An Entailment-based Ontology for Domain-specific Relations

Jonathan Berant, Ido Dagan, Idan Szpektor
Bar-Ilan University
Outline

• **Motivation**: Organizing information by semantic relations hierarchy
• Capturing semantic relations structure via *entailment graph*
• Approach:
  – Learning the entailment graph
  – Extracting propositions and projecting them over the entailment graph
  – Evaluation methodology
• Application within the FIRB project
Motivation

• A lot of work has been dedicated to retrieving information
• **An important challenge:** organize retrieved information for the user
• **Goal:** present the information in a set of documents in a succinct and semantically-structured manner
Information is expressed in Propositions

...Melatonin may cause drops in blood pressure...

...beta-blockers are less effective at preventing strokes than other types of blood-pressure medication...

...Caffeine can also raise blood pressure, in very high doses it can cause seizures...
Desideratum

Blood pressure (BP)

- Affects BP
  - Atenolol (35)
  - Caffeine (12)
  - Zantac (7)
  - Calcium (5)
  - Alcohol (4)
  - Melatonin (2)

- Raises BP
  - Caffeine (12)
  - Alcohol (4)

- Reduces BP
  - Calcium (5)
  - Melatonin (2)

- BP medication
  - Atenolol
  - Zantac (7)

Measuring unit
- mmHg (7)
Desideratum

Blood pressure (BP)

Affects BP

- Atenolol (35)
  - Caffeine (12)
  - Zantac (7)
  - Calcium (5)
  - Alcohol (4)
  - Melatonin (2)

Raises BP

- Caffeine (12)
  - Alcohol (4)

Reduces BP

- Calcium (5)
  - Melatonin (2)

BP medication

- Atenolol
  - Zantac (7)

Measuring unit

mmHg (7)

Caffeine raises BP → Caffeine affects BP
Desideratum

Blood pressure (BP)

Affects BP

- Atenolol (35)
- Caffeine (12)
- Zantac (7)
- Calcium (5)
- Alcohol (4)
- Melatonin (2)

Measuring unit

- mmHg (7)

Raises BP

- Caffeine (12)
- Alcohol (4)

Reduces BP

- Calcium (5)
- Melatonin (2)

BP medication

- Atenolol
- Zantac (7)

Caffeine is a BP medication → Caffeine affects BP
Desideratum

Blood pressure (BP)

Affects BP
- Atenolol (35)
- Caffeine (12)
- Zantac (7)
- Calcium (5)
- Alcohol (4)
- Melatonin (2)

Measuring unit
- mmHg (7)

Raises BP
- Caffeine (12)
- Alcohol (4)

Reduces BP
- Calcium (5)
- Melatonin (2)

BP medication
- Atenolol
- Zantac (7)

• The semantics is of **entailment** between relations
• And not **is-a**.
Rationale

- Most ontology learning work focused on concept taxonomies
- Existing ontologies (UMLS, YAGO, …) contain millions of concepts but just few relations
- Yet, the information lies in propositions
- Their presentation by entailment relations:
  - summarizes propositions
  - hierarchical, easily-navigable display
Relations and Predicates

• We want to present propositions by \textit{semantic relation} structure
• However, documents contain \textit{linguistic predicates}
• \textbf{Approach}: identify semantic relations structure based on linguistic predicates and their entailments
Entailment Graph

- **(Linguistic) Predicate**: A dependency path whose endpoints are variables

  

  ![Diagram](image)

  - subj
  - excavate
  - obj

- **Entailment Rule**: A pair of predicates \((R_1, R_2)\), such that \(R_1\) textually-entails \(R_2\)

  \[ X \text{ excavate } Y \rightarrow X \text{ discover } Y \]

- **Entailment Graph**: A graph whose nodes are predicates and edges are entailment rules
Entailment Graph – cont’

X<--dig up-->Y

X<--excavate-->Y

X<--discover-->Y
Entailment Graph – cont’
Overall Scheme

- **Input**: A domain-specific corpus annotated with domain concepts
- **Tasks**:
  1. Learn an entailment graph for the domain:
     i. Extract domain predicates from corpus (nodes)
     ii. Learn/acquire entailment rules graph (edges)
  2. Organize information by entailment graph:
     i. Extract propositions from target documents
     ii. Project them onto the entailment graph
Overall Scheme

• **Input**: A domain-specific corpus annotated with domain concepts

• **Tasks**:
  1. Learn an entailment graph for the domain:
     I. Extract domain predicates from corpus (nodes)
     II. Learn/acquire entailment rules graph (edges)
  2. Organize information by entailment graph:
     I. Extract propositions from target documents
     II. Project them onto the entailment graph

