Exact Methods for Computing Biological System Dynamics

Drew Endy (http://mit.edu/endy/)

Goals Covered Last Time

A. When to use computational methods?

B. When to use exact methods?

Goals for Today

C. What is the underlying physics models used by exact methods?

D. How can we use this model to compute when a reaction will take place?

25. **Question C.** To develop an exact discrete reaction event method we need to consider the underlying physics model that might be used to represent individual reactions events inside cells. [The method we' ll consider was developed by Daniel T. Gillespie and published in Journal of Computational Physics v22 p403-434 in 1976. You can find a one hour video of Dan talking about his work online; the structure of this lecture is based directly on his presentation. Here's the URL: http://streaming1.osu.edu/ramgen/media/mbi/mbi302.rm]

26. What sort of system are we talking about? Let's define a volume (V) containing N different types of molecules S (aka, species), with X number of molecules for each species.

27. How does the system change over time? They change via elementary chemical reactions. What is an elementary chemical reaction? An elementary chemical reaction is a single instantaneous physical event. Only three such events are really likely. (i) from nothing comes a molecule (i.e., spontaneous generation), (ii) one molecule changes into another molecule(s), and (iii) two molecule collide and react to produce something (or nothing).

28. How could we determine what reactions occur when? Well, we can consider all sorts of information. For example, consider the 3D position and velocity of all the molecules. This might lead us to use molecular dynamics: following every molecule in the system of interest as it moves about, bumping into other molecules, changing course, reacting every now and then. This would be great -- except that it's computationally impractical for most systems of interest. One reason for this is that most molecule-molecule collision events are non-reactive.

29. So, what if we didn't keep track of the non-reactive (i.e., bounce off) collisions?

30. If most collision events are non-reactive then it becomes possible to consider and accept an important assumption, that our system is spatially homogeneous. In other words, the jostling of the non-reactive collisions keeps everything well-mixed (on the time scales of the elementary reaction events that we care about).

31. As a result, we no longer need to keep track of the individual positions and velocities of molecules. Thus, our system is now defined as having N chemically active species, S in volume V, with X being the number of molecules for each species. Or, S_i in V, for i=1..N, w/ X_i = current numbers of molecules for each S_i. Thus, the state of the system is defined by the value of the array X and the behavior of the system is represented by how the values in X change over time.

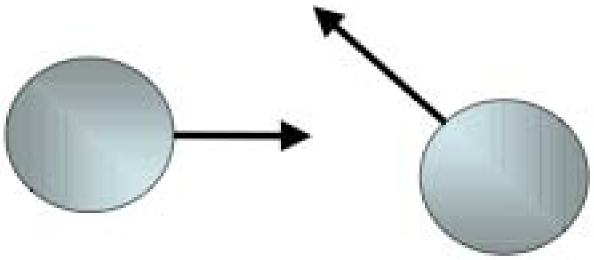
32. Another consequence of the above is that we only have to compute chemical reaction events, and not the non-productive molecular collisions that are going on all the time.

[Imagine modeling everybody in a business reception jostling about, when all you really care about is keeping track of whether or not two people exchange business cards or if somebody enters or exits the room].

33. So, how are we going to compute the events? It turns out that we need some model for telling us two things. First, what's the chance (or, probability) that a particular reaction will take place? Second, what's the consequence of a reaction taking place?

34. How do we model the probability, ***a***, that a particular reaction will take place? We need to compute (i) the chance that any individual reactions of a particular type will take place, ***c***, and (ii) the number of unique reactant combinations for each reaction type, ***h***.

35. How do compute ***c***? We need to use some physics. For example, for a second-order elementary reaction, we first determine the the chance that two molecules will collide and then we multiple this value by the chance that a reaction event will take place given a collision. This can be described via the equation on the next page:



$$c_{\upsilon}\delta t = V^{-1}\,\pi d_{12}^{\,2}\,(8kT/\pi m_{12})^{0.5}\,e^{(-\mu_{\upsilon}/kT)}\,\delta t$$

where:

C = average probability that a particular combination of molecules, u, will react
δt = time interval until the next reaction event occurs
V = volume of the system
d = effective collision diameter between molecules 1 and 2, equal to (d1+d2)/2
(8kT...) = average apparent molecule-molecule velocity via Maxwell-Boltzman
u = activation energy for reaction u

36. Basically, the above equation "says" that two molecules are moving about; that there is some "collision volume" over some infinitely small time slice (small enough so that the probability of more than one reactive collision is negligible); that the ratio of the collision volume to the total system volume defines the chance of a collision; and a standard Arrhenius term expresses the chance that the collision energy is above some level required to the reaction to take place.

