MARIO MAGUED MINA
UDS
25.06.2018
HOT AND ODD TOPICS IN SEMANTICS

METAPHORA

OUTLINE

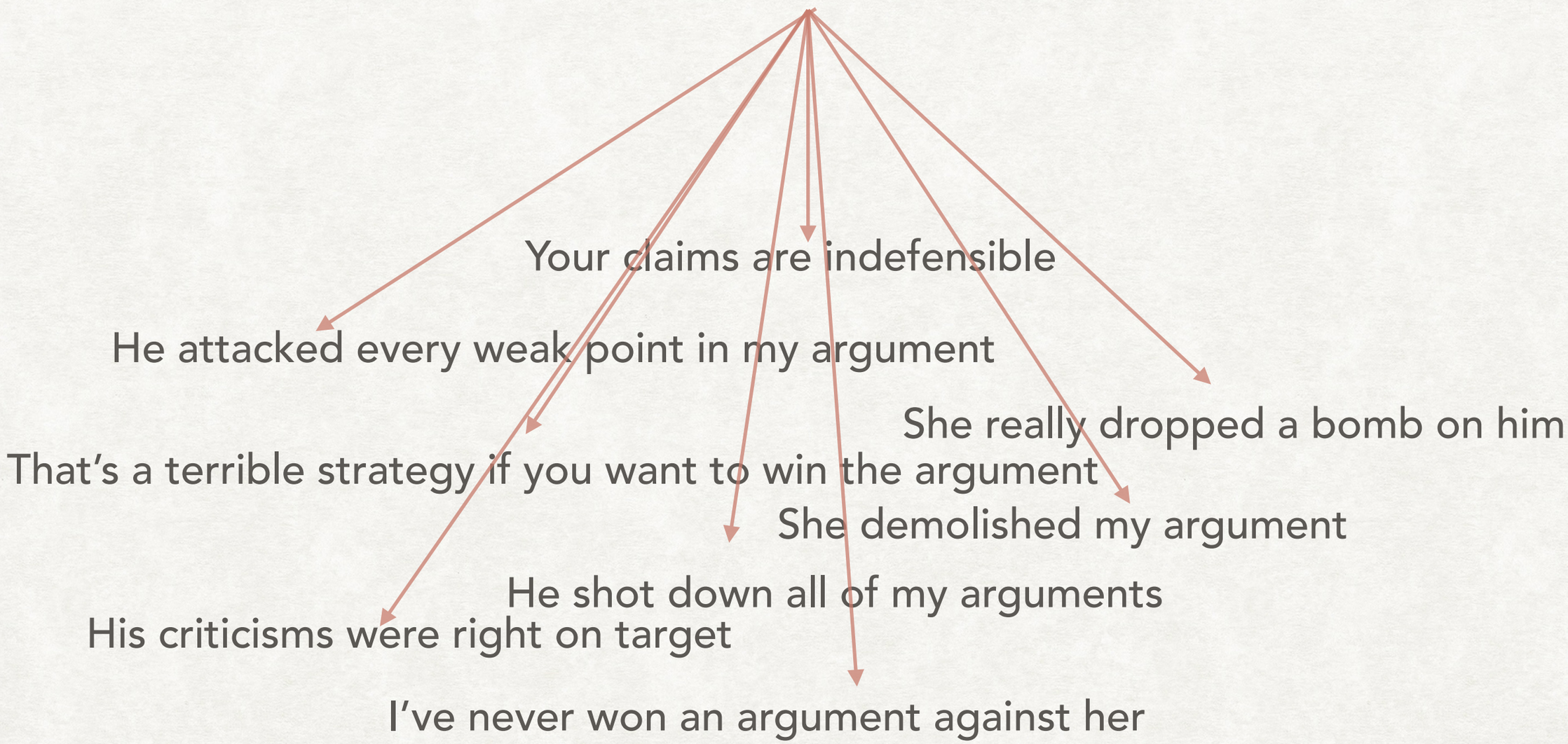
- **What exactly is metaphor or metaphora? What does linguistics have to say about it?**
- **Metaphor in computation**
  - Why does metaphor matter to us?
  - Design considerations
  - The task and techniques
- **Word Embedding and a WordNet Based Metaphor Identification and Interpretation (Mao, Lin and Frank(2018))**

WHAT DOES LINGUISTICS HAVE TO SAY ABOUT METAPHOR?

