UNIVERSITY OF MANNHEIM
Utrecht University

Funded by DFG Deutsche Forschungsgemeinschaft German Research Foundation
RATIO

Effect Graph: Effect Relation Extraction for Explanation Generation

by Jonathan Kobbe, Ioana Hulpuş, Heiner Stuckenschmidt

Example argument:
Legal abortions protect women.

Goal: Find possible explanations.
1. Abortions prevent the harm caused by giving birth and being pregnant.
2. Legal abortions strengthen the women's right to self-determination.
3. Abortions release women from the financial burden of raising a child.

Intuition: Explain an effect relation by chaining other effect relations.
Effect Relation: Legal abortions –[protect]→ women

Possible explanations with effect relations:
1. Abortions –[prevent]→ harm –[hurts]→ women
2. Legal abortions –[strengthen]→ self-determination –[empower]→ women
3. Abortions –[release]→ financial burden –[strain]→ women

Effect Relation Extraction & Graph Construction

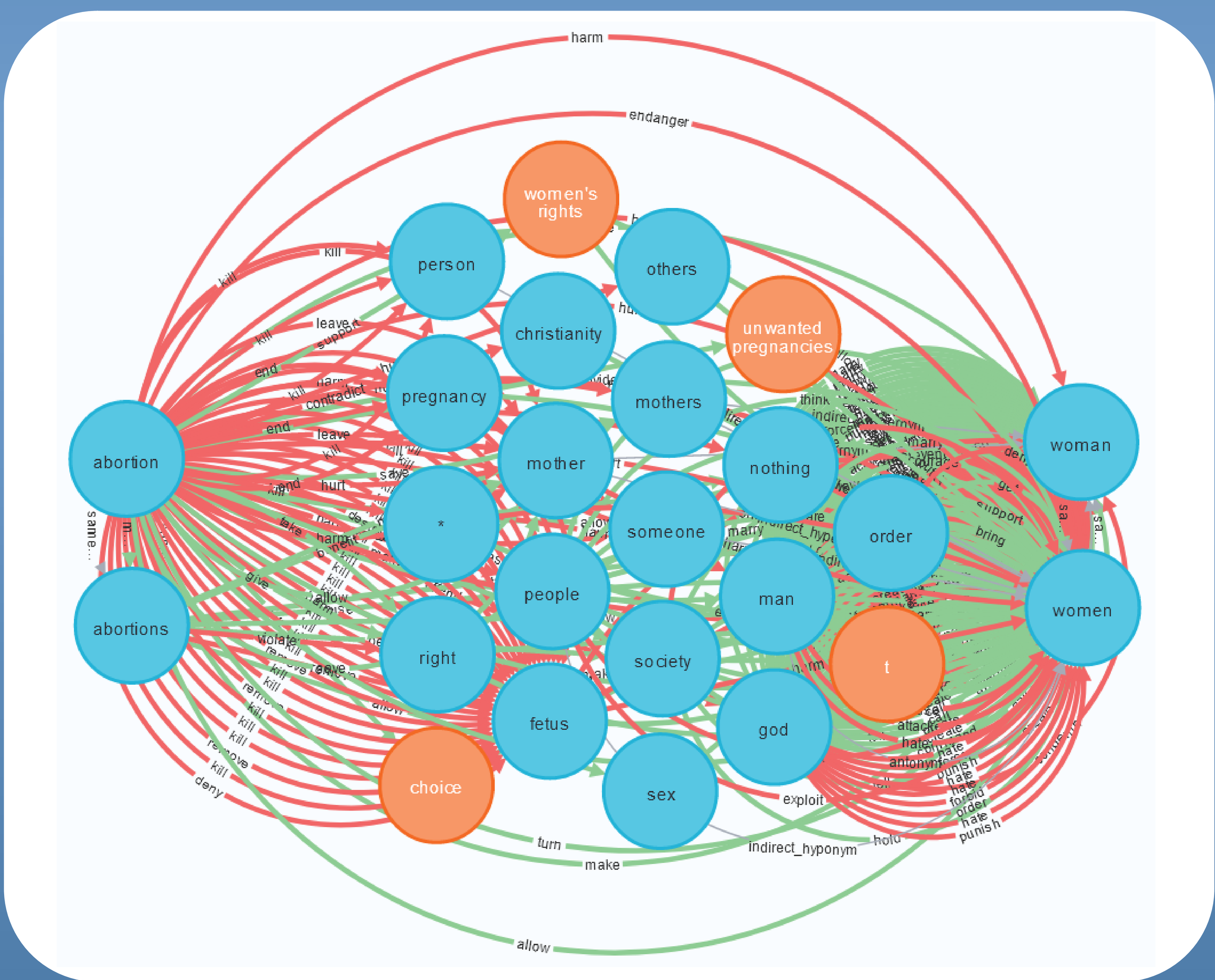
Text resources:
Online debate platforms & Wikipedia in simple English

Effect Lexicon (Rashkin 16, Kobbe 20)**:**
Positive Effect Words: increase, help, permit, cause, create, ...
Negative Effect Words: decrease, damage, forbid, ban, reduce, ...

