Semantic matching against a corpus: new applications and methods

Lucy H. Lin, Scott Miles, Noah A. Smith

27 April 2018

Introduction

How can we apply recent advances in broad-coverage modeling of sentential semantics?

$\Rightarrow$ Match natural language propositions against a corpus.

Possible end user...	**tracking occurrences of...**
Historian of science	“vaccines cause autism”
Political scientist	“immigrants are used as scapegoats for problems in society”
Public servant	“dealing with authorities is causing stress and anxiety”

Introduction

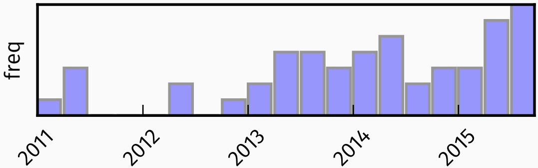
Proposition query:
"Dealing with authorities is causing stress and anxiety."

query corpus

Matched sentences:
"Unfamiliar bureaucratic systems are causing the majority of the stress."
"Those in charge of recovery are making moves to appease the growing anger among homeowners."

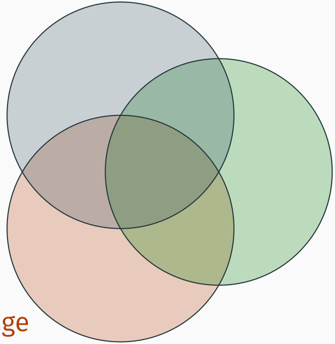
aggregate

Frequency across time:



Related to...

paraphrase (Dolan et al., 2004),
entailment (Dagan et al., 2006),
semantic similarity (Agirre et al., 2012)



information retrieval,
passage retrieval for QA
(Tellex et al., 2003)

dynamics of language
across a corpus
(e.g., Blei & Lafferty, 2006)

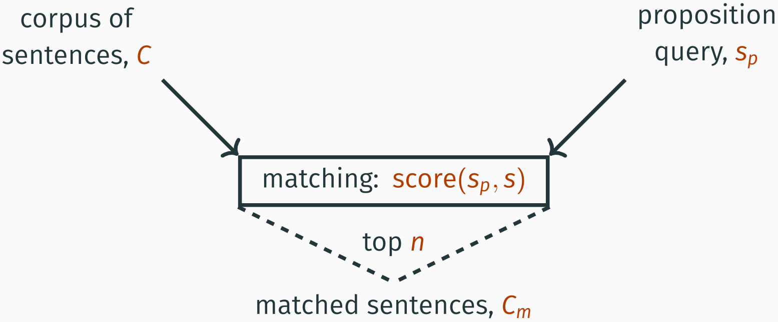
Outline

1. Introduction
2. Problem formulation
3. Matching exemplars from a codebook
   Domain: media framing of policy issues
4. Matching specific expert queries
   Domain: analysis of disaster recovery
5. Using semantic matching output for measurement
6. Discussion/conclusions

Outline

1. Introduction
2. Problem formulation
3. Matching exemplars from a codebook
   Domain: media framing of policy issues
4. Matching specific expert queries
   Domain: analysis of disaster recovery
5. Using semantic matching output for measurement
6. Discussion/conclusions

Problem formulation



where $\text{score}(s_p, s)$ should be high iff s expresses s_p

Outline

1. Introduction
2. Problem formulation
3. Matching exemplars from a codebook
   Domain: media framing of policy issues
4. Matching specific expert queries
   Domain: analysis of disaster recovery
5. Using semantic matching output for measurement
6. Discussion/conclusions

Matching exemplars from a codebook

Inputs:

- C: Media Frames Corpus (Card et al., 2015)
  - thousands of news articles on immigration (and other policy issues)
  - spans of text annotated with *framing dimensions*
- s_p: 30 annotation codebook examples
  e.g.: "immigration rules have changed unfairly over time"
  evokes the *fairness and equality* frame

Note: many ways to evoke a frame outside of the codebook.

