The Logics of Reliable Inquiry

Clark Glymour
Carnegie Mellon University

3 Parts

1. What is wrong with Bayesian epistemology?
2. Scientific illustrations of reliable learning—
   Conservation laws and elementary particles
   Learning chaos
   Cognitive neuropsychology
3. Introduction to a reliabilist framework for causal inference (to be continued=

Why I Am Not A Bayesian

- If you are computing probabilities
  - Its good to abide by the probability axioms
  - “Bayes theorem” is sometimes a good way to compute conditional probabilities—when a hypothesis specifies probabilities, or when there are frequencies that can be used to estimate probabilities.
- But the aim of inquiry is:
  - To discover new, interesting truths
  - Or, accurately to forecasts aspects of the future
  - Or, to explain the past

And, for that, probability is a valuable, ancillary tool, but Bayesian epistemology is not an aid and is sometimes a hindrance.

What Is Bayesian Epistemology About?

1. An historical account of the methods used in scientific discovery?
2. A psychological account of how people, and scientists in particular, reason—consciously or unconsciously?
3. A tool for inquiry?
4. A ‘reconstruction’?

Answer: Not 1, 2 or 3, and often not insightfully, 4.

An Historical Account?

Scientists who did not do Bayesian calculations to “confirm” hypotheses:

- Copernicus
- Kepler
- Newton
- Dalton
- Darwin
- Helmholtz
- Faraday
- Hertz
- Einstein

In fact, no scientist did until after 1970—before then, no conditional probability calculations were possible except for trivial problems.

A Psychological Account?

- There is no psychological evidence that individuals reason by Bayesian calculations except in trivial problems.
- Psychological experiments on individual judgement show
  - People use a variety of elementary heuristics in reasoning (more later if you want it)
  - Variety of individual responses in experiments is wildly inconsistent with Bayesian explanations of average behavior (more later if you want it)
  - Claims that people are “approximately Bayesian” are phony without a testable theory of how and in what sense they approximate (there is no such theory).

A Proposal for Practical Methodology?

Contemporary science (in genetics, astronomy, medical science, etc.) often involves terrabytes of data and superastronomical numbers of alternative hypotheses.

Bayesian methodology requires:

Keeping track of the posterior probabilities of each hypothesis as data accumulates.

Not computationally possible.

Instead, “Bayesians” must resort either to

- dogmatism (exclude almost all hypotheses) or
- heuristics (give up claim to reliability) or
- equivocation (do something and if it involves calculating conditional probabilities, call it “Bayesian.”
  - Some of the last is very good methodology.

Bayesian Reconstructions

- Logical and semantical relations determine entailment and synonymy relations among propositions
- Bayesian epistemology assigns numbers between 0 and 1 to the propositions.
- The numerical values respect the entailment relations– if p entails q, then Pr(p) is not greater than Pr(q)

The semantical relations do the most of the work—the numbers tag along

Example—

One form of scientific explanation of an empirical phenomenon is showing that it is a manifestation of another relationship that is already, independently, known.

Examples

Phenomenon

Solar years = cycles of anomaly + revolutions of longitude

becomes on heliocentric theory an instance of

Number of orbits of inner, faster body = number of overtakings of outer by inner + number of orbits of outer.

Phenomenon

Elementary substances combine in definite proportions by weight

becomes on Daltonian theory an instance of

Mass is additive: the mass of a body is the sum of the masses of its components.

These explanations have a structure

- *Theory* identifies quantities in the phenomenon with theoretical quantities or functions of theoretical quanitities—or with other measured quantities or functions of them.
- With these substitutions, the relation of the phenomenon becomes an instance of an independently known phenomenon—which may sometimes be a mathematical truth.
- Bayesian numerical assignments do not reveal the structure of the explanations.

Identifiability: another semantic relation in theory preference

Preference against theories that contain quantities that cannot be estimated from data using the theory + independently known constraints.

Dominant preference in econometrics through the 1980s (see Franklin Fisher, *The Identification Problem in Econometrics*)

Continuing objection to the chemical atomic theory from Wollaston (1810) to Ostwald (1910).

Bayesian epistemological reconstruction merely assigns numbers (e.g., prior degrees of belief) in accordance with the preference—*if one knew only the numbers, one would miss the point.*

Bayesian accounting

- Popper says:conjecture the *boldest* theory
  - Bayesians can adjust probabilities to agree with this, but that is no explanation of why and when it might or might not be a good thing to do.
- Lots of people *say* conjecture the *simplest* theory
  - Same point.

Bayesian Epistemology is Restrictive

There are simple idealized, reliably solvable learning problems that *Bayesian learners cannot reliably solve.*

Example–

***Background*--two possible worlds, one the nonnegative integers, the other the nonpositive integers**

In one, 0 is a first element. In the other, 0 is a last element.

World 1: 0,1,2,3.....

World 2; 3, 2,1,0

***Hypotheses*– there is a first element; there is a last element**

Given the background, these hypotheses are mutually exclusive.

***Data* consist of values of instances of**

Rx,z meaning *x is less than z*.

N. B.Which numbers x, z, etc. denote is hidden (otherwise the problem would be trivial).

Data from a world are presented in any order. In a data sequence from a world, every Rx,z fact about the world eventually comes to be known.

Discovery problem—for every admissable sequence of data from world 1, converge to the hypothesis that there is a first element, and

for every admissable sequence of data from world 2, converge to the hypothesis that there is a second element.

The problem is easily satisfied by a computable learner. (Choose a witness object S of world 1 such that no datum so far says any object is less than S, and a wtness object G of world 2 such that no datum so far says that G is less than any object. Conjecture the opposite of the world whose witness object has most recently changed.)

Bayesian Learners

Bayesian learner puts a prior over the two hypotheses, and conditional on each hypothesis a likelihood for any finite data sequence.

Pr(world 1), Pr(data | world 1), etc.

Bayesian learner conjectures the hypothesis (world 1 or world 2) with higher conditional probability given the data so far.

No such Bayesian learning solves the problem reliably.

But there is an obvious, computable learner that does so.

