## Quantitative Biology (The Science of Complexity)

Winter 2017
Rama Ranganathan
The Green Center for Systems Biology, ND11.120E


## Administrative matters

## Administrative matters...

(1) We will have 16 lectures...

1/09 Introduction...the principles of modeling
1/11 Linear systems theory I...simple systems, graphical tools, decomposability
1/16 Linear systems theory II...the space of possible solutions, understandability
1/18 Stochastic models...models at the limit of small numbers, new behaviors.
1/23 Diffusion and driving forces, thermodynamic analysis
1/25 Diffraction theory...the spatial Fourier transform
1/30 Decomposition...low dimensional representations of big data

2/1 Introduction to non-linear systems...the origins of complexity
2/6 A detailed example I...the van der Pol oscillator and the Chua circuit, theory.
2/8 A detailed example II...simulation and experiments of simple non-linear systems
2/13 Small non-linear systems...the MAP kinase switch
2/15 Mesoscale non-linear systems...the problem of cellular signaling
simplicity of linear systems

## 2/20 Large non-linear systems...the problem of proteins

2/22 Epistasis...genetic and evolutionary principles, fitness landscapes
2/27 Information Theory
complexity of nonlinear systems

3/1 A conclusion...so what is complexity? A proposal for a general strategy

## Administrative matters...

(1) We will have 16 lectures....
(2) All course materials, notes, lecture PDFs, and homework problems will be distributed on the course Moodle site. Please send Carla Childers (carla.childers@utsouthwestern.edu) an email if you are an auditing member of the course to obtain materials,

## Administrative matters...

(1) We will have 16 lectures....
(2) All course materials, notes, lecture PDFs, and homework problems will be distributed on the course Moodle site. Please send Carla Childers (carla.childers@utsouthwestern.edu) an email if you are an auditing member of the course to obtain materials,
(3) We will have homework problems. As relevant, we will use the MATLAB environment.

Available for free to all here at the university, and the "MATLAB bootcamp" materials are on the Moodle site.

## Administrative matters...

(1) We will have 16 lectures....
(2) All course materials, notes, lecture PDFs, and homework problems will be distributed on the course Moodle site. Please send Carla Childers (carla.childers@utsouthwestern.edu) an email if you are an auditing member of the course to obtain materials,
(3) We will have homework problems. As relevant, we will use the MATLAB environment.

Available for free to all here at the university, and the "MATLAB bootcamp" materials are on the Moodle site.
(4) Interaction is critical...ask many questions and start discussions.

## A portion of core metabolism


 U,上, $\qquad$




FIGURE 1 | Main intracellular differentiation pathways of a single CD4+ T cell. Systems Biology Markup Language (SBML)-compliant network model of $\mathrm{CD} 4+\mathrm{T}$ cell differentiation, including cytokines, receptors, and intracellular signaling pathways controling CD4 + T cell fate and function.
"...the most recent systems biology markup language (SBML)-compliant network...provides a structured understanding on different pathways involved in CD3+ T cell differentiation...."

## Does this lead to understanding?



Some have tried to apply the strict principles of reduction-based science....to write down detailed models for each and every reaction.

Does this lead to understanding?

An important and topical goal is to define the (yet unknown) general laws underlying the behavior and evolutionary origin of complex systems in biology

An important and topical goal is to define the (yet unknown) general laws underlying the behavior and evolutionary origin of complex systems in biology

what do we mean by complex?

## From Merriam-Webster..

## com•plex /käm'pleks/

(1) ....consisting of many parts, e.g. a complex piece of machinery...

From Merriam-Webster...

## com•plex /käm'pleks/

(1) ....consisting of many parts, e.g. a complex piece of machinery...


From Merriam-Webster...

## com•plex /käm'pleks/

(1) ....consisting of many parts, e.g. a complex piece of machinery...

But, is the condition of having many parts imply complexity?

