

Relational reasoning and generalization using non-symbolic neural networks

Atticus Geiger

Stanford, Stanford, California, United States

Alexandra Carstensen

Stanford, Stanford, California, United States

Michael Frank

Stanford University, Stanford, California, United States

Christopher Potts

Stanford University, Stanford, California, United States

Abstract

Humans have a remarkable capacity to reason about abstract relational structures, an ability that may support some of the most impressive, human-unique cognitive feats. Equality (or identity) reasoning has been a key case study of abstract relational reasoning. This paper revisits the question of whether equality can be learned in non-symbolic neural networks. We find that simple neural networks are able to learn basic equality with relatively little training data. In a second case study, we show that sequential equality problems (learning ABA sequences) can be solved with only positive training instances. Finally, we consider a more complex, hierarchical equality problem, and find that this task can be solved with either a vast amount of training data or pre-training on basic equality. Overall, these findings indicate that neural models are able to solve equality-based reasoning tasks, suggesting that essential aspects of symbolic reasoning can emerge from data-driven, non-symbolic learning processes.