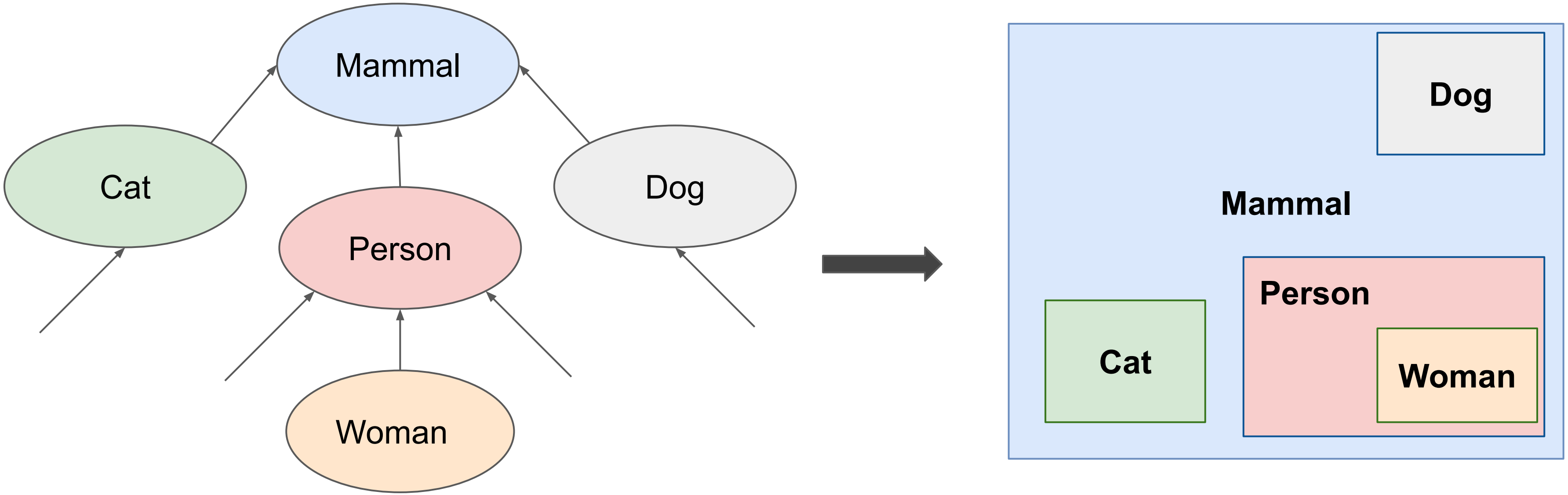
Representing Joint Hierarchies with Box Embeddings

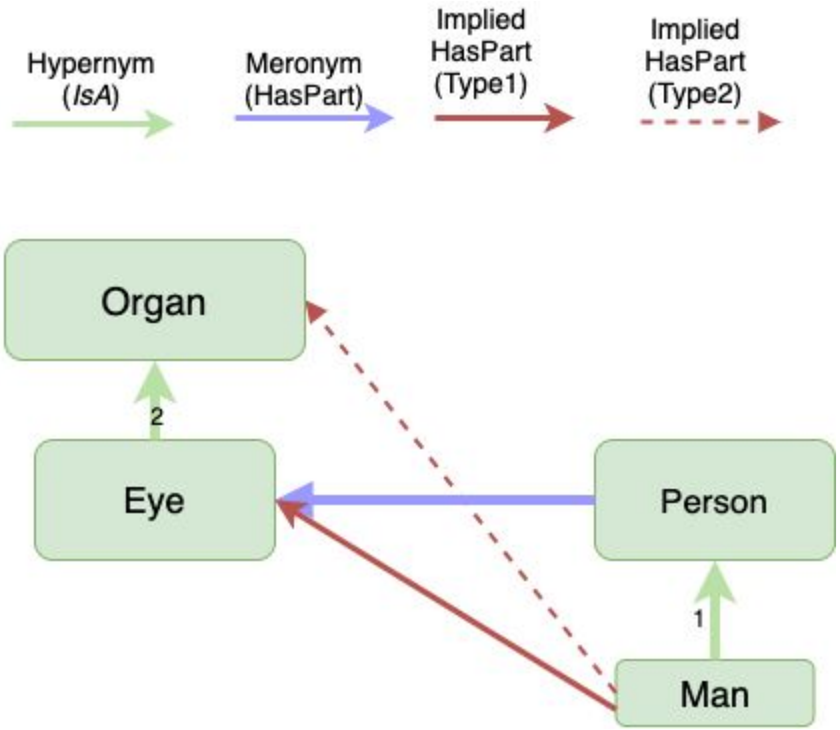
Dhruvesh Patel, Shib Sankar Dasgupta*, Michael Boratko, Xiang Li, Luke Vilnis, Andrew McCallum*

