AI Can Help Us Make Better Decisions

One of the main reasons I'm so interested in Artificial Intelligence (AI) is because we're terrible at making decisions and I believe AI can help.

Here's why:

1. We don't consider very many alternatives and tradeoffs.

   Should I order pizza or make a salad? Should I go to the gym or start a new season of Peaky Blinders? Do I hire the junior person willing to learn or the experienced person stuck in their ways? Should I move to Austin for the job I really want or stay here for the job I have? And should we invest more in taking care of the sick or in feeding the hungry?

   Of course we don't consider all the possible alternatives and tradeoffs for every decision that comes up—we'd never get anything done. AI is much better/faster at this.

   And if you think about it, we're just the consequence of a long sequence of decisions—made by us or other people. Those decisions explain how we got here, who we are, and what our potential is. Something so important can't be left so vulnerable to haste.

2. We don't think from first principles.

   Because we don't think from first principles, our emotions keep getting in the way. And because our emotions keep getting in the way, we can't make objective decisions.

   AI is objective.

   The obvious follow-up question: How can AI be objective if subjective/emotional humans need to develop it?

   An AI is just a bunch of "algorithms" (or "models") layered on top of one another. Those algorithm work on an explicit definition of fundamental rules and principles for how things work (or "first principles").

   First principals are fundamentally much easier to separate our emotions from.

   Example of first principles at play:

   - We don't manually program in every slight left or slight right an AI driven plan or rocket needs to take to get from point A to point B. Instead, we've programmed first principles in physics, like Newton's Laws of Motion and the acceleration of gravity.

   - Should we enforce rent controls to help those that can't afford housing? It's obviously a very emotional and divisive topic. However, the first principle at stake here is *price controls* and almost every econ textbook I've read and every economist I've ever talked to (or even any person that's just studied some economics) is clear about the effects of *price controls*: in the short term they increase demand and in the long term they decrease supply—which ultimately increase prices. In that context, it's easy to see how if price controls were implemented for housing, even if it helped some people in the short term, in the long term it would just make housing even more out of reach for those who already can't afford it (cue the housing problem in California).

3. We aren't consistent with our values.

   First principles thinking also leads to consistency. This is extremely important.

   When people are in the minority, one of their *first principle values* is individual liberty—the idea that each person should be able to live their lives in accordance with whatever values they have, as long as they're not hurting anyone else of course.

   But as soon as those same people become the majority, they realize *other people* aren't living their lives by their values and that *first principle value* gets traded in for social cohesion.

4. We don't admit our mistakes.

   There are two elements in AI called "modeling" (or "learning") and "reinforcement learning" that forces us to be explicit about how outcomes should be measured and the actions that should be taken according to those results.

   Have you ever heard of a politician propose a new government agency and say, "If [insert new agency] doesn't reach x, y, or z goal in the next 5 or 10 years we'll know we were wrong about our assumptions and will backtrack accordingly"?

   Why aren't people more explicit about their goals and plans to evaluate their progress? The answers are probably obvious.

   AI can of course make wrong decisions too. But these feedback mechanisms ensure accountability.

Objective Frameworks

We can develop objective frameworks to make decisions in every field.

My field, Computer Science, already has these frameworks. We use asymptotic space and time complexity formulas to measure our coding decisions and Service Level Indicators (SLIs), like availability, throughput and response time, to measure our system design decisions.

I love studying Economics. People think Economics is the study of scarcity but it's really the study of decision making. In a world with scarce (limited) resources, what we do with those resources is what matters. There are economic models, like the Production Possibilities Frontier Model, that provide an economic value answer to how we should allocate our time. There are even models like the Lorenz Curve to help us understand what to do about income inequality. These models provide a guiding framework to decide which issues we should focus on and how to allocate our time, energy, and physical resources to create a better world.

Ethics is why I dove into Philosophy and Value Theory. Did you know philosophers like Plato, Aristotle, and Aquinas actually provided means to measure how ethical ideas are (i.e., Devine Command Theory, Virtue Theory, Natural Law Theory, etc.)? If we want to decide whether to invest more in taking care of the sick or feeding the hungry purely on an ethical basis, we actually can.

Art used to make me so uncomfortable because a framework for deciding what's *good art* and what's *bad art* isn't as obvious. Until I realized that if we decide the purpose of art is to invoke an emotional response, we can measure the emotional response in the ways that it manifests. I think there are even objective ways to decide how *good* a piece of art is.