FORMAL DESCRIPTIONS OF DEVELOPING SYSTEMS: AN OVERVIEW

J. B. NATION

ABSTRACT. In this talk we will survey some significant results, methods and trends in mathematical modeling, in an attempt to define some of the questions to be explored in this workshop. We are particularly concerned with the formal description of biological systems. The examples, however, will be primarily drawn from the *mathematical* literature. A large and counterproductive gap remains between mathematics and the other sciences. Certainly one goal of this workshop is to narrow that gap.

1. The Scope of the Workshop

The official purpose of this workshop is

> to delineate the fundamental questions relevant to the study of the formal description, analysis, and modeling of developing systems, to critically examine approaches to these questions, to identify those approaches which seem to be most significant and fruitful, and to propose plausible methods and approaches to those questions that remain outstanding.

My assigned task is to parse that: to expand upon this charge by discussing some of the topics and problems that we will be addressing over the next few days.

The unofficial purpose of the workshop is to gather scientists from different fields - and different countries - for an exchange of ideas and expertise. Our success in this area will be measured not only in the improved perspective that we will each bring individually to our research, but also in the scientific cooperation and collaborations that will hopefully result.

The topic which brings us together is the *formal description of developing systems.* We should begin by determining what we mean by a *developing system.* A proper definition of the term might prove difficult, because its boundaries are diffuse. Our situation may be likened to an ancient kingdom: we know some areas that it definitely contains, and others that it definitely does not, but there is no carefully demarcated border. The following subject areas fall within the realm of developing systems. The

Date: November 15, 2002.
Key words and phrases. mathematical model, biological system.

references given are either to talks at this workshop (most of which are included in this volume), or to survey papers, or to recent work of particular interest.

- Population growth (Kapitza).
- Epidemiology (Hethcote [6]).
- Embryonic development (Melekhova, Stollberg).
- Biological growth (Coen, Nakielski, Zagórska-Marek).
- Evolution (Cohen, Skulachev, McNamara *et. al.* [10]).
- Ecology (Laws, Neuhauser [11]).
- Cosmology and geology (Esmer).
- Neurology and psychology (Bracha, Rand).
- Theory of learning and intelligence (Cucker and Smale [1], Grossberg [4]).
- Neural networks and artificial intelligence.
- Social institutions (Arion, Neves, Sergeev, Trofimova).

Surely this list is not complete, and there is considerable overlap between fields. An argument could be made for including proteomics and genomics (Karp [7]), which is certainly an important tool for the study of developing biological systems.

By a *formal description* we customarily mean a mathematical model or, more generally, a mathematical analysis. The mathematical tools which may be used in the description themselves constitute a developing system. Our toolbox includes the following items.

- Discrete dynamical systems and iterative models (Nation).
- Ordinary and partial differential equations (Wilkens).
- More general dynamical systems (Aubin, Brandts, Kopell [8], Sulis).
- Thermodynamics (Sergeev, Lieb and Yngvason [9]).
- Chaos and catastrophe theory (Rand).
- Statistical analysis.
- Fuzzy systems (Yager and Zadeh [13]).
- Ordered systems and closure operators (Ganter and Wille [3]).
- Simulation (Levkovich-Maslyuk, Prusinkiewicz).
- Genetic algorithms (Head).

This toolbox, like *The Magic Pudding*, is continually replenishing itself, adapting to its required applications.

At this point, an interesting question raises its head. How much of the above material can one person know? How much do you know? How about your colleague? Clearly, teamwork in science is increasingly necessary and beneficial. Promoting scientific cooperation, both international and interdisciplinary, is a major goal of this workshop.

2. The Use and Abuse of Models

The object of the modeling process is to find a mathematical structure, whose logic we understand, that mirrors the workings of a complex biological, chemical, physical or social process. A good model should be both

- *descriptive* or *explanatory*, that is, its behavior should reflect qualitatively the process in question, increasing our understanding of it, and
- *predictive*, in the sense that it can be supported or denied by experiment.

A model may, or may not, include a *simulation* of the process.

Each of us has seen useful, and not-so-useful, models. A colleage once facetiously described the modeling process as follows.

(1) Measure n parameters.
(2) Find a family of curves with n parameters that look roughly like your data.
(3) Fit the data.
(4) Declare that you now understand what is going on.