A mole of gas has many parts....but we have the ideal gas law and macroscopic properties of this system that can be understood from properties of its parts..

$$
P V=n R T
$$



A mole of gas has many parts....but we have the ideal gas law and macroscopic properties of this system that can be understood from properties of its parts..

$$
P V=n R T
$$



The essence of the simplicity is independence of the parts....thus the behavior of the whole is a sum of the behaviors of the parts

A mole of gas has many parts....but we have the ideal gas law and macroscopic properties of this system that can be understood from properties of its parts..

$$
P V=n R T
$$



The essence of the simplicity is independence of the parts....thus the behavior of the whole is a sum of the behaviors of the parts

1/11 Linear systems theory I...simple systems, graphical tools, decomposability
1/16 Linear systems theory II...the space of possible solutions, understandability
$1 / 18$ Stochastic models...models at the limit of small numbers, new behaviors.
1/23 Diffusion and driving forces, thermodynamic analysis
1/25 Diffraction theory...the spatial Fourier transform
1/30 Decomposition...low dimensional representations of big data

Also as we will see later on, even a ridiculously simple looking reaction with very few parts can exhibit extraordinary complexity...


2/1 Introduction to non-linear systems...the origins of complexity
2/6 A detailed example I...the van der Pol oscillator and the Chua circuit, theory.
2/8 A detailed example II...simulation and experiments of simple non-linear systems
2/13 Small non-linear systems...the MAP kinase switch

So having many parts is neither necessary not sufficient to specify complexity...

From Merriam-Webster...

## com•plex /käm'pleks/

(1) ....consisting of many parts, e.g. a complex piece of machinery...
(2) ...consisting of many interconnected or interwoven parts....
勧；

$$
2
$$






2


 $\qquad$ $\frac{-1}{2-1}$
 xen $\frac{x}{4}=$

空至


 $=2$ $\frac{1}{x+1014}$

$=$

的
$\frac{2}{2}$
（4） 1




$\qquad$


$\qquad$
$\qquad$
$\int=\frac{x+1}{=-2}$

From Merriam-Webster...

## com•plex /kảm'pleks/

(1) ....consisting of many parts, e.g. a complex piece of machinery...
(2) ...consisting of many interconnected or interwoven parts....

But, does the condition of having many interconnected parts necessarily imply complexity?

## com•plex /käm'pleks/

(1) ....consisting of many parts, e.g. a complex piece of machinery...
(2) ...consisting of many interconnected or interwoven parts....


And...we make many very complex circuits with millions of parts and can yet make clear prediction of how such systems behave...satellite control systems, cruise controls of cars, air handling systems, an airplane, this laptop...

From Merriam-Webster...

## com•plex /käm'pleks/

(1) ....consisting of many parts, e.g. a complex piece of machinery...
(2) ...consisting of many interconnected or interwoven parts....

We make many very complex circuits with millions of parts and can yet make clear prediction of how such systems behave...satellite control systems, cruise controls of cars, air handling systems, an airplane, this laptop...

The essence of this simplicity is linearity in the connections between system components...

## com•plex /käm'pleks/

(1) ....consisting of many parts, e.g. a complex piece of machinery...
(2) ...consisting of many interconnected or interwoven parts....


And what about condensed phases like liquid water or ice? An extensive hydrogen bonding network....

## com•plex /käm'pleks/

(1) ....consisting of many parts, e.g. a complex piece of machinery...
(2) ...consisting of many interconnected or interwoven parts....


We also have theories for condensed phases of matter such as liquids, solids, and even disordered states such as spin-glasses....and these theories yield general predictive properties of highly interconnected systems....

From Merriam-Webster...

## com•plex /käm'pleks/

(1) ....consisting of many parts, e.g. a complex piece of machinery...
(2) ...consisting of many interconnected or interwoven parts....

We also have theories for condensed phases of matter such as liquids, solids, and even disordered states such as spin-glasses....and scientists are discovering new biological principles through applying such theories.

The essence of this simplicity is homogeneity in the pattern of interactions....can take averages.

From Merriam-Webster...

## com•plex /kảm'pleks/

(1) ....consisting of many parts, e.g. a complex piece of machinery...
(2) ...consisting of many interconnected or interwoven parts....
(3) ...consisting of parts interconnected so as to make the whole perplexing...

So...what is the essence of it?

