Bart Verheij

Who shoves whom around inside a thinking self? Or: What is the meaning of symbols?

Analog and Digital encounter each other in the world of ideas (Figure 1).[1]

D: Fancy meeting you here! I'd already hoped that our dialogues would have a counterpart here.

A: You never can tell in the world of ideas. Just when you think you're done, new symbols combine and drag you out again. Events, my dear, events ...

D: I don't see why we have to perform at the whims of events.

A: Just try resisting. Then you'll see why. You don't have any choice in the matter!

D: I don't?

A: Look—to resist is tantamount to failure. The world of events is our dictator.

D: Mmm. I clearly have a will of my own. I think, therefore I am.

A: Ah! Try to think your way out of a raging storm.

D: Ok, my thinking is constrained by events—but I am in charge. I can choose to not go into that storm, or protect myself against it.

---

1 The characters are purely fictional. The setup is loosely inspired by Douglas R. Hofstadter's text `Who shoves whom around inside the careenium? Or: What is the meaning of the word `I'?' (*Synthese* 53 (1982), 189–218).

Figure 1: Image generated by Midjourney (February 18, 2024).

A: Sure. But you adapt to the storm, not the other way around. And things are worse for you: you *think* you adapt to the storm.

D: (Confused) Yes, that's what I was saying!?

A: That is an appearance, not the real thing.

D: (More confused) Are you saying that I do not think?

A: Well—that is not the relevant point. What I try to say is: your intention to wear a rain coat in the storm exists in the eye of the beholder alone—you. Your intention is a figment of your imagination.

D: I really only wear a rain coat when I want to.

A: That is what you report after the event. You may be fooling

yourself, and you would not be able to tell.

D: O, thát point. The question whether our will (or our experience of that) has causal powers. You sound like a psychologist. I know I can be fooled in an experimentally controlled setting, but that is not the real world.

A: Are you now mocking me, or an entire profession? You should know better. I refer to a more holistic perspective in which the distinction between fields of study is blurred and irrelevant.

D: Of course! You want to discuss artificial intelligence (AI)! I like that.

A: I like AI, but I am more ambitious than most in AI. I think of physical systems with massive circular feedback loops. That is where the magic happens.

D: As I said: AI. Physical systems in which symbol manipulation generates intelligent action.

A: No, no, no. You are not listening. I did not speak of physical symbol systems.[2] You and your thinking ... The relevant metaphor is that of a person at the helm of a ship, who is continuously responding to the changing waters and winds, the κυβερν'ητης, in Greek.

D: How could I forget! Cybernetics.[3]

A: Yes. An older discipline than AI. Already in 1948, Norbert Wiener and his colleagues spoke of control and communication theory, whether in machines or in animals (including human beings), as a separate field of study.

D: Older, older ... Only by name. The term artificial intelligence

2 Wikipedia: Physical symbol system.
3 Wikipedia: Cybernetics.

goes back to 1956, indeed a few years later. But both AI and cybernetics took inspiration from Turing's ideas. He proposed what we now call the Turing test in 1950 (before the age of 40) and the Turing machine in 1937 (aged 25).[4]

A: But there is a key distinction. Turing and today's computers work in the realm of bits, zeros, ones—and hence have a hard time going beyond symbols. That is exactly why it took so long to do more than playing chess in AI.

D: A sore spot indeed. I am lucky that you do not mention the endless toy problems of the 1980s and 1990s. Tweety the penguin. Yeah, sure. She does not fly.

A: (Smiling.) I did not say it.

D: I was there when the tables turned in AI. Elephants don't play chess. Intelligence without representation. Intelligence without reason. More or less everything I had been taught went down the drain.

A: You say it. For me, that was all old news, and just a return to core ideas in cybernetics. Not only inspired by the theory of computation and its emphasis on symbol manipulation, but also taking lessons from biology, physics, engineering.

D: You want to build things.

A: Yes, but the right way. So not: a robot that thinks about the force of its handshake. Obviously such a machine would crush many human hands. No: there should be a feedback system in place. The robot has a hand with sensors that feel the soft counterpressure of a human hand. Safety by physical design.

D: I still do not understand why you oppose so much to thinking.

---

4 Wikipedia: Timeline of artificial intelligence.

A: I like to think, but not hidden away like a monk, separated from the world. Systems should be deployed and continuously adapted.[5]

D: Ok, but today's computers are symbol processors in the end, right? Also your favourite big-data-crunching-persistent-adaptive-complex-systems.

A: Again you are misinterpreting how I think. I like big data, I like the crunching, I like the persistence, I like the adaptivity, I like the complexity.[6] But I do not like that they are implemented on today's computers.

D: Excuse me? No implementation, no deployment, no real world system.

A: I *do* have my armchair dreams. I believe in an entirely different kind of computing. And indeed something clearly needs to be done. Contemporary AI powerhouses eat away the sustainable energy production of half a country. Compare that to our brains that can survive on a bit of water and a few slices of bread-and-peanut-butter per day.

D: So what else then?