In fact, each of these steps occurs in modeling, but without an overall rationale nothing much will be accomplished.

Let us give some more useful guidelines for modeling. This list was distilled from conversations with Brian Fellmeth and Jes Stollberg of the Pacific Biomedical Research Center. (See also Fellmeth [2]).

- Let the data directly suggest model construction.
- Adopt provisional schemes to suggest experiments to test the scheme.
- Reject or revise models that fail the test.
- Don't overcommit to models that seem to work. In particular, be wary of carrying the model beyond the limits of its applicability.

One would hope that these caveats were unnecessary with this august crowd, but it never hurts to reiterate the basics.

3. A Simple Example

Let us digress to consider a familiar class of models. Furthermore, let us look at the example as a mathematician would, which is probably a different standpoint than that of say a biologist. Again, some of the ideas in this section originated in discussions with Brian Fellmeth.

Suppose we have a fixed quantity of a substance or substances which can exist in n different states or phases. The system has a given initial distribution of states, and is then allowed to approach equilibrium. There is a standard way to describe the evolution of this system under suitably general conditions (e.g., constant temperature). Let the vector $\mathbf{s}(t) = \langle s_1(t), \ldots, s_n(t)\rangle$ denote the amount of material in each state at time t, and for $i \neq j$ let k_{ij} denote the relative rate at which elements in state i

transform to state j. The evolution of this system is described by the initial value problem

$$\frac{d\mathbf{s}}{dt} = \mathbf{M}\mathbf{s}$$
$$\mathbf{s}(0) = \mathbf{s}_0$$

where

$$\mathbf{M}_{ij} = \begin{cases} k_{ji} & \text{if } i \neq j \\ -\sum_{\ell \neq i} k_{i\ell} & \text{if } i = j. \end{cases}$$

From physical considerations, we want each $k_{ij} \geq 0$ for $i \neq j$, and we have built in the conservation of mass in the diagonal elements, by making the column sums zero. Moreover, because the Gibbs free energy satisfies $G_i - G_j \propto \log k_{ij} - \log k_{ji}$, our system satisfies the following condition.

> $(*)$ For any cyclic permutation $(i_1 i_2 \ldots i_m)$ on $\{1, \ldots, n\}$, we have $k_{i_1 i_2} k_{i_2 i_3} \ldots k_{i_{m-1} i_m} k_{i_m i_1} = k_{i_1 i_m} k_{i_m i_{m-1}} \ldots k_{i_3 i_2} k_{i_2 i_1}$.

This condition says that the product of the rate constants around loops is the same in both directions, e.g., $k_{12}k_{23}k_{31} = k_{13}k_{32}k_{21}$.

The solution to our system is of course given by $\mathbf{s}(t) = e^{\mathbf{M}t}\mathbf{s}_0$. However, this elegant form of the solution is not very instructive. The following version, most of which is standard, gives a better description of the solution.

Theorem 1. *If* $\mathbf{M}$ *is as above, then the following hold.*

(1) $\mathbf{M}$ *is singular, so that zero is an eigenvalue.*
(2) *Every eigenvalue of* $\mathbf{M}$ *is real and nonpositive, i.e.,* $\lambda_i \leq 0$.
(3) *The solution of* $\frac{d\mathbf{s}}{dt} = \mathbf{M}\mathbf{s}$ *is of the form*

$$\mathbf{s}(t) = \mathbf{v}_1 + \sum_{j=2}^{n} e^{-\kappa_j t}\mathbf{v}_j$$

where $\lambda_j = -\kappa_j$ *for* $2 \leq j \leq n$ *are the eigenvalues of* $\mathbf{M}$*, and the* $\mathbf{v}_j$*'s are the corresponding eigenvectors, with* $\mathbf{v}_1$ *corresponding to the eigenvalue* 0.

Hidden in this solution is the fact that the scaling of the eigenvectors $\mathbf{v}_j$ depends on the initial vector $\mathbf{s}_0$. On the other hand, it is transparent that the solution tends to $\mathbf{v}_1$, which will represent equilibrium, and that it does so without oscillation (by part 2).

Proof. The matrix $\mathbf{M}$ is singular because its column sums are zero.