Hierarchies and Box Embeddings



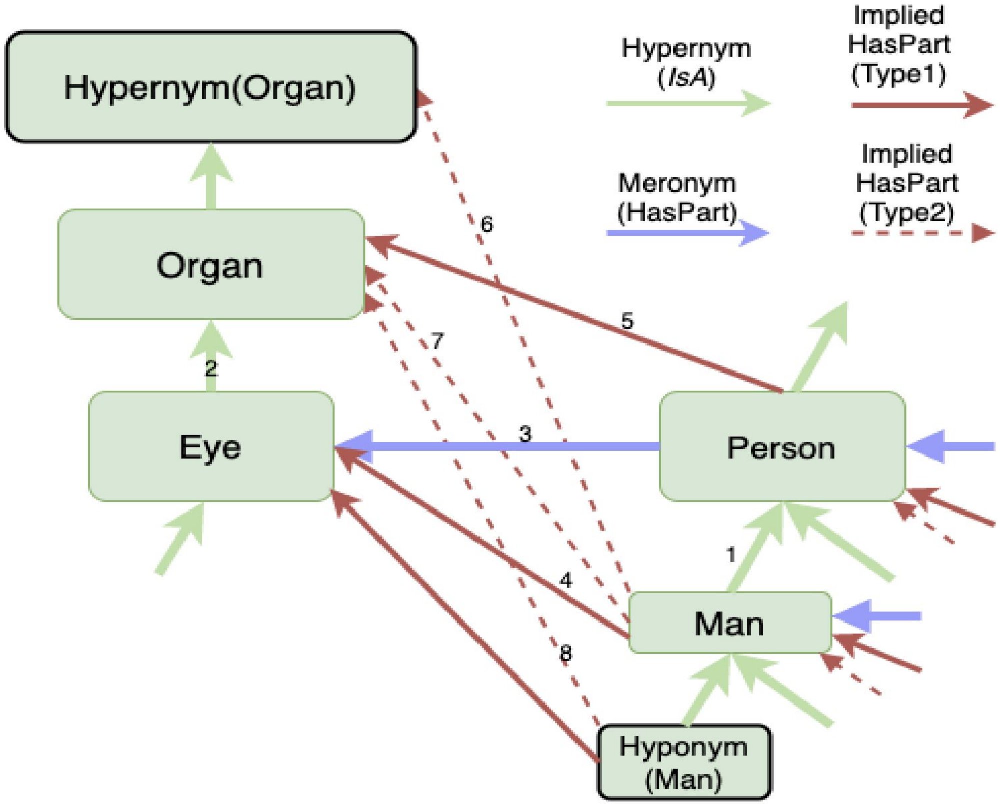
- Hierarchies are composed of **transitive** and **asymmetric** relations
- Box embeddings can naturally represent such relations using containment

