Symbol Grounding and Task Learning from Imperfect Corrections

Mattias Appelgren,
Alex Lascarides

INTRO

- We work in an Interactive Task Learning setting with the goal of teaching agents new tasks through interaction with a teacher
- Teachers aren't perfect and can act in ways the learner does not expect
- In this paper we extend an agent that learns from corrections to deal with a teacher that sometimes adds or misses corrections, which clashes with the agent's understanding of coherence

METHODS

1. An agent must learn a constrained tower building task by learning the constraints and how to ground the concepts populating them. E.g. "red blocks should be on blue blocks"
2. The simulated teacher will either miss or add corrections with some probability
3. Our agent uses its current knowledge of the world and its belief about the teacher's probability of failure to infer whether or not it should trust the teacher's feedback and update its knowledge

RESULTS

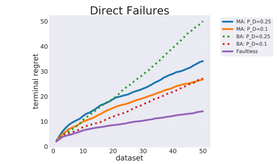
- If the agent ignores the fact that the teacher may be faulty then it will often fail to learn
- Our updated agent that takes into account the teacher's deficiencies reaches similar levels of proficiency as an agent that learns from a faultless teacher

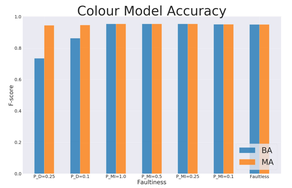
DISCUSSION

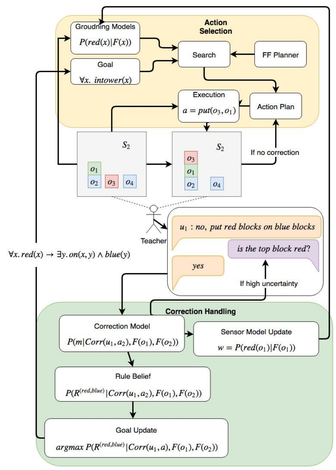
- Our model for learning from correction is based on concepts of coherence.
- We show that the models can be made robust to a suboptimal teacher simply by considering separately what the teacher said and what it means for something to be coherent

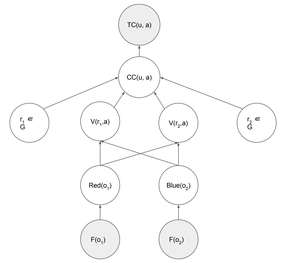
When considering separately what the teacher said and when it is coherent to say it, an agent can **learn symbol grounding** successfully even from **a teacher which makes mistakes**

Take a picture to **download** the **full paper**









THE UNIVERSITY of EDINBURGH
informatics