**Contributions:**
- Rigorous probabilistic semantics for ED.
- No engineered features. Model explains observed sufficient statistics.
- Very fast training - scalable to massive data.
- Easy to reproduce.
- Competitive/state-of-the-art performance.
- Generalizes well on different unseen datasets.

**Context - Entity interactions**

**Probabilistic Bag Of Hyperlinks model (PBOH)**

$$\log P(e|m, c) \propto \sum_{i=1}^{n} \log P(e_i|m_i) + \sum_{i, j \in \text{hyperlinks}} \lambda_{e_i, e_j} \frac{P(e_i)}{P(e_j)}$$

- **Mention - Entity compatibility**
- **Context - Entity interactions**
- **Entity - Entity coherence**

**Candidate selection:**
- First, top 64 based on $P(e|m)$
- Then, top 10 based on $P(e|m, c)$.

**Inference:**
- MAP inference w/ Loopy Belief Propagation: $e^* = \arg \max_{e \in e^*} P(e|m, c)$
- Fast empirical convergence (typically < 3 iterations, 400ms/doc)

**Future Research**

- Alleviate data sparseness using low-dimensional entity vector representations
- Joint mention detection and entity disambiguation w/ deep representations

**EXPERIMENTS (SEE PAPER)**

- On Gerbil platform of [1]:
  - State-of-the-art performance on 11/14 datasets
  - 2nd best on 2/14 datasets
  - 10 state-of-the-art competitors
- Very good generalization performance across many datasets
- Works also for short texts (e.g. micro-blog posts)
- More experiments in the paper

**REFERENCES**