The potential effectiveness of simulations versus phenomenological models

Michael D. Mauk

Computational models can potentially be a very effective way to understand how the brain processes information. However, their power depends on the location of artificial features and errors of omission within the simulation, and on whether it is possible for the model to fail.

In neuroscience, like all branches of science, the specific questions posed are no more important than the language or mode of expression used to phrase answers. The fabulous success of physics in the 20th century seems due in part to the power of mathematics as a mode of expression to understand the physical world. This inspired the notion of the 'unusual effectiveness' of mathematics. As neuroscientists, we should consider not only the important questions, but also the modes of expression that make answers unusually effective.

Simulations and neural function
I believe that well-constructed computer simulations are a mode of mathematical expression with similar potential for efforts to understand the operation of the brain. This belief is based on the premise that information processing is the fundamental function of neurons and brain systems. Each neuron receives inputs from other cells, processes those inputs, and generates outputs. Similarly, brain systems receive inputs from other systems and generate outputs according to their internal rules of information processing. Because this is how the brain accomplishes its tasks, brain function is ultimately best understood in terms of input/output transformations and how they are produced.

More so than any other approach, a well-constructed simulation makes contact with this input/output behavior, at many levels. A simulation composed of representations of conductances can evaluate how inputs to a patch of membrane influence membrane potential. A simulation built with representations of

Michael Mauk is at the Department of Neurobiology and Anatomy, University of Texas Medical School, Houston, Texas 77030, USA. e-mail: mmauk@nba19.med.uth.tmc.edu

neural compartments tests our understanding of how inputs to a neuron produce its outputs. Network simulations composed of representations of neurons ask how inputs to brain systems are processed to produce outputs to other systems. Thus by attempting to reconstitute input/output transformations from representations of known components, simulations help to evaluate how accurate and complete the current list of components is, and to identify key interactions between components. It is this inherent connection between simulations and neural information processing that gives simulations such potential utility. How can we best achieve the full measure of this potential?

The elegance of the Hodgkin and Huxley[1] (H&H) equations, which were the culmination of a quantitative characterization of voltage-gated conductances in squid giant axon, provides clues. The approach was simple: build mathematical representations of these conductances to produce differential equations predicting how membrane potential changes over time. These equations evaluated whether the properties of sodium and potassium conductances were sufficient to reproduce an action potential. That the answer was yes represents a great achievement in neurobiology. Although there remained much to learn, this understanding of the input/output properties of a patch of membrane was a milestone.

Given this notable success, it is interesting that the use of simulations is now generally held in low regard by many experimentalists. The ill feelings can be expressed reasonably well by the generic complaint, "Don't you just adjust parameters to get the result you want?". The measure of truth in this statement is the starting point for my second argu-

ment. All simulations are approximations whose realism is diminished to some degree by errors: errors of omission (lack of information), errors of commission (adding features to obviate shortcomings), or simplifications that provide computation convenience.

Confronting the impact of these errors is one key to discriminating well-constructed and useful simulations from those that frustrate. Although there is no formula that reveals the utility of each simulation, examining H&H illustrates an approach with the potential for unusual effectiveness. To my eye, there are at least two important lessons: simulations are only useful and well constructed if they can fail, and the location and type of errors in simulations is critically important.

Simulations versus P-models
The only useful hypotheses are those that can be disproved. Likewise, simulations are only useful when they can fail. This is one attribute that makes the H&H equations so elegant. There was no guarantee that the combined representations of conductances would accurately describe the voltage fluctuations of the action potential. Ensuring that a simulation produces the desired output with unwarranted features and parameter adjustments is like doing an experiment where only one outcome is possible. The results provide no information. Allowing the potential for failure is central to the fundamentally important distinction between building simulations and building phenomenological models.
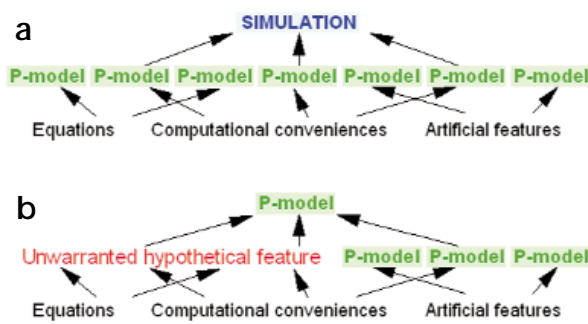
To simplify discussion, it will be useful to adopt a narrow definition for the phrase phenomenological model (P-model). For present purposes, P-model refers to any physical or mathematical representation of a neural process that

commentary

Fig. 1. The distinction between well constructed simulations and phenomenological models (P-model). (**a**) A simulation composed of a series of P-models. The ability of the P-models to reproduce the appropriate input/output function of their target component prevents the simulation's output from being influenced by the artificial ways in which the P-models are constructed. (**b**) Adding an unwarranted feature with no empirical support at the level of the P-models essentially turns an attempt at a simulation into a P-model. Thus, its success is ensured, and its potential for providing insight is greatly diminished.

produces the correct input/output behavior with no claims about biological validity. A black box that transforms patterns of inputs into the correct outputs for a particular neuron type would be a valid P-model of that neuron—independent of how artificial its internal workings might be. The equations for sodium and potassium conductances in H&H are P-models. Although they accurately describes the macroscopic behavior of these conductances, they do not capture how such conductances are generated—for example, they do not include discrete and stochastically governed openings of many channels.

