Multimodal KBs: Extraction & Completion

Sameer Singh
University of California, Irvine
Gray Vinyl Barstool
This sleek dual purpose stool easily adjusts from counter to bar height. The backless design is casual and contemporary which allow it to seamlessly accent any area in the home. The easy to clean vinyl upholstery is perfect when being used on a regular basis. The height adjustable swivel seat adjusts from counter to bar height with the handle located below the seat.

<table>
<thead>
<tr>
<th>Color Finish</th>
<th>Gray</th>
</tr>
</thead>
<tbody>
<tr>
<td>Style</td>
<td>Casual and Contemporary</td>
</tr>
<tr>
<td>Adjustable Height</td>
<td>Yes</td>
</tr>
<tr>
<td>Frame Material</td>
<td>Metal</td>
</tr>
</tbody>
</table>
Carles_Puyol
  isAffiliatedTo  Spain_national_under-18_football_team
  isAffiliatedTo  Spain_national_under-21_football_team
  isAffiliatedTo  Spain_national_under-23_football_team
  isAffiliatedTo  Catalonia_national_football_team

Carles_Puyol,  isAffiliatedTo,  ??
Carles_Puyol
isAffiliatedTo Spain_national_under-18_football_team
isAffiliatedTo Spain_national_under-21_football_team
isAffiliatedTo Spain_national_under-23_football_team
isAffiliatedTo Catalonia_national_football_team

Carles_Puyol, playsFor, ??

FC_Barcelona

Real_Madrid_CF
Information is in many modalities

Maybe we should be reasoning about all of these?

Time is ripe for doing multimodal stuff
Task 1: Attribute Extraction
Task 2: KB Completion
Outline

Attribute Extraction

KB Completion
Outline

Attribute Extraction

work with Robert Logan, Samuel Humeau, Mike Tung
Gray Vinyl Barstool
This sleek dual purpose stool easily adjusts from counter to bar height. The backless design is casual and contemporary which allow it to seamlessly accent any area in the home. The easy to clean vinyl upholstery is perfect when being used on a regular basis. The height adjustable swivel seat adjusts from counter to bar height with the handle located below the seat....
Gray Vinyl Barstool
This sleek dual purpose stool easily adjusts from counter to bar height. The backless design is casual and contemporary which allow it to seamlessly accent any area in the home. The easy to clean vinyl upholstery is perfect when being used on a regular basis. The height adjustable swivel seat adjusts from counter to bar height with the handle located below the seat.

{\{a_{ij}, v_{ij}\}\}  

- Color Finish: Gray
- Style: Contemporary
- Adjustable Height: Yes
- Frame Material: Metal
## MAE Dataset

Cleaned up crawl of retail products in the Diffbot Knowledge Graph

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Entities</td>
<td>2.25 million</td>
</tr>
<tr>
<td>Number of Images</td>
<td>4.172 million</td>
</tr>
<tr>
<td>Number of unique Attributes</td>
<td>2,114</td>
</tr>
<tr>
<td>Number of unique Values</td>
<td>15,380</td>
</tr>
<tr>
<td>Number of Attribute-Value Pairs</td>
<td>7.671 million</td>
</tr>
</tbody>
</table>
Challenges: Noisy, Open Domain

Real-world, Open-domain

Different Categories of Items

Redundancy and Typos

Values for `bluetooth_ver`:
- 4
- 4.0
- v4.0

Missing Images
Challenges: Weak Supervision

**Gray Vinyl Barstool**
This sleek dual purpose stool easily adjusts from counter to bar height. The backless design is casual and contemporary which allow it to seamlessly accent any area in the home. The easy to clean vinyl upholstery is perfect when being used on a regular basis. The height adjustable swivel seat adjusts from counter to bar height with the handle located below the seat.

Don’t know where the value *appears* in the input.

Don’t even know IF the attribute has a value in the images/text.
Data Quality Crowdsourcing Study

Only 46% of the queries answerable!

Top Text Attributes:
- Color
- Quantity
- Product Type
- Size
- Year
- Finish
- Warranty

Top Image Attributes:
- Color
- Product Type
- Shape
- Quantity
- Category
- Tool Type
- Exterior Color

Venn Diagram:
- Description: 65.7%
- Title: 7.4%
- Image: 18.9%
- Overlapping areas:
  - Description and Title: 1.4%
  - Description and Image: 1.7%
  - Title and Image: 3.5%
Evaluation

Noisy Data

• Great for training: massive, realistic, challenging
• Flawed for evaluation: better models might look worse*

Gold Evaluation Corpus

• Crowd-annotated with verified, *queryable* values
• 2,238 attribute-value pairs annotated so far, more coming
• Multiple values that mean the same are still a problem

Evaluation Metric, Accuracy@K

• Whether true *value* appears in the top-K predictions
• Required, since multiple values might be correct

* So far, results only on noisy data
Current (Baseline) Model

**Gray Vinyl Barstool**
This sleek dual purpose stool easily adjusts from counter to bar height. The backless design is casual and contemporary which allow it to seamlessly accent any area in the home. The easy to clean vinyl....