Joint Hierarchies



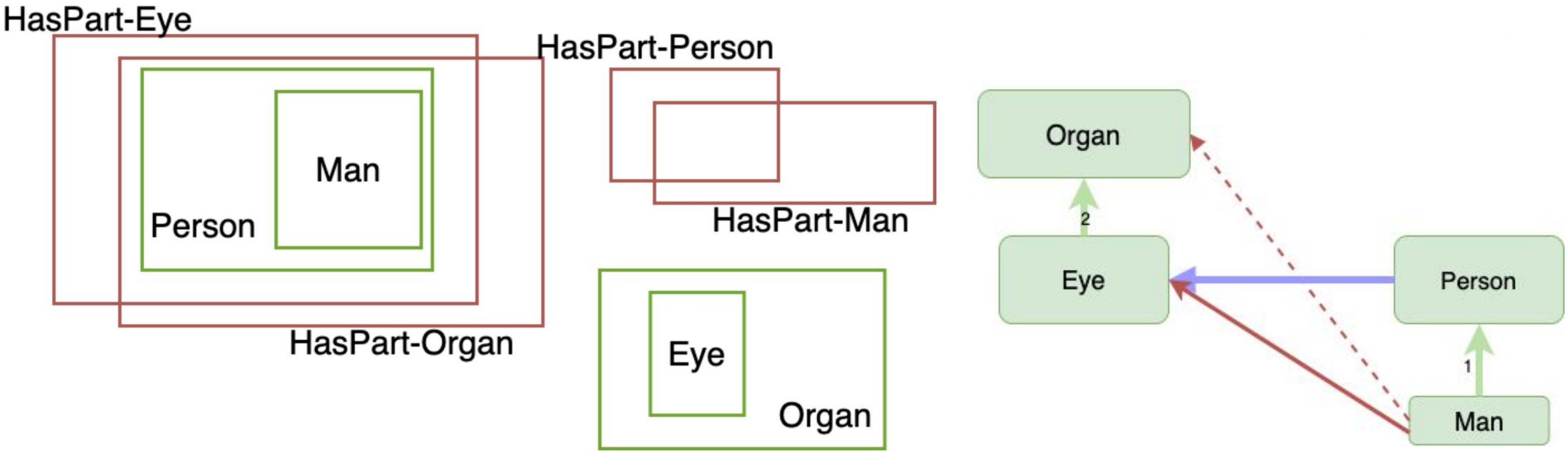
$$\textsc{IsA}(a,b) \wedge \textsc{HasPart}(b,c) \Rightarrow \textsc{HasPart}(a,c)$$
$$\textsc{IsA}(a,b) \wedge \textsc{HasPart}(c,a) \Rightarrow \textsc{HasPart}(c,b)$$