Matching exemplars from a codebook: scoring

Scoring function $f(s_p, s)$:

1. each sentence is the average of its word vectors
2. cosine similarity between two sentence vectors $\rightarrow$ score

We use two word vector variants (300D):

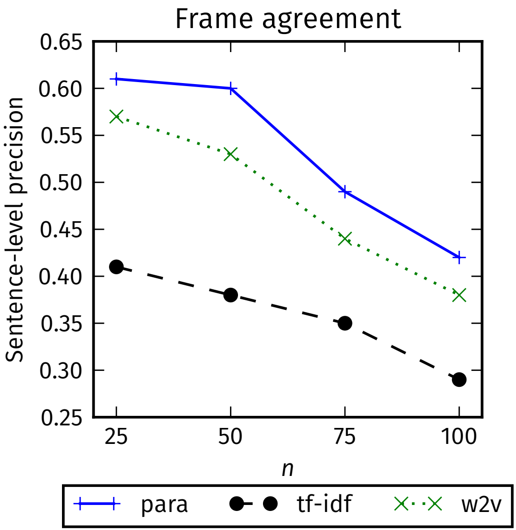
- paraphrastic word vectors (Wieting et al., 2016)
- word2vec (Mikolov et al., 2013) pretrained on Google News

Matching exemplars from a codebook: evaluation/results

How well does output align with corpus annotations?

Finding:

- paraphrastic > word2vec > tf-idf



Outline

1. Introduction
2. Problem formulation
3. Matching exemplars from a codebook
   Domain: media framing of policy issues
4. Matching specific expert queries
   Domain: analysis of disaster recovery
5. Using semantic matching output for measurement
6. Discussion/conclusions

Matching specific expert queries

Motivation: Researchers, public servants want to understand community challenges post-disaster.

Inputs:

- C: 982 NZ news articles after the 2010/2011 earthquakes
- s_p: 20 queries provided by domain expert,
  covering community wellbeing, infrastructure,
  and decision-making

  e.g.: "The council should have consulted residents before making decisions."

$\Rightarrow$ more fine-grained matching.

Matching specific expert queries: scoring

Scoring function $m(s, s_p)$:

$\Rightarrow$ dependency parses T, T_p

$\Rightarrow$ sequence of tree edit operations transforming T into T_p
(Heilman and Smith, 2010)

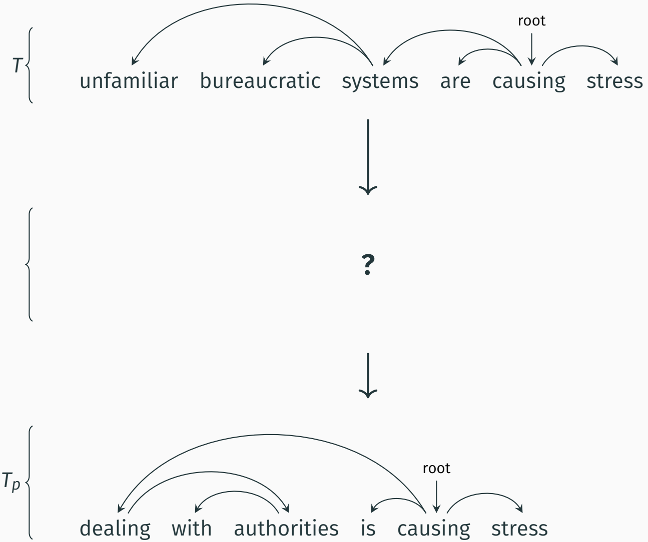
$\Rightarrow$ classify as match/non-match using tree edit sequence

Matching specific expert queries: scoring

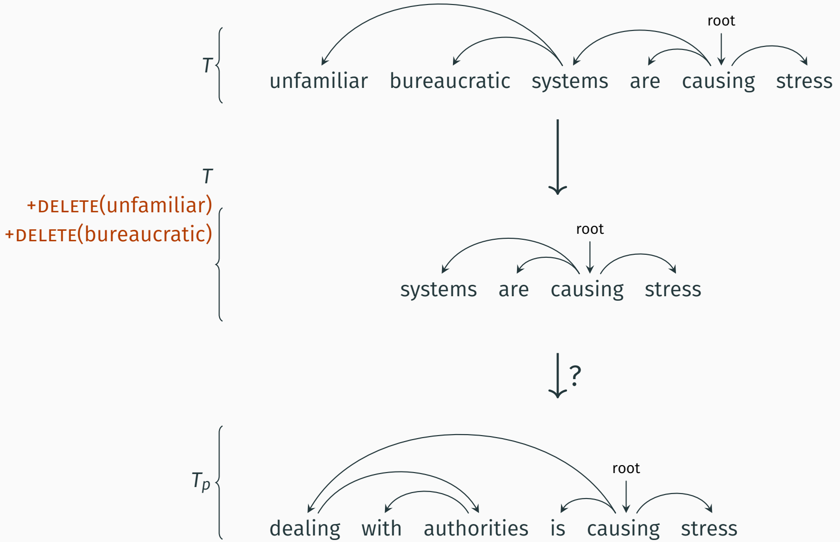
Scoring function $m(s, s_p)$:

- $\Rightarrow$ dependency parses T, T_p
- $\Rightarrow$ sequence of tree edit operations transforming T into T_p (Heilman and Smith, 2010)
- $\Rightarrow$ classify as match/non-match using tree edit sequence

Matching specific expert queries: scoring

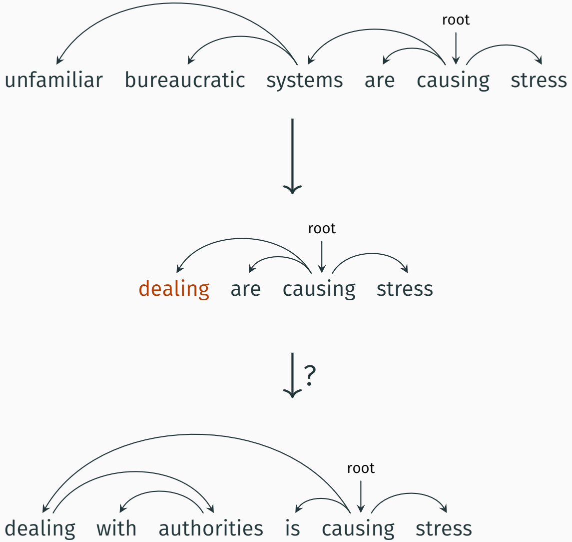


Matching specific expert queries: scoring



Matching specific expert queries: scoring

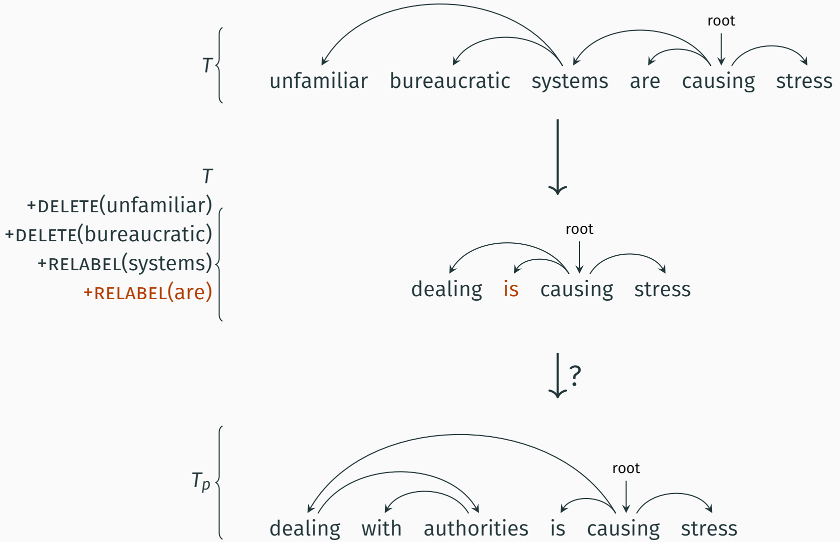
T



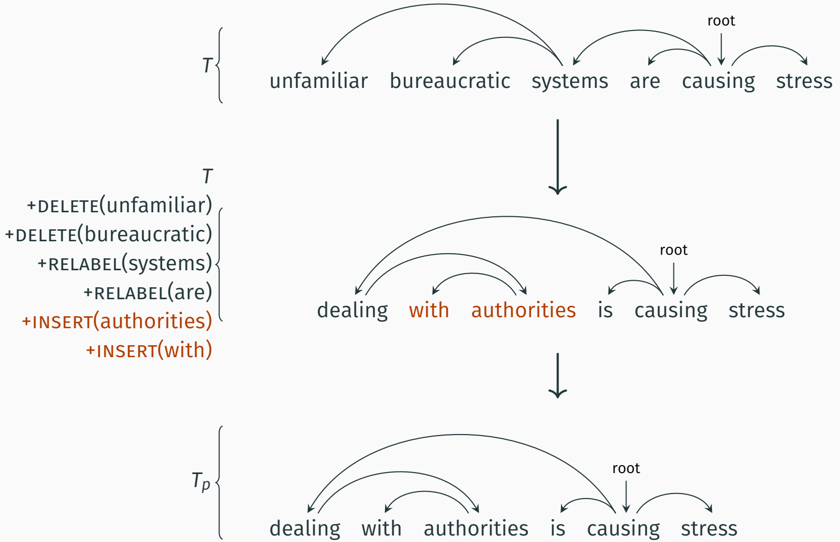
T
+DELETE(unfamiliar)
+DELETE(bureaucratic)
+RELABEL(systems)