To see that the eigenvalues are real, we define the symmetric matrix $\mathbf{B}$ by

$$\mathbf{B}_{ij} = \begin{cases} \sqrt{m_{ij} m_{ji}} & \text{if } i \neq j \\ m_{ii} & \text{if } i = j. \end{cases}$$

Using a theorem of H. Sachs [12] and the condition $(*)$, one can show that $\mathbf{M}$ and $\mathbf{B}$ have the same characteristic polynomial. Hence $\mathbf{M}$ has real eigenvalues.

The Gerschgoren Circle Theorem states that each eigenvalue λ of the matrix $\mathbf{M}$ satisfies $|\lambda - m_{ii}| \leq \sum_{j \neq i} |m_{ji}|$ for some index i. For our matrix this translates to $-2(\sum_{j \neq i} k_{ij}) \leq \lambda \leq 0$.

The remaining claims of the theorem are standard. □

Typically, all we can measure experimentally is one component of the solution, say the first:

$$s_1(t) = v_{11} + \sum_{j=2}^{n} e^{-\kappa_j t} v_{j1}.$$

From this we want to determine

(1) What is n?
(2) What are the time constants κ_j?

The current methods for doing so, using Laplace or Fourier transforms, are not very robust or satisfactory (see, e.g., [2]). They do allow you to obtain at least a lower bound for n, the number of states, which is an important component of any model.

On the other hand, for complex systems we might want to allow an infinite number of discrete states, or a continuous state-space. Moreover, it is not clear that the condition $(*)$ applies to developing systems. For that condition corresponds to a form of conservation of energy, which need not apply locally in developing systems. The same comment could be made regarding the conservation of mass.

The form of the most direct continuous analogue of this situation is straightforward. Given a non-negative function $m(x, y)$ on some domain, the corresponding equation would be

$$\frac{\partial \mathbf{s}(x,t)}{\partial t} = \int m(x,z)\mathbf{s}(z,t)\,dz - \mathbf{s}(x,t)\int m(u,x)\,du$$

where $\mathbf{s}(x, 0)$ is given. It is again easy to write down a formal solution to this equation, but what is required is an analysis of its behavior. This is an example of an old problem where basic, and potentially useful, work remains to be done.

What has been described is a class of models, based on rather minimal assumptions, which can be used in a variety of situations. Some modifications may be appropriate for developing systems. Nonetheless, I would claim that there is some virtue in this "top-down" approach, as opposed to inventing separate models for each particular application. Or, put more plainly, general mathematical systems may have wide applicability.

4. A Plug for Ordered Structures

Another example of this type is the application of thermodynamical principles to different situations. Professor Sergeev is speaking in this tradition, and the survey article by Lieb and Yngvasson [9] gives a nice exposition of the formal development of entropy and the second law of thermodynamics. However, the theory outlined in [9] includes a strong, and to me rather strange, assumption. The theory is based on a quasi-order (a reflexive, transitive relation) on a space of states. The relation that state X precedes state Y means that it is possible for the system to go from state X to state Y. The assumption called the *Comparison Hypothesis* is made: All pairs of states are comparable, i.e., either X precedes Y or Y precedes X (or both). In other words, the quasi-order induces a total order on the equivalence classes of states under precedence.

The total order assumed in this Comparison Hypothesis seems somewhat artificial and archaic. Nature may contain partial orders, which need not be linear orders. Note the successful use of lattice theory in data structures (as in Ganter and Wille's Formal Concept Analysis [3]) and social choice theory (see Johnson and Dean [5] for a recent update and references).

There seem to be two orders of business.

(1) Assumptions of linear order should be questioned, and where appropriate, replaced.
(2) A dynamic theory of ordered sets or lattices may be necessary to describe evolving and developing systems.

Just as we have learned the limitations of using only linear dynamics, so we must learn the limitations of linear order.

5. Mathematics and Science: C. P. Snow Revisited

Finally, let us consider the interaction of two largely distinct and non-overlapping cultures: mathematics and science.

The mathematics establishment is conservative in nature.

