Probabilistic Embedding of Knowledge Graphs with Box Lattice Measures

Luke Vilnis*, Xiang Li*, Shikhar Murty, Andrew McCallum
College of Information and Computer Sciences, UMass Amherst

Problem Statement

Embedding methods which enforce order or lattice structure over concept space, such as Order-Embeddings (OE), are a natural way to model transitive relational data.

- However, OE learns a deterministic knowledge base, limiting the expressive power.
- The probabilistic extension, POE (Probabilistic Order Embeddings) provides calibrated probabilities, but lacks the ability to model negative correlations found in real-world knowledge.

We propose Box Lattices to capture negative correlations, and retain the benefits of a probabilistic model, such as multi-variable queries.

Method: Box Lattices

Motivation

- The lack of negative correlation of POE come from the fact that cones always have an overly-large intersection.
- To remedy this, we propose to learn box representations to gain back an extra degree of freedom.

Definition

- We define our box embedding using two vectors in \([0, 1]^*\), \((x_m, y_m)\) representing the minimum and maximum at each coordinate

\[
\begin{align*}
  x \land y &= \max(x, y) \\
  x \lor y &= \min(x, y)
\end{align*}
\]

- Surrogate function when positive examples disjoint

\[
\begin{align*}
  p(a \land b) &= p(a) + p(b) - p(a \lor b) \\
  &\geq p(a) + p(b) - p(a \land b)
\end{align*}
\]

- Trained by weighted cross-entropy loss

Experiments

WordNet

- Hypernym binary prediction task on WordNet, same as OE.
- Aggregated conditional probabilities (CPD) based on co-occurrence counts for ancestors.
- Examples of negatively correlated variables produced by the model:

<table>
<thead>
<tr>
<th>term1</th>
<th>term2</th>
</tr>
</thead>
<tbody>
<tr>
<td>craftswoman.n02</td>
<td>shark.n03</td>
</tr>
<tr>
<td>homogenized.milk.n01</td>
<td>apple.juice.n01</td>
</tr>
<tr>
<td>tongue.depressor.n01</td>
<td>paintbrush.n01</td>
</tr>
<tr>
<td>dusk.morn.n01</td>
<td>bathing.cap.n01</td>
</tr>
<tr>
<td>skywriting.n01</td>
<td>transcript.n01</td>
</tr>
</tbody>
</table>

Experiments (continued)

2D Embedding of a Toy Lattice

- A small hand constructed ontology over 19 concepts.
- Aggregated from synthetic examples to get marginal probabilities and the full conditional probabilities.
- Train with weighted cross-entropy.

Flickr Entailment Graph

- 45 million Flickr image captions pairs.