

Thesis Project Portfolio

**CRUGS: A Language Dataset for Compositional and Relational
Understanding Using Geometric Shapes**

(Technical Report)

Care Ethics in Apple v FBI
(STS Research Paper)

An Undergraduate Thesis

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Kevin Ivey

Spring, 2023

Department of Computer Science

Table of Contents

Sociotechnical Synthesis

CRUGS: A Language Dataset for Compositional and Relational
Understanding Using Geometric Shapes

Care Ethics in Apple v FBI

Prospectus

Sociotechnical Synthesis

In my technical work, I explored the reasoning capabilities of large language models (LLMs) on progressively challenging tasks. Meanwhile, my science, technology and society (STS) work studied how Apple acted in response to the FBI's demands to unlock an encrypted iPhone. While my technical work assessed the capabilities of an algorithmic technology, my STS work studied the impact of an algorithmic technology on society. Despite being different in nature, they are connected through the usage of algorithmic technology that is available for public use.

LLMs have been shown to exhibit a number of capabilities; however, these capabilities need to be further studied before being used for real-world tasks. My technical work explored the capabilities of LLMs on compositional reasoning tasks. My work consisted of two phases: 1) Developing a synthetic benchmark consisting of geometric shapes in a 2-D canvas; and 2) Evaluating several recent LLMs on the developed benchmark. The synthetic benchmark contained natural language descriptions of geometric shapes in a canvas along with progressively harder tasks that ask various aspects about the canvas. Additionally, the benchmark contained the ground truth answer for each task. To evaluate the LLMs, each model was given a description of the canvas and the corresponding task from the benchmark. The model's response was recorded and scored against the ground truth answer. By evaluating against the entire set of tasks in the benchmark, the model's overall capabilities were assessed.

My STS research explored the case of "Apple vs. FBI" through the lens of care ethics. Through an analysis of legal case documents and Apple's letter to customers, I explored the relationship between Apple and its customers. I claimed that Apple holds an advantageous position in its relationship with its customers. As such, it has a duty to care for its customers. By refusing to aid the FBI in unlocking the encrypted iPhone, Apple demonstrates three aspects of

care: attentiveness, responsibility, and competence. Therefore, Apple acted morally by adhering to the tenets of care ethics.

Working on my technical and STS project in tandem allowed me to better understand the future impact of novel technologies. While my technical work allowed me to explore the capabilities of state-of-the-art LLMs, my STS work reminded me of the unintended impact that technology can have in the future. In the case of LLMs, a number of ethical issues have already been raised despite their recency. While encryption and LLMs are different in nature, both technologies are widely available to the public and have uses beyond their initial purpose. Studying the role of encryption in an ethical case has allowed me to consider the future societal impact of my technical work involving LLMs.