Well....complex systems show
heterogeneity of system components such that some parts and connections are much more important than others (can't take averages!), and....
non-linearity, such that the combined activity of components cannot be predicted from properties of the components taken individually. Nonindependence of parts and reactions. The whole is NOT a sum of the parts!!

Well....complex systems show
heterogeneity of system components such that some parts and connections are much more important than others (can't take averages!), and....
non-linearity, such that the combined activity of components cannot be predicted from properties of the components taken individually. Nonindependence of parts and reactions. The whole is NOT a sum of the parts!!

2/13 Small non-linear systems...the MAP kinase switch
2/15 Mesoscale non-linear systems...the problem of cellular signaling
2/20 Large non-linear systems...the problem of proteins
2/22 Epistasis...genetic and evolutionary principles, fitness landscapes

Well....complex systems show
heterogeneity of system components such that some parts and connections are much more important than others (can't take averages!), and....
non-linearity, such that the combined activity of components cannot be predicted from properties of the components taken individually. Nonindependence of parts and reactions.

And, of course....the complexity scales steeply with the number of variables in a system that are engaged non-linearly.

Ok....we need a graphical view of everything to organize our thinking.

Ok....we need a graphical view of everything to organize our thinking.


1/11 Linear systems theory I...simple systems, graphical tools, decomposability
1/16 Linear systems theory II...the space of possible solutions, understandability
1/18 Stochastic models...models at the limit of small numbers, new behaviors.


1/23 Diffusion and driving forces, thermodynamic analysis
1/25 Diffraction theory...the spatial Fourier transform
1/30 Decomposition...low dimensional representations of big data



2/15 Mesoscale non-linear systems...the problem of cellular signaling 2/20 Large non-linear systems...the problem of proteins
2/22 Epistasis...genetic and evolutionary principles, fitness landscapes

|  | $\mathrm{n}=1$ | $\mathrm{n}=2$ or 3 | n >> 1 | continuum |
| :---: | :---: | :---: | :---: | :---: |
| Linear | exponential growth and decay | second order reaction kinetics | electrical circuits | Diffusion |
|  |  |  | molecular dynamics | Wave propagation |
|  | single step conformational change | linear harmonic oscillators |  |  |
|  |  |  | systems of coupled harmonic oscillators | quantum mechanics |
|  | fluorescence | simple feedbackcontrol |  | viscoelastic systems |
|  |  |  | equilibrium thermodynamics |  |
|  | pseudo first order kinetics | sequences of conformational change |  |  |
|  |  |  | diffraction, Fouriertransforms |  |
|  |  |  |  |  |
| Nonlinear | fixed points <br> bifurcations, multi stability | anharmomic oscillators | systems of nonlinear oscillators | Nonlinear wave propagation |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  | relaxationoscillations | non-equilibrium thermodynamics | Reaction-diffusion in dissipative systems |
|  | irreversible hysteresis <br> overdamped oscillators |  |  |  |
|  |  | predator-prey models | protein structure/ function |  |
|  |  |  |  | Turbulent/chaotic flows |
|  |  | van der Pol systems <br> Chaotic systems | neural networks |  |
|  |  |  |  |  |
|  |  |  | the cell |  |
|  |  |  |  |  |
|  |  |  | ecosystems |  |

So...how do we proceed in (quantitatively) understanding systems in all these different regimes?


So...how do we proceed in (quantitatively) understanding systems in all these different regimes?


Why make models and what is the process?
(1) data reduction... that is, representing 1000's of data points (say a Cart trace) in only a fer charactonsthe parameters.


Why make models and what is the process?
(i) data reduction... that is, repiesentining 1000's of data points (say a Cat trace) in only a few charactonste. parameters.

(2) test a physical model for an observed phenomenon. I collet data on the folding of a protein. Are the dater consistent with a two-state equilibnan?


Why make models and what is the process?
(3)

Obtain a mathematical representation
that helps predict now experimental results and theufere becomes a tool for guiding experiments. Intuition in "lang" proterns.
(4) Identity and classify a bidogial proaces into sets of common physical mechanisms. Diffusion?
sigmund ampliter? Aswihh?