WHAT DOES LINGUISTICS HAVE TO SAY ABOUT METAPHORS

BASIC NOTIONS

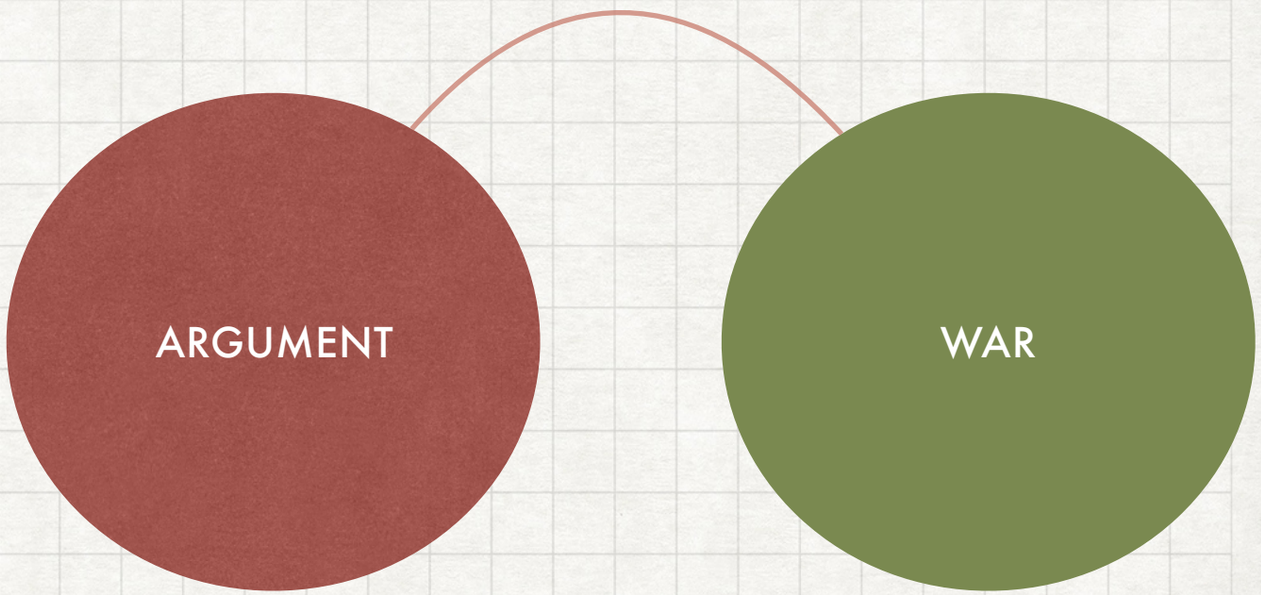
ARGUMENT is WAR



WHAT DOES LINGUISTICS HAVE TO SAY ABOUT METAPHORS

BASIC NOTIONS

- Conceptual mapping between two domains
- Conceptual structures govern our world view
- Visible through language
- Use one to talk about the other
- Culturally influenced, even individual



WHAT DOES LINGUISTICS HAVE TO SAY ABOUT METAPHORS

A "WONKY" METAPHOR

- *(regarding someone) Their brain is pure cashmere!*

WHAT DOES LINGUISTICS HAVE TO SAY ABOUT METAPHORS

A "WONKY" METAPHOR

"Since I don't understand sports, I have no metaphors to use for our business situation."

WHAT DOES LINGUISTICS HAVE TO SAY ABOUT METAPHORS

METAPHOR IN HUMAN COGNITION

- Deliberate flouting of rules
  - Psycholinguistic evidence shows that the details are a bit fuzzy
  - RT vs. EEG experiments
- Insight for computation
- Theta-role constraint

“

According to Sullivan(2007, 2013), conceptual metaphors are realised in language via mapping semantic roles in the source frame onto roles in the target frame. This suggests that both the source and the target domain impose a set of constraints on the roles that can be mapped, and thus on the syntactic options for expressing the metaphorical meaning.

— Shutova (2015)

”

WHAT DOES LINGUISTICS HAVE TO SAY ABOUT METAPHORS

SOME "WONKY" METAPHORS

- *Inspiration is a spark, you have to water it so that it would grow*

METAPHOR AND COMPUTATION

METAPHOR AND COMPUTATION

SURE, METAPHORS ARE INTERESTING, BUT WHY DOES IT MATTER TO US?

