Turtle-like Geometry Learning: How Humans and Machines Differ in Learning Turtle Geometry

Sina Rismanchian, Shayan Doroudi, Yasaman Razeghi

University of California, Irvine, srismanc@uci.edu

Abstract

The remarkable progress made by large language models (LLMs) and, more broadly, cutting-edge generative artificial intelligence (AI) has prompted researchers to evaluate these models in various fields where humans typically excel, like mathematics and science question answering (Team et al. 2023). The extent to which the reasoning and learning processes in large language models (LLMs) parallel human learning remains a contentious topic among researchers. Here, we aim to examine the strengths and shortcomings of these models in performing visual programming tasks, specifically in the domain of Turtle Geometry (Abelson and DiSessa 1986).

Recent work in the field of psychology shows that infants possess two different systems in the domain of geometry: a *form* system which is used for object detection and a *place* system that is used when navigating (Spelke and Kinzler 2007; Dillon 2023). Furthermore, researchers have shown that humans tend to prioritize the place system over the form system when perceiving abstract geometric shapes (Lin and Dillon 2023). We believe such findings resonate with the ability of young children (and adults) to learn Turtle geometry, a form of geometry that can be explored through programming, in contrast to Euclidean geometry as traditionally taught in schools. Indeed, recent work even suggests the human mind might recognize shapes in terms of procedural programs in a Turtle geometry-like language (Sablé-Meyer et al. 2022). The problem of examining the abilities of LLMs in performing Turtle geometry tasks is significant as visual programming is a niche human ability that draws on both perception and problem-solving. The answer to this question can enhance our knowledge about the fundamental characteristics of these models.

We hypothesize that state-of-the-art generative AI multimodal models such as GPT4-V, lack the human ability to visualize and procedurally generate abstract shapes and patterns—and to quickly learn how to do this. To test this hypothesis, we (1) curated a dataset of images generated with Turtle programming and subsequently tested a large-language model in its ability to generate programs for those shapes and (2) ran preliminary human subject experiments with graduate students who have a background in programming and mathematics but are not familiar with Turtle geometry, to see how quickly they learn to create some of the shapes in our dataset.

When tested on our dataset, we found that while the model is able to produce coherent Python code (using the Turtle module) that runs without errors, it is often unable to write programs that can recreate shapes that are found to be easily learned by human learners. Our experiments with students show three significant differences between how humans and machines learn in the domain of visual programming: (1) humans can adopt the navigation system in the domain, while transformer models mostly rely on object recognition in their architecture, (2) humans can generate subgoals to engage in active learning strategies such as trial and error that help them learn from feedback while on the other hand, current deep neural networks are passively fed by input-output pairs, and (3) humans can hold different forms of visual abstractions from a single shape and choose the one that eases their programming experience while to best of our knowledge, these models lag behind humans in abstraction tasks (Moskvichev, Odouard, and Mitchell 2023) and we have not seen any reports on the flexibility of these models in holding different abstractions. We conclude by bringing up different research questions that are key to understanding how human-like learning in the domain of Turtle geometry works and how AI research can be informed by these findings.

References

Abelson, H.; and DiSessa, A. 1986. *Turtle geometry: The computer as a medium for exploring mathematics*. MIT press.

Dillon, M. R. 2023. Divisive language. 10.31234:1706.03762.

Lin, Y.; and Dillon, M. R. 2023. We Are Wanderers: Abstract geometry reflects spatial navigation. *Proceedings of the National Academy of Sciences of the United States of America*, 110: 35.

Moskvichev, A.; Odouard, V. V.; and Mitchell, M. 2023. The ConceptARC Benchmark: Evaluating Understanding and Generalization in the ARC Domain. *arXiv preprint arXiv:2305.07141*.

Sablé-Meyer, M.; Ellis, K.; Tenenbaum, J.; and Dehaene, S. 2022. A language of thought for the mental representation of geometric shapes. *Cognitive Psychology*, 139: 101527.

Spelke, E. S.; and Kinzler, K. D. 2007. Core knowledge. *Developmental science*, 10(1): 89–96.

Team, G.; Anil, R.; Borgeaud, S.; Wu, Y.; Alayrac, J.-B.; Yu, J.; Soricut, R.; Schalkwyk, J.; Dai, A. M.; Hauth, A.; et al. 2023. Gemini: A family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*.