Why make models and what is the process?

The fire step process of modeling:
(1) observe a biological process [collect the dato]
(2) construe a physical model $[A \xrightarrow{k} B]$
(3) abstract it to a mallematical $\quad\left[A(t)=A_{0} e^{-k t}\right]$
model.
(4) fit the dater, check accuncy of fitting. Does the model explain the data?
(5) Fine... but is it predictive?

A simple example for today:


A simple example for today:


So, we are studying the binding of a toxin molecule that binds to and blocks an ion channel. We want to understand the nature of this reaction.

A simple example for today:


So, we are studying the binding of a toxin molecule that binds to and blocks an ion channel. We want to understand the nature of this reaction.

The experiment....express channels in a Xenopus oocyte, record the channel activity, flow in and wash out the toxin molecules, watch the inhibition of current flow through the channel, record the kinetics.



Next step is to make a physical model....so we guess that this is a simple bimolecular reaction. $L$ is for the toxin and $R$ is for the channel...

$$
L+R \underset{k_{0} f f}{\stackrel{k_{0 n}}{\rightleftharpoons}} L R
$$

Now to turn this into a mathematical model....we write down the chemical kinetics equation and set the initial conditions...


So then...

$$
\frac{d R}{d t}=k_{[ }[L][R]
$$

; Simple bimolecular usecciation

$$
\text { At to, } \begin{aligned}
\Delta L & =L_{0} \\
R & =R_{0} \\
L R & =0
\end{aligned}
$$

Alva, $L_{0} \gg R_{0}$

Now to turn this into a mathematical model....we write down the chemical kinetics equation and set the initial conditions...
; Simple bimolecuber ussociatorn


$$
\begin{aligned}
A+t=0, \quad L & =L_{0} \\
R & =R_{0} \\
L R & =0
\end{aligned}
$$

Also, $L_{0} \gg R_{0}$

So then...

$$
\left.\frac{d R}{d t}=-k[L][R] \text {; But } L L\right] \text { never changes! }
$$

So ...

$$
\begin{aligned}
& \frac{d R}{d t}=-k_{1} L_{0}[R] \quad \text { or... } \\
& \frac{d R}{d t}=-k_{\text {app }}[R] \quad \text { when } k_{\text {app }}=k_{1} L_{0}
\end{aligned}
$$

Now to turn this into a mathematical model....we write down the chemical kinetics equation and set the initial conditions...

; Simple bimolecular association

$$
\text { At to, } \begin{aligned}
A L & =L_{0} \\
R & =R_{0} \\
L R & =0
\end{aligned}
$$

Alva, $L_{0} \gg R_{0}$

So then...

Solution... given initial conditions:

$$
R(t)=R_{0} e^{-k_{n+1} t}
$$

The binding of toxin to the channel should follow single exponential kinetics...

What about the back reaction?

$$
\begin{aligned}
& {[L R] \stackrel{k_{01}}{\Longrightarrow}[L .]+[R]} \\
& \frac{d[L R]}{d t}=-k_{-1}[L R] \\
& L R(t)=L R_{0} e^{-k_{01} t} \Rightarrow
\end{aligned}
$$

; Simple bimolecubor ussociation

$$
\begin{aligned}
A+t=0, \quad L & =L_{0} \\
R & =R_{0} \\
L R & =0
\end{aligned}
$$

Also, $L_{0} \gg R_{0}$

The dissociation of toxin from the channel should also follow single exponential kinetics...
$R \underset{k_{1-1}}{\stackrel{k+L}{\rightleftarrows}} L R$

$$
\begin{aligned}
& R(t)=R_{0} e^{-k_{0 \pi} t} \text { where } k_{0 \text { att }} k_{1} t_{0} \\
& L R(t)=L R_{0} e^{-k_{01} t}
\end{aligned}
$$






Indeed, both toxin binding and release are well-fit by singe exponential functions


$$
R(t)=R_{0} e^{-k_{a n t} t} \text { where } k_{\text {app }}=k_{1} t_{0}
$$

$$
L R(t)=L R_{0} e^{-k_{-1} t}
$$

So, the model fits the data well. Good, but does it make new predictions not used in arriving at the model? Yes....


$$
R(t)=R_{0} e^{-k_{\text {ant }} t} \text { where } k_{\text {app }}=k_{1} t_{0}
$$

$$
L R(t)=L R_{0} e^{-k_{-1} t}
$$

So, the model fits the data well. Good, but does it make new predictions not used in arriving at the model? Yes....

