Would a Digital Brain Have a Mind?

Neville Holmes, University of Tasmania

Certain recent events caused me to doubt whether I know my own mind or not. Let me explain.

Last week, the first of our academic year, all first-year students in our degree program underwent a supervised test in which they pull an old computer to pieces and put it back together again. We give this test to put a healthy disrespect for digital circuitry—which is, at heart, only carefully polluted sand—into each student's mind as early as possible. We intend this disrespect to counter the superstition, held both by naïve students and by members of the public susceptible to media persuasion, that digital machinery has much in common with the human brain.

Yesterday, I went to a lunchtime philosophy club lecture titled "Why the Body Is the Mind." Because some of the discussion related to consciousness, I recalled Giorgio Buttazzo's article, "Artificial Consciousness: Utopia or Real Possibility?" (*Computer*, July 2001, pp. 24-30). The juxtaposition suggested a strange contrast between computing people, who see mental capabilities in machines because they do not appreciate how complex the human brain is, and philosophers, who see complexities in the human mind because they do not appreciate that the brain and the computer share some simple and fundamental properties.

However, not being a philosopher, I find it difficult to be confident that I understand them when they discuss the mind. This uncertainty leaves at least three possibilities:

Although the human brain and the digital computer share many traits, simulating the mind in silicon may be impossible.

- either I understand the basic nature of my mind and the philosophers don't,
- vice versa, or
- we share the same understanding but express it in mutually incomprehensible language.

Later, I read Bob Colwell's provocative essay "Engineering, Science, and Quantum Mechanics" (*Computer*, Feb. 2002, pp. 8-10). Toward his essay's end, Colwell reported of entanglement theory that "the [photon's] wave function's actual point of collapse is when a conscious mind perceives the results" and that the collapse was caused by "the synapses of our brains, acting in concert to form our minds, at the instant we detected the photon."

Suddenly, I felt alone, isolated, out of my depth, and fearfully vulnerable. What follows is meant to enlist your sympathy and rebuild my confidence.

THE MIND AS PROCESS

According to my *Macquarie Dictionary* (Macquarie Library Pty. Ltd., Sydney, 1981; http://www.macquariedictionary.com.au), the principal meaning of *mind* is "that which thinks, feels, and wills, exercises perception, judgment, reflection, etc., as in a human or other conscious being: *the processes of the mind*." The principal definition of *process* is "a systematic series of actions directed to some end."

An action requires an actor, presumably the mind in this case. Why not the brain? The Macquarie defines the *brain* as "the … nerve substance that fills the cranium of man and other vertebrates; centre of sensation, body coordination, thought, emotion, etc." Why so coy? Where is the mind in this brain?

The computer works like the mind-as-actor in that it functions as a device that processes data—conventional representations of facts or ideas. The circuits carry out the computer's processing by copying, transmitting, and transforming these data.

The mind, given the Macquarie definition, processes thoughts, feelings, intentions, perceptions, judgments, reflections, and so on. Neurons and glial cells process neural and hormonal representations of sensations past and present.

Although the idea of the mind as distinct from the brain has a natural appeal, defining the mind as an actor distinct from the brain invokes an unnecessary, even deceptive, dualism. Our mind thinks, feels, perceives, and so on, whereas our brain merely exists between our ears. The distinction can be useful and productive.

The problem lies in defining the mind as an *actor* rather than an *action*. If we regard the mind as the thinking *process*, it becomes distinct from the brain and becomes the systematic

series of actions the brain takes as it processes sensations and reconstructs memories.

We can thus view the brain as substance, the mind as process. Likewise, we can view the computer as substance and its computations as process. The brain exists materially, while the mind arises as a property of changes within the brain. Similarly, the computer exists materially, while computation occurs as a property of changes within the computer. So far, so simple.

SOURCE OF CONFUSION

Consciousness seems to be the confusing factor. We associate self-awareness and identity with consciousness. Buttazzo writes, "Because we cannot enter another being's mind, we cannot be sure about its consciousness." This theme recurs in writings on the philosophy of the mind. But if a mind *is* a process, it's meaningless to talk about anything entering it.

Processes can only be perceived, thus inferring the operation of another mind from such perception must surely be sufficient. If anyone argues that we need certainty, we can counter that no one can be certain of anything, as the "brain in a vat" argument shows (http://www.artsci.wustl.edu/~philos/MindDict/cartesianskepticism.html).

Cycle time versus data rate

In speculating about the effect of cycle time on artificial consciousness, Buttazzo poses a curious question, "If consciousness emerges in an artificial machine, what will time perception be like to a simulated brain that thinks millions of times faster than a human brain?" What does "be like" mean here? In any case, the question confuses cycle time with data rate—the stupendous parallelism of the brain makes the cycle time of our present digital computers irrelevant.

However, Buttazzo speculates that the world might seem to slow down for a simulated brain as perhaps it does for a fly, "thus giving the fly plenty of time to glide out of the way" of a swatting hand.

Let's apply some numbers to this speculation. The fly and I have much the same kind of neural signaling system. The average local fly measures about 10 mm, and I am roughly 200 times that length. My reaction time is about one-tenth of a second. A submillisecond reaction time for a fly is thus not at all mysterious, nor would much shorter reaction times in a digital computer be in any way puzzling. Where then is consciousness in all this?