I believe the distinction between P-models and simulations is the fundamental ingredient to realizing the full potential of computational neuroscience. As illustrated clearly by H&H, building a P-model is not particularly informative, whereas building simulations can be especially so. There was little to be learned from the particular mathematical implementation that H&H used to represent voltage-dependent conductances. Because they were intended only as mathematical tools to produce the correct input/output behavior, the ingredients of the P-model did not need to reflect the underlying biological processes. For example, the conductances could have been described in lookup tables. Thus, like experiments with only one possible outcome, the ability to build these P-models meant little mechanistically. The information derived from combining these P-models to build a simulation of membrane potential. Although the P-models of the conductances could not fail, the simulation of membrane potential could.

In this way, the H&H equations illustrate a key property of well-constructed simulations: they are composed of appropriate interactions among valid P-models (**Fig. 1a**). Because failure was not explicitly excluded by the construction of the simulation, successes can be useful in the obvious ways that H&H's were. That is, they indicate that the current level of characterization is sufficient to describe the behavior of the system, and they can also lead to new hypotheses.

Whereas the value of successes is relatively clear, in many ways the value of simulations is determined by how failures are used. Failures can yield information, pointing to key experiments that are needed or to shortcomings of working hypotheses. When simulations can be fixed by adding a biologically plausible feature, they can be the source of specific and testable hypotheses. In contrast, adding hypothetical or biologically unjustified features simply makes the simulation a P-model (**Fig. 1b**). When success is ensured in ways that do not guide experiment, neither successes nor failures are particularly informative.

Indeed, often just trying (and failing) to build a simulation can provide useful information. One often begins with the impression that more than enough is known to build a simulation. Then follows a disappointing but informative realization that key pieces of information are missing that provides fertile ground and focus for new experiments.

Thus, the H&H equations illustrate the distinctions between building simulations and P-models. In constructing the latter, nothing is learned that was

not already known. It is simply a device designed to produce a target input/output function. This clearly does not imply that all models are useless. There is great utility in simple models asking if a particular kind of coupling, oscillation or plasticity rule can in principle explain an observation. However, to engage the full potential of computational neuroscience requires appreciating the scope and purpose of these simple models, and especially requires discriminating simulations from P-models.

The type and location of errors

All simulations contain errors of some kind or features chosen mostly for computational convenience (like a lookup table) that create an unwanted separation between the simulation and biological reality. However, the distinction between P-models and simulations also illustrates the importance of the location of these types of errors in simulations.

I have argued that adding hypothetical features without empirical justification to overcome limitations is a weak approach. Because they ensure proper output, they simply turn a simulation into a P-model (**Fig. 1b**). The impact of errors of omission, where information is lacking or too complex to implement, is completely different. Because errors of omission do not add function, they do not preclude the possibility of obtaining useful information. Errors of omission are just part of the game. Adding features as computer power permits or as they become known engages the ongoing question, "Are enough features known about this system to explain its mode of operation?". In contrast, errors of commission largely interfere with gaining information from failures and trivialize the significance of successes.

Finally, the artificiality of simulations can be considered a sort of error. Because the brain is not built from equations, lookup tables or simplified representations like integrate-and-fire neurons, there can be a strong sense of disconnection from biology. I suggest that because brains are a biological implementation of information processing, using good P-models to build simulations can maintain a strong connection to neural function. As long as P-models produce accurate input/output transformations, the artificial nature of their construction is largely

irrelevant with respect to information processing and the output of the simulation (**Fig. 1**).

Imagine that each neuron in a brain were replaced with an artificial mechanical or silicon replica that exactly matched the input/output function of the original. This would be a different, non-biological brain, but it would accomplish the same information processing. A good simulation of a brain system is essentially trying to accomplish the same thing. It might be built using integrate-and-fire representations of neurons. This method can accurately reproduce membrane potential given the leak and synaptic conductances that apply, but it uses completely artificial means to generate action potentials. Yet, if enough is known about the relevant synaptic conductances and connectivity, integrate-and-fire representations can be reasonable P-models of target neurons. From these P-models, which are fraught with errors of omission and

hypothetical features, a simulation can address how the cells and synapses of the target brain structure process information to produce input/outputs transformations. Thus, as in real estate, location is everything. Artificial features that go into the construction of good P-models can have minimal impact on the ability of a simulation to address information processing. Artificial capabilities added to P-models, or artificial features added at the level that interacts with the P-models (**Fig. 1b**) can do great damage to the biological relevance of a simulation.

In sum, I believe that there is a special and potentially powerful relationship between computer simulations and analysis of brain function, and that we tap into this power best when certain pitfalls are avoided. These ideas have important implications for consumers of computational neuroscience. The astute reader should appreciate that although building a P-model can be an important means toward an end, it is

not in itself an end. Building a P-model of a neuron, synapse or other neural component is similar to developing a new protocol or technique, appropriate perhaps as a methods publication. When Hodgkin and Huxley built P-models of voltage-dependent sodium and potassium conductances in squid axon, it was a nice piece of technical work. When they combined these P-models into a simulation of voltage across a small patch of membrane, with many errors of omission but no errors of commission, it was an elegant piece of neuroscience.

1. Hodgkin, A. L. & Huxley, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol. (Lond.)* **117**, 500–544 (1952).