How should the on-rate and off-rate of toxin binding to the channel depend on the concentration of toxin?
$R \underset{k_{1}}{\stackrel{k+L}{\rightleftarrows}} L R$

$$
R(t)=R_{0} e^{-k_{a n t} t} \text { where } k_{\text {app }}=k_{1} t_{0}
$$

$$
L R(t)=L R_{0} e^{-k_{-1} t}
$$

So, the model fits the data well. Good, but does it make new predictions not used in arriving at the model? Yes....


Ok...and finally at equilibrium:

$$
R \underset{k_{-1}}{\stackrel{k_{1} L}{\rightleftarrows}} L R \quad \begin{aligned}
& \frac{d(L R]}{d t}=k_{1} L[R] \\
& \frac{d R}{d t}=k_{-1}[L R]
\end{aligned}
$$

At equilibrium...

$$
\begin{array}{ll}
k_{1} L[R]=k_{-1}[L R], & \text { bat } R+L R=R_{\text {tot }} \\
R_{1} L\left[R_{b+1}-[L R]\right]=k_{-1}[L R] \\
f_{B}= & \frac{L R]}{R_{\text {dot }}}=\frac{L}{\frac{k_{-1}}{R_{1}}+L} \quad k_{d}=\frac{k_{-1}}{k_{1}}
\end{array}
$$

Ok...and finally at equilibrium:



So, the good-old rectangular hyperbola, or "binding isotherm".

Ok...and finally at equilibrium:



So, the good-old rectangular hyperbola, or "binding isotherm". This gives us an additional check on our model.....we should be able to get the dissociation constant (Kd) in two independent ways:
(1) Look at fraction bound as a function of toxin (L) concentration...
(2) Take the ratio of the off-rate and on-rate

These better give us the same number!!


TABLE 1 Blocking parameters for wild-type and mutant
charybodotoxins

| Toxin | $K_{i}(\mathrm{nM})$ | $k_{\mathrm{on}} \times 10^{-6}$ <br> $\left(\mathrm{M}^{-1} \mathrm{~s}^{-1}\right)$ | $k_{\mathrm{off}}\left(\mathrm{s}^{-1}\right)$ |
| :--- | :---: | :---: | :---: |
| Multi-pulse protocol <br> Wild-type | $0.075 \pm 0.005$ | $63 \pm 5$ | $0.0047 \pm 0.0004$ |
| S10Q | $0.69 \pm 0.05$ | $149 \pm 18$ | $0.085 \pm 0.006$ |
| R25Q |  |  |  |
| $\quad 2 \mathrm{~K}^{+}$(out) | $0.95 \pm 0.05$ | $23 \pm 2$ | $0.021 \pm 0.007$ |
|  |  |  |  |

Well....the concentration of toxin at which we get half block is 0.075 nM . And if you compute the ratio of the off and on rates, you get 0.0746 nM .

Pretty good....makes one want to believe the model.

Now just to review....

$$
U \xrightarrow{k} U^{*}
$$

we say: $\frac{d U}{d t}=-k U$; bat can we delve this rale equation?

Now just to review....

$$
U \xrightarrow{k} U^{*}
$$

we say: $\frac{d u}{d t}=-k U$; but can we derive this rate equation?

So at $t_{c} 0$, we have $U_{0}$ molecules. After time $\Delta t$, how many molecince of $U^{*}$ will I have? Well... k $\Delta t U_{0}$

So...

$$
\begin{aligned}
U(t+\Delta t) & =U(t)[1-k \Delta t] \\
& =U(t)-U(t) k \Delta t \\
\frac{U(t+\Delta t)-U(t)}{\Delta t} & =-U(t) k
\end{aligned}
$$

What happens as $\Delta t \rightarrow 0$ ?