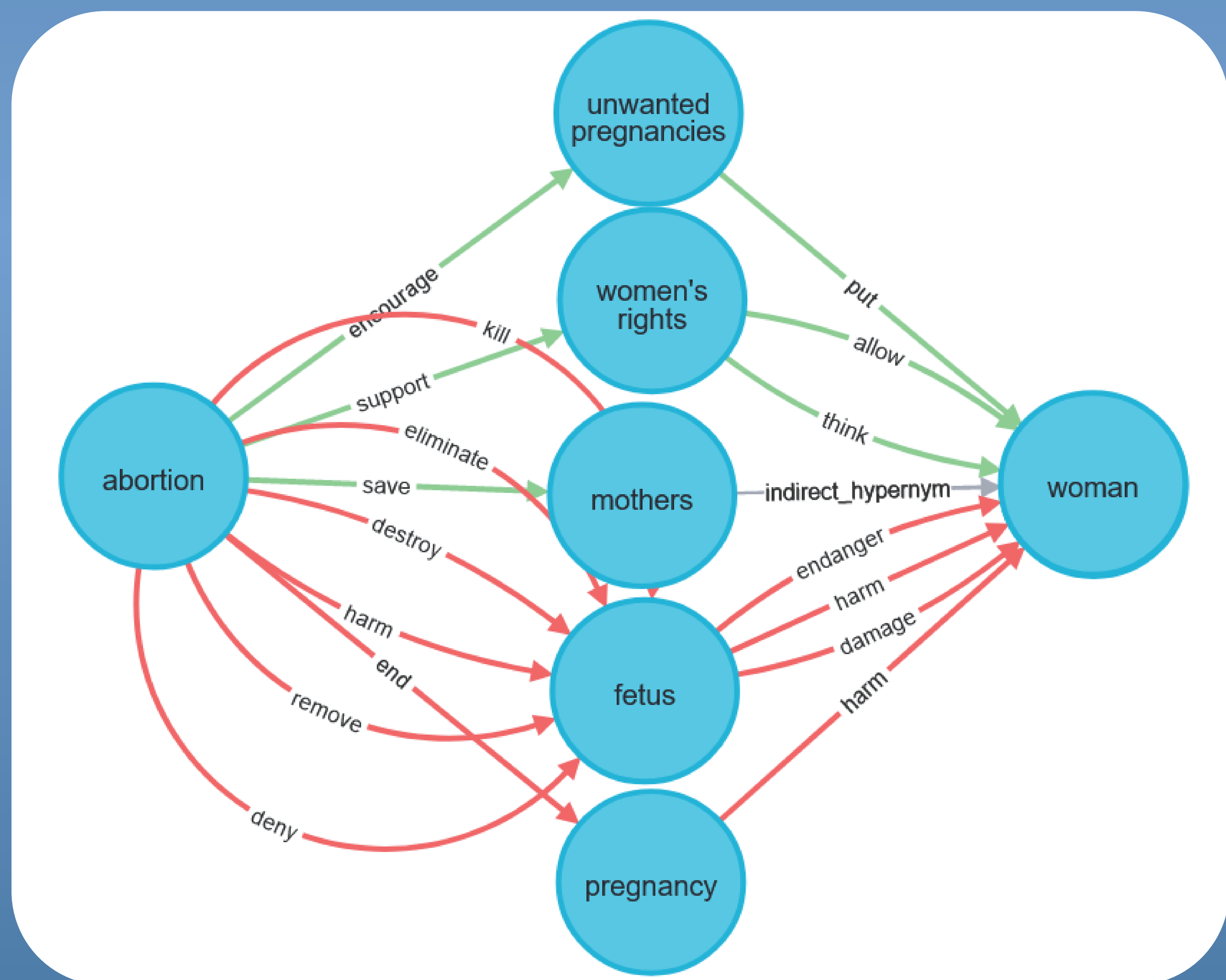
Extract Subject-Predicate-Object (SPO) triples using dependency parsing where:
- P is an effect word.
- S and O are noun phrases linking to Wikipedia entries.

→ The lemmas of S and O become ***nodes*** in the graph which are connected by a directed edge labeled with P.

Explanation Generation



Consistency Checking & Node Ranking:
High degree in subgraph,
Low degree in effect graph



Evaluation of the Effect Graph

Evaluation of the Effect Relation Extraction

By the subtasks of Al-Khatib 20:

Subtask	Al-Khatib	*EREx*
Relation Classification *(macro F1)*	***0.79***	0.65
Relation Type Class. *(macro F1)*	0.77	0.77
Subject Identification *(accuracy)*	0.69	***0.71***
Object Identification *(accuracy)*	0.28	***0.35***

The evaluation was done on the crowd-sourced dataset of Al-Khatib 20. Since the train-test split is unknown to us, we evaluated on the full dataset.

Evaluation of the Final Graph

Baseline: Stanford CoreNLP's OpenIE system for finding SPO triples, instead of dependency parsing.

Precision:
Given a random edge of the graph, how likely was it extracted correctly?
→ We randomly select 250 edges per graph
→ Each edge is annotated by 3 crowd-workers and myself:

Statement: Throughout history, nuclear weapons have killed many innocents.
Relation: history (may) **negatively** affect → innocents
Assuming the statement is correct, is the Relation also correct?

Agreement is low: Fleiss 0.15 for the workers,
Spearman 0.56 between their aggregated label and mine

Precision	OpenIE	*EREx*
Crowd Annotations	0.70	***0.80***
Expert Annotations	0.34	***0.54***

Recall:
Given any effect relation, how likely is it contained in the graph?
→ We manually extracted effect relations for 180 arguments.
→ "known" means these arguments were included when building the graph.

Recall	OpenIE	*EREx*
known	***0.14***	0.09
unknown	0.09	0.06

Rashkin16: H. Rashkin, S. Singh and Y. Choi: Connotation Frames: A Data-Driven Investigation. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016.
Al-Khatib 20: K. Al-Khatib, Y. Hou, H. Wachsmuth, C. Jochim, F. Bonin and B. Stein: End-to-End Argumentation Knowledge Graph Construction. In Proceedings of the AAAI Conference on Artificial Intelligence. 2020.
Kobbe 20: J. Kobbe, I. Hulpuş, and H. Stuckenschmidt: Unsupervised stance detection for arguments from consequences. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2020 .