Temporality and Modality in Entailment Graph Learning

Liane Guillou

with Sander Bijl de Vroe, Mohammad Javad Hosseini, Miloš Stanojević, Nick McKenna, Mark Johnson, Mark Steedman
1. Introduction

2. Learning Entailment Graphs

3. Temporality

4. Modality

5. Summary
Introduction
The SEMANTAX Project

- Aim: Learn entailments between predicates from raw text

**Example: buy → own**

Google *bought* YouTube for $1.65 billion. Google *owns* YouTube and it has proven to be an amazingly successful purchase.

- Use entailment information in downstream applications:
  - Question Answering
  - Knowledge Graph population
• Question: Did Arsenal \textcolor{red}{play} Man United last night?
Recognising Textual Entailment and Question Answering

- Question: Did Arsenal play Man United last night?

<table>
<thead>
<tr>
<th>Match Report</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Arsenal beat Man United 1-0”</td>
</tr>
</tbody>
</table>
• Question: Did Arsenal play Man United last night?

Match Report

“Arsenal beat Man United 1-0”

• To answer, we must know the entailment relation:

beat → play

TeamA beats TeamB → TeamA plays TeamB
Learning Entailment Graphs
Entailment Graphs

- Nodes: predicates (e.g. play, win, lose)
- Edges: entailment relations
- For multiple type pairs, e.g. ORG-ORG for sports teams
Entailment Graphs

- Nodes: predicates (e.g. play, win, lose)
- Edges: entailment relations
- For multiple type pairs, e.g. ORG-ORG for sports teams
- Learned from large corpora of multi-source news text
  - Authors use different language to describe the *same* event
- **Unsupervised method** of Hosseini et al. (2018)
Learning Entailment Relations

• Learning signal: Distributional Inclusion Hypothesis (Dagan et al., 1999; Geffet and Dagan, 2005):

  A predicate $p$ entails another predicate $q$ if for any context in which $p$ can be used, $q$ may be used in its place.

### Example: co-occurrences with argument pairs

<table>
<thead>
<tr>
<th></th>
<th>Arsenal-Man U.</th>
<th>Arsenal-Chelsea</th>
<th>Chelsea-Spurs</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>win</strong></td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td><strong>play</strong></td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>...</td>
</tr>
</tbody>
</table>

• Distribution of **win** is *included* in distribution of **play**

• Compute similarity between **win** and **play**
Challenge: Spurious Entailments

- This can fail for some highly correlated, contradictory relations: win → lose etc.

**Example: co-occurrences with argument pairs**

<table>
<thead>
<tr>
<th></th>
<th>Arsenal-Man U.</th>
<th>Arsenal-Chelsea</th>
<th>Chelsea-Spurs</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>lose</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>win</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>play</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>...</td>
</tr>
</tbody>
</table>

- How can we avoid learning spurious entailment relations?
• This can fail for some highly correlated, contradictory relations: win → lose etc.

Example: co-occurrences with argument pairs

<table>
<thead>
<tr>
<th></th>
<th>Arsenal-Man U.</th>
<th>Arsenal-Chelsea</th>
<th>Chelsea-Spurs</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>lose</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>win</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>play</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>...</td>
</tr>
</tbody>
</table>

• How can we avoid learning spurious entailment relations?

1. **Temporality**: Compare eventualities that happen *at the same time*
• This can fail for some highly correlated, contradictory relations: win → lose etc.

Example: co-occurrences with argument pairs

<table>
<thead>
<tr>
<th></th>
<th>Arsenal-Man U.</th>
<th>Arsenal-Chelsea</th>
<th>Chelsea-Spurs</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>lose</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>win</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>play</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>...</td>
</tr>
</tbody>
</table>

• How can we avoid learning spurious entailment relations?