- Relevant for several Tasks
  - MT
  - Opinion Mining
  - Sentiment Analysis
  - IR
- Linguistic Metaphor is not explicit/misleading
- Ubiquitous, existing in ⅓ of utterances according to corpora

METAPHOR AND COMPUTATION

DESIGN CONSIDERATIONS

- Level of analysis
  - Conceptual
  - Linguistic
    - Conventionality
    - Syntactic constructions
  - Extended
  - Metaphorical Inference
- Applicability
- The task and techniques

METAPHOR AND COMPUTATION

DESIGN CONSIDERATIONS:LEVEL OF ANALYSIS

- Conceptual level
  - Underlying cognitive mechanism
  - Detecting source and target domains
  - Detecting domains via word-clusters(Shutova and Sun 2013), LDA (Heintz et al. 203)

METAPHOR AND COMPUTATION

DESIGN CONSIDERATIONS:LEVEL OF ANALYSIS

- Linguistic level
  - Central to most studies
  - Level of conventionality
    - Metaphor is a highly productive phenomenon
    - *They had to prune the workforce*
      (maybe not so accurate in this setting)
    - *Employers reaped enormous benefits from cheap foreign labour*
    - *He works for the local branch of the bank*
    - *There is a flourishing black market there*

METAPHOR AND COMPUTATION

DESIGN CONSIDERATIONS:LEVEL OF ANALYSIS

- Linguistic level
  - Syntactic constructions covered
  - Verbal and adjectival very common (e.g. *My car drinks gasoline, she is such a sweet person*)
  - Copula construction (*Bob is a brick*)
- Level of annotation
  - Sentence, relation, word

METAPHOR AND COMPUTATION

DESIGN CONSIDERATIONS:LEVEL OF ANALYSIS

- Extended metaphor
  - Metaphor in discourse
  - Chains of linguistics manifestations in one discourse
  - Same conceptual coupling
  - Interesting for discourse analysis
    - Underlying cultural and moral models

METAPHOR AND COMPUTATION

DESIGN CONSIDERATIONS:LEVEL OF ANALYSIS

- Metaphorical inference
  - Inferential mappings projected between domains
  - New knowledge derived from mappings
  - Emotional response
  - Highly relevant for study of metaphor
  - Not really popular for computational systems

METAPHOR AND COMPUTATION

DESIGN CONSIDERATIONS:APPLICABILITY

- Applicability
  - The whole idea is to support other NLP applications
  - Easily integrated
  - Able to run on large quantities of text
  - Not domain-specific
  - Deals with all syntactic constructions

METAPHOR AND COMPUTATION

DESIGN CONSIDERATIONS:TASKS AND TECHNIQUES

- Tasks: Identification vs. interpretation
- Different techniques
  - Selectional preference violation (remember the rule flouting for earlier?)
  - Topical structure of text
  - Supervised classification
  - Clustering
  - Web search

METAPHOR AND COMPUTATION

TECHNIQUES: SELECTIONAL PREFERENCES

- First implementation by Wilks (1975)
  - *My car drinks gasoline*
  - Intuition: detect violation of preference or rule flouting
  - Shutova (2013) extended to data-driven method to detect preferences from text
  - Did not perform well (0.17 precision, 0.55 recall)
  - Problematic aspects with technique

METAPHOR AND COMPUTATION

TECHNIQUES: SELECTIONAL PREFERENCES

- Topical structure
  - Metaphora has to do with different domains
  - One would expect to see this in topics
  - Requires extended discourse which is not always available

METAPHOR AND COMPUTATION

TECHNIQUES: SELECTIONAL PREFERENCES

- Supervised classification
  - Train classifier on manually annotated data
  - What features are indicative of metaphor?
  - Lexical, syntactic, semantic information
  - Animacy and concreteness

WORD EMBEDDING AND WORDNET BASED METAPHOR IDENTIFICATION (MAO, LIN AND FRANK 2018)

MAO, LIN AND FRANK (2018)

OUTLINE

- Introduction: Intuition
- Method: Hypotheses and Word Embeddings using Skip-gram and CBOW
- Results
- Application to MT
- Discussion

MAO, LIN AND FRANK (2018)

INTRODUCTION:INTUITION

- Take a sentence
- Go word for word and separate it (target word) from rest
- A candidate list is generated using hypernyms and synonyms of target word (from WordNet)
- Get best fit word using embeddings (literal sense)
  - Candidate list member with highest similarity to context
- Calculate similarity between best fit and target word
- Threshold determines metaphoricity

