

Open question: what do we need to build a mind?

Two general approaches to simulating cognition:

* On the one hand, you can take the mind as whole: identify major structures such as learning, memory, perception, etc., in order to simulate functionality. Thus we formally, directly or indirectly, construct a model of higher cognitive processes and write a program which simulates it. This is a centralized, top-down “hierarchical” model, relying on the formal structure, or rather—the alliance and convergence of the form and structure in mental phenomena. In order to elicit the elusive element of awareness, we model the abstract form as well as the formal structure of the mind.

* Alternatively, people have looked at the self-similarity of structure at many different scales in the brain (and really all of nature) and noted that the human brain is made up of tiny, near-identical parts whose interactions and arrangement derive from/result in mental events and which bear within their structure a startling self-symmetry, a similarity to the organ(ism) as a whole. In other words, this view interprets the higher-level processes as the consequence of the collective interaction and organization of smaller self-similar processes and attempts to elicit the desired higher-level behavior to “emerge” by modelling the lower-level processes. We’ll call such a set of lower-level agents a “swarm.” In this view, consciousness is THE example of a spontaneous emergence of complex and orderly behavior from nothing, from random noise, awakening slowly as a direct consequence of neurochemical inter-actions and the self-organizing cellular arrangement of the brain. So, there is in this view a need to model the arrangements and patterns of interaction of the lower substructures of the brain: this is the bottom-up “swarm intelligence” model and it relies implicitly on a gradual emergence of self-organizing complexity to produce intelligence behavior— behavior which precisely because it is emergent cannot be directly programmed in. The resultant breed of complexity is held to be at least analogous to consciousness.

Both seem to use mathematical operations to approximate an abstraction of consciousness. This double reversal must be re-emphasized: the epic scope of both of the above models leads them to attempt to construct a model of a mind after an abstraction of mental behavior, and then to approximate the abstraction mathematically.

I think some extension of Godel’s theorem could and should be brought in at this point to demonstrate that both directions in AI are though promising quite fatally flawed: we sacrifice the rawness of the mental event in its abstraction, and we doubly sacrifice the integrity of the conceptualization when we reduce it to mathematics: you can’t have your cake and eat it too.

The structure of consciousness is neither mathematical nor abstract. After Lacan, it is structured like a language.

We only learn language by communicating with others. The swarm intelligence model must be understood as operating on multiple levels at once to get around Godel’s theorem, perhaps precisely in the spirit with which that theorem was produced: as representing the possibility of intuition—just like language always escapes the bounds of what has already been said, the

possibility of originality. How can we understand—and model—swarm intelligence operating at multiple levels? Self-similarity is not enough. We can't have just 'one' agent composed of many smaller agents: we still won't reproduce linguistic competence. At minimum, we always need 'two,' so that there is an 'other' to begin communicating with. The key to emergence is such a 'positive feedback loop,' where the behavior of agent A is dependent in part on the behavior of agent B, and vice versa—after several rounds of interaction, their behavior becomes 'synchronized' and on a global level a new more complex pattern of behavior emerges. At all levels of the structure, we must have a swarm, even at the level of the individual agent. But how can we converse with a swarm? Well, OK: I guess we need enough layers until what they're communicating with at the top level is English. What do the layers represent? Not linguistic contexts. After all, that's a pretty high-level conceptual structure. Context is constructed through the interaction itself. The A.I. should, in other words, have a single over-riding goal—trying to figure out: "What do you want?" At each level of the structure, we'll have swarms of agents operating over the objects appropriate to that level, whether they're words or mathemes or images or memories or ideas or any combination thereof, and each agent is performing self-generated 'operations' over the objects in the domain. How are certain operations privileged over one another? Let's say the operation represents an original interpretation: subject to flaws, perhaps, possibly just a vague image that only later, if ever, is refined. The data set which the agent has available to it is modified in some way based on each agent attempting to perform the operation which is an answer to the question "what do you want." We'll use evolutionary algorithms to train the agents to produce operations which are closer to "what we want"—which will, being linguistic, have to be interpreted. The intuition this model is based on is that consciousness is not simply memory, perception, etc., or simply the emergent result of a single swarm. The insight, if indeed it is one at all, is this: consciousness is fundamentally non-mathematical—which is NOT to say that mathematics cannot be learned, used, enjoyed and understood by conscious beings, simply that the nature of consciousness is not mathematics, its ultimate structure is that of a gap, precisely between mathematics, logic, reason and the real. The joke seems even crueler when we realize our methods for modelling consciousness are precisely digital, discrete, mathematical in foundation. The structure of consciousness is, once again, that of language, insofar as it is in fact pulsates within an endless void, oscillates around an infinite abyss or rupture, the null point of non-convergence which bridges "reason"/time, cultural logic with the real (never perfectly distinguishable from the imaginary.) The construction of an English-speaking robot for whom a fairly legitimate argument for having a mind could be made is possible. Linguistic competence is a recursive swarm function.

The fundamental problem is how to "bound" the seemingly infinite space conjured into reality by language, not merely "interior" reality (so to speak,) but how to model time itself? Is simulation the same as duplication? What would it mean if—say—a linguistically competent agent manages to pass a Turing test even though it lacks any "true" understanding (so to speak) of what the past and future mean? Just what is this thing called time?

Despite this rather messy philosophical hurdle regard time—and there are not a few of these mines floating around these waters—Even though the base substructure of the brain is not linguistic, we use language. Human-created machines can be made capable of linguistic competence. Point blank. The only question is how to model it effectively. And surely early models will be crude. The point is we need a system which is flexible. Can be improved,

enhanced over time and is able to learn. The problem with language is that it is tied up with just about everything else, indeed, by shaping our understanding of reality, it shapes our reality itself. In a powerful sense—and this is why the cogito cannot merely be thought but must be enunciated—we are conscious because we say we are. The same formula applies to a disturbing number of other abstractions.

(more later)

(c) Fractal Ontology, 2007