T_p

Matching specific expert queries: scoring



Matching specific expert queries: scoring



Matching specific expert queries: scoring

Scoring function $m(s, s_p)$:

$\Rightarrow$ dependency parses T, T_p

$\Rightarrow$ sequence of tree edit operations transforming T into T_p (Heilman and Smith, 2010)

$\Rightarrow$ classify as match/non-match using tree edit sequence

- logistic regression (39 features), LSTM
- trained on SNLI (entail vs. neutral/contradiction)

Matching specific expert queries: scoring

For reasonable runtime:

1. f: word-vector-based matching on C
   $\Rightarrow$ obtain top k matches (C_f)
2. m: entailment-based model on just C_f
   $\Rightarrow$ obtain top n matches (C_m)

Four combinations of f, m:

- w2v, LR
- w2v, LSTM(w2v)
- para, LR
- para, LSTM(para)

Matching specific expert queries: evaluation

User study:

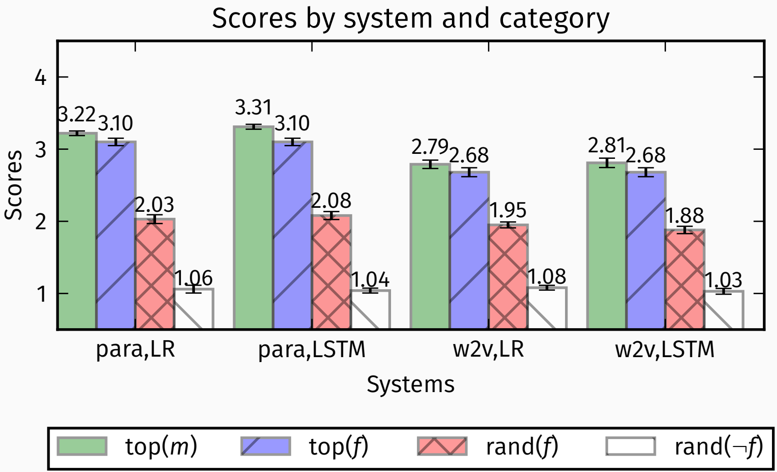
- surveyed 20 emergency managers
- output of the four f, m combinations obtained per query
- candidate sentences rated 1-5
  (Krippendorf's $\alpha = 0.784$)

Matching specific expert queries: evaluation

Example s_p: **There is a shortage of construction workers.**

Score	Example candidate
1	The quarterly report for Canterbury included analysis on Greater Christchurch Value of Work projections.
3	The construction sectors workload was expected to peak in December.
5	Greater Christchurchs labour supply for the rebuild was tight and was likely to remain that way.

Matching specific expert queries: study results



Findings: paraphrastic > word2vec;

m is useful and LSTM > LR if using paraphrastic vectors

Matching specific expert queries: evaluation

Other feedback:

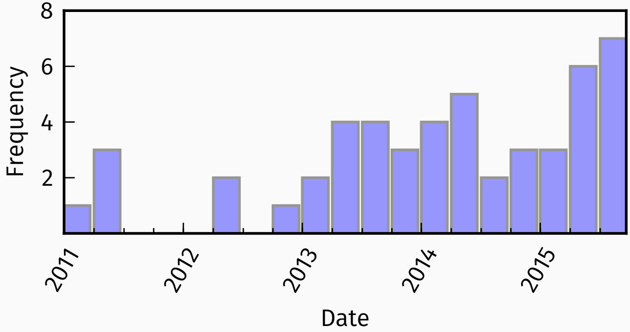
- 17/20 respondents interested in a way to match ideas in news or other text corpora
- Half of respondents interested in follow-up study, with their own idea queries (in progress!)

Outline

Semantic measurement

Example measurement:

s_p: “Dealing with authorities is causing stress and anxiety.”



Outline

1. Introduction
2. Problem formulation
3. Matching exemplars from a codebook
   Domain: media framing of policy issues
4. Matching specific expert queries
   Domain: analysis of disaster recovery
5. Using semantic matching output for measurement
6. Discussion/conclusions

Discussion

Idea complexity:

- queries that are too general/too specific
- user guidance for writing good queries?