MAO, LIN AND FRANK (2018)

INTRODUCTION:INTUITION

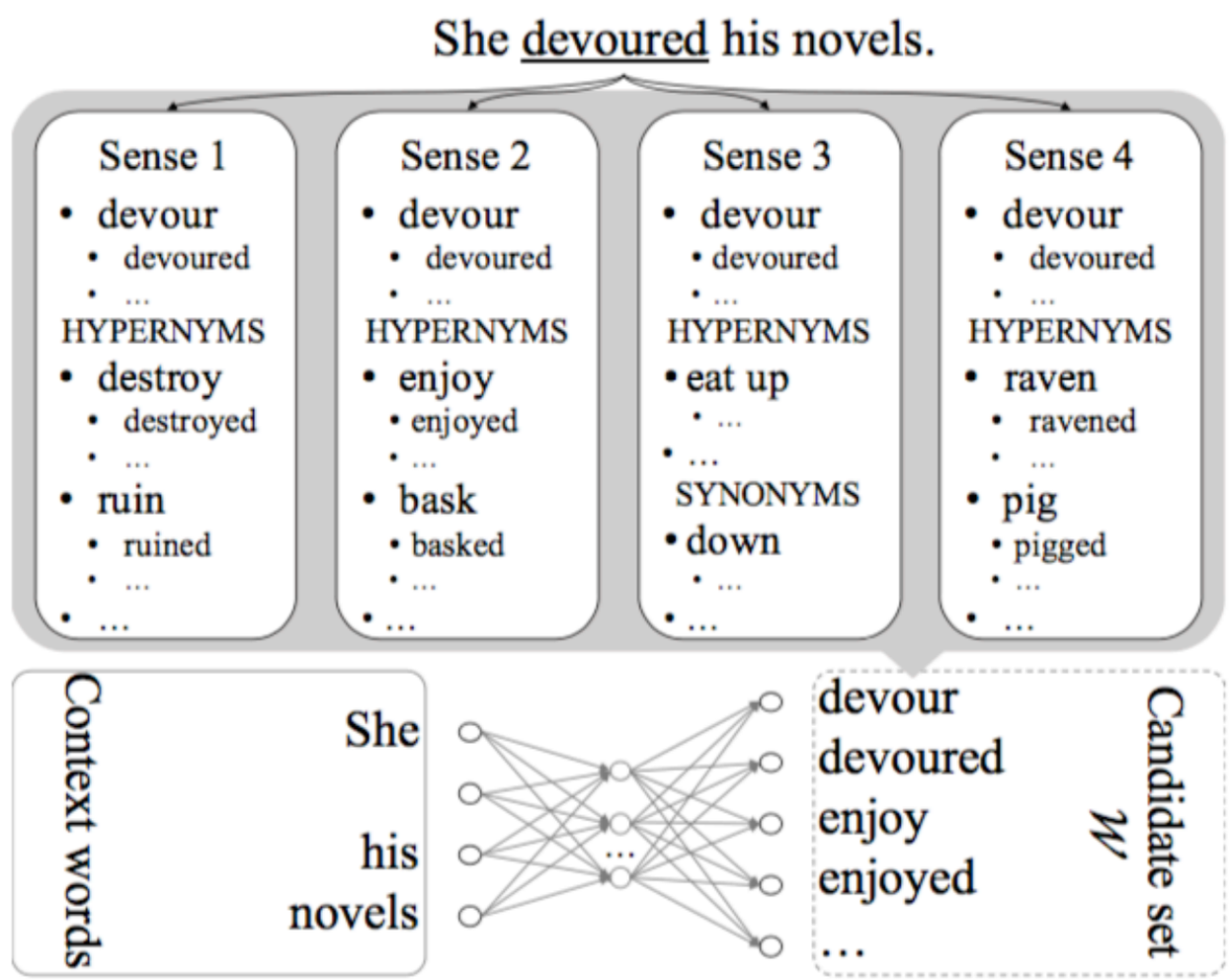


Figure 3: Given CBOW trained input and output vectors, a target word of *devoured*, and a context of *She* [ ] *his novels*, $cos(v^o_{devoured}, v^i_{context}) = -0.01$, $cos(v^o_{enjoyed}, v^i_{context}) = 0.02$.