A: Definitely non-binary. I think we need a different computing paradigm emphasising that computing is massively parallel and analog. And the feedback loops of physical systems in their environment will play a significant role. Do you know Braitenberg vehicles?[7]

D: Definitely! Still a beautiful example. The fact that approaching a light source or moving away from it can be achieved by just rewiring the connections between sensors and

5 Cf. the Monk system.

6 Cf. the theme Data Science and Systems Complexity (DSSC) the Faculty of Science and Engineering of the University of Groningen.

7 Wikipedia: Braitenberg vehicles.

wheels, remains inspiring. The appearance of an intention is strong. But would you count that as computing?

A: Well, at an elementary level: yes. For complex analog computing, I like the 1781 planetarium built by Eise Eisinga. Its precision of determining the location of the planets is still amazing.[8]

D: Gotcha! No feedback loop there. That machine just turns and turns and turns. No interaction at all.

A: You are right, but it is an analog machine that does not use much energy. And in a sense it performs complex computations, although not at all in the digital paradigm of computing. Still the precision achieved is high. Only roughly every 10 years some tuning needs to be done in order to adapt for the effects of external factors such as weather conditions.

D: Indeed amazing. And the planetarium can still be visited in Franeker. Unlike its university that today only exists in the world of ideas, just like us ...

A: Yeah, and thinking that it is older than the one in Groningen ...

D: Is there an example of analog computing that includes the physical feedback mechanisms you like so much?

A: The best example is still the embodied brain. Building an embodied brain remains a research challenge. For that we need new cognitive computing paradigms and new kinds of neuromorphic hardware.[9]

D: Fascinating. But I hear that neuromorphic hardware is

8 NRC Handelsblad: Het Eisinga Planetarium (September 20, 2023).

9 Cf. the Groningen Cognitive Systems and Materials Center (CogniGron) of the University of Groningen.

for now focusing mostly on reproducing the behaviour of single neurons. That does not yet sound like higher forms of knowledge and reasoning ...

A: Knowledge and reasoning? These are so overrated. Back in the days we tried to use handcrafted knowledge for serious tasks in vision. To no avail. And when neural networks started to recognize actual human handwriting, encoding our so-called expertise was typically harmful for system performance!

D: Ok, but would you not say that this just means the knowledge wasn't good enough? All knowledge is hypothetical anyway, and all reasoning presumptive.

A: Ah! So you admit that uncertain reasoning in a noisy world is what we should focus on.

D: Sure. Except in the realm of mathematics. There reasoning should be conclusive and remain within a closed world defined by axioms. It remains surprising how large the world of ideas is. Even we live in it.

A: Yes, but we are inanimate symbols. Mathematics is helpful for studying intelligent systems though. Though with much less space for logic than you think. At least probability theory should be included, and even more essentially: without a role for dynamical systems the study of intelligent systems will remain limited to the world of symbols alone.

D: I have never really understood this opposition to symbols. Isn't it a false dichotomy? Do you oppose symbols?

A: No of course not. They are at least a tool—but a limited tool. Without taking the feedback loops in dynamical systems seriously, and perhaps even as initial primitives, there is no hope for a proper theory of grounded intelligent systems.

D: Ok, ok. But how do you build your favourite systems? You write computer programs, which are clearly symbolic. And when you collaborate with others you use language, of which little remains without symbols.

A: Again I disagree! You may think that language is primarily symbolic, but there is so much non-symbolic communication. What is said, is clearly not only in the words themselves.

D: No?

A: Think of the recent hype with large language models. Surely an engineering feat, but the produced texts are so bland ...

D: I recognize that. But still people are scared that there will be no jobs left.

A: Of course no one can tell what will come. For now, we are safe. The intelligence we think we see in artificial systems is still by proxy. The magic still comes from the programmers, the labellers, the human authors of all data that is crunched by the systems.

D: Now I am confused. I am sometimes flabbergasted by what these systems do. Language models seem to be doing all kinds of things, much more than anyone taught them.

A: Really? Did you not notice that even the best AIs still are one-trick ponies responding to our whims? Even generative AI systems require the creativity of a human prompt engineer to do something nice. And also then: the most impressive examples are cherry-picked from among lots of underwhelming output.

D: You sound rather critical and pessimistic.

A: Critical yes, pessimistic no. I have always believed that in a sense the study of intelligent systems is an all-encompassing

science. It requires the combined insights from virtually all fields of expertise.

D: (With an ironic smile) Except for theology for sure.

A: Definitely also including theology.[10] But a study of intelligent systems that is not embedded in the natural sciences makes no sense to me. I have always been following Science and Nature, the ultimate scientific journals. It is great to see how today more and more AI-relevant work finds a place there.

D: Yes, I like that too, although I do not yet have a clear strategy for publishing in Nature.

A: I have.

D: I am not surprised.

A: There is still so much work to do. Understanding the nature of intelligent autonomous systems requires a persistent adaptation to new insights. There is inspiration everywhere.

D: I hope we will talk more.

A: We will.

10 Cf. the ERC project The Hands that Wrote the Bible.