: discourse markers, terms, sentence length, topic signatures, etc.
  - Models: HMM, AdaBoost, SVM, etc.

[Nenkova and McKeown, *Automatic Summarization*, Foundations and Trends in Information Retrieval, 2011]

Text-based Features: Node Centrality

- Nodes: words/sentences/etc
- Edges: relations between nodes
  - Many types of relations: structural up to linguistic
  - Applications: word sense disambig., summarization, plot analysis, etc.
- Central node: maximally connected to all other nodes
  - Centrality: as a measure of the influence of a node wrt the information flow over the graph
- Various measurements of node centrality
  - Concise overview via NLP perspective in
  - Basic approach: in-degree centrality
    - Number of edges terminating in a node
    - Normalized by the maximum degree

[Navigli and Lapata, *Graph Connectivity Measures for Unsupervised Word Sense Disambiguation,* International Joint Conference on Artificial Intelligence, 2007]

Text-based Features: Node Centrality

- Eigenvector centrality
  - Basic idea: not all edges are of equal importance
  - Score nodes wrt the importance of their edges

  PageRank (*PR*)

$$PR(m) = \frac{(1-k)}{|V|} + k \sum_{(m,n)\in E} \frac{PR(n)}{\text{outdegree}(n)}$$

  - Sum over the edges of node *m*: *(m,n)* ∈ *E*
  - Outdegree: number of edges leaving a node
  - *1-k*: prob. to randomly select a node scored with *1/|V|*

  Hypertext Induced Topic Selection (HITS)

$$H(m) = \sum_{(m,n)\in E} A(n) \quad ; \quad A(m) = \sum_{(m,n)\in E} H(n)$$

  - Hub and authority value for node *m*: *H(m)* and *A(m)*
  - Good hub: node pointing to many good authorities
  - Good authority: node pointed by many good hubs

[Brin and Page, *Anatomy of a Large-scale Hypertextual Web Search Engine*, WWW '98]
[Kleinberg, *Authoritative Sources in a Hyperlinked environment,* ACM-SIAM '98]

Text-based Features: Node Centrality

- Closeness centrality
  - Node is important if it is close to other nodes – aka Key Player Problem

$$\mathrm{KPP}(m) = \frac{\sum_{n \in V : n \neq m} 1/d(m,n)}{|V| - 1}$$

    - V: number of nodes
    - $d(m,n)$: shortest distance between nodes m and n
- Betweeness centrality
  - Node is important if it is involved in many paths (compared to total paths)

$$B(m) = \sum_{x,y \in V : x \neq m \neq y} \frac{\sigma_{xy}(m)}{\sigma_{xy}}$$

    - V: number of nodes
    - σ_{xy}: number of shortest paths from node x to node y
    - $\sigma_{xy}(m)$: number of shortest paths from x to y passing through node m
    - Normalize by $(|V|-1)(|V|-2)$

[Borgatti, *Identifying Sets of Key Players in a Network,* Conference on Integration of Knowledge Intensive Multi-Agent Systems, 2003]

[Freeman, *Centrality in Networks: I. Conceptual Clarification,* Social Networks, 1979]

Text-based Features: Script Summarization

- Script summarization: select chain of scenes representing movie’s most important content
  - Scene: unit of action associated with one place/action

```
We can’t get a good glimpse of his face, but
his body is plump, above average height; he
is in his mid 30’s.  Together they easily
lift the chair into the truck.
                 MAN (O.S.)
        Let’s slide it up, you mind?
CUT TO:
INT. THE PANEL TRUCK - NIGHT
He climbs inside the truck, ducking under a
small hand winch, and grabs the chair.  She
hesitates again, but climbs in after him.
                    MAN
          Are you about a size 14?
                 CATHERINE
                (surprised)
                   What?
Suddenly, in the shadowy dark, he clubs her
over the back of her head with his cast.
```

  - Scene boundaries: available in the script via discourse markers

[Gorinski and Lapata, *Movie Script Summarization as Graph-based Scene Extraction*, HLT-NAACL, 2015]

Text-based Features: Script Summarization

- Script represented as $M(S_n, C_m)$
  - Set of scenes $S_n=\{s_1, s_2, \ldots, s_n\}$; Set of characters $C_m=\{c_1, c_2, \ldots, c_m\}$
- Set $S_k=\{s_1, s_2, \ldots, s_k\}$ of ordered, consecutive scenes

$$S^* = \mathrm{argmax}_{S_k \subset S_n} Q(S_k) \quad ; \quad Q(S_k) = \lambda_1 P(S_k) + \lambda_2 D(S_k) + \lambda_3 I(S_k)$$