Computable Bayesians

- Must be infinitely and uncomputably dogmatic—simply an application of Turing's Theorem.
- Are not logically omniscient. So how is the "confirmation" of a new theory by old, already believed evidence it entails to be accomodated?

$Pr(\text{Old evidence}) \sim 1;$

$Pr(\text{Old evidence} \mid T) \sim 1,$

$$Pr(T \mid \text{old evidence}) = \frac{Pr(T)\, Pr(\text{Old evidence} \mid T)}{Pr(\text{Old evidence})}$$

$\sim Pr(T)$

Generalized Framework

- Introduce background assumptions
  - Exclude some possible worlds
- Various critteria of success
  - Verify, Refute, Decide
  - By a deadline, with certainty, in the limit, gradually

Generalized Framework

- Important criteria for this problem
  - Verification (Refutation) in the limit: stabilize to “true” (“false”) iff the hypothesis is true (false)
    - Similar to Bayesian convergence given appropriate prior probabilities
  - Gradual verification (refutation): Real-valued conjectures converge to 1.0 (0.0) iff the hypothesis is true (false)

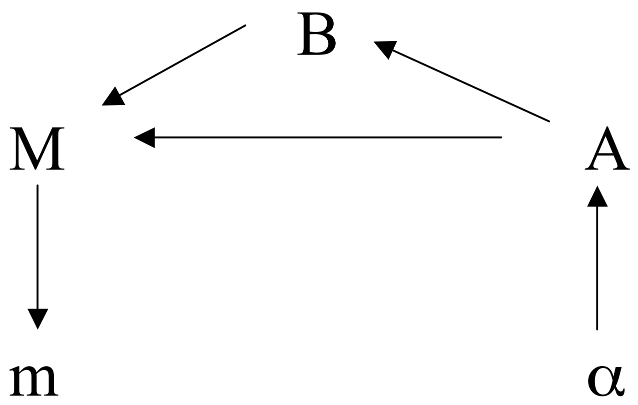
Reliabilism in Practice

1. Methodological disputes in classical neuropsychology resolved by a Popperian rule that is reliable.
2. We can learn whether a system is chaotic provided it is one-dimensional, deterministic, and our measurements are free of error.
3. There is a most efficient learning strategy for conservation laws in particle physics—and it sometimes requires the postulation of unobserved fundamental particles.
4. There are reliable, feasible strategies for obtaining causal relations from passive observations—and, with strengthened assumptions, for discovering the existence and mutual influences of unobserved causes.

Reliabilism in Practice: Bayesian epistemology is gratuitous

- A simple example: Classical Cognitive Neuropsychology; Try to infer features of cognitive organization from combinations of disabilities of brain damaged persons.
- Apasias, Agnosias
- Broca, Meynert, Wernicke, Freud, Lichtheim, etc.
- Continues to this day (Farah, Shallice, Caramazza).

Lichtheim's Diagram for Aphasia



Freud (1891) to Farah (1990)

Same theory (connectionist)

Same kind of data (individual deficit profiles, some group data)

Same kind of argument (by example and non-example)

Some methodological questions

- What can be learned from individual data?
- Group data versus individual data
- Can neural net models explain every conceivable pattern of deficits?

Practice and philosophical slogans about scientific inference disagree. For example:

Popper: Make bold conjectures and try to falsify them.

Neuropsychological practice: If a model predicts the existence of a deficit combination not yet found, count it against the model.

How should we think about inference?

1. What is the goal? (find out something about the normal architecture, M)

2. What are the background assumptions?

3. What methods of data acquisition and analysis will *reliably* reach the goal, given the background assumptions?

Themes from Computational Learning Theory

* Set of hypotheses that will be considered

*Each hypothesis determines a collection of data sequences

*No two hypotheses determine the same data sequence

*Inference procedures are functions from initial segments of data sequences to hypotheses

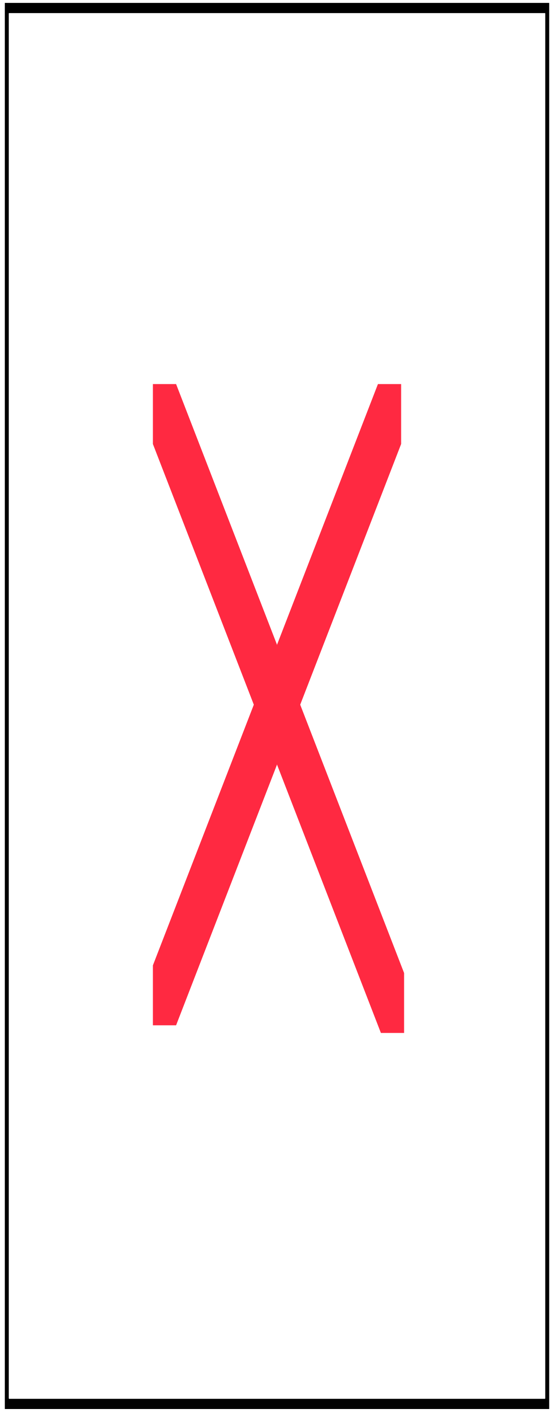
*Procedure is reliable for a data sequence if it makes at most a finite number of erroneous inferences

*Procedure is reliable for a hypothesis if it is reliable for each of the data sequences for the hypothesis.

