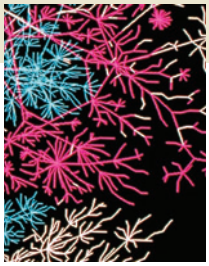

A Turing-like test for biological modeling

David Harel

In 1950, Alan Turing proposed an 'imitation game,' for determining whether a computer is intelligent[1]. A human interrogator, the candidate computer and another human are put in separate rooms, connected electronically. Alice, the interrogator, doesn't know which is the human and which the computer, and has a fixed amount of time to determine their correct identities by addressing questions to them. The computer has to make its best effort to deceive Alice, giving the impression of being human, and is said to pass the Turing test if after the allotted time Alice doesn't know which is which. Succeeding by guessing is avoided by administering the test several times. Here, I argue that a variant of this idea, but with a Popperian twist, is applicable to the computerized modeling of natural systems, particularly in systems biology.

The grand challenge

Many characteristics of man-made systems, especially those termed reactive by computer scientists, are central to the dynamics of biological systems too: heavy concurrency (simultaneity), event-driven and time-dependent behavior, cause-effect phenomena and distributed control. These occur from the atomic level, via the molecular and the intra- and intercellular levels, to full organisms and even entire populations, suggesting that biological systems can be modeled (that is, reverse-engineered, simulated and analyzed) using methods and tools developed for the engineering of complex man-made systems. Recent results from small-scale modeling efforts have been extremely encouraging[2–6].

Most modeling efforts are partial, intended to capture some limited phenomena or mechanism, or a small part of a larger system. There is often a particular goal for the modeling, with specific laboratory observations relevant to it and specific behaviors for simulation and checking. The ultimate motivation for such modeling efforts could be making carefully considered predictions that lead to particular experiments. This is analogous to building a small system from a very clear set of requirements. A different approach is to aim at modeling a complete biological system, such as an organ or a full multicellular animal (see ref. 7, in which I propose to model the *Caenorhabditis elegans* nematode).

Such a comprehensive 'grand challenge' is extremely nontrivial and by its very nature is intended to take modeling to the limit: let's model (reverse-engineer) a worm or an elephant similarly to the way we engineer a chemical plant or an F-15 fighter jet. The challenge is to construct a realistic model, true to all known facts, which is smoothly extendable as new facts are discovered. It would feature a three-dimensional, animated graphical front end and would enable interactive multilevel probe-able simulations of the animal's development and behavior. The underlying computational framework would be mathematically rigorous, but would also be intuitive enough for biologists to enter newly discovered data themselves. The model would also support alternative theses reflecting disagreement among the scientists, to observe and compare their effects at run time.

As I have argued previously[7], achieving this goal, even for a very modest organism like *C. elegans*, would require enormous amounts of interdisciplinary work, both in the computational and analysis realms and in the accumulation, assimilation and formalization of the biological data itself. It raises numerous difficulties, for which no solutions are known at present. The good news is that this is typical of many grand challenges, both past and future; like putting a man on the moon, proving Fermat's last theorem or solving cancer. One of the characteristics of a 'good' long-term challenge is that, if successful, the results would be spectacular, but even if it is not completely successful, many fundamental and useful results will have been achieved along the way. In our case, a comprehensive *in silico* model of an entire living organism would constitute an unprecedented tool, allowing researchers to see and understand life in ways not otherwise possible, triggering and helping predict behavior, filling gaps and correcting errors, suggesting hitherto unimagined experiments and much more. It is not my intention here to try to convince the reader of the virtues of such an effort, but many benefits can easily be imagined.

SPL/PHOTO RESEARCHERS, INC.

Alan Turing, who in 1950 proposed a test to determine a computer's intelligence.

Measuring success

Still, what does being successful mean? How do we know when we are done, labeling the model valid? And are we ever done? It is one thing to build a computerized model that looks good and captures some desired, but limited, features of a biological system for some specific research reason, but quite another thing to claim to have constructed a valid model of a full organism, using all that is known about it. In

David Harel is at The Weizmann Institute of Science, Computer Science and Applied Mathematics, PO Box 26, Rehovot, 76100 Israel. e-mail: dharel@weizmann.ac.il

limited modeling, one has in mind a manageable set of laboratory observations—analogous to requirements in engineering man-made systems—so that one essentially knows what to test for. The challenge in comprehensive modeling is so great and the required multileveled amounts of detail and their intercombinations so vast, that it is no longer clear how to tell when one is done or what it means for one's model to be valid.