MAO, LIN AND FRANK (2018)

METHOD: HYPOTHESES

- Hypothesis 1: Metaphor can be identified if sense in context and literal sense are from different domains
- Hypothesis: Literal prevalence in data

MAO, LIN AND FRANK (2018)

METHOD: WORD EMBEDDINGS

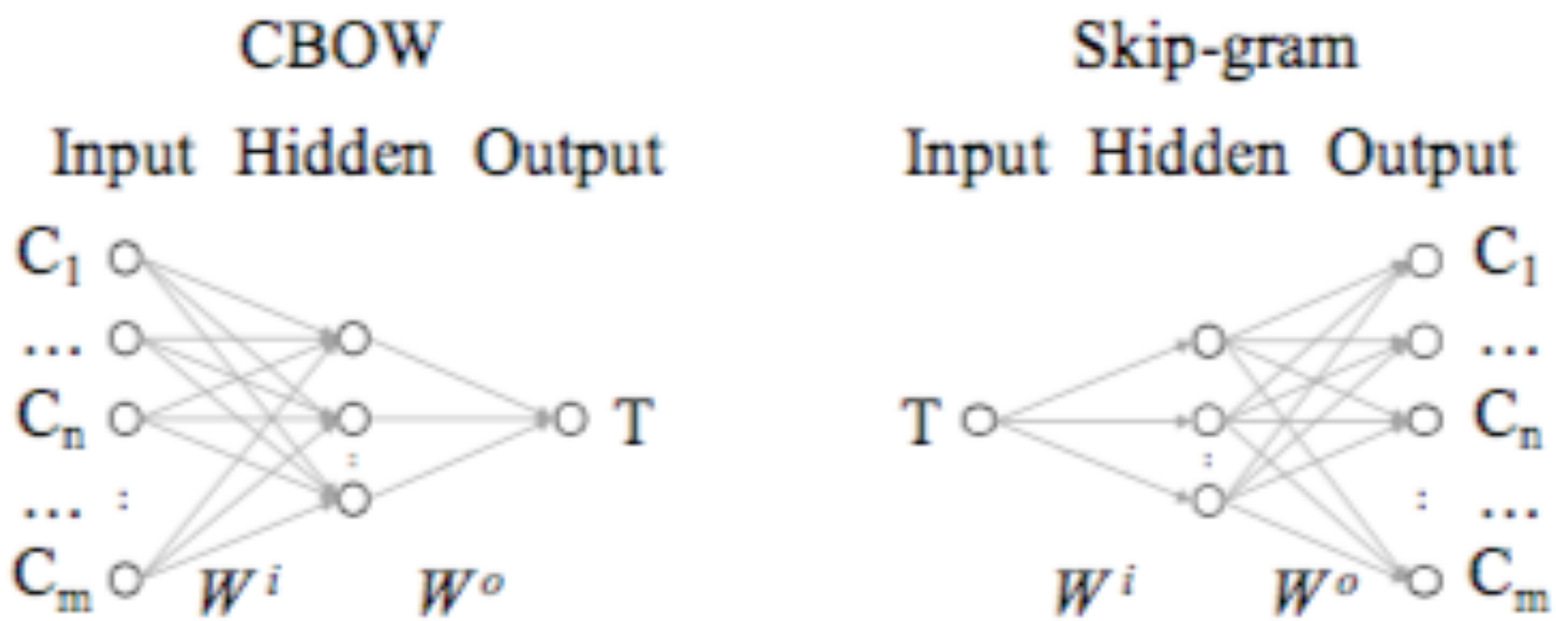


Figure 1: CBOW and Skip-gram framework.

MAO, LIN AND FRANK (2018)

METHOD: WORD EMBEDDINGS

- Continuous Bag of Words (CBOW)
  - Maximises probability of centre word given context
    - Cats ____ dogs
  - Idea is that it averages out word (using number of context words) vectors of context to obtain context vector
  - Outputs expected centre word
- Skip-gram
  - Basically the same but the other way around
  - _____ and _____
- Input vectors and output vectors

MAO, LIN AND FRANK (2018)

METHOD: WORD EMBEDDINGS

- Both input and output vectors were explored
- Mixing input and output vectors to compute similarity expected to be efficient

MAO, LIN AND FRANK (2018)

METHOD: WORD EMBEDDINGS

Drink orange juice

Drink apple juice

- Inputs for *orange* and *apple* similar
- Outputs for *orange* and *apple* similar
- Other way around for *orange/apple* and context

MAO, LIN AND FRANK (2018)