**Our focus**
Learning an Entailment Graph
Alzheimer's disease (AD) is now the fourth leading cause of death in adults. It is estimated that 4.5 million Americans and eight million more people worldwide have it. Age is the biggest risk factor for Alzheimer's disease.
Predicate Extraction
Predicate Extraction

1. Parse the corpus

- diabetes
- is controlled by
- a careful diet
Predicate Extraction

1. Parse the corpus
2. Extract and normalize dependency paths with **domain-specific concepts** as arguments
3. Replace arguments with variables

```
1. X control Y
2. ...
3. ...
```

**diabetes** is controlled by a careful diet
Predicate Extraction

1. Parse the corpus
2. Extract and normalize dependency paths with domain-specific concepts as arguments
3. Replace arguments with variables

Vitamin D is a controller of cell growth

1. X control Y
2. ...
3. ...
Predicate Extraction

1. Parse the corpus
2. Extract and normalize dependency paths with domain-specific concepts as arguments
3. Replace arguments with variables
4. Predicates: frequent dependency paths

Vitamin D is a controller of cell growth

1. X control Y (189)
2. ...
3. ...
Acquiring Entailment Rules

**Methods**

- Manual resources:
  - WordNet
  - Argument Mapped WordNet
- Distributional similarity
  - DIRT
  - Balanced Inclusion
- Meta-similarity
- Graph constraints
Augmenting WordNet-based Inference with Argument Mapping

Szpektor and Dagan (TextInfer, 2009)

• WordNet is a widely used machine-readable lexical database organized by meanings (synsets)

  – S1: buy, purchase (obtain by purchase; acquire by means of a financial transaction)

  – S2: bribe, corrupt, buy, grease one's palms (make illegal payments to in exchange for favors or influence)
Augmenting WordNet-based Inference with Argument Mapping
Szpektor and Dagan (TextInfer, 2009)

- WordNet is a widely used machine-readable lexical database organized by meanings (synsets)
  - **S1**: `buy`, `purchase` (obtain by purchase; acquire by means of a financial transaction)
  - **S2**: `bribe`, `corrupt`, `buy`, `grease one's palms` (make illegal payments to in exchange for favors or influence)

- WordNet contains inferential relations between predicates
WordNet Inferential Relations for Predicates

- “IBM bought Cognos for $5 billion”
WordNet Inferential Relations for Predicates

- “IBM bought Cognos for $5 billion”

- Who acquired Cognos?
WordNet Inferential Relations for Predicates

- "IBM bought Cognos for $5 billion"

- Who acquired Cognos?
  - $buy_{v}$ $\rightarrow$ $acquire_{v}$:

  IBM $bought$ Cognos for $5$ billion $\rightarrow$
  IBM $acquired$ Cognos for $5$ billion
WordNet Inferential Relations for Predicates

• “IBM bought Cognos for $5 billion”

• Who acquired Cognos?
  \[
  \text{buy}_v \rightarrow \text{acquire}_v : \quad \text{Substitutable Relations}
  \]

  IBM bought Cognos for $5 billion \rightarrow
  IBM acquired Cognos for $5 billion
WordNet Inferential Relations for Predicates

• “IBM bought Cognos for $5 billion”

• Who \textit{acquired} Cognos?
  \begin{itemize}
  \item buy$_v$ $\rightarrow$ acquire$_v$ : \textit{Substitutable Relations}
  \end{itemize}

  \begin{itemize}
  \item IBM \textit{bought} Cognos for $5$ billion $\rightarrow$
  \item IBM \textit{acquired} Cognos for $5$ billion
  \end{itemize}

• How much was \textit{paid} for Cognos?
WordNet Inferential Relations for Predicates

- "IBM bought Cognos for $5 billion"

- **Who** acquired Cognos?
  - `buy_v` → `acquire_v`:
    - **Substitutable Relations**
      - `IBM bought` Cognos for $5 billion →
      - `IBM acquired` Cognos for $5 billion

- **How much was paid for Cognos?**
  - `buy_v` → `pay_v`:
    - `IBM bought` Cognos for $5 billion →
    - `IBM paid` Cognos for $5 billion
WordNet Inferential Relations for Predicates

• “IBM bought Cognos for $5 billion”

• Who **acquired** Cognos?
  - **buy** → **acquire**: Substitutable Relations
    - IBM *bought* Cognos for $5 billion →
    - IBM *acquired* Cognos for $5 billion

• How much was **paid** for Cognos?
  - **buy** → **pay**:
    - IBM *bought* Cognos for $5 billion →
    - IBM *paid* Cognos for $5 billion
    - IBM *paid* $5 billion for Cognos
WordNet Inferential Relations for Predicates

• “IBM bought Cognos for $5 billion”

• **Who acquired Cognos?**
  - \( \text{buy}_v \rightarrow \text{acquire}_v \): **Substitutable Relations**
    