To address this question, we must clarify what we mean by modeling a creature based on all that is known. We must decide upon the model's levels of detail, so that we don't find ourselves having to deal with quarks or quantum effects. Moreover, we cannot hope to ever find out everything there is to know about a complex organism, even after limiting the levels of discourse. A computerized simulation can at best be based on the knowledge available at a given point in time and will have to take liberties with what we don't know yet or never will. For example, the model can be made to fill in certain dynamics by carrying out the biological equivalent of the movie industry's 'morphing' technique. In any case, this begs the question of how to tell when all that we do know has indeed been captured.

Here is where validating the model can be likened to the Turing test, but with a Popperian twist: a comprehensive model of a full organism will be deemed valid/complete/adequate if it cannot be distinguished from the real thing by an appropriate team of interrogators. This idea raises many subtle questions and may attract opposition on many grounds, which this short essay will not attempt to address. The reader is referred to Turing's original paper[1], which discusses several issues that are relevant here, too.

Modifications to Turing's test

If we were to apply the idea in Turing's paper to validate biological models, what types of modifications to the original test would we have to implement? First, to prevent us from using our senses to tell human from computer, Turing employed separate rooms and electronic means for communication. In our version, we are not simulating intelligence but development and behavior. Consequently, our 'protection buffer' will have to be quite more complex—intelligent, in fact! It would have to limit the interrogation to be purely behavioral and to incorporate means for disguising the fact that the model is not an actual living entity. These would have to include neutral communication methods and similar-looking front-ends, as in Turing's original test, but also means for emulating the limitations of actual experimentation. A query requiring three weeks in a laboratory on the real thing would have to elicit a similarly realistic delay from the simulating model. Moreover, queries that cannot be addressed for real at all must be left unanswered by the model, too, even though the main reason for building models in the first place is to generate predictive and work-provoking responses even to those.

Second, our test is perpetually dynamic, in the good sprit of Popper's philosophy of science[8]. A computer passing the Turing test can be labeled intelligent once and for all because, even if we take into account the variability of intelligence among average humans, we don't expect the nature and scope of intelligence to change much over the years. In contrast, a model of a worm or a fly that passes our test can only be certified valid or complete for the present time. New research will repeatedly refute that completeness, and the model will have to be continuously strengthened to keep up with the advancement of science. The protection buffer will also have to change as advances are made in laboratory technology (but, interestingly, it will have to be made weaker, since probing the model and probing the real thing will become closer).

Third, our interrogators can't simply be any humans of average intelligence. Both they, and the buffer people responsible for 'running' the real organism and providing its responses to probes, would have to be experts on the subject matter of the model, appropriately knowledgeable about its coverage and levels of detail. In the *Caenorhabditis elegans* proposal, for example, these would have to be members of the worm research community.

Clearly, this modified test is not without its problems, and is not proposed here for immediate consideration in practice. Still, it could serve as an ultimate kind of certification for the success of what appears to be a worthy long-term research effort. Variations of the idea are also applicable to efforts aimed at modeling and simulating other kinds of natural systems.

1. Turing, A.M. *Mind* **59**, 433–460 (1950).
2. Ciobanu, G. & Rozenberg, G. (eds.). *Modeling in Molecular Biology* (Springer-Verlag, Berlin, 2004).
3. Kam, N. *et al.* In *Proceedings of the 1st International Workshop on Computational Methods in Systems Biology. Lecture Notes in Computer Science*, vol. 2602. (ed. Priami, C.) 4–20 (Springer-Verlag, Berlin, 2003).
4. Priami, C., Regev, A., Silverman, W. & Shapiro, E. *Information Process. Lett.* **80**, 25–31 (2001).
5. Efroni, S., Harel, D. & Cohen, I.R. *Genome Res.* **13**, 2485–2497 (2003).
6. Fisher, J. *et al. Proc. Natl. Acad. Sci. USA* **102**, 1951–1956 (2005).
7. Harel, D. *Bull. Eur. Assoc. of Theor. Comput.Sci.* **81**, 226–235 (2003).
8. Popper, K.R. *The Logic of Scientific Discovery* (Hutchinson, London, 1959).