λ_1, λ_2, λ_3: weights

  - $P(S_k)$: scene progression, i.e., preserve story coherence
    - Basic idea: include scenes that follow a scene of important character
  - $D(S_k)$: scene diversity, i.e., avoid redundancy
    - Basic idea: compute the diversity between two scenes s_i and s_{i+1}
  - $I(S_k)$: scene importance, i.e., selection of important scenes
    - Basic idea: scene importance as the proportion of important characters appearing in it
  - Identification of main characters: centrality-based wrt script graph

[Gorinski and Lapata, *Movie Script Summarization as Graph-based Scene Extraction*, HLT-NAACL, 2015]

Text-based Features: Semantic-Affective Mapping

- Assumption: the affective score of a word can be expressed as a linear combination of the affective scores of seed words weighted by semantic similarity and trainable weights a_i
  - Affective dimensions: valence, arousal, dominance
  - Example for valence

$$\hat{v}(t) = a_0 + \sum_{i=1}^{N} a_i \, v(w_i) \, d(w_i, t)$$

t : a word or n-gram (token) not in the affective lexicon

$w_1 \ldots w_N$: seed words

$v(.)$: valence rating of a word or n-gram

a_i : weight assigned to seed w_i

$d(w_i, t)$: semantic similarity between word w_i and token t

[Malandrakis et al., *Distributional Semantic Models for Affective Text Analysis,* IEEE Transactions on Audio, Speech, and Language Processing, 2013]

[Turney and Littman, *Unsupervised Learning of Semantic Orientation from a Hundred-Billion-Word Corpus,* arXiv preprint cs/0212012, 2002]

Text-based Features: Semantic-Affective Mapping

Order	w_i	$v(w_i)$	a_i	$v(w_i) \times a_i$
1	mutilate	-0.8	0.75	-0.60
2	intimate	0.65	3.74	2.43
3	poison	-0.76	5.15	-3.91
4	bankrupt	-0.75	5.94	-4.46
5	passion	0.76	4.77	3.63
6	misery	-0.77	8.05	-6.20
7	joyful	0.81	6.4	5.18
8	optimism	0.49	7.14	3.50
9	loneliness	-0.85	3.08	-2.62
10	orgasm	0.83	2.16	1.79
-	w_0 *(offset)*	1	0.28	0.28

Results from COGNIMUSE: Text Only

- Classify documentary subtitles as salient vs. not salient
  - Ground truth annotations wrt all modalities
  - Here: exploit only text (subtitles - English)
- Text-derived features
  - Lexico-syntactic (PoS classes, features related to stylistics)
  - Word informativeness (variant of TF-IDF)
  - Word centrality
  - Word affective scores (based of semantic-affective mapping)
  - Word saliency scores (modified semantic-affective mapping)
- Dataset: travel documentaries

