Why Should Mathematical Modeling In Biology Be Even Possible?

Winfried Just
Department of Mathematics, Ohio University

March 27, 2018

An example of mathematical modeling in biology. Based on Oduro, Grijalva, Just; 2018 and 20??

Chagas disease is transimitted primarily through the bites of insect vectors called triatomines. These can infest housing units by migrating either from sylvatic areas or from infested units. Insecticide spraying remains an important control measure. But:

- The effect of the insecticide wears off after some time, so that previously treated units may become reinfested.
- Insecticide is costly, in monetary terms as well as toxicity to humans and potential for evolution of resistance to it.
- Treatment may only be partially successful in the sense that some insect colonies may survive inside a treated unit and re-emerge after the effect of the insecticide has worn off.

Biological Problem: How to maintain a low endemic equilibrium of infestation levels with as little insecticide use as possible?

A Mathematical Problem: The above implicitly assumes that an endemic equilibrium always be approached. Is this true?

Outline of the Imperfect Treatment Model

- Hosts are housing units. Their total number m is constant.
- A unit is in compartment **I** (infested) if it contains a viable insect colony.
- An uninfested aka susceptible unit in compartment **S** can become infested by migration of insects from a sylvatic area or form a unit in compartment **I**.
- Only infested units receive insecticide treatment at rate r.
- With probability $1-\alpha$, treatment will be unsuccessful. Then Some insects survive treatment and temporarily hide within the treated unit while the insecticide remains effective. In this case, treatment will move the unit to compartment $\mathbf{R_1}$.
- If the treatment is successful and destroys the colony, it will move the unit to compartment $\mathbf{R_2}$.
- As the effect of the insecticide wears off, units move out of compartments $\mathbf{R_1} \cup \mathbf{R_2}$ at rate w.

Flow chart of the model

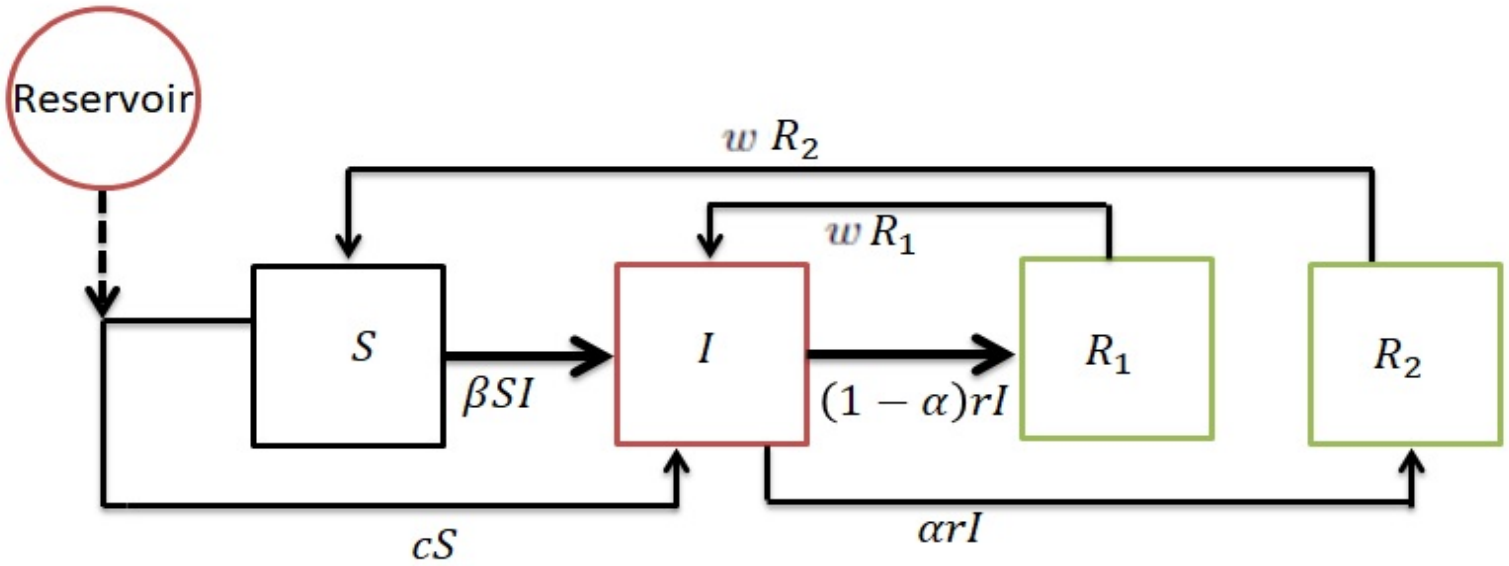


Figure: The Imperfect Treatment Model

The Imperfect Treatment Model

Here the variables S, I, R_1, R_2 will denote the numbers of units in the compartments with analogous names. The DEs are:

