

The Task Task: Creative problem generation in humans and language models

Junyi Chu

Harvard University, Cambridge, Massachusetts, United States

Jennifer Hu

Harvard University, Cambridge, Massachusetts, United States

Tomer D. Ullman

Harvard University, Cambridge, Massachusetts, United States

Abstract

Machine creativity is on the rise. Recent studies find that large language models achieve human performance on common psychological tests of creativity, which often pose a given problem and ask for novel or unusual solutions. But can AI go beyond producing solutions for given problems, to creatively propose new problems? We present the Task Task, a novel test that asks participants to come up with creative problems. In this test, we assess the ability of humans and GPT-4 to design challenge tasks for a game show. We evaluated proposed tasks using crowdsourced subjective creativity ratings, as well as computational measures of linguistic complexity and semantic content. We found that GPT-4 achieves similar scores as humans on creativity, originality, and judgments of how fun or difficult the tasks are. However, model-generated output tends to be shorter and connect more semantically distant concepts. We discuss implications and future directions for the psychology of creativity.