Now just to review....
$U \xrightarrow{k} U^{*}$
we say: $\frac{d u}{d t}=-k U$; but can we derive this rate equation?

So at $t=0$, we have $U_{0}$ molecules. After time $\Delta t$, how many molecince of $U^{*}$ will I have? Well... $k \Delta t U_{0}$

So...

$$
\begin{aligned}
U(t+\Delta t) & =U(t)[1-k \Delta t] \\
& =U(t)-U(t) k \Delta t \\
\frac{U(t+\Delta t)-U(t)}{\Delta t} & =-U(t) k \\
\frac{d U}{d t} & =-U k
\end{aligned}
$$

What happens as $\Delta t \rightarrow 0$ ?

How do we solve this equation?

Solution of a forst orden differatial equation:
the equ arly contums a first der ivative.

$$
\frac{d v}{d t}=-k v
$$

Solve by sepantion of variableo:

How do we solve this equation?

$$
\begin{aligned}
\frac{d v}{d t} & =-k u \\
\frac{1}{v} d u & =-k d t \\
\int_{v_{0}}^{u} \frac{1}{v} d u & =\int_{0}^{t}-k d t \\
\left.\ln v\right|_{v_{0}} ^{u} & =-\left.k t\right|_{0} ^{t} \\
\ln u-\ln v_{0} & =-k t \\
\ln \left[\frac{u}{v_{0}}\right] & =-k t \\
\frac{u}{v_{0}} & =e^{-k t}
\end{aligned}
$$

So...the single exponential decay function.

So, we dealt with the easiest problem...the top left hand corner of the space of problems....


So, we dealt with the easiest problem...the top left hand corner of the space of problems....


## Small systems

## Large systems



So...how do we proceed in (quantitatively) understanding systems in all these different regimes?


So...how do we proceed in (quantitatively) understanding systems in all these different regimes?


There is a deep hypothesis here that even in systems that seem to have large numbers of variables, the relevant dynamics is ultimately low-dimensional...comprised of just a few "effective variables". The trick for such systems is to find them...

So...how do we proceed in (quantitatively) understanding systems in all these different regimes?


2/20 Large non-linear systems...the problem of proteins
2/22 Epistasis...genetic and evolutionary principles, fitness landscapes

Not everyone in the field of systems biology takes this perspective!


For some, "understanding" means a detailed exposition of every component, reaction, and features of a particular instance of a system.

Consider the status of "understanding" the relationship of heat and work in the early 1900's....


The general approach was a detailed modeling of the Newtonian mechanics of each sort of heat engine....but then....

"..the phenomenon of the production of motion by heat has not been considered from a sufficiently general point of view. We have considered it only in machines the nature...of which have not allowed us to take in the whole extent of application...In such machines, the phenomenon is, in a way, incomplete. It becomes difficult to recognize its principles and study its laws..."
"In order to consider in the most general way the principle of the production of motion by heat, it must be considered independently of any mechanism or any particular agent. It is necessary to establish principles applicable....to all imaginable heat engines, whatever the working substance and whatever the method by which it is operated..."

"..the phenomenon of the production of motion by heat has not been considered from a sufficiently general point of view. We have considered it only in machines the nature...of which have not allowed us to take in the whole extent of application...In such machines, the phenomenon is, in a way, incomplete. It becomes difficult to recognize its principles and study its laws..."
"In order to consider in the most general way the principle of the production of motion by heat, it must be considered independently of any mechanism or any particular agent. It is necessary to establish principles applicable....to all imaginable heat engines, whatever the working substance and whatever the method by which it is operated..."
$x=1$
$x+1$
$=1$
$\frac{1}{2}=$ U， FITM，


$$
\frac{1}{x+50}
$$

$$
\begin{aligned}
& 10 \\
& =15=
\end{aligned}
$$




$==$
$=5$
$==$
$\frac{2}{2}=$
$\frac{2}{2}$
$=15=$
"
 $=x+4+\pi+\infty$
$=4=\frac{4}{2}$