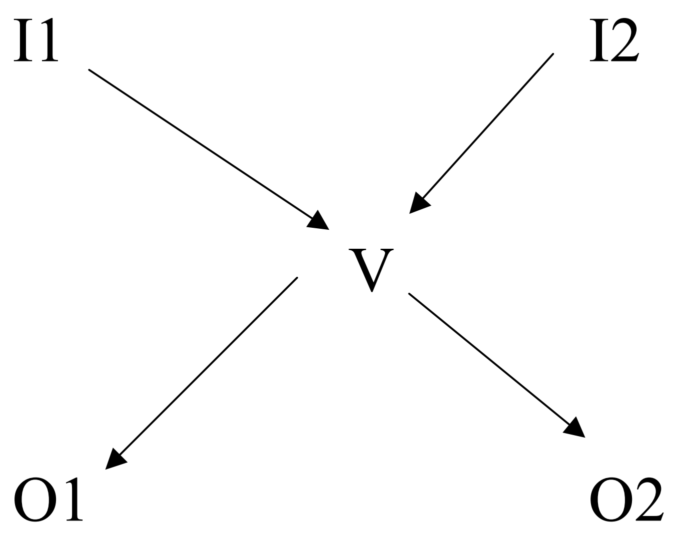
*Procedure is reliable for the problem if it is reliable for each hypothesis.

Example problem, loosely based on M. Farah, ***Visual Agnosia*, 1990**

Six alternative models, represented graphically:



Representation



I = no input

I = normal input

O = no output

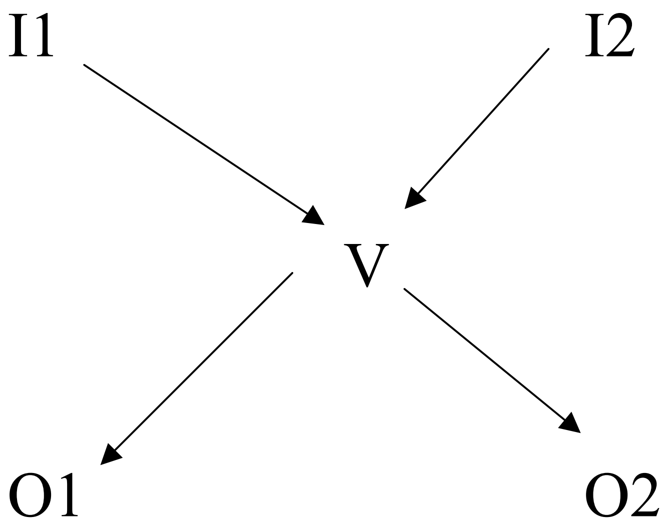
O = normal response

O = abnormal response

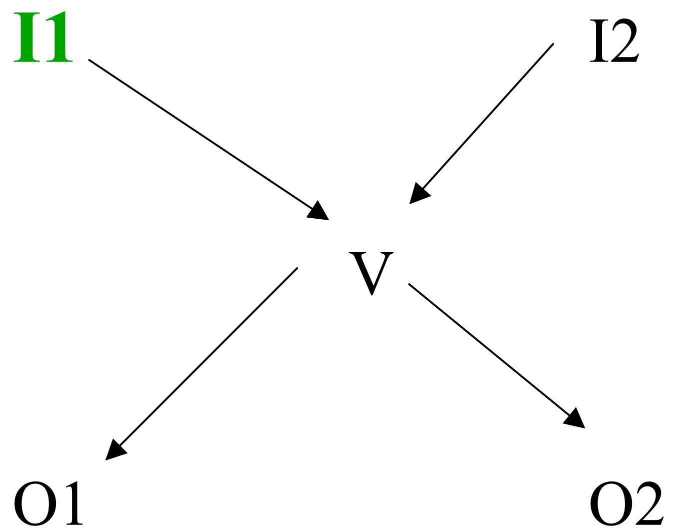
Big Assumption

If any pathway from an input to an output is damaged, an abnormal response is given to the input.