RESULTS

- Compared to several strong baselines
  - Unsupervised word embedding based model (Shutova et al. 2016)
  - Context2Vec
- Dataset
  - Taken from Mohammad et al. (2016)
  - 1230 literal and 409 metaphorical sentences
  - Tested for sentence level and phrase level (subject verb and verb object pairs)
    - 331 literal 316 metaphoric
  - 40/40 for development and selecting a threshold

MAO, LIN AND FRANK (2018)

RESULTS

	Method	P	R	F1
Phrase	Shutova et al. (2016)	0.67	0.76	0.71
	Rei et al. (2017)	**0.74**	0.76	**0.74**
	SIM-CBOW$_{I+O}$	0.66	0.78	0.72
	SIM-SG$_{I+O}$	0.68	**0.82**	**0.74***
Sent.	Melamud et al. (2016)	0.60	0.80	0.69
	SIM-SG$_{I}$	0.56	**0.95**	0.70
	SIM-SG$_{I+O}$	0.62	0.89	0.73
	SIM-CBOW$_{I}$	0.59	0.91	0.72
	SIM-CBOW$_{I+O}$	**0.66**	0.88	**0.75***

Table 1: Metaphor identification results. NB: * denotes that our model outperforms the baseline significantly, based on two-tailed paired t-test with $p < 0.001$.

“

Skip-gram is trained by using a centre word to maximise the probability of each context word, whereas CBOW uses the average of context word input vectors to maximise the probability of the centre word. Thus, Skip-gram performs better in modelling one-word context, while CBOW has better performance in modelling multi-context words.

— Mao et al. (2018)

”

MAO, LIN AND FRANK (2018)

RESULTS

Table 8
Identification systems: Measures used and results. * The results of Gedigian et al. (2006) should be interpreted with a reference to the performance of an *all metaphor* baseline attaining 0.92. ** Turney et al. (2011) report results on the verb dataset in terms of F-score and on the adjective dataset in terms of accuracy. LM stands for linguistic metaphor and CM for conceptual metaphor.

System	Precision	Recall	F-score	Acc	Lim.	Open
Mason (2004)	–	–	–	0.77	–	✔
Birke and Sarkar (2006)	–	–	0.54	–	–	✔
Gedigian et al. (2006)	–	–	–	0.95*	✔	–
Krishnakumaran, Zhu (2007)	–	–	–	0.58	✔	–
Shutova et al. (2010)	0.79	–	–	–	–	✔
Li and Sporleder (2010)	–	–	0.75	0.78	–	✔
Turney et al. (2011)	–	–	0.68**	0.79**	–	✔
Neuman et al. (2013)	0.71	0.43–0.97	–	–	✔	–
Dunn (2013)	–	–	0.58	–	–	✔
Tsvetkov et al. (2013; 2014)	0.78	0.79	0.78	–	–	✔
Mohler et al. (2013)	0.56	0.93	0.7	–	✔	–
Hentz et al. (2013)	0.54	0.64	0.59	–	✔	–
Hovy et al. (2013)	0.7	0.8	0.75	0.75	–	✔
Wilks et al. (2013)	0.57	0.82	0.67	–	✔	–
Strzalkowski et al. (2013)	–	–	–	0.71	✔	–
Shutova and Sun (2013)	0.65 (LM); 0.69 (CM)	0.61 (CM)	–	–	–	✔
Gandy et al. (2013)	0.76 (LM); 0.65 (CM)	0.82 (LM)	–	–	✔	–
Li et al. (2013)	0.65–0.73	0.52–0.66	0.58–0.69	–	–	✔
Shutova (2013)	0.68	0.66	0.67	–	–	✔

MAO, LIN AND FRANK (2018)

RESULTS

- Whole point is to support other tasks
- Application to MT (English to Chinese)
- Sentence level metaphor identification
- From the test set, 200 literal and 200 metaphoric sentences were taken, 50 and 50 selected at random
- Run through system
  - If literal, no change
  - If metaphorical, target word is replaced with best fit
  - Fed to Google translate and Bing

MAO, LIN AND FRANK (2018)

RESULTS

- 5 native speakers of Chinese were enlisted
- Asked to evaluate if translation could represent its sense within translated sentence
- Note: just translation of target word was taken into account
- Translation accuracy is used as a metric
- Label taken as true if more than half agree

MAO, LIN AND FRANK (2018)