$$\begin{aligned}
\frac{dS}{dt} &= -\beta IS - cS + wR_2 \\
\frac{dI}{dt} &= \beta IS + cS - rI + wR_1 \\
\frac{dR_1}{dt} &= (1-\alpha)rI - wR_1 \\
\frac{dR_2}{dt} &= \alpha rI - wR_2.
\end{aligned} \tag{1}$$

Note that $R_2 = m - S - I - R_1$. Thus we can reduce (1) to:

$$\begin{aligned}
\frac{dS}{dt} &= -\beta IS - cS + w(m - S - I - R_1) \\
\frac{dI}{dt} &= \beta IS + cS - rI + wR_1 \\
\frac{dR_1}{dt} &= (1-\alpha)rI - wR_1.
\end{aligned} \tag{2}$$

Our major findings when $c > 0$:

- This model has a unique equilibrium $EE = (S^*, I^*, R_1^*)$.
- EE is both locally and **globally asymptotically stable.**
- This equilibrium is endemic, which means that $I^* > 0$.
- Thus it is not feasible to eradicate infestation with insecticide treatment alone.
- For any fixed parameters $\alpha, \beta, c, w > 0$, when m is sufficiently large, the Imperfect Treatment Model exhibits a dual-rate effect, which implies that the long-range cost of maintaining an equilibrium with I^* below a threshold that is deemed tolerable can be decreased by initially highly aggressive interventions in villages that have have high levels of infestation.

These findings are analogous to our previous findings for the Basic Model where all treatments were assumed successful, as well as for some generalizations of the Basic Model in a different direction.

How did we prove global asymptotic stability?

- Proving uniqueness and local asymptotic stability of the equilibrium $EE = (S^*, I^*, R_1^*)$ required substantial work, but was possible by direct calculations and analysis of the eigenvalues of the Jacobian.
- However, we still needed to rule rule out trajectories that would not approach any equilibrium.

Recall the Imperfect Treatment Model

Since $R_2 = m - S - I - R_1$, we could reduce the model to 3 DEs:

$$\begin{aligned}
\frac{dS}{dt} &= -\beta IS - cS + w(m - S - I - R_1) \\
\frac{dI}{dt} &= \beta IS + cS - rI + wR_1 \\
\frac{dR_1}{dt} &= (1 - \alpha)rI - wR_1.
\end{aligned}$$

For the Basic Model (which can be obtained by setting $\alpha = 1$), we had only 2 variables and could use Dulac's Criterion together with the Poincaré-Bendixson Theorem.

This is not feasible here. Now what?

How did we prove global asymptotic stability?

Not directly feasible, to be more precise.

- We first observed that all trajectories approach the 2-dimensional invariant region $\Omega_=$ of the state space where $\alpha R_1 = (1-\alpha)R_2$.
- Next we used Dulac's Criterion to rule our periodic orbits inside $\Omega_=$.
- The Poincaré-Bendixson Theorem then implies that all trajectories that start in $\Omega_=$ approach the equilibrium.
- Finally, we deduced global asymptotic stability of (S^*, I^*, R_1^*) on the entire state space from these results.

Recap: What have we done in our modeling?

The real biological system is complex: It involves complicated interactions between a huge number of agents (housing units, insects, the molecules of the insecticide) in many different environments (villages).

We distilled the situation into a model with very few variables with that postulates fairly simple interactions between these variables.

We were able to analyze this model by even further reducing the number of variables.

The model gave us some important insights and makes predictions of potential importance for public policy.

These predictions turned out to be remarkably robust under changes of details of the model.

Were we just lucky???

Not *exceptionally* lucky, anyway

The entire field of mathematical biology is based on the assumption that with enough skill, effort, and external funding researchers will usually get lucky in similar ways as we did.

Empirical evidence supports this assumption.

But why should we expect it to be true?

