

Progress in building a machine that can ask interesting and informative questions

Anselm Rothe

New York University

Brenden Lake

New York University

Todd Gureckis

New York University

Abstract: Asking creative questions is a hallmark of human cognition. In comparison, machine learning systems that attempt to mimic this ability are still extremely limited (e.g., current chatbots ask questions based on preprogrammed routines). In the present work, we developed a computational model of question generation. Based on a corpus of questions collected from online participants playing an information-seeking game, we designed a “grammar of questions.” The grammar is powerful enough to represent all human questions we collected and thus defines the “question space.” Given a particular context (game scenario), people are more likely to ask (generate) some questions than others. Our computational model predicts these likelihoods, that is, a probability distribution over the question space. In addition, the model can generalize to novel contexts. Key model ingredients are informativity, compositionality, and length of a question.