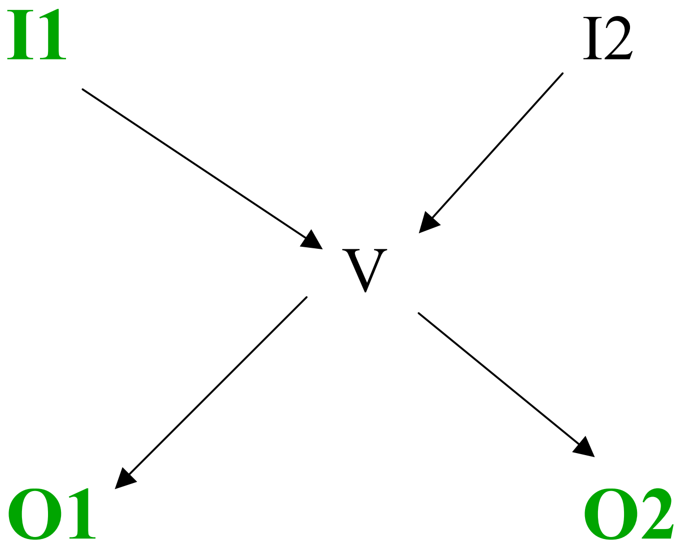
Normal Functioning



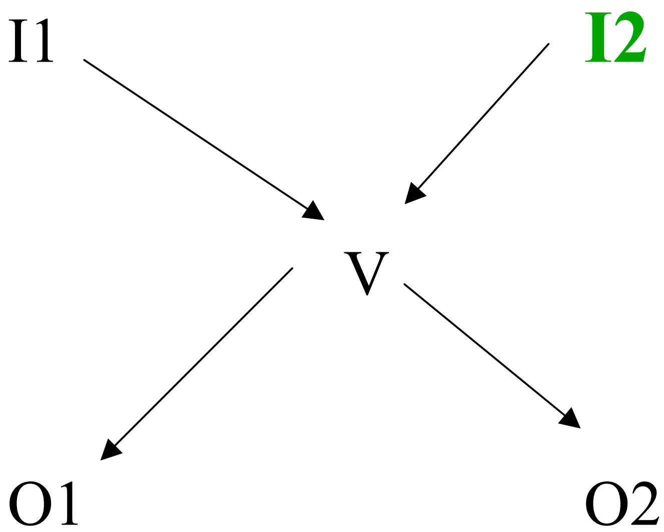
Normal Functioning



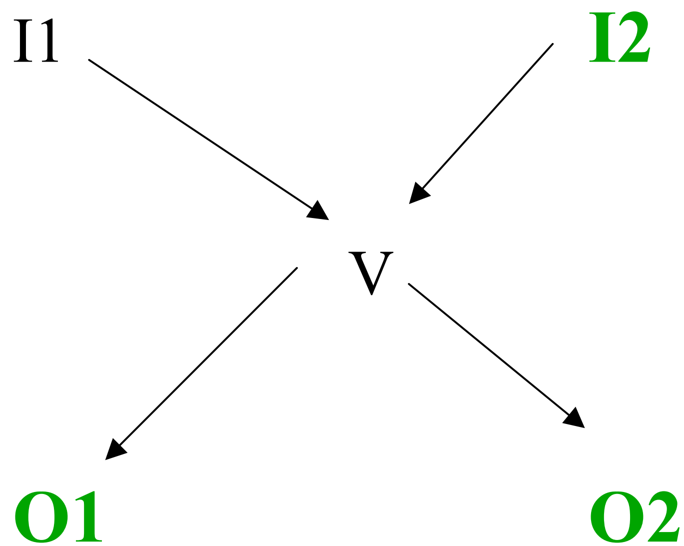
Normal Functioning



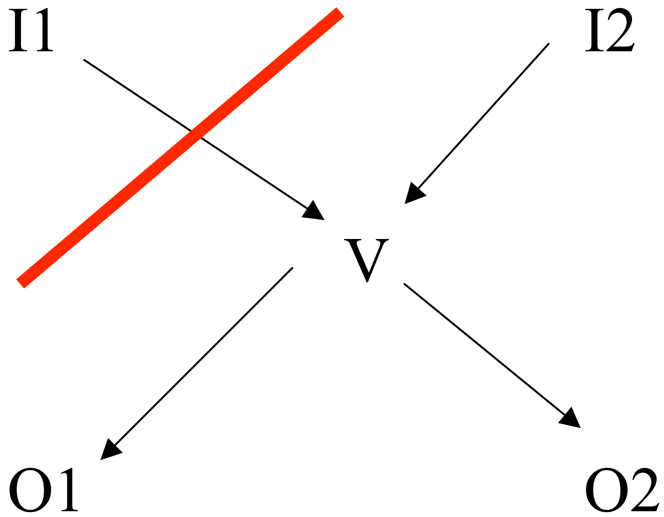
Normal Functioning



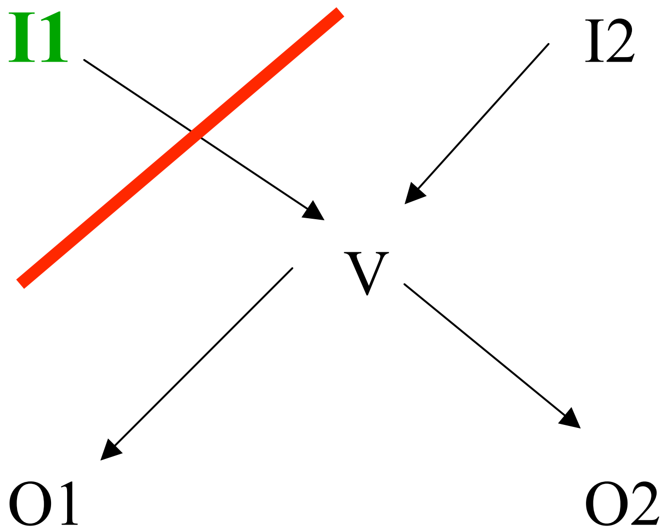
Normal Functioning



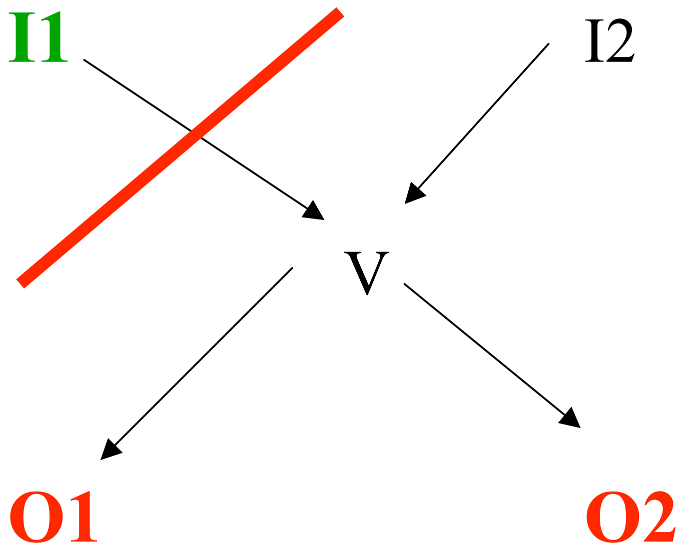
Abnormal Functioning



