The Language-Turn Metaphor and AGI

The term linear in the traditions of mathematical programming and hence machine learning doesn’t have the same meaning as linear in the tradition of physics. So when ML folks speak of non-linearity, it is not the same as non-linear that physicists speak of. I read this latest article and it was obviously apparent. “They’re “linear” because the only allowable power is exactly 1 and graphs of solutions to the equations form planes.”

Dynamical equations in physics usually have the power of 2. A non-linear equation in physics is one that typically does not have a closed-form analytic solution. The dynamics of a non-linear system is that it feedbacks into itself.

So when ML folks speak of non-linearity, they mean something entirely different from what physicists mean. It does not mean that there is feedback in the system. Unfortunately, somewhere along the way people have conflated the two given the common use of the word non-linear. To be clear, Machine Learning and thus Deep Learning draw their vocabulary from two fields. On is mathematical programming where you find optimization methods. The second, which is usually pigeoned holed, is probabilistic methods.

Mathematical programming ultimately is about constraint satisfaction. Probabilistic methods are methods to reason about aggregate phenomena. They aren’t the same thing, but the discipline’s tradition is that the metaphors used come from these two fields.

But when we speak of biological brains, there’s an entirely different vocabulary. Also when we speak of the physics of non-linear systems and complexity, it’s also a different vocabulary. So there’s a lot of impedance mismatches in the conversation.

Adjacent fields like computer science, genetics, neuroscience, different kinds of psychologies (cognitive, ecological, enactive) all have different vocabularies. There are a lot of ideas that are out there that are overlooked because they are actually misunderstood by others.

I come from a tradition of physics and computer science. Mathematical programming isn’t foreign, but probabilistic models of the kind that Bayesians speak about are foreign. Mathematically, there’s a resemblance with Statistical Mechanics, but metaphorically it is different.

But when one encounters the horrific complexity that is intrinsic in biology, then even the powerful metaphors that exist in physics and computer science have their limitations. But a big mistake is to favor one kind of metaphor over another. Norbert Weiner favored a physics metaphor. His cybernetic systems were analogs of dynamical systems. Then came AI, which took the opposite metaphor, which was inspired by digital computers and mathematical logic.

But biology is neither a dynamic system nor a digital system. It is in fact both! It is dynamic because it has to meet with the physical world. It is digital because this is what ensures its repeatability.

What I would like to point out that we have to be very careful about the metaphors that we employ to describe cognition. These metaphors are indoctrinated in us through years of practice in a specific tradition of study.

Wittgenstein said, “Philosophy is a battle against the bewitchment of our intelligence by means of our language.” We have the same problem when we attempt to express our understanding of biology and ultimately brains. We don’t understand general intelligence because of the limits of our language. “The limits of my language mean the limits of my world.” according to Wittgenstein.

Wittgenstein introduced the ‘language-turn’ to philosophy. The language-turn was introduced into biology via biosemiotics. Recently, the language-turn has had an immense contribution to the field of Deep Learning in the impressive capabilities of language models like GPT-3. Note that philosophers use the term ‘linquistic-turn’. This is unfortunate since linguists ever since Chomsky have been using a terrible metaphor for language!

Metaphors that are intuitive to grasp by humans are the kinds that we can assimilate via our daily experiences. The fortunate thing about language is that we are immersed in it and thus we have an intuitive understanding of its nuances.

Metaphors are like models. All metaphors are wrong, but some are more useful than others. I propose then that the metaphor of language be the primary metaphor used to understand general intelligence. In a previous life, I was involved in interoperable protocol design specifically in B2B protocols. I explored the question “what makes for frictionless protocols?”

This question is ultimately a question about language. How are humans able to coordinate by communicating in ambiguous ways? Chomsky’s error was to straight jacket language into rigid grammar. To further boost his argument, he hypothesized that humans have a genetic mutation that he called merged that gifted us with compositional language capabilities. That, we have special circuitry that allows us to do language.

Of course, Chomsky is wrong. What gives human language is power is in its ambiguity. Just as visual ambiguity can lead to different interpretations, language ambiguity also leads to different interpretations. The source of human creativity is because of this inconsistency. The source of any innovation is what is lost in translation. This you might call spontaneous symmetry breaking. That is when a crystalized idea loses its rigidity and becomes more fluid to alternative interpretations.

Darwinian natural selection would not make progress in the absence of ‘lost in translation’ mechanisms.