1. **Temporality**: Compare eventualities that happen *at the same time*
2. **Modality**: Exclude eventualities that are *uncertain* to happen
Temporality
Overview: Temporality

- **Problem**: observed *spurious entailments* between disjunctive outcomes e.g. win → lose

- **Approach**: use information about when events took place to refine the learning process

- Initial experiments focused on sports domain

- Later experiments applied the technique to the general domain
Arsenal - played and lost against - Man United 1-3 (25/01/2019)
Arsenal - played and beat - Man United 2-0 (10/03/2018)
Arsenal - played and tied with - Man United 1-1 (30/09/2019)

Aim: Learn entailments: win/lose → play
Avoid learning spurious entailments: win → lose
Adding Temporal Information

- Extract binary relations with eventuality **start/end times**:

  \[
  \text{arg1} - \text{predicate} - \text{arg2} \quad \text{time interval}
  \]

  Arsenal - tied with - Man United  \( (30/09/19, 30/09/2019) \)

- Two temporal information sources:
  - Document creation date
  - Automatically resolve temporal expressions in the text, e.g.

  **Manchester United vs. Arsenal | 30th September 2019**

  Manchester United and Arsenal played to a 1-1 draw in a sloppy, rain-soaked match at Old Trafford **on Monday** night.
Relation Extraction (with MoNTEE)

- Article Sentences
- CoreNLP (Tokenisation, NER, POS tagging)
- CCG Parser (Extract CCG trees and dependencies)

Per sentence:
- Construct dependency graph
- Extract unary relations (traverse graph from verb node to argument)
- Extract binary relations (combine unaries)

Per relation:
- Type relation arguments (AIDAlight, FIGER mapping)
- Add time interval (SUTime / document date)

Tagged relations
Arsenal beat Man United last night.
Relation Extraction (with MoNTEE)

1. Article Sentences
   - CoreNLP (Tokenisation, NER, POS tagging)
   - CCG Parser (Extract CCG trees and dependencies)

Per sentence
- Construct dependency graph
- Extract unary relations (traverse graph from verb node to argument)
- Extract binary relations (combine unaries)

Per relation
- Type relation arguments (AIDALight, FIGER mapping)
- Add time interval (SUTime / document date)

Tagged relations

Arsenal beat Man United last night.

beat

Arsenal
Relation Extraction (with MoNTEE)

Article Sentences

CoreNLP
(Tokenisation, NER, POS tagging)

CCG Parser
(Extract CCG trees and dependencies)

Per sentence

Construct dependency graph

Extract unary relations
(traverse graph from verb node to argument)

Extract binary relations
(combine unaries)

Per relation

Type relation arguments
(AIDALight, FIGER mapping)

Add time interval
(SUTime / document date)

Tagged relations

Arsenal beat Man United last night.

beat

Man United

2
Relation Extraction (with MoNTEE)

Bijl de Vroe et al. (2021)

Arsenal beat Man United last night.
Relation Extraction (with MoNTEE)

Bijl de Vroe et al. (2021)

Arsenal beat Man United last night.

(beat.1, beat.2)

Arsenal: sports_team
Man United: sports_team
Time: 17/11/2021 - 17/11/2021
• Filter out co-occurrence counts where there is no temporal overlap between events

Filtering Algorithm

```python
for a in argPairs:
    for p in predicates:
        for q in predicates:
            count_{p, q} += co-occur(a, p, q);
            filteredCount_{p, q} += temporal_overlap(a, p, q);
```

**co-occur**: count of predicate $p$, given that $q$ also occurs with argPair $a$

**temporal_overlap**: number of events of $p$ that overlap with any event of $q$ (for argPair $a$)
Two matches between Arsenal and Man United:

- Arsenal played against and beat Man United (10/03/2018)
- Arsenal played against and lost to Man United (25/01/2019)