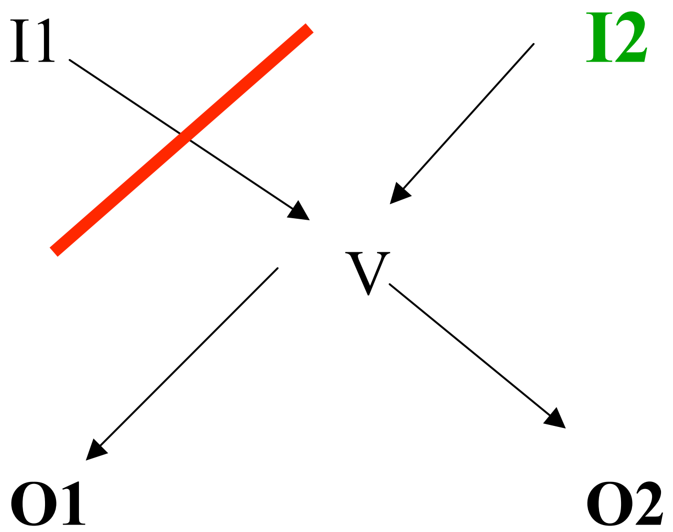
Abnormal Functioning



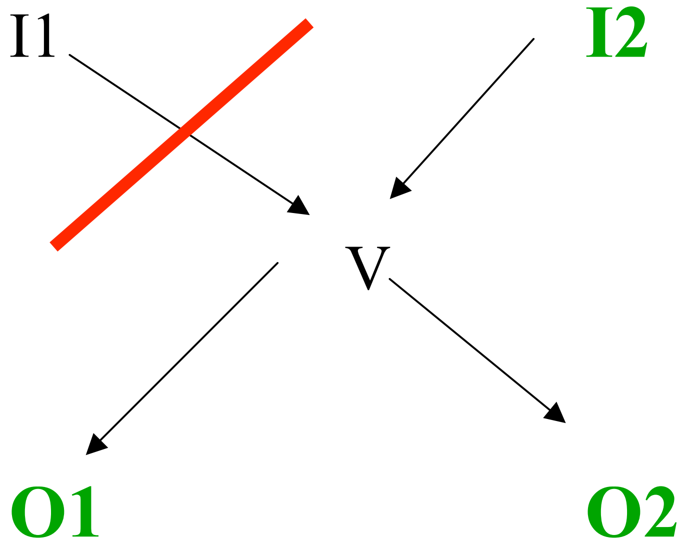
Abnormal Functioning



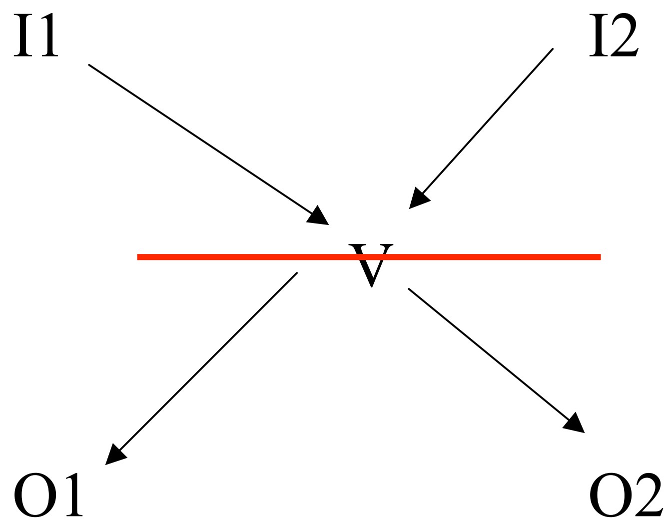
Abnormal Functioning



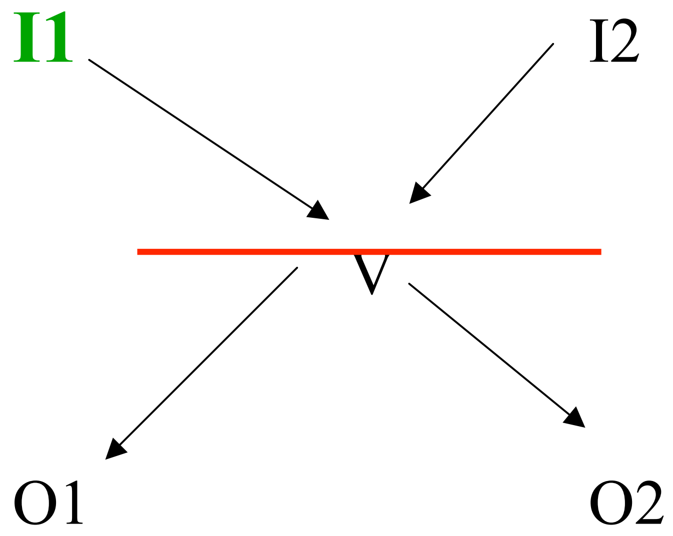
Abnormal Functioning



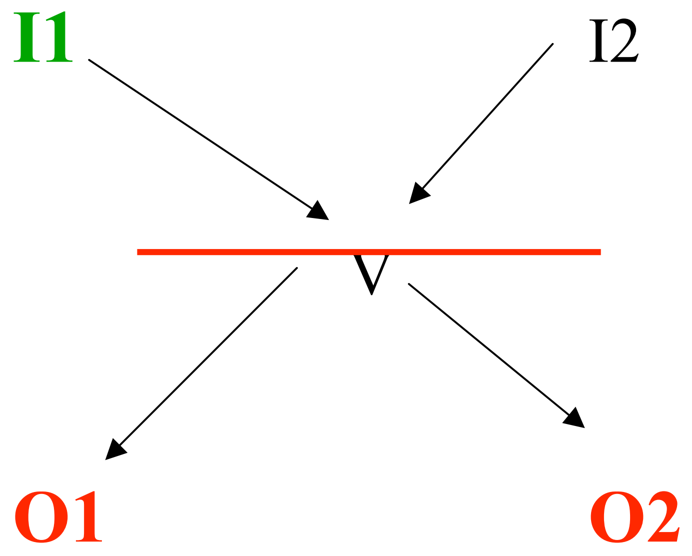
Abnormal Functioning



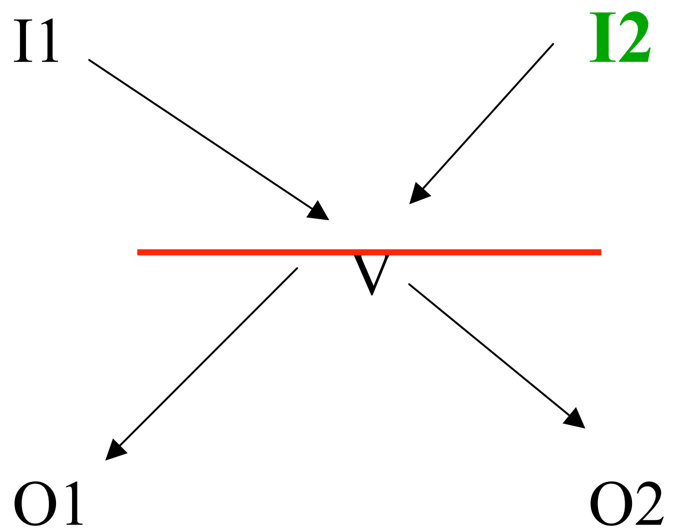
Abnormal Functioning



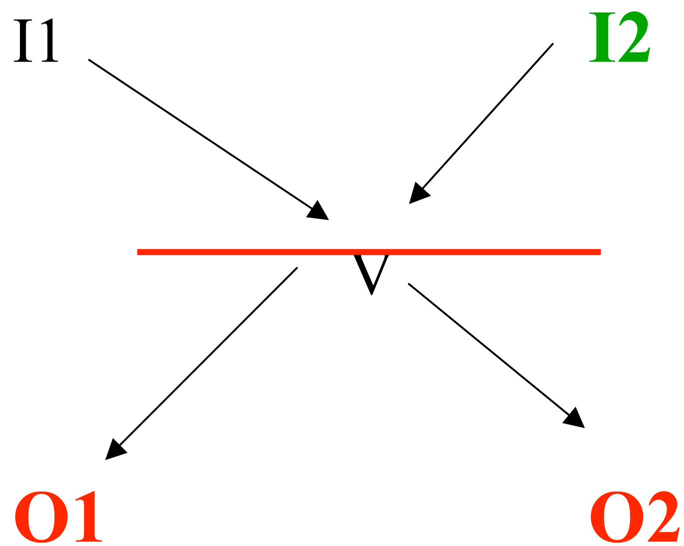
