Weekly report

Aug. 19 – Aug. 23

Progress

- Improve the system on the augmented dataset
- Build the evaluation metric and analyze the performance
- Deploy the web application

Scheme

- Build scene graph from the input text. Subject-action-object triplets are captured based on the dependency tree. The associated properties may not necessarily be same as the ones in the database because of the variety of phrasing.
- Entity grounding. Bind each noun to an entity keyword in the database. Properties (subject or object) are determined according to the priors. Triplets are thereupon rectified. GloVe is used to measure similarity.
- Action intension detection. Determine if a verb involves a strong or weak interaction, based on the priors.
- Combination enumeration
  - **Subjects**: Two or more character subjects can be grouped.
  - **Objects**: Objects associated with strong interactions are bound with their subjects while objects associated with weak interactions can be exchanged to another subjects.
- Propose layer candidates. For each triplet (or group), find several layers in the database containing all or part of the mentioned entities, ranked by entity sets mIoU, verb relatedness and keyword frequency priors. ConceptNet is utilized.
- Global optimization. Find the layer combination with maximized inter-category layer relatedness, based on the collocates in the picture base. This step can be replaced by a picture reasonability discriminator.

Problem

- Entities in the description are not included in the layer base.
- Entities in the picture are not mentioned in the description. Background and most accessories will never be explicitly mentioned.
- Tokens referring to the same object mismatch when binding words in description to keywords
- Improper partition. Whether or not to group two entities in the same layer is ambiguous.
- Measure word not considered
- Query efficiency

Skills

- Sophisticated text processing techniques
- Deploy web applications
- Related tools
  - Hunspell
  - Flask and Jinja

To do

- Implement the modified pipeline
- Improve the query efficiency
- Add interactions to the web application