



# Focused Entailment Graphs for Open IE Propositions

Bar-Ilan University

Omer Levy, Ido Dagan, Jacob Goldberger

## 1. Motivation

### Open IE

- Unsupervised information extraction.
- No pre-defined schema.
- **Lacks structure!**

aspirin **relieves** headache  
 ≠  
 aspirin **treats** headache

## 2. Approach

*Add structure to Open IE!*

### Which structure?

Build **entailment graphs** of propositions.

### Why entailment?

- Merges equivalent propositions.
- Induces specific-to-general structure.

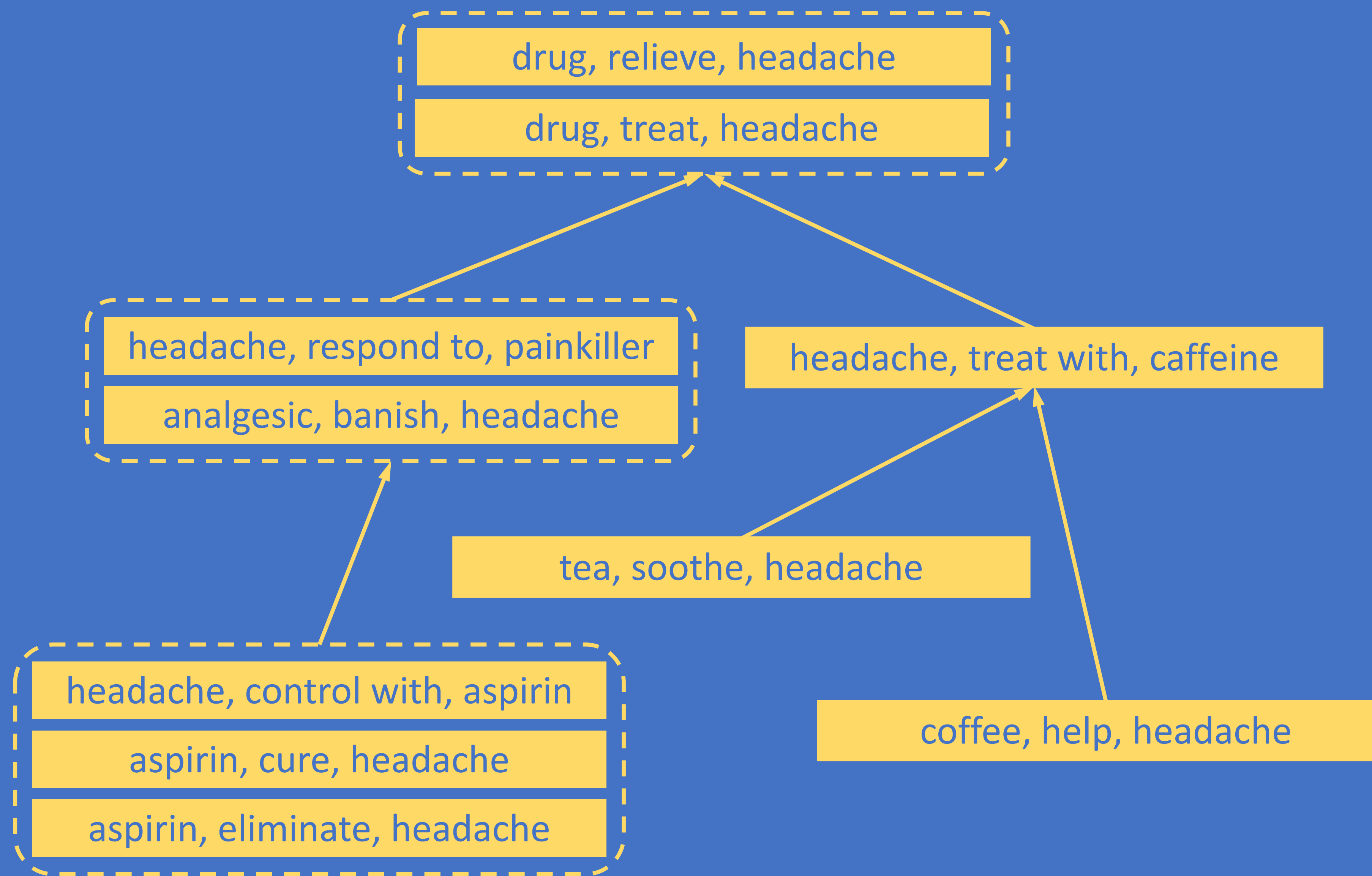
## 3. Findings

*Reality is not quite what we thought...*

- Predicate entailment is **extremely context-sensitive** and **situational**.

antibiotic **cures** infection  
 antibiotic **kills** infection  
 antibiotic **is needed for** infection

- Current lexical-semantic resources **do not capture lexical inferences** within **real** proposition entailment.



Proposition entailment graph focused on **headache**. Dashed boundaries denote cliques (equivalence).

## 4. Dataset

### Data Acquisition

1. Extract open IE propositions from Google books (syntactic ngrams).
2. Collect propositions focused on a single topic (**nodes**).
3. Annotate entailments (**edges**).

### Volume

- 30 annotated entailment graphs
- ~200 propositions/graph (average)
- **1,500,000 entailment judgments**
- 8% positive (entailing) pairs

**Publicly available!**

## 5. Method

### Lexical Inference Paradigm

- Train classifier over lexical entailment features: WordNet, DistSim, etc.

### Previous Methods

- Use WordNet for distant supervision.
- **However, WordNet relations do not correlate well with lexical inference** in the context of propositions.

### Our Method

- Models components separately.
- Propagates **direct supervision** from data to lexical components.

## 6. Results

- Our method outperforms extensions of previous methods.
- Better to model predicates and arguments separately.
- Direct supervision is better than distant supervision.
- Learning predicate entailment from WordNet cripples performance.

## 7. Future Work

- Capture **context-sensitive** predicate inferences in proposition entailment.
- Increase coverage of argument entailment knowledge resources.
- **Scale up** to general (non-focused) entailment graphs of propositions.
- Explore **richer proposition structures** (modifiers, modalities, coordination).

## Contact Us

Omer Levy [omerlevy@cs.biu.ac.il](mailto:omerlevy@cs.biu.ac.il)  
 Ido Dagan [dagan@cs.biu.ac.il](mailto:dagan@cs.biu.ac.il)  
 Jacob Goldberger [goldbej@eng.biu.ac.il](mailto:goldbej@eng.biu.ac.il)

Bar-Ilan University  
Israel