Are there some important **mathematical reasons** why complex biological systems **cannot escape** fairly simple structures so that:

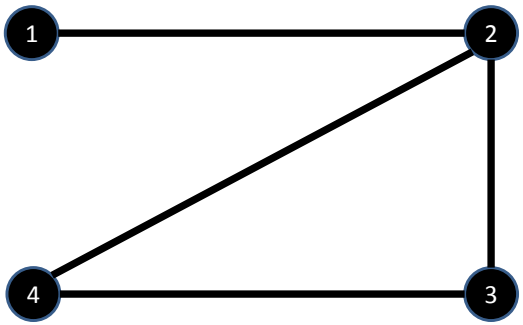
1. These structures can be described with relatively few variables with fairly simple interactions.
2. The behavior of these structures is fairly robust under moderate modifications of the real system.
3. "Reasonable questions of interest" can be answered in terms of these structures.
4. Mathematical biologists have a chance of discovering them.

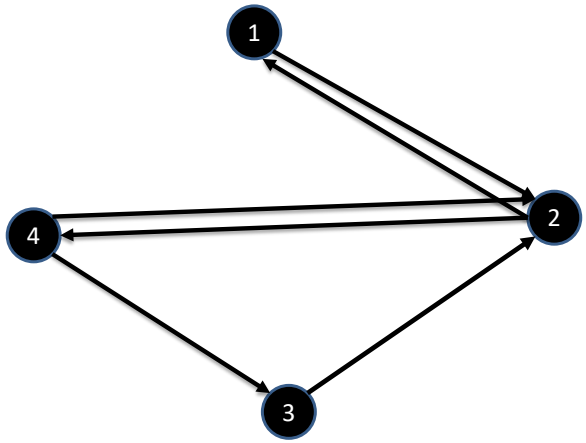
Graphs and digraphs

A graph is a pair $G = (V, E)$ such that V is a set of vertices and E is a set of unordered pairs $\{v, w\} \subset V$ with $v \neq w$ called edges.

A directed graph aka digraph is a pair $D = (V, A)$ such that V is a set of vertices and A is a set of ordered pairs $\langle v, w \rangle$ of vertices $v \neq w$ called arcs.

The structure on the left is a graph; the structure on the right is a digraph:



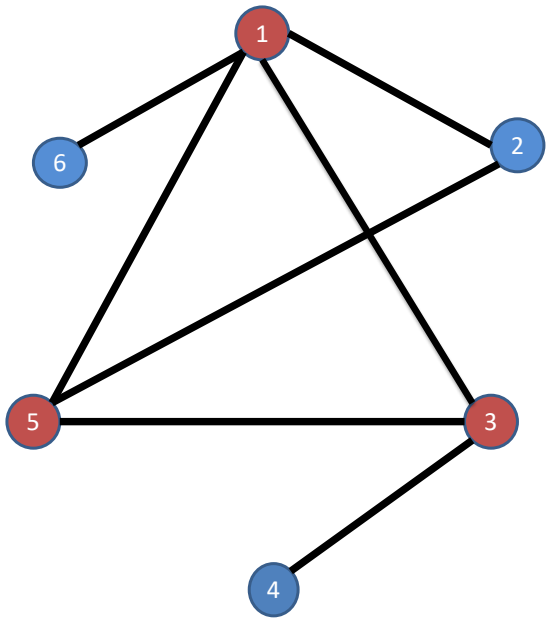


Simple structures in graphs: Cliques and independent sets

Let $G = (V, E)$ be a graph. A clique in G is a subset $C \subseteq V$ such that $\{v, w\} \in E$ for all $v, w \in C$ with $v \neq w$.

An independent set in G is a subset $I \subseteq V$ such that $\{v, w\} \notin E$ for all $v, w \in I$.

In the graph below, the set $\{1, 3, 5\}$ forms a clique; the set $\{2, 4, 6\}$ is independent.



The impossibility of avoiding simple structures: Ramsey's Theorem

Theorem (Ramsey)

For every $n > 0$ there exists a smallest positive integer $R(n)$ such that every graph with at least $R(n)$ vertices contains either a clique of size $\geq n$ or an independent set of size $\geq n$.

How fast does $R(n)$ grow?

- $\liminf_{n\to\infty} R(n)^{1/n} \geq \sqrt{2}$.
- $\limsup_{n\to\infty} R(n)^{1/n} \leq 4$.
- **Problem:** (Erdős, 1947, with prizes offered for the solution)
  Does $\lim_{n\to\infty} R(n)^{1/n}$ exist? ($100)
  If $\lim_{n\to\infty} R(n)^{1/n}$ does exist, what is its value? ($250)

Random (di)graphs

Instead of studying particular graphs or digraphs, we may want to assume that for a fixed vertex set V the edges of G or the arcs of D are randomly drawn from a certain probability distribution.

For example, when V has size n and each of the $\binom{n}{2}$ potential edges of G is randomly and independently included with connection probability p, then we obtain (instances of) Erdős-Rényi graphs.