    IBM **bought** Cognos for $5 billion
    IBM **acquired** Cognos for $5 billion

• **How much was paid for Cognos?**
  - \( \text{buy}_v \rightarrow \text{pay}_v \):
    
    IBM **bought** Cognos \text{for $5 billion}
    IBM **paid** Cognos for $5 billion
    IBM **paid** \$5 billion for Cognos
WordNet Inferential Relations for Predicates

• “IBM bought Cognos for $5 billion”

• Who acquired Cognos?
  – \( \text{buy}_v \rightarrow \text{acquire}_v \) : Substitutable Relations
    
    IBM \textit{bought} Cognos for $5 billion \rightarrow
    IBM \textit{acquired} Cognos for $5 billion

• How much was paid for Cognos?
  – \( \text{buy}_v \rightarrow \text{pay}_v \) : Non-substitutable Relations
    
    IBM \textit{bought} Cognos for $5 billion \rightarrow
    IBM \textit{paid} Cognos for $5 billion
    IBM \textit{paid} $5 billion for Cognos
Using WordNet for Predicate Inference

- Non-substitutable rules need argument mapping
  - buy → pay: $X$ buy $Y$ for $Z$ → $X$ pay $Z$ for $Y$
  - buy ↔ buying: $X$ buy $Y$ ↔ $X$’s buying of $Y$
Using WordNet for Predicate Inference

- Non-substitutable rules need argument mapping
  - buy $\rightarrow$ pay: $X$ buy $Y$ for $Z$ $\rightarrow$ $X$ pay $Z$ for $Y$
  - buy $\leftrightarrow$ buying: $X$ buy $Y$ $\leftrightarrow$ $X$'s buying of $Y$

Entailment Rules
Using WordNet for Predicate Inference

- Non-substitutable rules need argument mapping
  - buy → pay: \[ X \text{ buy } Y \text{ for } Z \rightarrow X \text{ pay } Z \text{ for } Y \]
  - buy ↔ buying: \[ X \text{ buy } Y \leftrightarrow X \text{ 's buying of } Y \]

Entailment Rules

- Argument mapping not specified in WordNet
  - Common practice: only substitutable relations used
    - synonyms, hypernyms

\[ \rightarrow \text{WordNet relations are not fully utilized} \]

(~9000 non-substitutable relations)
AmWN Representation
Highlights

• Extending unary entailment rules
  – $X$ buy $\rightarrow$ $X$ pay ;  buy $Y$ $\rightarrow$ pay for $Y$ ;  buy for $Z$ $\rightarrow$ pay $Z$
  – Rule templates are parse fragments with a single variable

Argument functional roles

$X_{\text{subj}}$’s acquisition $\rightarrow$ $X_{\text{subj}}$ acquire  \hspace{1cm} (Google ‘s acquisition of Deja)

$X_{\text{obj}}$’s acquisition $\rightarrow$ acquire $X_{\text{obj}}$  \hspace{1cm} (Deja ‘s acquisition by Google)

Predicate subcategorization frames

\textit{The window broke} intransitive vs. \textit{John broke} transitive the window

$X_{\text{subj}}$ break intransitive $\rightarrow$ $X_{\text{obj}}$ was damaged transitive
AmWN in the Entailment Graph

AmWN
1. \( X_{\text{subj}} \xrightarrow{\text{raise}_{\text{transitive}}} X_{\text{subj}} \xrightarrow{\text{affect}_{\text{transitive}}} \)
2. \( \xrightarrow{\text{raise}_{\text{transitive}}} Y_{\text{obj}} \xrightarrow{\text{affect}_{\text{transitive}}} Y_{\text{obj}} \)

Entailment-graph

- \( X_{\text{subj}} \xrightarrow{\text{raise}} Y_{\text{obj}} \xrightarrow{\text{affect}} Y_{\text{obj}} \)
## Distributional Similarity

<table>
<thead>
<tr>
<th>X affect Y</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
<td>#</td>
<td></td>
</tr>
<tr>
<td>insulin</td>
<td>metabolism</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Zantac</td>
<td>BP</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X treat Y</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
<td>#</td>
<td></td>
</tr>
<tr>
<td>Zantac</td>
<td>BP</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Diet</td>
<td>diabetes</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Distributional Similarity

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
<td>#</td>
</tr>
<tr>
<td>insulin</td>
<td>metabolism</td>
<td>7</td>
</tr>
<tr>
<td>Zantac</td>
<td>BP</td>
<td>4</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
<td>#</td>
</tr>
<tr>
<td>Zantac</td>
<td>BP</td>
<td>9</td>
</tr>
<tr>
<td>Diet</td>
<td>diabetes</td>
<td>2</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>
### Distributional Similarity