> **The brain exists materially, while the mind arises as a property of changes within the brain.**

Whence consciousness?

Is a fly conscious? Well, it's aware to the extent it can often dodge a swat—its perceptual neural system alerts it to the swatting hand so that it can dart out of harm's way. But what part of the fly's nervous processing is aware, and thus to some degree conscious?

Awareness must emerge at least from the transformation of perception into the intent or neglect of an action. The transformation of sensation into perception can be unconscious because it can be automatic: We remain, for example, cheerfully oblivious to the dramatic data compression our retinas carry out. The transformation of intention into motion can similarly be unconscious: We do not consciously stimulate each individual muscle in our mouth, throat, and chest as we speak.

For the fly, we might imagine that its nervous system functions like a computer system: Its central processor "consciously" forms intentions on the basis of perceptions that its peripheral sensory system "unconsciously" produces, then its peripheral motor system "unconsciously" puts those intentions into effect.

Cache as cache can

We can transfer this analogy to the human nervous system to explain how consciousness arises from it, except that we have a more complex central processor than does a fly and much more occupies our minds than mere perceptions and intentions. We have extensive memory traces from which neural processes can reconstruct pseudoperceptions that pass through our consciousness. Intentions cannot be based practically on reconstructing all our possible memories at once. Part of our brain processes perceptions together with relevant pseudoperceptions to derive intentions. The processing of this area within the forebrain must be closely allied to our consciousness.

Human consciousness therefore strongly resembles the processing that a digital computer's central processor and its associated main store cache carry out. The cache brings relevant data close to where the current computation can use them. In this sense, then, our present-day computers *are* conscious, *if* their CPU has a cache.

VIVE LA DIFFÉRENCE

Observing that "The human brain has about 10^{12} neurons, and … 10^{15} synapses," Buttazzo calculates that, using artificial neural networks, "Simulating the human brain requires 5 million Gigabytes" of data storage. Moore's law suggests that digital computers will have main stores of that capacity by 2029, although Buttazzo carefully qualifies this observation by adding that it "refers only to a necessary but not sufficient condition for the development of an artificial consciousness."

Using a digital computer to simulate an artificial neural network that simulates the human brain does not seem the best approach. Given the neural parallelism to be modeled, using analog circuits to directly implement neural networks would seem a better alternative, one that might bring the feasibility date well forward, if research could divert the circuit manufacturing industry to this cause. But this begs the question of whether artificial neural networks can be made comparable with real ones.

In an artificial neural network, each node, or neuron, has an activation value that the network passes forward to other nodes through connections, or synapses. The synapse to each forward node has an associated weight that modulates the incoming activation value's effect on the forward node's activation. The weights can be adjusted in various ways likened to "learning." An artificial neural network's nodes mimic the classical neuron—but very roughly—with an axon down which the repetition rate of an action potential, the *spike,* passes an analog value, dependent on the activation of the neuron's main body.

Disregarding as a mere production problem attaining 10^{12} neurons and 10^{15} synapses in analog circuitry, where do artificial neural networks fall short of the real thing?

- The activation of a classical neuron is not an arithmetic sum of the synaptic effects. Rather, a complex process involving the intervals between action potentials at individual synapses and their relative timings between synapses determines the activation. The more neurons are studied the more such complexity is revealed.
- The human nervous system contains many different kinds of neurons and many kinds of glial cells. The glial cells provide more than support because they signal and have synapses just as neurons do.
- Action potentials alone do not control nervous signaling. Graded potentials and hormonal signaling also play a part, as does the great variety of different neurotransmitters and hormones.

Creating an artificial consciousness does not require simulating these complexities. But human consciousness lies far beyond any presently contemplated artificial one. We have, for example, developed a highly complex and utterly human consciousness of our physical bodies. Likewise, we have a highly developed consciousness of other people and of our society, whose collective consciousness shapes our development as humans.

Obvious parallels exist between the brain and digital computers. To fulfill their responsibility to themselves, and to others who might be misled by journalistic hyperbole, computing professionals should have well-founded

> **An artificial neural network's nodes mimic the classical neuron— but very roughly.**

opinions about the extent of these parallels. The profession should refrain from applying humanistic names to its mechanistic endeavors, and it must be conscious always of the essential differences between people and computers.

Perhaps in 50 or 100 years, our machines will acquire a humanlike consciousness and intelligence. But such machines will be utterly different from the puerile imitations we now have or can realistically design. Getting to such machines will raise professional and philosophical issues quite different from those that reflect on the nature of what human and machine consciousness can generate.

And what about the mind and entanglement theory? If a CPU and its cache possess consciousness, we could leave looking at photons to them. In any case, I've now put entanglement theory into the back of my mind under the shade of the tree in Bishop Berkeley's quad (http://homepages.which.net/ ~radicalfaith/search/a3.htm). ■

Neville Holmes is an honorary research associate at the University of Tasmania's School of Computing. Contact him at neville.holmes@utas. edu.au; http://www.comp.utas.edu.au/ users/nholmes/prfsn.