<table>
<thead>
<tr>
<th>predicate</th>
<th>count</th>
</tr>
</thead>
<tbody>
<tr>
<td>play against</td>
<td>2</td>
</tr>
<tr>
<td>beat</td>
<td>1</td>
</tr>
<tr>
<td>lose to</td>
<td>1</td>
</tr>
</tbody>
</table>
Two matches between **Arsenal** and **Man United**:

**Arsenal** played against and **beat** **Man United** (10/03/2018)

**Arsenal** played against and **lost to** **Man United** (25/01/2019)

<table>
<thead>
<tr>
<th>predicate</th>
<th>count</th>
<th>entailment pair</th>
<th>regular</th>
<th>filtered</th>
</tr>
</thead>
<tbody>
<tr>
<td>play against</td>
<td>2</td>
<td>beat → play against</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>beat</td>
<td>1</td>
<td>lose to → play against</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>lose to</td>
<td>1</td>
<td>beat → lose to</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>lose to → beat</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Temporal Similarity Measures

- *Similarity measures* determine whether predicates in the graph entail each other

- Temporal measures inspired by BINC (Szpektor and Dagan, 2008)
  - Directional component: Weed’s precision
  - Symmetrical component: Lin’s similarity

<table>
<thead>
<tr>
<th>Measure</th>
<th>Directional</th>
<th>Symmetrical</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. BINC BINARY</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>T. BINC RATIO</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>T. WEED’S PRECISION</td>
<td>✓</td>
<td>X</td>
</tr>
</tbody>
</table>

- Baseline: BInc (atemporal)
Evaluation

- Previous work evaluated on Levy/Holt (Levy and Dagan, 2016; Holt, 2018)
  - Does not test for antonymous non-entailments e.g. win $\not\rightarrow$ lose
  - Poorly balanced: many paraphrases, few directional examples

- Semi-automatically constructed two new datasets:
  - Sports - sports domain
  - ANT - general domain
Sports Entailment Dataset

- Manually construct paraphrase clusters: win, lose, tie, play from predicates in the training data
- Automatically construct entailment pairs according to patterns:

<table>
<thead>
<tr>
<th></th>
<th>win</th>
<th>lose</th>
<th>tie</th>
<th>play</th>
</tr>
</thead>
<tbody>
<tr>
<td>win</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>lose</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>tie</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>play</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

1 = entailment
0 = non-entailment

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>directional entailment 1</td>
<td>defeat $\rightarrow$ face</td>
<td>272</td>
</tr>
<tr>
<td>antonym 0</td>
<td>beat $\not\rightarrow$ fall to</td>
<td>446</td>
</tr>
<tr>
<td>directional non-entailment 0</td>
<td>play $\not\rightarrow$ win</td>
<td>272</td>
</tr>
<tr>
<td>paraphrase 1</td>
<td>defeat $\leftrightarrow$ outplay</td>
<td>322</td>
</tr>
</tbody>
</table>

Guillou et al. (2020)
- Extract *antonymous predicate pairs* and *synonyms* from Wordnet
- For each antonym pair (A1, A2), identify a set of predicates (E) entailed by all elements in U(A1, A2)
- Automatically construct entailment pairs according to patterns:

<table>
<thead>
<tr>
<th>A1</th>
<th>A2</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

1 = entailment
0 = non-entailment

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>directional entailment 1</td>
<td>acquitted $\rightarrow$ accused</td>
<td>1,465</td>
</tr>
<tr>
<td>antonym 0</td>
<td>acquitted $\nrightarrow$ convicted</td>
<td>1,800</td>
</tr>
<tr>
<td>directional non-entailment 0</td>
<td>accused $\nrightarrow$ convicted</td>
<td>1,465</td>
</tr>
<tr>
<td>paraphrase 1</td>
<td>acquitted $\leftrightarrow$ absolved</td>
<td>1,570</td>
</tr>
</tbody>
</table>
Experiments

Data:

- NewsSpike: multi-source news corpus, 0.5M articles, spanning ~6 weeks (Zhang and Weld, 2013)
- Extract relation triples. Approx. 19% time-stamped with SUTime