Abnormal Functioning



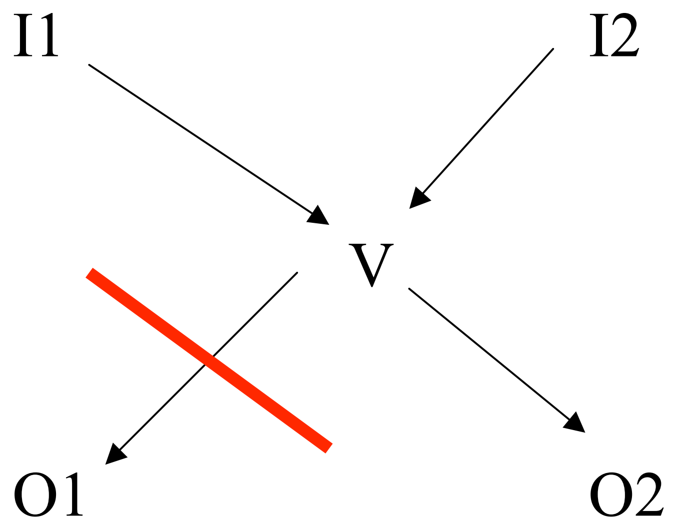
Abnormal Functioning



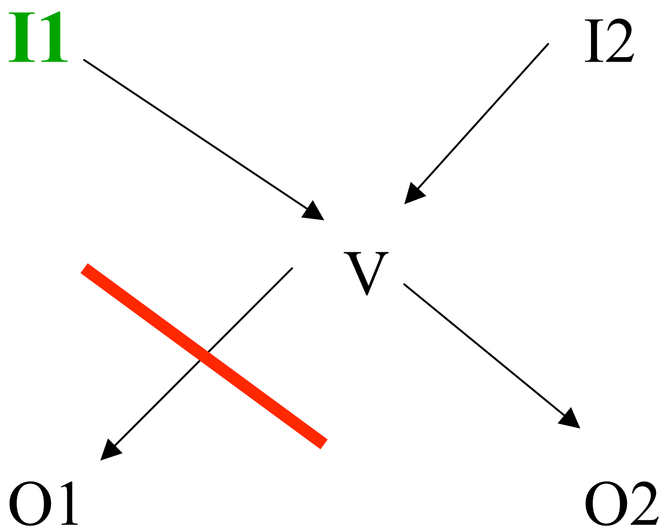
Abnormal Functioning



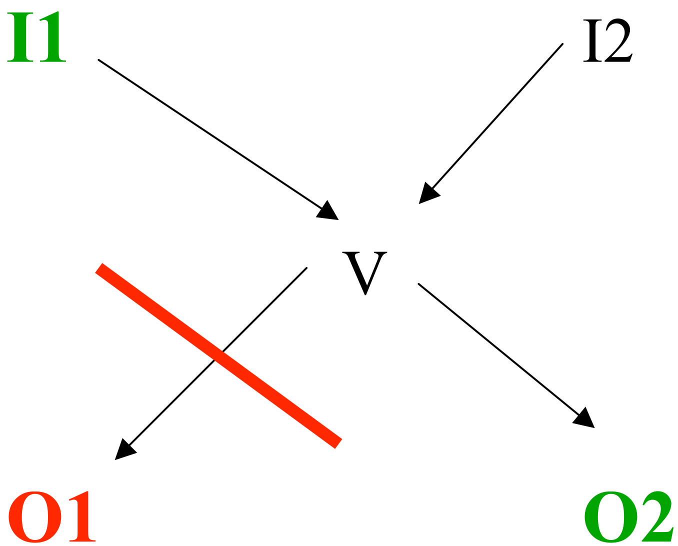
Abnormal Functioning



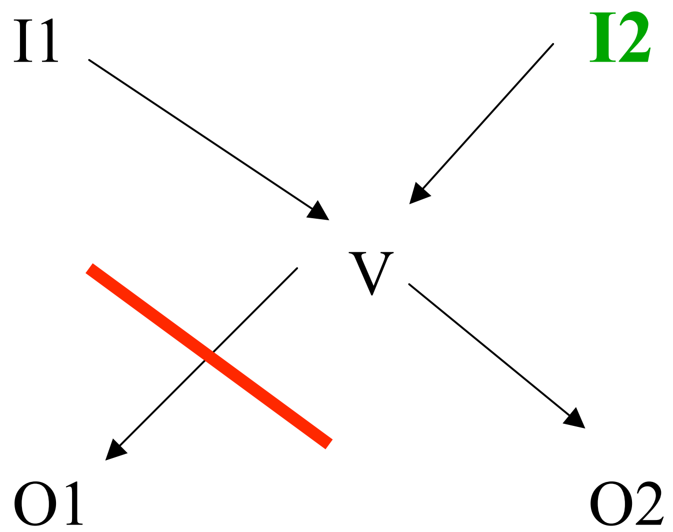
Abnormal Functioning



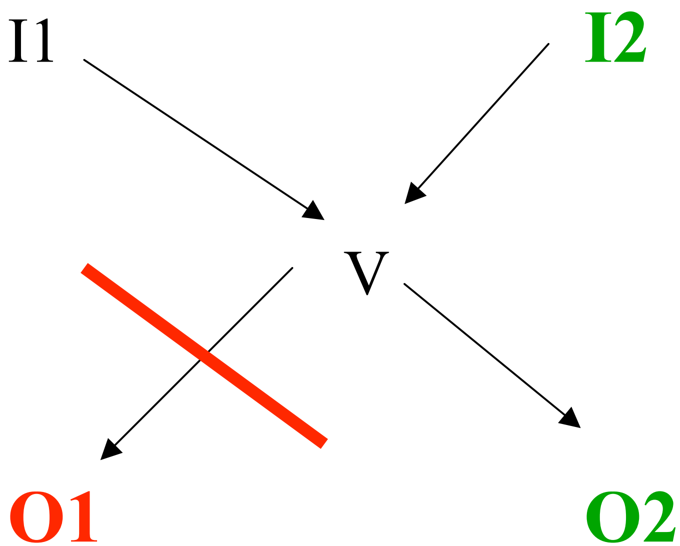
Abnormal Functioning



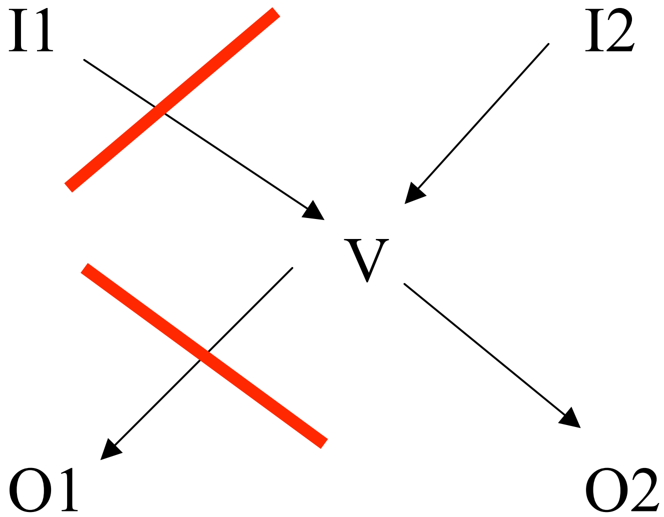
