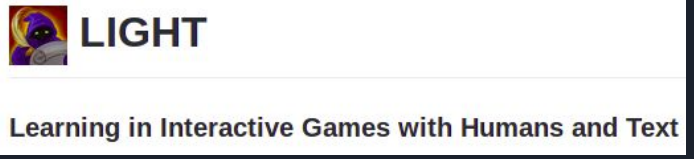
Learning to Speak and Act in a Fantasy Text Adventure Game

Jesús Javier Calleja
Nerea Losada

Index

1. LIGHT environment
2. Related work
3. Learning methods
4. Results
5. Conclusion
6. Demo



- The environment is based in a text adventure game where agents can perceive, emote and act and in the meantime talk to other agents.
- Implemented using PyTorch library.
- ParlAI framework.

Description of the environment: LIGHT

- Large-scale and configurable text adventure environment
  - To research on learning grounded language and actions
- It features both humans and models as integrated agents by a multi-player fantasy MUD (multi-user dungeon)
- Crowdsourced: locations, objects and their affordances, characters and their personalities, dialogues and actions
- To train agents and to evaluate them in situ

Related work

LIGHT framework

allows learning from:

- actions AND
- two-way dialogue

Other existing simulations

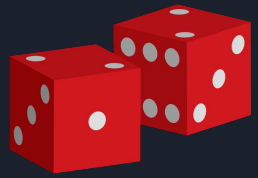
- actions OR
- two-way dialogue

The language: based on the crowdworkers that participated

- inherit properties of natural language: ambiguity and coreference

Learning methods

Baselines



- Information Retrieval
- Starspace
- FastText

Models

Transformer based models

- Transformer Memory Network (Generative model)
- BERT Bi-Ranker
- BERT Cross Ranker

Results

Method	Test Seen Dialogue R@1/20	Test Seen Action Acc	Test Seen Emote Acc	Test Unseen Dialogue R@1/20	Test Unseen Action Acc	Test Unseen Emote Acc
Random baseline	5.0	12.2	4.5	5.0	12.1	4.5
IR baseline	23.7	20.6	7.5	21.8	20.5	8.46
Starspace	53.8	17.8	11.6	27.9	16.4	9.8
Transformer MemNet	70.9	24.5	17.3	66.0	21.1	16.6
BERT-based Bi-Ranker	**76.5**	42.5	25.0	70.5	38.8	25.7
BERT-based Cross-Ranker	74.9	**50.7**	**25.8**	69.7	51.8	28.6
Human Performance*	*87.5	*62.0	*27.0	*91.8	*71.9	*34.4

	Dialogue R@1/20	Action Acc	Emote Acc
BERT-based Bi-Ranker	76.0	38.7	25.1
actions+emotes only	58.6	18.3	10.6
dialogue only	68.1	39.4	23.6
dialogue+action+emote	73.2	40.7	23.1
dialogue+persona	73.3	41.0	26.5
dialogue+setting	70.6	41.2	26.0
dialogue+objects	68.2	37.5	25.5

Conclusion

- Analyzed a variety of models and their ability to take advantage of the grounding information

LIGHT platform

Agents can act and speak in a rich and diverse environment of locations, objects and other characters