RESULTS

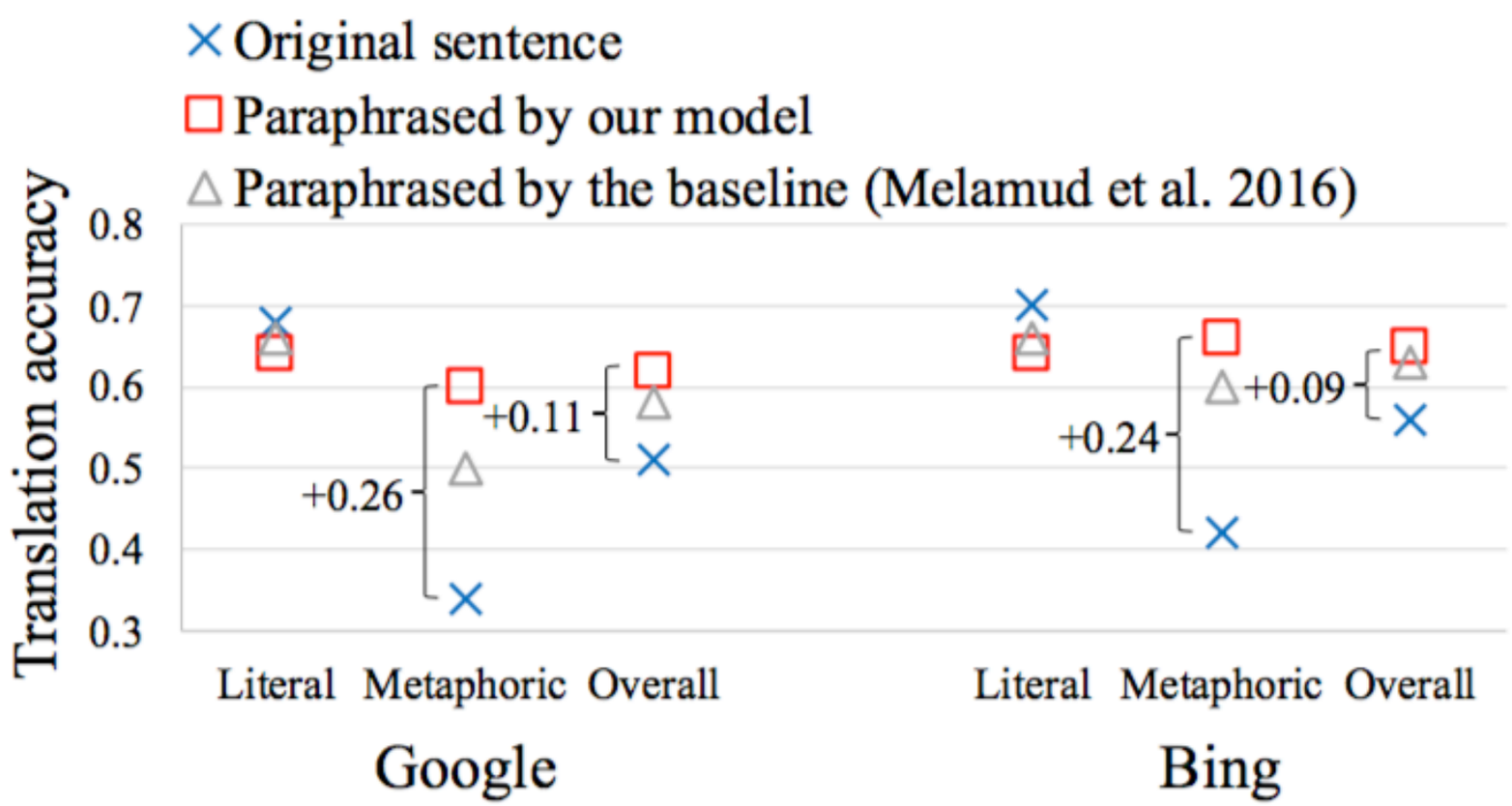


Figure 5: Accuracy of metaphor interpretation, evaluated on Google and Bing Translation.

MAO, LIN AND FRANK (2018)

DISCUSSION

- Linguistic
  - All constructions
  - More novel
- Applicability
  - The whole idea is to support other NLP applications
  - Easily integrated
  - Able to run on large quantities of text
  - Not domain-specific
  - Deals with all syntactic constructions

Probably???

MAO, LIN AND FRANK (2018)

DISCUSSION

- Takeaway
  - Taking metaphor into account improves NLP application performance
  - Input-output vectors

QUESTIONS AND WHATNOT?