- Graduate education in mathematics focuses heavily on 19th and early 20th century mathematics. Much of this material is irrelevant to a career in modern mathematics, whether pure or applied.
- Research support is comparatively minimal, especially in pure mathematics. Research results remain a major factor in academic promotion and tenure, and it is usually a factor in personal and professional satisfaction. However, most mathematicians make their living teaching undergraduate mathematics.
- Most mathematicians work in specialized areas, with very little incentive to diversify or consider applications. There is merit in an individual pursuing a long-term, very specific, research program. But when an entire mathematics department becomes isolated, then it is time to re-establish communication with the rest of the world.
- Publication standards are very different. Mathematics papers must be precise and correct. You must say exactly what you mean, be precise and define every

term. Papers with major errors, undefined terms or no proofs are not allowed in principle. (This can lead to frustration when mathematicians try to read papers from other fields.)

- This is reflected in publication policy. Mathematicians often wait six months to a year for referee reports; publication takes at least another additional year. On the other hand, a good mathematics paper may remain relevant for decades.

Moreover, it is safe to assume that most mathematicians have only minimal training in the sciences.

Of course, there is an active research community in biomathematics, biostatistics, biophysics and bioinformatics. This community is well-represented here. Similarly, there are active groups in mathematical social sciences and economics. Unfortunately, these fields remain isolated from the mainstream of mathematics. This will not be changed by retraining older mathematicians: we need to train a new generation. This process must include attracting some of our best students to applied mathematics at a relatively early age. The new Mathematical Biosciences Institute funded by the National Science Foundation at Ohio State University is a big step in the right direction.

We have left a large factor out of this discussion: computers and computer science. The effect of increased computing power and availability for both mathematics and science is evident. However, historically, computer science has developed largely independently of mathematics and science. In many universities, computer science represents a distinct and isolated department, to the detriment of all concerned. This should not be allowed to happen with biomathematics and modeling, or to mathematics for the social sciences.

The fault is not all with mathematics. Scientists from various disciplines need to learn the mathematical language. The mathematical training of scientists, outside of chemistry and physics, is often minimal. Most scientists learn mathematics up to about the year 1700 plus a little statistics. That won't cut it. Future generations of scientists must have a better, and more thorough, mathematical education.

The interaction between mathematics and science is a developing relationship. It would be interesting to try to model the developing relation between mathematics and the other sciences. Our task here is simpler, merely to promote that relationship. I wish you all a stimulating and successful workshop.

References

[1] F. Cucker and S. Smale, *On the mathematical foundations of learning*, Bull. Amer. Math. Soc. **39** (2002), 1–50.

[2] B. Fellmeth, *Kinetic study of ionic currents in the crayfish giant axon using the voltage clamp technique*, Ph.D. Dissertation, University of Hawaii, 1981.

[3] B. Ganter and R. Wille, Formal Concept Analysis, Springer-Verlag, Berlin, 1999.

[4] S. Grossberg, *Linking Mind to Brain: The Mathematics of Biological Intelligence*, Notices Amer. Math. Soc. **47** (2000), 1361–1372.

[5] M. Johnson and R. Dean, *Locally complete path independent choice functions and their lattices*, Math. Soc. Sci. **42** (2001), 53–87.

[6] H. W. Hethcote, *The mathematics of infectious diseases*, SIAM Rev. **42** (2000), 599–653.

[7] R. M. Karp, *Mathematical challenges in genomics and molecular biology*, Notices Amer. Math. Soc. **49** (2002), 544–553.

[8] N. Kopell, *We Got Rhythm: Dynamical Systems of the Nervous System*, Notices Amer. Math. Soc. **47** (2000), 6–16.

[9] E. H. Lieb and J. Yngvason, *A guide to entropy and the second law of thermodynamics*, Notices Amer. Math. Soc. **45** (1998), 571–581.

[10] J. McNamara, A. Houston and E. Collins, *Optimality models in behavioural biology*, SIAM Rev. **43** (2001), 413–466.

[11] C. Neuhauser, *Mathematical challenges in spatial ecology*, Notices Amer. Math. Soc. **48** (2001), 1304–1314.

[12] H. Sachs, *Abzählung von Wäldern eines gegebenen Typs in regulären und biregulären Graphen, I*, Publ. Math. Decebren. **11** (1964), 119–134.

[13] R. Yager and L. Zadeh, eds., An Introduction to Fuzzy Logic Applications in Intelligent Systems, Kluwer Acad. Pub., Boston, 1992.

Department of Mathematics, University of Hawaii, Honolulu, HI 96822, USA
E-mail address: jb@math.hawaii.edu