AI and Empiricism

Even though the mathematics underlying Large Language Models (LLM) such as ChatGPT and Claude is well-known advanced linear algebra and a little calculus, the statement has been made that exactly how the models work is not understood. In fact, an article in MIT Technology Review (The new biologists, Will Douglas Heaven, Vol 129, January/February 2026) describes a situation in which a companion LLM was built solely to interpret and describe the actions of the primary LLM. So, the question is why can we not understand the workings of the LLM and why are the results sometimes surprising if we understand the math behind it?

In my opinion, this dilemma exists because the LLM is operating in a vector space the dimension of which is far beyond the ability of most humans to visualize. Hence, the output of the LLM becomes an empirical question. That is, theoretical analysis of the input-to-output operation is not possible because that many dimensions, say, 1000 or so, cannot be visualized. So, the best we can do is catalog the output.

The book, Flatland: A Romance of Many Dimensions (Edwin A Abbott, originally published 1884, Seeley & Co, London) provides an analogy. Flatland is composed of only two dimensions. So, if a three-dimensional object such as a baseball were to pass through, it would appear as a small hole, expanding to the diameter of the baseball, gradually shrinking and then disappearing. The inhabitants could catalog the various effects of three-dimensional objects passing through Flatland but could not analyze them much further since they have no concept of and cannot visualize the additional dimension. So, they are limited to “here comes another one” empiricism.

With the LLM we are similarly limited to “let’s try it and see” empiricism. We are a bit better off than the Flatlanders since we understand there are more dimensions and know the math. However, we stumble when attempting to analyze the calculation in all dimensions. It is just too much data for most of us to handle. Hence, we become empiricists or perhaps phenomenologists, catalog the outputs and sometimes marvel and sometimes frown at the results.

Perhaps, this is the point where Artificial Intelligence (AI) begins. That is, if we have it all worked out in our heads and only use modern processing power to handle the drudgery of calculation, that is hardly what we think of a AI. On the other hand, if we do not have it all worked out and we require the processing power of a modern computer to generate the result, it is AI that is generating this result while we empiricists look on.

Disclaimer: I suppose there are those among us that can visualize many, many dimensions; even though, I have not met anyone like that, as far as I know. When I use the term “we”, I specifically exclude those extraordinary individuals.