The analogous construction when the vertex set is partitioned into subsets V_i for $i = \{1, 2, \ldots, r\}$ and an arc $\langle v, w \rangle$ from V_i to $w \in V_j$ is included with connection probability p_{ij} gives (instances of) multitype Erdős-Rényi digraphs.

Here p, p_{ij} may or may not depend on n.

We may then investigate "typical" properties that such such (di)graphs will have with probability approaching 1 as $n \to \infty$.

If a given biological network can be assumed to be roughly "typical" in this sense, then it would be "robust" with respect to such properties.

Random graphs and Ramsey numbers

Theorem

Let $f : \mathbb{N} \to \mathbb{N}$ be a function such that $\lim_{n\to\infty} f(n)\sqrt{2^{-n}} = 0$.

Then for connection probability $p = 0.5$, with probability approaching 1 as $n \to \infty$, an Erdős-Rényi graph with $f(n)$ vertices will have neither a clique nor an independent subset of size n.

It follows that $\liminf_{n\to\infty} R(n)^{1/n} \geq \sqrt{2}$.

Can this result be proved without using random graphs?

If not, this would presumably mean that the lack of certain simple structures (large cliques or independent sets) in graphs can only be caused by randomness; anything constructed more purposely would contain such structures.

Finding a **constructive** proof that $\liminf_{n\to\infty} R(n)^{1/n} > 1$ is another **open problem.** Paul Erdős offered a prize of $100 for it.

So is Random = Maximally Unstructured?

Not at all! At least not in terms of biological modeling.

Recall our introductory example. The modeling relies on the implicit assumption of a lot of independent events (an insect does or does not migrate to an unifested housing unit, does or does not survive insecticide treatment, etc.) The effect of individual such events at a time t can be represented at the level of the agents by an indicator random variables and the variables that we used in our model represent sums of many suitable such random variables. The DEs then represent presumed dynamics of their mean values.

Many models in mathematical biology rely on variables that in effect represent statistics about the states of the agents, and the resulting models rely implicitly on assumptions of randomness in the interactions of agents.

Do there exist complex systems that simultaneously lack describable structure at the agent level and approximability of their statistics by suitable probability distributions?

Simple models of neuronal networks

Consider a class of neuronal network models $N = (n, D)$, where:

- n is the number of neurons.
- $D = (V, A)$ is a digraph with vertex set of size n.
- $\langle w, v \rangle \in A$ signifies that neuron w **can** send firing input to neuron v (via some synaptic connections).
- Time t proceeds in discrete steps.
- At any time t, neuron v can be in state $s_v(t) = 0$ (fire) or $s_v(t) = 1$ (rest).
- If $s_v(t) = 0$, then $s_v(t+1) = 1$
  (the neuron must go through a refractory period).
- If $s_v(t) = 1$ & $\exists w\ s_w(t) = 0$ &$\langle w, v \rangle \in A$, then $s_v(t+1) = 0$.
- If $s_v(t) = 1$ & $\neq \exists w\ s_w(t) = 0$ &$\langle w, v \rangle \in A$,
  then $s_v(t+1) = 1$.
  Thus a neuron will fire in the next step if, and only if, it is currently at rest and receives some firing input.

Simple structures in our simple models

The class of neuronal network models $N = (n, D)$ on the previous slide is a subclass of a broader class of such models that has been introduced and investigated by Terman, Ahn, Borisyuk, Smith, Wang, Just and is currently studied by Rabi K.C.

A neuron v in such a model is minimally cycling if
$(s_v(0), s_v(1), s_v(2), s_v(3), s_v(4), \dots) = (0, 1, 0, 1, 0, 1, \dots)$ or
$(s_v(0), s_v(1), s_v(2), s_v(3), s_v(4), \dots) = (1, 0, 1, 0, 1, 0, \dots)$.

An autonomous set (for the initial state) $AU \subseteq V$ consists of neurons that are all minimally cycling and receive firing input from other neurons in AU whenever they are ready to fire.

Notice that the dynamics on an autonomous set lacks complexity and is totally unresponsive to inputs from other part of the network. It reminds me of echo chambers in social networks.

Does Random $\Longrightarrow$ Unstructured at the agent level?

