UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Cell-based models

Tamulonis, C.V.T.

Publication date
2013

Citation for published version (APA):
Tamulonis, C. V. T. (2013). *Cell-based models*. [Thesis, externally prepared, Universiteit van Amsterdam].

7

Conclusions

Over the course of this thesis we have reviewed various techniques for modeling cells as individual mechanical units that can move, collide, deform and adhere to one another. We covered a range of scales and complexity, from simple point-like cells to complex polygonal representations. In some cases, it was assumed that these cells have simple programs that dictate their behavior, causing a self-organized system to emerge. We have attempted to show that cell-based models are useful in situations where traditional continuum mathematical modeling approaches do not apply due to the heterogeneity or non-smoothness of the system. We then explored two cases in depth: the gastrulation of *Nematostella vectensis* and pattern formation in cultures of gliding filamentous cyanobacteria. Despite the disparate nature of these systems, we showed how a simple mechanical approach could be used to model both cases.

Hopefully cell-based models will continue to scale to ever larger and more complex systems. In the previous chapter we demonstrated how a 3D model

could be used on realistic time and space scales to simulate 10^5 trichomes. However, each trichome consisted of a polyline with tens to hundreds of vertices. By modeling individual cells instead of whole trichomes, we could in principle simulate 10^6 cells using a simple point or ellipsoid representation on a single workstation. This would be sufficient for a full scale model of the lifecycle of *D. discoideium*, with the same number of cells as found in real slime molds. A single model could be used for all the phases, from aggregation to crawling slug to the fruiting body, with an excellent level of detail. Likewise, a full scale model of the life cycle of myxobacteria could be possible. Another interesting prospect is *C. elegans*. With only around one thousand somatic cells, and a trove of published experimental data, including a complete cell fate map, a full scale cell-based model of *C. elegans*, perhaps of the Honda-Voronoi type, could become the first whole multicellular organism computational model, complete with a true nervous system and differentiated organs.

In each of these cases, the cell mechanics in the models has to be sufficiently realistic for any model to succeed, certainly a non-trivial task. In particular, good models of cell-cell adhesion are still a challenge. One reason why the Cellular Potts Model has been so successful is that it simulates this often essential modeling component very well. But for complex polygon models, simulating differential adhesion phenomena such as cell sorting is still a challenge. This is a pity, since complex polygon models give the researcher great flexibility in modeling cells and aggregates of any shape. A good cell-cell adhesion model for complex polygons would go a long way towards providing a more general framework for all cellular systems.

Given that reasonable cellular mechanics models can be applied to the chosen model geometry, the next step would be to calibrate the parameters of the model such that it can be shown that the model behaves well in various equilibrium configurations. For example, if we initially arrange the cells such that they form a *D. discoideum* crawling slug, we must ensure that for some set of chosen parameters, and in the absence of any outside forces, the slug keeps its shape and does not distort, fall apart or that the cells do not redistribute themselves and destroy any internal structure. Another important calibration step is to ensure that the rheological properties of the cell aggregate are similar to real aggregates, so that the virtual aggregate responds to stress and strain in a measurably similar way to the real system (e.g. 60).

A full scale computational model of any of these systems would be a

boon to researchers in the field. Such a model would serve as a sort of "puppet" of the real organism, but one that is mechanically similar to the real thing. Researchers could animate this puppet with any models of cell behavior they choose. These behavioral models would "pull the strings" and allow researchers to experiment with the system with both low level mechanistic/molecular models and phenomenological models at a higher level of abstraction, where one can study the logic of cell regulation and cell-cell cooperation, without first having to unravel the extreme complexity of how that logic is implemented at the molecular level. Cells could perhaps begin to be understood in logical terms, and not just as a set of quantitative relationships between chemical components. Consider electronics: a processor could be described quantitatively as electrical signals through transistors and being transformed. However such a low level description would be a mind boggling way in which to describe how a microprocessor works. Instead, microprocessors are described logically, as boolean functions. The description at this logical level is much easier to grasp than raw signaling and allows simple functional modules to be described. These modules in turn are composed to form more sophisticated modules, and so forth. If biological systems could also be decomposed into hierarchy of functional sub-systems, each described using an agent-based model, this might help bridge the divide between genotype and phenotype.