*So far, concatenation works best*
Extraction Results

Performance on Attribute Extraction

- Hits@1
- Hits@5

Performance metrics for different modalities:
- Baseline
- Images
- Text
- Text+Images
Multimodal Attribute Extraction

**Task:** Given text and images about an entity, extract attributes

**Dataset:** Massive, diverse, open-domain dataset

**Evaluation:** Curated, small, held-out dataset

**Baseline:** Shows the challenge, and promise, of the task

[https://rloganiv.github.io/mae/](https://rloganiv.github.io/mae/)
Outline

Attribute Extraction

work with Robert Logan, Samuel Humeau, Mike Tung

KB Completion
Outline

Attribute Extraction

KB Completion
Outline

Attribute Extraction

KB Completion

work with Pouya Pezeshkpour, Liyan Chen
Knowledge Base Completion

Entity Prediction

Link Prediction
Knowledge Base Completion

Scoring Function

<table>
<thead>
<tr>
<th>Model</th>
<th>Score $\psi_r(e_s, e_o)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESCAL [21]</td>
<td>$e_s^T W_r e_o$</td>
</tr>
<tr>
<td>SE [3]</td>
<td>$| W_r^L e_s - W_r^R e_o |_p$</td>
</tr>
<tr>
<td>TransE [1]</td>
<td>$| e_s + r_r - e_o |_p$</td>
</tr>
<tr>
<td>DistMult [34]</td>
<td>$\langle e_s, r_r, e_o \rangle$</td>
</tr>
<tr>
<td>ComplEx [33]</td>
<td>$\langle e_s, r_r, e_o \rangle$</td>
</tr>
<tr>
<td>ConvE</td>
<td>$f(\text{vec}(f([e_s; r_r] * \omega)) W) e_o$</td>
</tr>
</tbody>
</table>

Table from Dettmers, et al. (2017)
Restrictions in the Model

Each object has a vector representation:
- Limits number of objects
- Large number of parameters
- Is not compositional (doesn’t generalize)

What about other kinds of objects?
- Dates and Numbers: should generalize
- Text: Names and Descriptions
- Images: Portraits, Posters, etc.
Multimodal KB Embeddings

- Object → Encoder → $v_0$
- Entity → Lookup → $v_0$
- Images → CNN → $v_0$
- Text → LSTM → $v_0$
- Numbers, etc. → FeedFwd → $v_0$
Multimodal KB Embeddings

All kinds of objects modeled directly (not as “features”)
Augmenting Existing Datasets

<table>
<thead>
<tr>
<th>MovieLens-100k-plus</th>
<th>YAGO3-10-plus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relations</td>
<td>13</td>
</tr>
<tr>
<td>Users</td>
<td>943</td>
</tr>
<tr>
<td>Movies</td>
<td>1682</td>
</tr>
<tr>
<td>Posters</td>
<td>1651</td>
</tr>
<tr>
<td>Ratings</td>
<td>100,000</td>
</tr>
<tr>
<td>Relations</td>
<td>37 → 45</td>
</tr>
<tr>
<td>Entities</td>
<td>123,182</td>
</tr>
<tr>
<td>Structure Triples</td>
<td>1,079,040</td>
</tr>
<tr>
<td>Numbers (Years)</td>
<td>1651</td>
</tr>
<tr>
<td>Descriptions</td>
<td>107,326</td>
</tr>
<tr>
<td>Images</td>
<td>61,246</td>
</tr>
</tbody>
</table>
Link Prediction Results

Hits@1 for Yago3-10+

- DistMult
- ConvE

- Links
- +Numbers
- +N+Text
- +N+T+Images
Predicting Multimodal Values

Predicting Numerical Values in Yago3-10+

RMSE

Links  +Text  +Images  +Text+Images
Predicting Multimodal Values

“Die Hard”

“Action, Thriller”

“The Band Wagon”, “Underground”, “Roseanna’s Grave”


“Drama, War”, “Action, Drama, War”, “Comedy, Drama, War”
Multimodal KB Completion

Task: Given graph and other kinds of objects, predict links

Dataset: Extended existing datasets to introduce benchmarks

Baseline: Unique, multimodal model gets impressive gains

Decoding: Some promising results for decoding the objects

Code and datasets coming soon!
Multimodal AKBC

Important to do now!

Introduced two multimodal tasks
• Attribute Extraction
• KB Completion

Datasets, metrics, and models

Future Work:
Tell us what you will do!
Thank you!
sameersingh.org
sameer@uci.edu
@sameer_