Anaphoric Phenomena in Situated Dialog: A First Round of Annotations

Sharid Loáiciga

Simon Dobnik, David Schlangen

CLASP – Department of Philosophy, Linguistics and Theory of Science
University of Gothenburg

October 16th, 2022

Annotation of Situated Dialog

► We extend the ARRAU guidelines to annotate the texts of the *Tell-me-more* corpus.

Poesio, Camilleri, Carretero-Garcia, & Artstein (2021). Arrau 3 annotation.
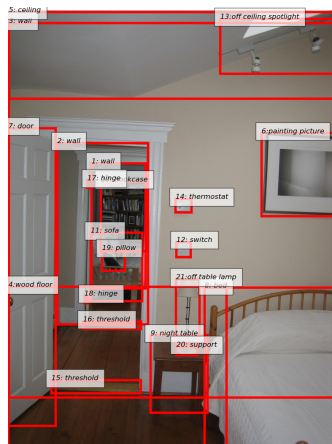
Situated Dialog Data

Tell-me-more

1. it's a bedroom scene with the bed partially visible
2. the bed has a curved wooden headboard with slots like a fence
3. there is framed art hanging above the bed
4. to the left of the bed is a door, which is open
5. there is a small square nighstand next to the bed which has a lamp on top of it

Situated Dialog Data

Tell-me-more: 5,701 image-document pairs. Images and BBs come from the ADE dataset.

1. it's a bedroom scene with the bed partially visible
2. the bed has a curved wooden headboard with slots like a fence
3. there is framed art hanging above the bed
4. to the left of the bed is a door, which is open
5. there is a small square nighstand next to the bed which has a lamp on top of it

Ilinykh, Zarrieß & Schlangen (2019). Tell me more: A dataset of visual scene description sequences. INLG.

3

Annotation Tool

4

This Release

- Single annotations for 500 documents ($\approx 10\%$ of total data).
- Double annotations for 50 documents for validation.

`https://doi.org/10.5281/zenodo.7084861`

IAA

- Krippendorff's α 0.56 $\pm$ 0.23,
  according to Passonneau (2004), *Computing reliability for coreference annotation*, LREC.

1. Mostly [[this room]] is [[a bed]].

2. [[There]] is [[a lamp on a small white nightstand next to the bed.]]

3. [[The bed]] has [a [light blue bed skirt]] and [[white comforter]] and [[4 white pillows]].

4. [[There]] is [[a blue dresser with a lamp] on [it]] .

5. [[There]] is [[a full length window with vertical shades]].

Interesting Multimodal Items

Collective entities

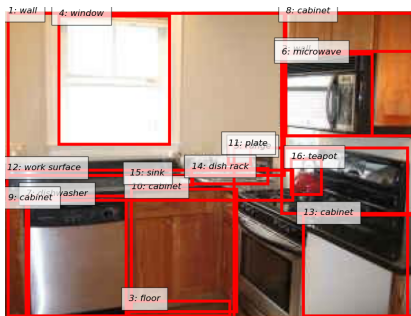`cardinality: unique or group to account for collective entities.`

1. There is **four chair red laquer dining set** shown in the image.
2-3 [...] 4. The pools is surrounded by green shrubbery.
5. [...]

Interesting Multimodal Items

Almost every document starts with seemingly `old` information.

1. **It**'s a well-lit kitchen with stainded wooden cupboards. 2. There's a microwave mounted over **the stove**, which has a red tea kettle on it. 3. **The appliances** are black and stainless steel in **the kitchen**. 4. **The countertops** look like they're black granite. 5. The window has sunlight streaming in and it's very brightly light.
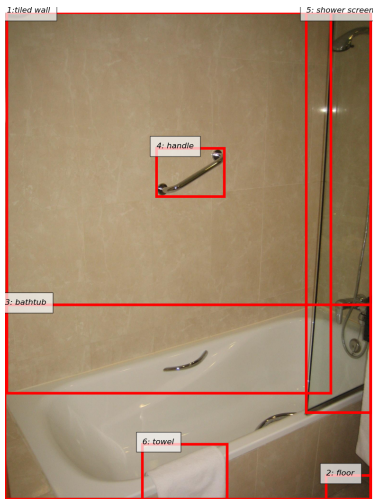
Interesting Multimodal Items

We distinguish between predefined visible objects and visible objects without a predefined ID.

```
- on image:  yes|no
  - if yes:
    - bounding box:  yes|no
      - if yes:
          choose label
```

1. This is a picture of a bathtub. 2. The tub is white. 3. The wall and *base of the tub* are brown. 4. [...] 5. [...]

Conclusions

▶ Annotation makes theoretical advances possible.

▶ Extending the annotation of coreference to the V&L domain is crucial to understand the relationship between reference and coreference.

▶ Although small, the corpus is useful for linguistic studies, validation of automatic methods linking entities to the image, and as an example for the annotation of more complex datasets.

Thank you