Unsupervised Induction of Frame-Semantic Representations

Ashutosh Modi  amodi@mmci.uni-saarland.de
Ivan Titov  titov@mmci.uni-saarland.de
Alexandre Klementiev  aklement@mmci.uni-saarland.de

In frame semantics, a conceptual structure describing a situation, object, or event along with associated properties and participants is used. For a collection of sentences:

1. Identify predicates
2. Identify arguments
3. Label predicates with frames
4. Label arguments with roles

**Frame Induction**
- Cluster predicates (each cluster is a frame)
- Key Signals
  - Related to Levin classes

**Role Induction**
- Associate arguments with syntactic signatures (argkeys)
- Cluster argkeys (each argkey cluster is a role)
- Key Signals
  - Argkeys with similar argument fillers clustered together
  - Most roles occur once (per predicate occurrence)

Our unsupervised model simultaneously induces frames (clusters of predicates) and roles (clusters of argkeys) by exploiting distributional context.

**Evaluation**

**Qualitative Evaluation**
- Induced Frames: FrameNet frames corresponding to the verbs
- Occurrence counts

**Quantitative Evaluation**
- METRICS
  - Purity (PU): Extent to which predicted cluster occurrences share the same gold label
  - Collocation (CO): Extent to which gold label is assigned to a single cluster
  - F1: Harmonic mean of PU and CO

**The Model**

For each frame we generate a prior partition of argkeys while \( f \sim \text{Unif}(0, 1) \)
and \( r \sim \text{Unif}(0, 1) \).

```
for every role \( r \), for every occurrence of frame \( f \):
  if \( r \) \#\# \( f \):
    GenArgument(\( f \), \( r \):
      \( k_f \sim \text{Unif}(1, \ldots, n_f) \)
      \( \eta_f \sim \text{Beta}(\alpha_f, \beta_f) \)
      \( \delta_f \sim \text{Beta}(\alpha_f, \beta_f) \)
      \( \theta_f \sim \text{Beta}(\alpha_f, \beta_f) \)
      \( \phi_f \sim \text{Beta}(\alpha_f, \beta_f) \)
```

Assumptions
- We consider verbal predicates only
- Every verb belongs to a single frame, i.e. we do not model polysemy
- Priors encode sparsity of selectional preferences and predicates over frames

Approximate MAP inference
- Initialize with one frame per predicate
- Iteratively merge frames (greedy search)
- Each frame merge involves greedy role alignment

Related to Titov and Klementiev, ACL'11, EACL'12
Can be extended to share alternation patterns across frames (as in EACL'12)
Can be extended to induce cross-cutting clusters of argument fillers and multi-word expressions (as in ACL'11)