#### Table: X affect Y

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
<td>#</td>
</tr>
<tr>
<td>insulin</td>
<td>metabolism</td>
<td>7</td>
</tr>
<tr>
<td>Zantac</td>
<td>BP</td>
<td>4</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

#### Table: X treat Y

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
<td>#</td>
</tr>
<tr>
<td>Zantac</td>
<td>BP</td>
<td>9</td>
</tr>
<tr>
<td>Diet</td>
<td>diabetes</td>
<td>2</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

The similarity score $Lin(u, v)$ can be calculated as:

$$Lin(u, v) = \frac{\sum_{f \in F_u \cap F_v} [w_u(f) + w_v(f)]}{\sum_{f \in F_u} w_u(f) + \sum_{f \in F_v} w_v(f)}$$
## Distributional Similarity

<table>
<thead>
<tr>
<th>X affect Y</th>
<th>X treat Y</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>X</strong></td>
<td><strong>Y</strong></td>
</tr>
<tr>
<td>insulin</td>
<td>metabolism</td>
</tr>
<tr>
<td>Zantac</td>
<td>BP</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

\[
\text{Precision}(u,v) = \frac{\sum_{f \in F_u \cap F_v} w_u(f)}{\sum_{f \in F_u} w_u(f)}
\]
### Distributional Similarity

#### X affect Y

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>insulin</td>
<td>metabolism</td>
<td>7</td>
</tr>
<tr>
<td>Zantac</td>
<td>BP</td>
<td>4</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

#### X treat Y

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zantac</td>
<td>BP</td>
<td>9</td>
</tr>
<tr>
<td>Diet</td>
<td>diabetes</td>
<td>2</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

\[
BInc(l, r) = \sqrt{Lin(l, r) \cdot Precision(l, r)}
\]
Meta-similarity measure

1. Create a feature vector of similarity scores for every pair of predicates

2. Construct a training set using AmWN entailments (cf. Snow et al., 2005)

3. Train a classifier and classify new candidate entailments

<table>
<thead>
<tr>
<th>Measure</th>
<th>score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin</td>
<td>0.5</td>
</tr>
<tr>
<td>Precision</td>
<td>0.9</td>
</tr>
<tr>
<td>BInc</td>
<td>0.7</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Label</td>
<td>1</td>
</tr>
</tbody>
</table>
Graph-based Constraints

• Graph should contain cliques of synonyms
• Strong connectivity components with sparse edges are suspicious
• The score of an edge should reflect the score of the set of edges it induces (cf. Snow et al., 2006)
• These scores can be features for the above classifier
Organizing Propositions by the Entailment Graph
Projection onto the Graph

Corpus
Exercise affects BP (1)
Caffeine increases BP (1)
Lead raises BP (1)
Calcium reduces BP (5)
Corpus
Exercise affects BP (1)
Caffeine increases BP (1)
Lead raises BP (1)
Calcium reduces BP (5)
Projection onto the Graph

- **affect**
  - **Exercise**
  - **Increase, Raise**
    - **Caffeine Lead**
  - **reduce**
    - **Calcium**

**Corpus**
- Exercise affects BP (1)
- Caffeine increases BP (1)
- Lead raises BP (1)
- Calcium reduces BP (5)
Projection onto the Graph

Corpus
Exercise affects BP (1)
Caffeine increases BP (1)
Lead raises BP (1)
Calcium reduces BP (5)
Evaluation

1. For a sample of domain concepts
   I. Choose the $k$ most frequent predicates
   II. Manually construct the gold standard entailment graph for these predicates
   III. Compare the learned graph to the gold standard graph.
Evaluation – cont’

- Graphs can be compared by:
  - edges (or induced edges)
  - proposition instances or mentions
FIRB Project

- Use this framework in the archeological domain
- The archeological corpus is annotated for archeological concepts by the FBK group
- Extract propositions to populate info boxes about domain concepts
FIRB Project – cont’

Tower of David

<table>
<thead>
<tr>
<th><strong>Location:</strong></th>
<th>Jerusalem</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Foundation:</strong></td>
<td>1st century BCE</td>
</tr>
</tbody>
</table>
| **Rulers:**      | - Hasmonean kings  
|                  | - King Herod  
|                  | - Romans  
|                  | - Muslims |

It's an iconic symbol of Jerusalem, located on a hill overlooking the city.
Conclusions

• Structured presentation of information is an important challenge
• An entailment-graph can be used to put propositions in a hierarchy of relations
• Proposed approach:
  – Construct entailment graph for a domain
  – Project extracted propositions over it
• Challenging future work…
Conclusions

• Structured presentation of information is an important challenge
• An entailment-graph can be used to put propositions in a hierarchy of relations
• Proposed approach:
  – Construct entailment graph for a domain
  – Project extracted propositions over it
• Challenging future work…

Thank you!