NEW RULES FOR SIMULATION AND ANALYSIS IN ECOLOGY, EPIDEMIOLOGY, AND ELSEWHERE

Clarence Lehman and Adrienne Keen

Computation supplied by the Minnesota Supercomputer Institute

A. What just happened?

1. An invisible "phase change" has recently reshaped computation, in only a few years.
2. Vast memories allow new modelling methods—still being invented and deployed.
3. We apply such methods to epidemiology and ecology, outlined here.

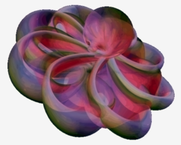

Figure 1. Outside-In, a step in the mathematical process of turning the sphere inside-out, without creases, shown here merely for symbolic imagery. In a sense, the methods describe here turn time-scheduling inside out, without creases. [Depicted at www.geom.uiuc.edu/ graphics/ pix/ Video_Productions/ Outside.In]

B. What is different?

1. Saving limited computer memory, a long-time goal of software design, is no longer crucial.
2. Instead, judiciously "wasting memory" in ways that speed up processing is the new goal.
3. This involves turning our view of modelling inside out.

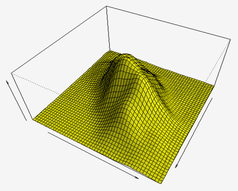

Figure 2. Microscopic entities such as molecules are modelled as continuous, while macroscopic individuals are modelled as discrete. But between the continuous and the discrete lie meso-scale entities such as viruses that are so numerous they cannot be modelled individually, yet so important that just one can alter the course of the simulation. These are maintained as probability clouds that materialize only when they encounter macroscopic individuals. Processing the meso-scale is still in development.

C. New rules

1. Model continuous processes in the usual way, with small finite time steps in the simulated present.
2. Determine when discrete events will occur and cast their fate into the simulated future.
3. Fashion a world of probability clouds between the continuous and discrete, that start in the present and drift into the simulated past.
4. Seek maximal-speed procedures for all major steps—"Order-1" procedures, see Part E.

Table 1. Scales of attention in modelling natural systems.

Scale	*Continuity*	*Model*	*Examples*
Macro-	Always continuous	Differential equations operating in the present	Molecules and smaller
Micro-	Always discrete	Individuals cast forward into the future	Trees and seeds, birds and bees, people
Meso-	Collapsing from continuous to discrete	Probability functions receding into the past	Virus, bacteria, spores, and larger

D. Choosing times in the future

1. Deciding the future is substituting cumulative probability distributions for density functions.

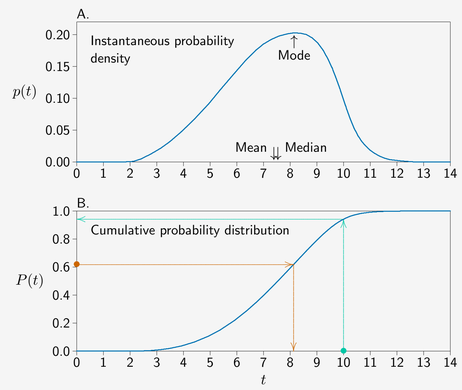

*Figure 3. Hypothetical progression from time of infection until the host becomes infectious. **A.** The probability density function, $p(t)$, showing the fraction of individuals who become infectious in the time interval $t \pm \Delta t$. **B.** The corresponding cumulative distribution function, $P(t)$, showing the fraction of individuals who will have become infectious by time t. This is how "intrinsic rates" convert to times of future events (red arrow).*

2. Scheduling future events separates into "intrinsic rates" and "extrinsic probabilities".
3. Intrinsic rates involve individuals, setting future times when events are destined to arise.
4. Extrinsic probabilities involve groups, picking which future events will have an effect.
5. These probabilities can be greater than one—more than certain—when they represent events which already happened. To be avoided in simulations.

E. Order-1 procedures

1. Order-1 procedures access tables of 60 million individuals as fast as tables of 60 individuals.
2. Our time management is Order-1 for scheduling new events, canceling old events, and fetching the next event in sequence.
3. Our group management is Order-1 with respect to the number of individuals, and approximately Order-1 with respect to the number of groups.
4. We run 6×10^7 individuals—the population of the UK—in a multi-compartment disease model through 30 years in approximately 90 seconds on each Itasca processor. Runtime is proportional to the number of individuals.
5. We have applied these methods to large-scale TB and HIV models.

F. The flow of time

There are no "time steps" here, only "inter-event times." Computer processing is spent only on events themselves, not checking for possible events in each time step. The simulation steps are these:

1. Select the earliest event and note its time, t_{earliest}.
2. Set the present time, $t = t_{\text{earliest}}$, and remove the event from the schedule.
3. Dispatch the event if its extrinsic probability allows. Process everything associated with the event, adding new events and removing old ones.
4. If more events remain and the ending simulated time has not been reached, repeat from Step 1.
5. Otherwise terminate successfully.

G. Equation vs equation-free

1. The time-honored method of modeling starts with the microscopic details and abstracts into mathematical equations for a macroscopic view.
2. Then the mathematical equations are simulated on a computer.
3. The alternative is to simulate the microscopic details directly, without reducing to equations.
4. This makes many new things possible for complex systems. Derivatives and eigenvalues can still be computed, even without equations.

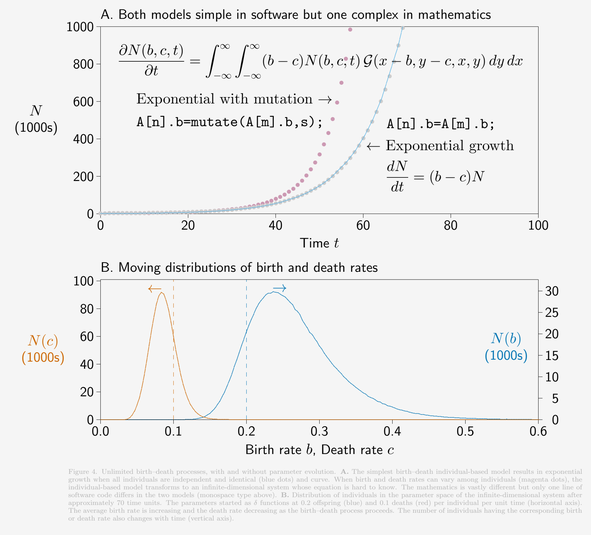

*Figure 4. Unlimited birth-death processes, with and without parameter evolution. **A.** The simplest birth-death individual-based model results in exponential growth when all individuals are independent and identical (blue dots) and curve. When birth and death rates can vary among individuals (magenta dots), the individual-based model transforms to an infinite-dimensional system whose equation is hard to know. The mathematics is vastly different but only one line of software code differs in the two models (monospace type above). **B.** Distribution of individuals in the parameter space of the infinite-dimensional system after approximately 70 time units. The parameters started as 4 functions at 0.2 offspring (blue) and 0.1 deaths (red) per individual per unit time (horizontal axis). The average birth rate is increasing and the death rate decreasing as the birth-death process proceeds. The number of individuals having the corresponding birth or death rate also changes with time (vertical axis).*

Exploiting fate, chance, and procrastination can speed up simulations by orders of magnitudes.