Experiments:

1. Sports: Temporal info source: doc date / time expressions / both
2. Sports: Add a uniform temporal window: N days
3. General: Add a dynamic per-predicate window with TacoLM (Zhou et al., 2020)

Evaluation:

- Compare using AUC score: area under precision-recall curve
- Points on the curve = different entailment score thresholds

Guillou et al. (2020)
## Results: Temporal Information Source

<table>
<thead>
<tr>
<th>Similarity measure</th>
<th>timexOnly</th>
<th>docDateOnly</th>
<th>timexAndDocDate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rec &lt; 0.1</td>
<td>&lt; 0.75</td>
<td>&lt; 0.75</td>
</tr>
<tr>
<td>BInc</td>
<td>0.072</td>
<td>0.471</td>
<td>0.471</td>
</tr>
<tr>
<td>T. BInc Ratio (PMI)</td>
<td>0.051</td>
<td>0.051</td>
<td><strong>0.493</strong></td>
</tr>
<tr>
<td>T. BInc Binary (PMI)</td>
<td>0.058</td>
<td>0.081</td>
<td>0.489</td>
</tr>
<tr>
<td>Weed’s Pr (Count)</td>
<td>0.061</td>
<td>0.440</td>
<td>0.440</td>
</tr>
<tr>
<td>T. Weed’s Pr (Count)</td>
<td>0.067</td>
<td>0.120</td>
<td>0.449</td>
</tr>
</tbody>
</table>

- **Sports** subset: `BASE` (directional entailment 1 + antonym 0)
- Uniform temporal window size: 5 days
- \( r \) = recall threshold reached by all similarity measures
Exp 1: Sports BASE Subset

Settings: timexAndDocDate, 5 day window, evaluate on BASE subset

BASE: directional entailment 1 + antonym 0

Conclusion: Temporal filtering is **beneficial** in separating out events
Exp 1: Sports DIRECTIONAL Subset

Settings: timexAndDocDate, 5 day window, evaluate on DIRECTIONAL

DIRECTIONAL: dir. entailment 1 + dir. non-entailment 0

Conclusion: Temporal info helps us learn directional entailments
Exp 2: Uniform Temporal Window Size

Settings: timexAndDocDate, evaluate on Sports BASE subset

Conclusion: Window size is important

Question: Why two peaks for each temporal similarity measure?
### Exp 3: Dynamic Temporal Window

**Evaluate on:** ANT dataset

<table>
<thead>
<tr>
<th>Window Method</th>
<th>ANTR Base</th>
<th>ANTR Directional</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Uni.</td>
<td>Dyn.</td>
</tr>
<tr>
<td></td>
<td>Uni.</td>
<td>Dyn.</td>
</tr>
<tr>
<td>Similarity measures:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weed’s Pr (Count)</td>
<td>0.181</td>
<td>0.181</td>
</tr>
<tr>
<td>T. Weed’s Pr (Count)</td>
<td>0.164</td>
<td>0.180</td>
</tr>
<tr>
<td>BInc (PMI)</td>
<td>0.161</td>
<td>0.161</td>
</tr>
<tr>
<td>T. BInc (Ratio PMI)</td>
<td>0.144</td>
<td>0.161</td>
</tr>
</tbody>
</table>

**Conclusions:**

- Adding a **dynamic per-predicate window doesn’t help**, but brings performance in line with the atemporal method.
- The **atemporal formulation of the DIH** is appropriate for the general domain.
Analysis

• Effect of temporal filtering is greater for the sports domain (than the general domain):
  • Antonym pairs are a) observed and b) temporally disjoint more often in Sports
  • Some areas in the general domain (e.g. legal news) could benefit from temporal filtering

• SUTime is not enough: limited number of time expressions + partial time information