Theorem (Just, Ahn, Terman 2008; Rabi K.C. 2018)

Consider a family of multitype Erdős-Rényi digraphs D with connection probabilities $p_{ij} = \lambda \frac{\hat{c}_{ij}}{n}$, where the number r of types and the values $\hat{c}_{ij} > 0$ are fixed, while λ is a scaling parameter, and D has an arbitrarily large number of vertices of n. Then for every $\varrho < 1$ there exists $\lambda_\varrho > 0$ such that for any fixed $\lambda > \lambda_\varrho$, with probability approaching 1 as $n \to \infty$, a generic initial state in the model $N = (n, D)$ will have an autonomous set of size $> \varrho n$.

This theorem holds for the larger class of neuronal networks mentioned on the previous slide.

Question: To what other updating rules for the dynamics of individual agents does this result generalize?

Question: Can the result be generalized to a *bona fide* mathematical version of "too much connectivity will cause the population to split up into a few echo chambers"?

Can we get away from probability distributions?

Theorem (Simplified Version)

Consider a random digraph D drawn from a suitable distribution. Then for every $\varrho < 1$, as long as the connection probabilities for this distribution are large enough, with probability approaching 1 as $n \to \infty$, a generic initial state in the model $N = (n, D)$ will have an autonomous set of size $> \varrho n$.

- For Erdős-Rényi D with $p = 0.5$, we get $\varrho = 1$.
- Rabi's result covers a larger class of random digraphs than in the version of the theorem on the previous slide.
- Rabi is still working on further generalizing the results to even more probability distributions.
- **Question:** But can we perhaps dispense with probability distributions altogether and prove a version that would hold for all digraphs that have enough arcs?

ε-regularity

Consider V of size n and a partition of V into sets V_i of size n_i for $i = 1, 2, \ldots, r$.

For $X, Y \subset V$, let $|E(X, Y)|$ denote the number of edges between vertices in X and and vertices in V.

In a multitype Erdős-Rényi graph with this partition, we would expect, for randomly chosen $X \subset V_i$ and $Y \subset V_j$, that $|E(X, Y)| \approx p_{ij}|X||Y|$ with probability close to 1.

Definition

Let $G = (V, E)$ be any graph, and let $\varepsilon > 0$. A pair (V_i, V_j) of disjoint sets of vertices is ε-regular if for all $X \subset V_i$ and $Y \subset V_j$ with $|X| \geq \varepsilon|V_j|$ and $|Y| \geq \varepsilon|V_j|$ the following inequality holds:

$$\left| \frac{|E(X, Y)|}{|X||Y|} - \frac{|E(V_i, V_j)|}{|V_i||V_j|} \right| \leq \varepsilon.$$

Szemerédi's Regularity Lemma

Lemma (Szemerédi; 1978)

For every $\varepsilon > 0$ and there exists an integer $r(\varepsilon)$ such that if $G = (V, E)$ is any graph, there exists a partition of V into r pairwise disjoint subsets V_i for some $r \leq r(\varepsilon)$ such that

- *All except at most εr^2 of the pairs (V_i, V_j) for $1 \leq i < j \leq r$ are ε-regular, and*
- *$|V_i| - |V_j| \leq 1$ for all $1 \leq i < j \leq r$.*

- This lemma basically says that **any** graph with sufficiently many edges essentially behaves like a multitype Erdős-Rényi random graph on most of the vertex set, where the number of types r has a universal upper bound $r(\varepsilon)$ that only depends on our error tolerance ε.
- Thus in a way, the lemma says that no large enough graph can be "entirely nonrandom."
- One caveat: The number $r(\varepsilon)$ grows horribly fast as $\varepsilon \to 0^+$.

Some (open) questions on Szemerédi's Lemma

- **Conjecture:** Szemerédi's Lemma has a natural generalization to digraphs.
- **Question:** Under what additional conditions would a partition as in Szemerédi's Lemma for digraphs be sufficient to deduce the existence of a large autonomous set as in Rabi's theorem for connectivities that are mutitype Erdős-Rényi digraphs?
- **Problem:** Are there analogues of Szemerédi's Lemma for (certain types of) discrete-time dynamical systems on networks, as conceptualized by graphs or digraphs?

References

- Oduro, B, Grijalva, MJ, Just, W (2018); Models of disease vector control: When can aggressive initial intervention lower long-term cost? *Bulletin of Mathematical Biology,* First online February 05, 2018.
- Oduro, B, Grijalva MJ, Just W (20??); A model of insect control with imperfect treatment. *Ready for submission.*
- Just, W, Ahn, S, Terman D (2008); Minimal attractors in digraph system models of neuronal networks. *Physica D* **237**, 3186–3196.
- Szemerédi, E (1978), Regular partitions of graphs. In: *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976).* *Colloq. Internat. CNRS* **260**, Paris: CNRS, 399–401.