Joint Hierarchies

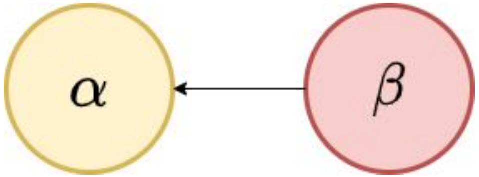


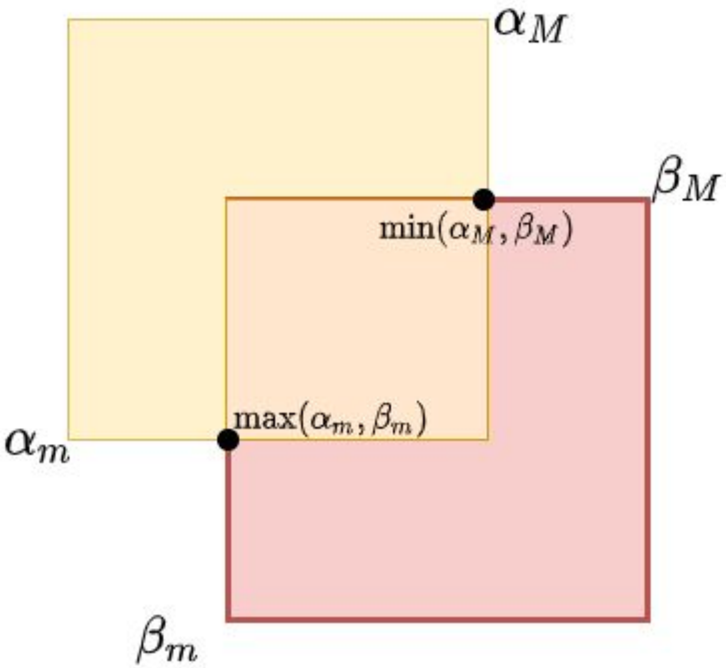
Joint Hierarchy in Box Embedding Space

Has-Part and **Is-A Box** for **each** Entity



Optimization



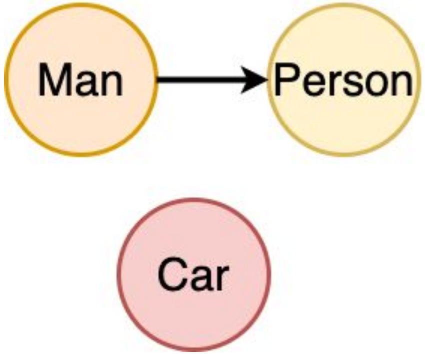


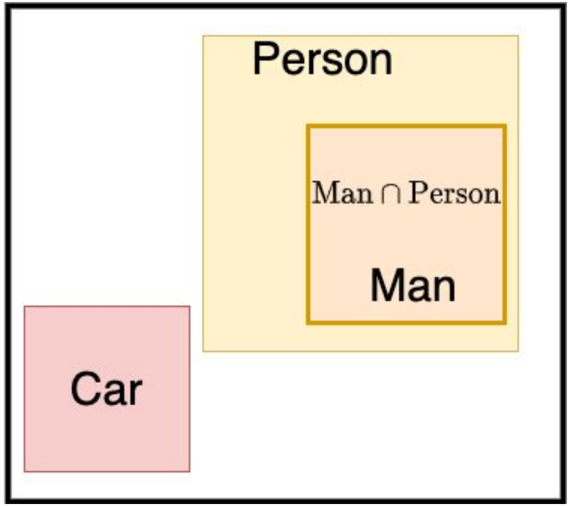
$$\Pr(\alpha|\beta) = \frac{\text{Vol}(\alpha \cap \beta)}{\text{Vol}(\beta)}.$$

$$\text{Vol}(\alpha \cap \beta) := \prod \text{softplus}_t(\min(\alpha_{M,i}, \beta_{M,i}) - \max(\alpha_{m,i}, \beta_{m,i}))$$

Optimization

$$\Pr(\alpha|\beta) = \frac{\mathrm{Vol}(\alpha \cap \beta)}{\mathrm{Vol}(\beta)}.$$

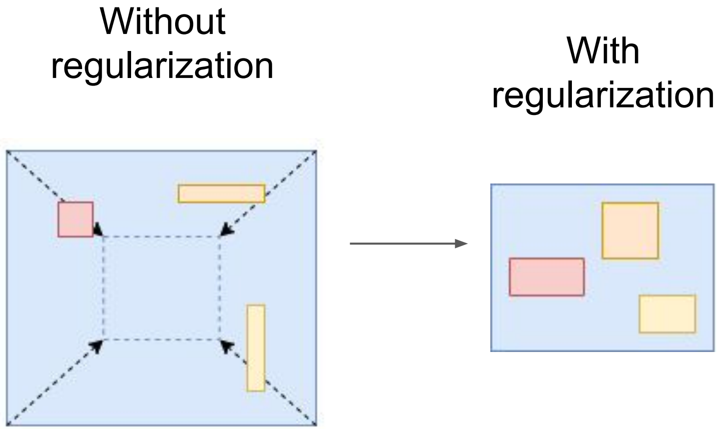




$$L = \sum_{i}^{N} -y_i \log p_i - (1 - y_i) \log(1 - p_i)$$

Regularisation

- **Using Volume Penalty**
  - Penalize when size of box is greater than a fixed volume



$$L = \sum_{i}^{N} -y_i \log p_i - (1 - y_i)\log(1 - p_i) + \boxed{\sum_{j}^{N_e} \mathbb{I}_{[\mathrm{Vol}(\alpha^{(j)}) > \tau]} \ \mathrm{Vol}(\alpha^{(j)})}$$

Regularization loss

Individual Hierarchies: Datasets

	Transitive Closure	Transitive Reduction	Validation (pos:neg)	Test (pos:neg)
Hypernym	84363	661127	28838/288380	28838/288380
Meronym	9678	30333	5164/51640	5164/51640

- Hypernym graph is more tree like.
- Meronym is less tree like:
    - contains a lot of connected components and
    - nodes with multiple parents.