This approach was recently employed by Karr et al. (192), who developed a complete cell model of the bacterium *Mycoplasma genitalium*, including its entire genome. Creating a model of this complexity was only possible by dividing the model into sub-models, and working on each separately and then composing the modules until they obtained a complete model. Fortunately, computer programming languages are built around the notion of functional sub-modules (such as functions, objects and libraries), and unlike most mathematics, modules can be composed as much as desired since computers have no problem digesting complexity, whereas purely mathematical models become easily unworkable as models grow.

As computational biology expands, and models such as Karr et al.'s (192) become more commonplace, it is natural that programming languages designed specifically for cell models will arise, since there is a tendency for new computer languages to emerge when existing languages are awkward to use for a given class of problem. Agarwal's (193) cell programming language is an early example of this, but did not gain much recognition. On

the other hand CompuCell, based on the Cellular Potts model, is a well developed cell-based modeling environment designed to ease model development, albeit with the inherent limitations of the CPM. CellML and FieldML are XML-based markup languages for describing ODE and PDE models of biological systems. They standardize a format for these types of models so that researchers can easily exchange models without the substantial difficulty of implementing the models themselves. While none of these is a complete biological programming language, it is tempting to think that eventually a range of languages covering different levels of abstraction might evolve, from the mechanistic to ecological and societal levels. Having formal languages that can be successively compiled (in the computer science sense) into longer and lower level descriptions is an exciting prospect for theoretical biology, since this paradigm has been very successful in computer science for tackling complexity. By progressively describing biological systems through languages of higher and higher levels of abstraction, the complexity of these systems would be compressed to an extent that it becomes humanly comprehensible. On the other hand, biological systems, although clearly information processors, are so different from current technology that studying biological systems through computers might lead to new computing paradigms that are fundamentally different from the current approaches, which stem from the basic Turing machine and other equivalent theoretical computers. Witness, for example, Adleman's (194) DNA-based setup for solving the traveling salesman problem, involving operations of a very different nature than the basic arithmetic and memory of digital computers.

How far could the scaling of cell-based models go? One important property of these models is their linear computational and space complexity. If this property continues to hold, we can assume that a one hundred billion cell model (the human brain) will require on the order of ten thousand times more computational capacity and memory than the currently feasible one million cell model (on a single workstation). Fortunately, computational capacity has grown exponentially since the development of the first microprocessor ("Moore's Law") and if the trend continues, a ten thousand fold increase in capacity would occur in a mere fourteen years. Furthermore, cell-based models are well suited for scaling with the current mode of increasing computational capacity by increasing the number of computing units operating in parallel, since the interactions between cells are local and confined to a small neighborhood around each cell and each cell can be processed

independently of global conditions. But even if such an approach is feasible, representing every single cell in a large system is probably not the best approach for all modeling problems. Traditional continuum models are often more appropriate when modeling large number of identical cells with predictable behaviors forming smooth, "continuous" tissues, like the heart and the growth of sponges and corals. Even if the dynamics at the cellular level are complex and the bulk behavior of cells is difficult to generalize to a continuum model, multi-scale models, which employ a different model for different time and space scales of the system, can be used to bridge the gap between the cell and macroscopic scales. Still, cell-based modeling is a natural choice for those systems in which cell individuality plays pivotal roles, such as cancer and embryogenesis.

Cell-based modeling certainly holds many interesting challenges for the future. In the short term, I would highly encourage researchers to attempt whole organism models, starting with the slime mold and roundworm. A whole organism model could constitute a virtual laboratory that is inexpensive, fast to execute simulations and also offers complete control over every aspect of the organism. Virtual experiments would allow rapid prototyping of new ideas and help target effective experiments. Modeling is generally a rewarding experience, and modeling of whole organisms should certainly capture the imaginations of researchers. Personally, I would like to see more computer scientists working in computational biology since computational approaches have proven to be the only effective way of tackling the complexity of biological systems, and any cross-over between computer science and biology would enrich both fields. In what has already been dubbed "the Century of Biology", we may find that computer science and biology are kindred spirits.