Abnormal Functioning



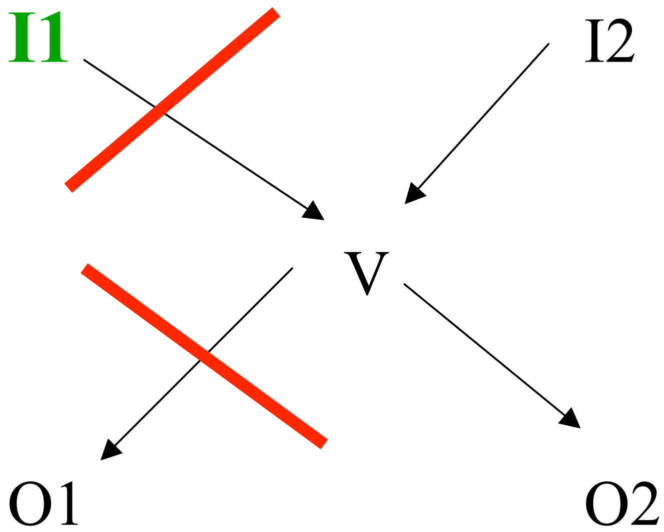
Abnormal Functioning



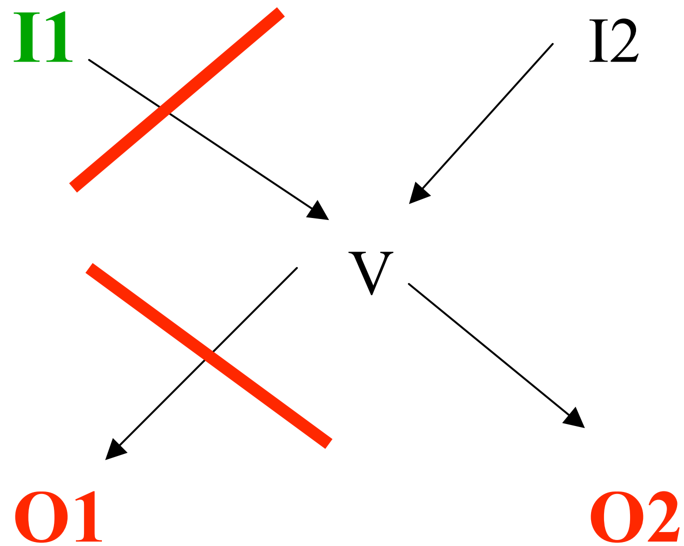
Abnormal Functioning



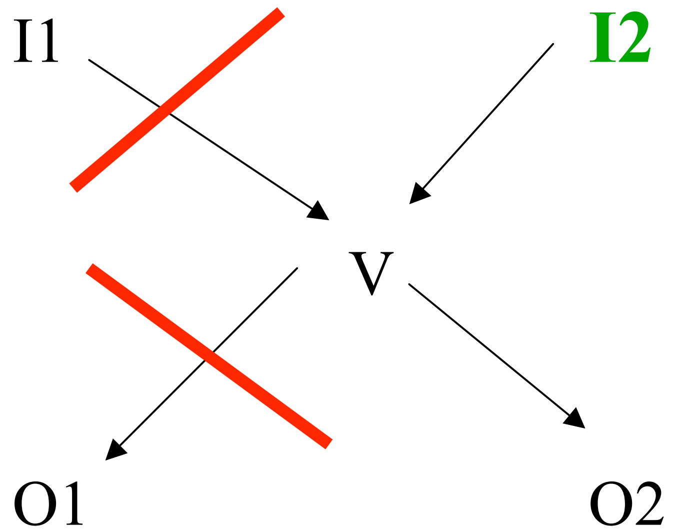
Abnormal Functioning



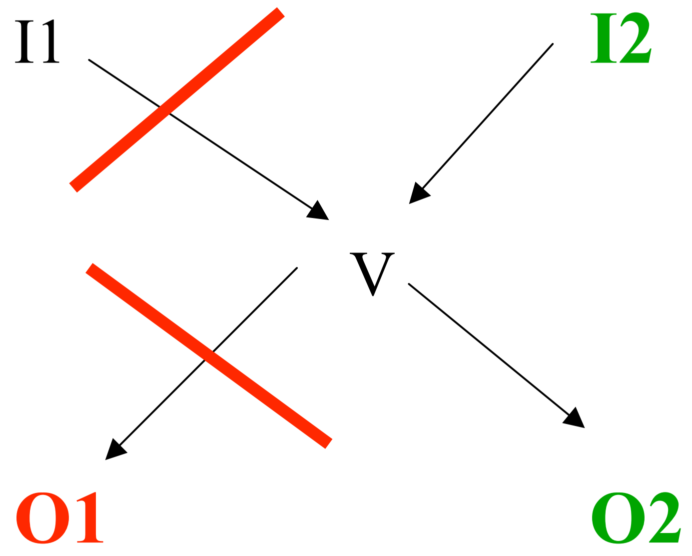
Abnormal Functioning



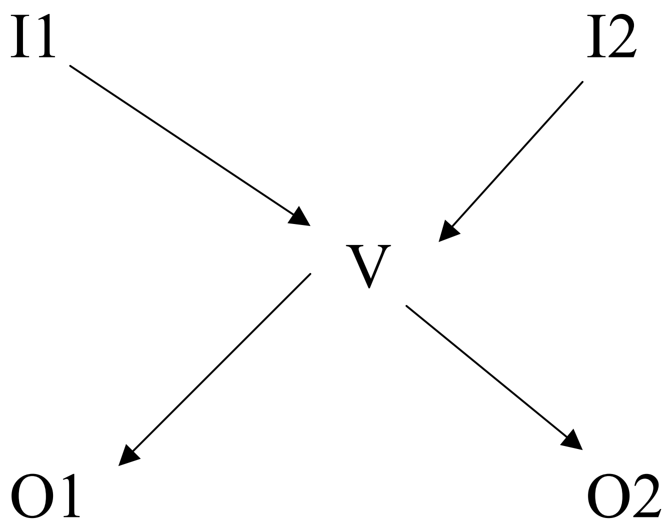
Abnormal Functioning



Abnormal Functioning



