

Do you agree that your current mental process is a result of the following exhaustive list of factors: (A) DNA (B) chemical reactions in your body (C) environment affecting those reactions (D) sensations and experiences (E) quantum fluctuations? For the sake of simplicity I would like subsume A and B into the category “starting point” and the same for C, D, and E into “randomness”.

A human’s thought process begins when their brain grows into existence based on their “starting point” (the initial configuration due to DNA, with chemistry and physics determining how the system behaves after that). As a human lives, their thought process is altered as the brain adapts (in some way we don’t necessarily know) to external stimuli. Not experiencing anything and sleeping are valid stimuli. This is essentially the process of incorporating “randomness” (from the environment, perception, and quantum stuff) and altering some aspect of the brain’s architecture (probably neural connections). We don’t know *exactly* how the above works (although we have a pretty good idea at the single-neuron level) but we don’t need to. Simply knowing that DNA plus external stimuli effectively result in the final thought process is enough.

Somehow, creativity results from this process. Heavily simplified, creativity is coming up with an action that most people would not normally consider in a given situation. Since a person’s thought process is completely determined by the factors we listed (we don’t know how, but we know what the inputs are) then creativity must arise as a result of these inputs. Therefore, creativity is a result of DNA (pre-programmed behavior), the laws of physics (how the system actually works), and a bunch of external stimuli (essentially randomness). Creativity is completely explained by these factors, because the factors completely explain a thought process, and because creativity is contained by the thought process.

A single individual has unique of (possessed by literally no other person in human history): DNA, life experiences, and quantum fluctuations. I would thus posit that creativity is merely a result of the thought process created by these unique factors encountering and considering a certain external stimulus. This unique thought process, created by unique factors, then generates an idea. Sometimes this idea is the same as those generated by other thought processes, since humans share huge parts of their DNA and major aspects of external stimuli. Sometimes, however, this unique thought process ends up at a thought no one else would have had.

The same applies to goals. Since goals are contained by the thought process, the factors that completely explain the thought process must completely explain “goals” or “agency”. Having goals is therefore a result of DNA and randomness. We have many pre-programmed goals such as “eat” and “mate” that form our “starting point” but over the course of our life the external stimuli we encounter and internalize (essentially “randomness”) result in additional goals unique to the individual. As the unique combination of DNA and experiences can create an original thought viewed as “creativity” so can the unique combination create unique goals.

Goals or “agency” are therefore the result of some “starting point” which is then tempered by selectively internalizing outside “randomness.” And so I ask you. Why is an algorithm’s “starting point” of “optimize this function” and “randomness” of “generate random numbers that change the function I’m optimizing” any different than what a human brain does? The answer to me is quite obviously that they aren’t. Just because the AI cannot say to you “I want to paint a picture

of a dog” but rather would say “I am minimizing this ridiculously long function” is just a question of semantics. In fact, I would easily argue the AI’s more general optimization function can more completely describe goals than our limited use of language can.

Both the human thought process and the AI have a starting point and then update their internal model based on external randomness filtered through that model. The only difference is that the human thought process can communicate itself to us in words we understand because it shares our starting point. It is theoretically possible for an AI to communicate such goals in the same way as us. We just haven’t found the combination of “starting point” and “randomness” that leads to that yet. Finding an AI able to communicate in a human way is the traditional goal of most AI research, to effectively create an artificial human. For this endeavor, understanding precisely how neurology maps to psychology is incredibly important. This describes an unimaginably small subset of all possible AI.

But viewing humans as “special” just because we don’t fully understand all of the inner workings of the brain is ridiculous. Given that physics works the same in the brain as it does in a computer, it is the height of hubris to expect that AI can’t replicate our thought process given similar “starting point” and “randomness”. It is an even greater height of hubris to believe that an AI is only amazing, interesting, or useful if it exhibits the decidedly human characteristics we have attempted to define as “sapience” or “agency.”

But my overall point, based on our previous debates, is that there is nothing fundamentally different between humans and AI. Both are literally given pre-programmed goals (a function for AI; survival instincts for humans) which are then updated based on external stimulus (random numbers, video feed, audio feed, etc. for an AI; perceived experiences for a human) and eventually result in “goals” and “agency”. The only difference is that humans are designed from the start to become humans and achieve “human goals” and “human agency.” But the fact remains that these attributes were achieved by starting from a defined program and then incorporating random noise into that program.