Entities/coreference: e.g., “Cera”

Sentence-related issues:

- surrounding context invalidates a match
- potential match spread across a sentence boundary

In conclusion...

In this talk, we:

- Demonstrated viability of semantic matching methods in two different domains
- Performed a user study to establish end user interest
- Motivated future work on semantic matching/measurement applications

Thanks!

(contact: lucylin@cs.washington.edu)

(more slides)

Tree edit classifier details

Original model (Heilman and Smith, 2010):

- Extract 39 integer features from tree edit sequence: sequence length, counts of edit types
- Logistic regression (LR) $\rightarrow m(s_p, s)$

Tree edit classifier details

New variation: input tree edit sequence into a LSTM

Each operation in the sequence is vectorized as:

- One-hot encoding of the operation type
- Word vector Δ between the sentences pre- and post-operation
    - insert $\rightarrow$ word embedding of new word
    - relabel $\rightarrow$ difference between word embeddings
    - delete $\rightarrow$ negated word embedding of deleted word

Tree edit classifier details

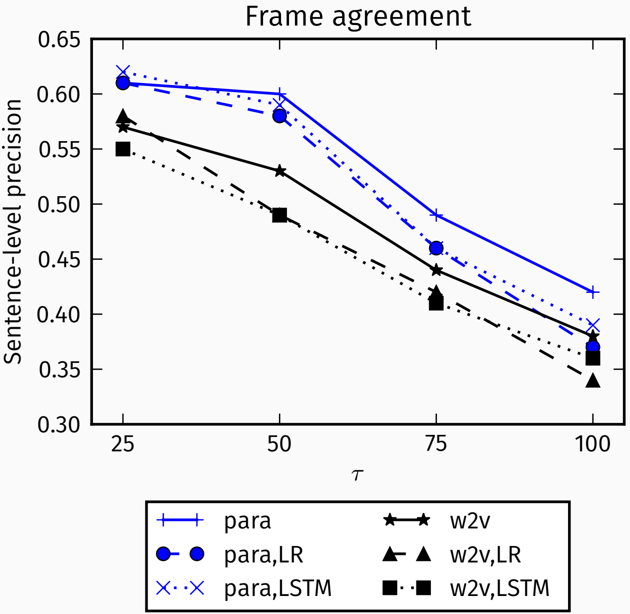
Training: SNLI corpus (Bowman et al., 2015)

- 570k pairs of sentences
- labels: entailment, contradiction, neutral
- e.g.,: “A soccer game with multiple males playing.” *entails* “Some men are playing a sport.”

Mapping to our problem:

- $s \rightarrow$ premise, $s_p \rightarrow$ hypothesis
- match $\rightarrow$ entailment,
  non-match $\rightarrow$ contradiction/neutral

Media frames results (w/tree edit models)



(Some) media frames queries

Punishments should be softer on immigration. (crime+punishment)

Immigrants are taking over the country. (cultural identity)

Immigrants work for less money, driving the wages down for domestic workers. (economic)

Immigrants aid law enforcement by acting as witnesses. (health+safety)

Right to work does not mean right to cross national borders. (legality)

It would be immoral to turn our backs on those in need. (morality)

Businesses have a legitimate interest in lobbying for immigration issues. (politics)

The public supports immigration rights. (public sentiment)

Immigrants drive up the cost of living. (quality of life)

Disaster recovery queries

Residents are frustrated by the slow pace of recovery.

The repair programme is on schedule to be completed.

Money for repairs is running out.

The council should have consulted residents before making decisions.

Mental health rates have been rising.

Dealing with authorities is causing stress and anxiety.

Most eligible property owners have accepted insurance offers.

Confidence in Cera has been trending downwards.

Water quality declined after the earthquakes.

The power system was fully restored quickly.

Disaster recovery queries

Cera missed several recovery milestones.

Prices levelled off as more homes were fixed or rebuilt.

People are suffering because they've lost the intimacy of their relationships.

Coordination between rebuild groups has been problematic.

Few people said insurance companies had done a good job.

Having the art gallery back makes the city feel more whole.

Scirt has spent less money than predicted.

Traffic congestion was severe due to road repairs.

Some of the businesses forced out by the earthquake are returning.

Some of the burden on mental health services is caused by lack of housing.