On the Mathematics of Learning and Intelligence

Asvin G

October 17, 2025

I have a mathematical background in algebraic geometry, logic, computation, and number theory. For the last two years, I have focused on developing a mathematical understanding of learning and intelligence, initially motivated by AI alignment. My interests have since broadened, leading to the following research questions.

1. **The Geometry of Proof Space.** I hypothesize that a geometric structure on the space of proofs explains which proofs are discoverable in practice. The space of proofs branches exponentially, making naive search intractable, yet humans prove complex theorems with limited search. How? I believe a key component is our ability to compute and update 'belief' strengths in statements—a richer function than binary provability. This belief function, updated via partial evidence and consistency, could also quantify a statement's 'interest.' This is an active area of my research.

2. **A Mathematical Theory of Learning.** The mathematical foundations of deep learning are specific to neural networks and heavily analytic. However, learning is a more general phenomenon, prolific across biology and beyond, and existing theory fails to fully explain the empirical success of deep learning. Inspired by Turing, I aim to develop a new 'synthetic,' combinatorial theory of learning systems to define what it means 'to learn' generally. I have made progress on a model, inspired by neuroscience, that rigorously defines 'goals' for such systems. This framework appears compatible with modern AI techniques, particularly reinforcement learning.

3. **The origin of consciousness as a learning tool.** I find Graziano's attention schema theory of consciousness intriguing and motivating. A future goal is to extend the model in the previous bullet point to include 'self-and-other-models', and to locate the importance of these models for the dual purposes of learning from experience and for making decisions in uncertain situations. Once again, I find myself drawing inspiration from both biology and machine learning in constructing these mathematical models.

The overarching goal of this research is to understand learning systems well enough to reason meaningfully about their capabilities and constraints.