

Visually-Grounded Dialogue Models: Past, Present, and Future

Raquel Fernández
University of Amsterdam, The Netherlands
raquel.fernandez@uva.nl

Abstract

The past few years have seen an increasing interest in developing neural-network-based agents for visually-grounded dialogue, where the conversation participants communicate about visual content. I will start by discussing how visual grounding can be integrated with traditional task-oriented dialogue system components. Most current work in the field focuses on reporting numeric results solely based on task success. I will argue that we can gain more insight by (i) analysing the linguistic output of alternative systems and (ii) probing the representations they learn. I will also introduce a new dialogue dataset we have developed using a data-collection setup designed to investigate linguistic common ground as it accumulates during visually-grounded interaction.