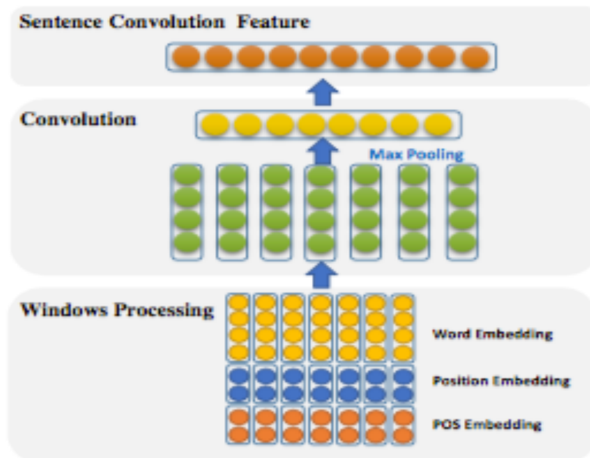
Results from COGNIMUSE: Text Only

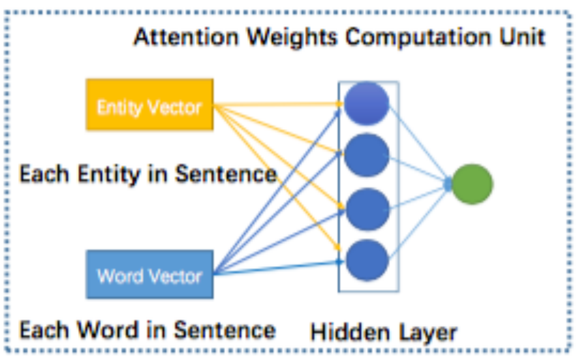
- Examples from documentary about London
  - Word informativeness
    - High: “queen”, “backpack”, “wine”
    - Mid/Low: “London”, “city”, “beer”
  - Word centrality
    - High: “London”, “wine”, “music”
    - Mid/Low: “backpack”, “Westminster”, “Brittania”

- Summary of experimental findings
  - Top performance: all features (fuse classifiers via majority voting)
  - Best perf individual features: word informativeness, lexico-synt.
  - Cues from other modalities needed for improving performance

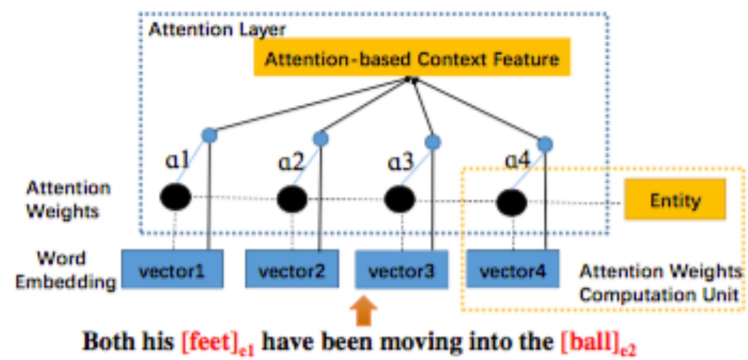
Models: Attention-based Deep Neural Networks

- Attention-based convolutional NN for relation extraction





(a) Attention weights computation unit.



(b) Attention layer network.

[Shen and Huang, *Attention-Based Convolutional Neural Network for Semantic Relation Extraction*, COLING, 2016]

Models: Attention-based Deep Neural Networks

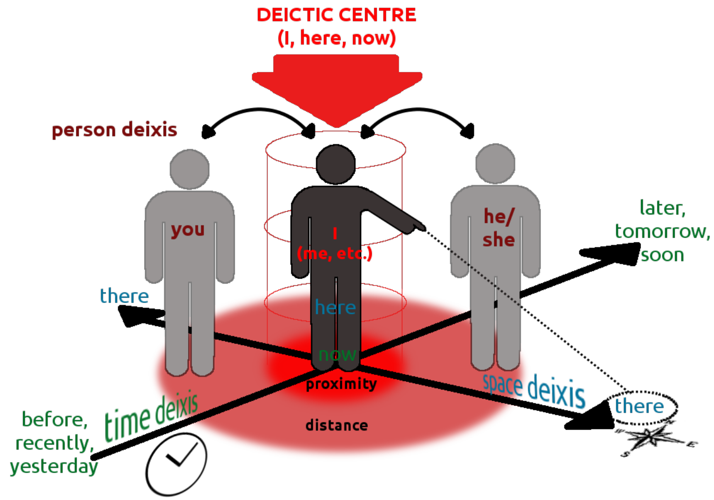
- Neural nets (NN): successfully applied for capturing several linguistic phenomena, e.g., X but Y, negation, etc.
  - Addressing such phenomena is essential for analyzing complex linguistic structures, e.g., phrases and sentences
- Recursive NN (RNN): applied over sentence syntactic trees
  - Capture structural information: from word- to phrase-level
- Example: Bidirectional RNN (two computation phases)
  - Upward (bottom-up), and downward (top-down)
  - Use case: Stanford Sentiment Treebank (movie reviews)

Structural attention mechanism also incorporated for the selection of informative tree nodes

[Kokkinos and Potamianos, *Structural Attention Neural Networks for Improved Sentiment Analysis*, EACL, 2017]

Grounding Language Understanding

- Often language needs additional information to be understood
  - From contextual environment involving other modalities
  - Also referred to as the symbol grounding problem
- Example: deixis (person, spatial, temporal)



By Wesn - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=23637208

[Harnad, *The Symbol Grounding Problem*, Physica D: Nonlinear Phenomena, 1990]
[Lyons, *Deixis, Space and Time*, Semantics, 1977]