• Speculation about events:
  • Conditionals (e.g. “If Arsenal win”)
  • Modals (“I still expect Arsenal...”)
  • Incorrect future predictions (“Arsenal will win”)
  • Counterfactuals (“Had Arsenal won,...”)

  is especially common in the sports domain and can result in conflicting evidence e.g. if Arsenal actually lost
Conclusions and Future Work

• Results (Exp 1) are promising, but we rely heavily on document creation date - temporal expressions are sparse
  ➡ We need an accurate way to temporally locate all eventualities

• Essential to add a temporal window around time intervals (Exp 2)

• Adding temporality is beneficial in the Sports domain (Exp 3)
  • Especially for directional entailments
  ➡ Reinterpret the DIH to include time

• The atemporal formulation of the DIH is appropriate for the general domain (Exp 3)
Modality
Linguistic Modality

<table>
<thead>
<tr>
<th>Category</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modal operator</td>
<td>Protesters <em>may</em> have attacked the police</td>
</tr>
<tr>
<td>Conditional</td>
<td><em>If</em> protesters attack the police...</td>
</tr>
<tr>
<td>Counterfactual</td>
<td><em>Had</em> protesters attacked the police...</td>
</tr>
<tr>
<td>Propositional attitude</td>
<td>Journalists <em>said</em> that protesters attacked the police</td>
</tr>
</tbody>
</table>

Essential for downstream tasks: Question Answering and Knowledge Graph population

→ Also useful for Entailment Graph Learning?

Guillou et al. (2021)
Method

- Learn entailment graphs from different training sets:
  - Only predications \textit{asserted} as actually happening?
  - A mixture of \textit{asserted} and \textit{modalised} predications?

- Extract binary relations using MoNTee (Bijl de Vroe et al., 2021)
  - Binary relations: \texttt{arg1}-predicate-\texttt{arg2} e.g. \texttt{Spurs}-beat-\texttt{Arsenal}
  - Tag binary relations as: modal operator, conditional, counterfactual, propositional attitude

Guillou et al. (2021)
MoNTEE: Modality Tagging

Dependency graph

Lexicon

Identify trigger nodes (doubts, will)

Check for path: trigger → verb (win)

Select and apply tag (lexicon + precedence rules)

Tagged relation MOD_(Labour; win; election)

Bijl de Vroe et al. (2021)
Lexicon

530 entries composed from:

- Modality Lexicon (Baker et al., 2010)
- Reporting verbs (Fay, 1990)
- Conditionals (Somasundaran et al., 2007)
- Conflicting event outcomes (Guillou et al., 2020)
- WordNet synonyms / antonyms (Miller, 1995)

<table>
<thead>
<tr>
<th>Lemma</th>
<th>Category</th>
<th>POS-tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>shall</td>
<td>MOD</td>
<td>MD</td>
</tr>
<tr>
<td>conceivably</td>
<td>MOD</td>
<td>RB</td>
</tr>
<tr>
<td>impossible</td>
<td>MOD</td>
<td>JJ</td>
</tr>
<tr>
<td>as long as</td>
<td>COND</td>
<td>RB</td>
</tr>
<tr>
<td>reckon</td>
<td>ATT_THINK</td>
<td>VB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Experiments

- **Data:** NewsSpike, approx. 0.5M articles (Zhang and Weld, 2013)
- **Models:**

<table>
<thead>
<tr>
<th>% Data</th>
<th>Modalised predications present?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asserted</td>
<td>85</td>
</tr>
<tr>
<td>BaselineLarge*</td>
<td>100</td>
</tr>
<tr>
<td>BaselineSmall</td>
<td>85</td>
</tr>
</tbody>
</table>

* Equivalent to (Hosseini et al., 2018)
Evaluation

Datasets:

- Levy/Holt: general domain, 18,407 entailment pairs

  medicine \textit{kills} disease \rightarrow medicine \textit{treats} disease

  medicine \textit{treats} disease \nRightarrow medicine \textit{kills} disease

- Sports Entailment Dataset: sports, 718 entailment pairs

  Spurs \textit{beat} Arsenal \rightarrow Spurs \textit{play against} Arsenal

  Spurs \textit{beat} Arsenal \nRightarrow Spurs \textit{lose to} Arsenal

Metric:

- AUC score: area under precision-recall curve
- Points on the curve = different entailment score thresholds
Conclusion: **Ignoring** modality is beneficial in the general domain
Conclusion: Removing modalised predications is beneficial in the sports domain

Guillou et al. (2021)
Why does BASELINESMALL outperform ASSERTED on Levy/Holt?

- Perhaps this is caused by size and/or coverage of the graph

<table>
<thead>
<tr>
<th></th>
<th>Nodes</th>
<th>Edges</th>
<th>% Levy/Holt predicates found all examples</th>
<th>% Levy/Holt predicates found directional</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASELINE_LARGE</td>
<td>334K</td>
<td>72.7M</td>
<td>63.06</td>
<td>70.29</td>
</tr>
<tr>
<td>BASELINESMALL</td>
<td>277K</td>
<td>58.4M</td>
<td>61.13</td>
<td>69.29</td>
</tr>
<tr>
<td>ASSERTED</td>
<td>254K</td>
<td>46.3M</td>
<td>58.51</td>
<td>67.92</td>
</tr>
</tbody>
</table>

- However, this pattern also holds in the sports domain (where ASSERTED performs best)
  ➡️ Further investigation required...

Guillou et al. (2021)
Analysis: Examples

• Why is ignoring modality helpful in the general domain?
  • Perhaps modals are often used when the prior probability of the main predicate is high

<table>
<thead>
<tr>
<th>Acquisition of Dell by Michael Dell</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Feb 5th 2013:</strong> “...founder and CEO Michael Dell and investment firm Silver Lake Partners will buy Dell.”</td>
</tr>
<tr>
<td><strong>Feb 6th 2013:</strong> “So Michael Dell and a private equity group have bought Dell and taken it private.”</td>
</tr>
</tbody>
</table>

• Why does removing modalised predications help in the sports domain?
  • Match outcomes are widely speculated upon, but highly uncertain

<table>
<thead>
<tr>
<th>Seattle vs. Atlanta</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Jan 10th 2013:</strong> “The popular opinion on this game seems to be Seattle beating Atlanta because...”</td>
</tr>
<tr>
<td><strong>Jan 14th 2013:</strong> “Falcons come back to beat Seahawks.”</td>
</tr>
</tbody>
</table>
Conclusions

- Overall, uncertain predications constitute a valuable learning signal for Entailment Graphs
- Removing modalised predications can help in specific domains, e.g. sports
Future Directions

- Contextualised modality / uncertainty detection
- How can we use modal information to retain what is beneficial vs. not?
  - Identify specific sub-domains
  - Retain data under different epistemic strengths e.g. “undoubtedly” vs “unlikely”
- Explore entailments between modal predicates (+temporality?)
  - if beat → play, then also play → MODBeat (precondition)
  - if buy → own, then also MODbuy → MODown (consequence)
Summary
Summary

• Experimented with:
  • Using temporal information to avoid learning spurious entailments such as win → lose
  • Ignoring modality vs. removing modalised predications

• Conclusion: Temporality and modality can provide a benefit in Entailment Graph learning
  ➡️ But we should pay attention to the domain

• Future Directions:
  • Contextualised modality / uncertainty detection
  • Robust temporal location of eventualities, within document and cross-document
Other projects within the group:

- Multivalent Entailment Graphs (McKenna et al., 2021)
- Cross-lingual Entailment Graphs (English + Chinese) (Li et al., 2022)
- Smoothing Entailment Graphs with Language Models (McKenna and Steedman, 2022)
- Incorporating Entailment Graphs for link prediction in Knowledge Graphs (Hosseini et al., 2021)
Questions?


