graphbrain selected use cases

extracting inter-actor conflict networks from large news headline corpora

a

graphbrain is an open-source platform implementing **semantic hypergraphs**, a knowledge representation model that is intrinsically recursive and accommodates the natural hierarchical richness of natural language – see http://graphbrain.net
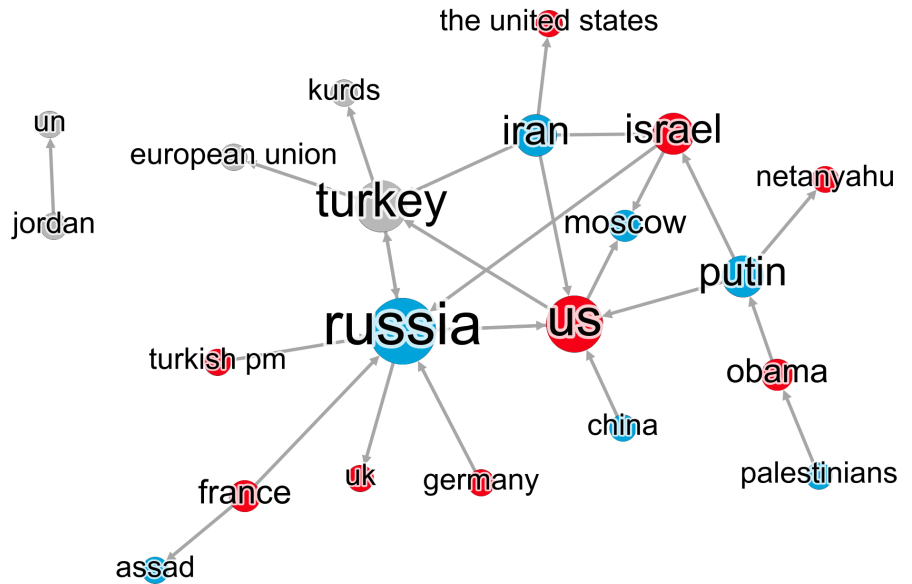
Goals

Map actor conflicts and factions

Extract networks of conflict between actors and identify topics of conflict

Typical empirical domain

News articles

NEWS	**News headers**
	Hypergraph
	Coreference resolution
	Frequent patterns
	Rule inference
	Rule testing
	Conflict extraction
	Network mapping

"Obama warns Putin against strikes in Syria."

```
(warns/P obama/C putin/C
 (against/T
  (in/B strikes/C syria/C)))
```

```
putin/C
≡ (+/B president/C putin/C)
≡ (+/B vladimir/C putin/C)
```

```
(warns/P */C */C */S)
```

```
(warns/P ORIG/C TARG/C
 (against/T TOPIC))
```

```
(warns/P china/C us/C
 (against/T
  (on/B attack/C syria/C)))
```

Methodology & technical workflow

The corpus consists of all news headers shared between January 2013 and August 2017 on the *r/worldnews* subreddit. These were **parsed into a semantic hypergraph (SH)**.

graphbrain includes a **coreference resolution algorithm for SH** that was used to identify different designations for the same actor. For example: « Putin », « President Putin » and « Vladimir Putin » all refer to the same person.

The **most common hyperedge patterns** were computed, allowing for a researcher to identify common expressions of conflict from which rules could be derived. These **rules were iteratively refined** by applying them to the entire hypergraph and then inspecting the validity of the results.

Once found satisfactory, the rules enabled the extraction from the corpus of thousands of cases of conflicts between actors, associated with topics. The **implicit taxonomy of SH** concept representation allows for topics to be considered at different levels of generality, for example « attacks in Syria » or just « Syria ».

Finally, the above network was generated for conflicts between actors around the topic of « Syria ». A simple alliance detection algorithm was used to identify two factions, depicted in red and blue.

Telmo Menezes menezes@cmb.hu-berlin.de

Camille Roth roth@cmb.hu-berlin.de