[Aside: Note that the above assumes a sort of hard-sphere gas approximation for the inside of a cell. Is this what the inside of a cell is like? Such issues can be considered here by changing the above collision model].

37. Next, how do we determine the number of unique reactant combinations, ***h***, for any given type of reaction? This depends on the type of the reaction. If the reaction is a first order reaction then the number of unique reactant combinations is equal to X, the number of molecules of that particular reactant in the system at the time. If the reaction is heterogeneous second order (A+B -> something), then the number of unique reactant combinations is the product of X_A and X_B. If the reaction is a homogeneous second order (A+A -> something), the it's X*(X-1)/2. And so on. Finally, if the reaction is a zero order (e.g., spontaneous generation) then ***h*** equals 1 (everything will depend on the chance that a reaction happens as there is no formal accounting for the substrate).

38. Finally, how do we model the consequence of a reaction taking place? Easy! We use the reaction stoichiometry. For example, if the reaction is "A turns into B," then every

time this reaction occurs we decrease the number of A by one and increase the number of B by one, and so on.

39. Now we have some way of computing that a reaction event will take place in some very small slice of time. Or, in Gillespie's notation... ***a*** = ***c********h***, as defined above. Or, borrowing from one of his slides:

$$S_1 + S_2 \underset{c_2}{\overset{c_1}{\rightleftharpoons}} 2S_1 \quad \Rightarrow \quad \begin{cases} a_1(\mathbf{x}) = c_1 x_1 x_2, & \boldsymbol{\nu}_1 = (+1, -1, 0, \ldots, 0) \\ a_2(\mathbf{x}) = c_2 \dfrac{x_1(x_1 - 1)}{2}, & \boldsymbol{\nu}_2 = (-1, +1, 0, \ldots, 0) \end{cases}$$

40. **Question D.** Finally, how can we use this model to compute when a reaction will take place?

41. We've some probability that a particular reaction will take place over time. This leads to an exponential distribution for the time at which a reaction will take place. That is, the reaction is likely to take place at some point in time; if it does, great; if it doesn't, then the probability that it will take place in the future is the chance that it's not yet taken place multiplied by the chance that it will happen now; and so on.

42. From this exponential, if we want to compute one instance of a time that a particular reaction will next take place, we can use the following equation...

$$\tau = (1/a)*\ln(1/r)$$

where τ is a randomly sampled time that a reaction next takes place

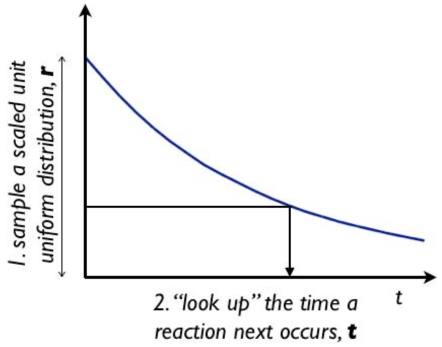
a is the reaction propensity (chance of reaction)

r is a random number drawn from a unit uniform distribution

43. Say what? OK. I'm sure a mathematician somewhere is screaming but here's how I think about this. Solve the above equation for ***r***. You get...

$$r = \exp[-a*\tau]$$

Then, graphically, what this leads to is... the first sample along the Y-axis of an exponential probability density function using a uniform distribution scaled to max. PDF value at = 0. Then, lookup the value of for any value of ***r***. See the example on the next page.



44. Big picture, what does this mean? It means that each run of our computational method will be unique, because we are sampling a uniform PDF to calculate the times that reactions will next occur.

45. Next time we'll think about how to wrap all this math up into a useful computational algorithm. This turns out to be pretty easy and will give us a chance to revisit the essentials of the material from the first two classes.