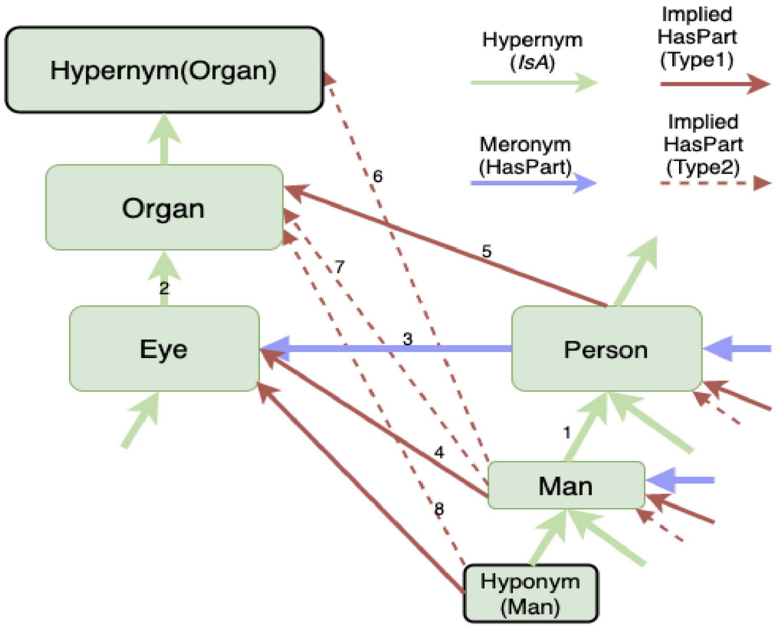
Individual Hierarchies : Results

	Hypernym				Meronym			
Transitive Closure Edges	0%	10%	25%	50%	0%	10%	25%	50%
Order Embedding	43.0%	69.7%	79.4%	84.1%	69.7%	74.1%	77.3%	81.0%
Poincaré Embedding	28.9%	71.4%	82.0%	85.3%	44.7%	73.6%	84.9%	88.0%
Hyperbolic Entailment Cones	32.2%	85.9%	91.0%	94.4%	49.70%	83.2%	88.4%	92.8%
Box Embeddings (w/o regularization)	45.4%	72.6%	81.5%	89.2%	**83.4%**	87.2%	88.7%	92.6%
Box Embeddings (Our Method)	**60.2%**	**90.0%**	**92.7%**	**94.7%**	80.1%	**91.4%**	**93.8%**	**94.3%**

- The F1 score of binary classification on the unseen test edges with a fixed set of random negatives.
- The hierarchies are modelled separately.

Joint Hierarchy: Results

Embedding Model	F1 score
Poincaré Embeddings	43.80%
Hyperbolic Entailment Cones	44.00%
TransE	57%
ComplEx	60.61%
Order Embeddings	**68.50%**
Box embeddings	68.10%



- F1 score of binary classification on the test edges of the Joint Hierarchy (all the red edges in the figure) with a fixed set of random negatives.

Conclusion

- We show that the regularized box embeddings can **learn to represent a tree-like hierarchical relation** graph with far fewer edges from the transitive closure.
- We also show that the box embeddings are **not restricted to strictly tree-like structures**.
- We propose a method to **model multiple hierarchical relations** jointly in a single embedding space.
- In all cases, our proposed method outperforms or is at par with all other embedding methods.

*Source code and processed data : https://github.com/iesl/Boxes_for_Joint_hierarchy_AKBC_2020