### What about predicting smarter-than-human AI from lines on a graph?

#### **Lines on a graph don’t help because nobody knows which point on the vertical axis corresponds to everything getting weird.**

One class of successful predictions involves taking a straight line on a graph, one that has been steady for many years, and predicting that the straight line continues for at least another year or two. It doesn't always work, to be clear, sometimes the line just changes; but it often works; it is a case where people make successful predictions in practice.

The great trouble with this method is that often what we really want to know is not “Out of curiosity, how high up will this line on the graph be by 2027?” Often, what we want to know is more, “What happens qualitatively, if this line does keep going up? Where on the graph would that event happen?” And then that part turns right back into a hard prediction problem.

We don't know what happens if you throw exactly 10^27 operations at training a transformer network, and we don't think anyone else knows either; people have been bad at that kind of prediction historically. How well will the resulting AI be able to play chess? We can’t guess, just like how nobody was able to guess how well GPT-4 was going to play chess before just trying it. Almost certainly, what 10^27 operations can do exactly, depends on which algorithms the compute is being poured into; and new algorithms get invented over time. So we have not spent any of this book extrapolating lines on a graph to predict exactly when somebody will throw 10^27 operations at training an AI. If you really want a book like that, somebody else will happily sell you one, and everything in it besides the graph itself will be invalid.

Some calls are just very hard to make, as far as we can tell. The book focuses on the easy calls.