Possible Deficit Combinations from Normal Structure



{I1, O1,O2; I2, O1,O2} {I1, O1,O2; I2, O1,O2}

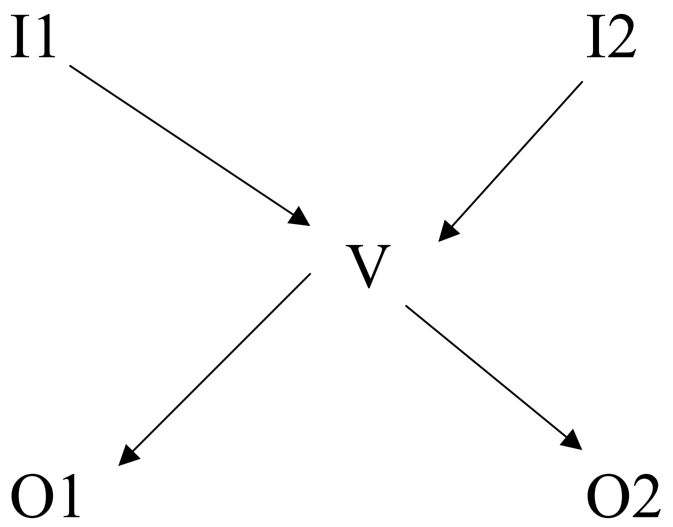
{I1, O1,O2; I2, O1,O2}

{I1, O1,O2; I2, O1,O2} {I1, O1,O2; I2, O1,O2}

{I1, O1,O2; I2, O1,O2} {I1, O1,O2; I2, O1,O2}

{I1, O1,O2; I2, O1,O2} {I1, O1,O2; I2, O1,O2}

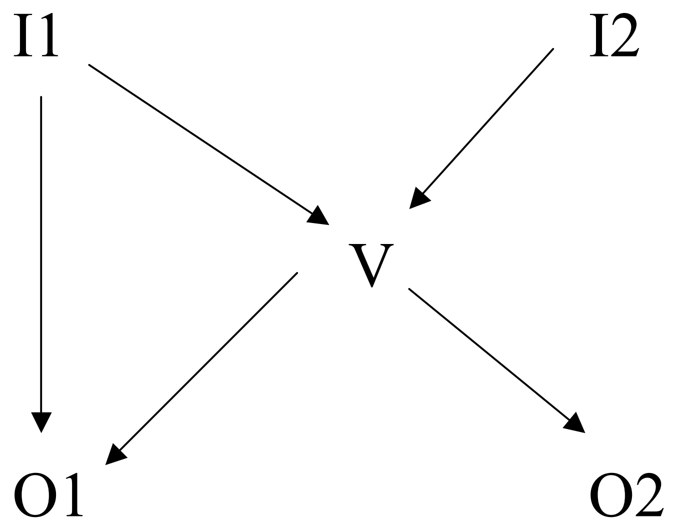
Impossible Deficit Combinations from Normal Structure



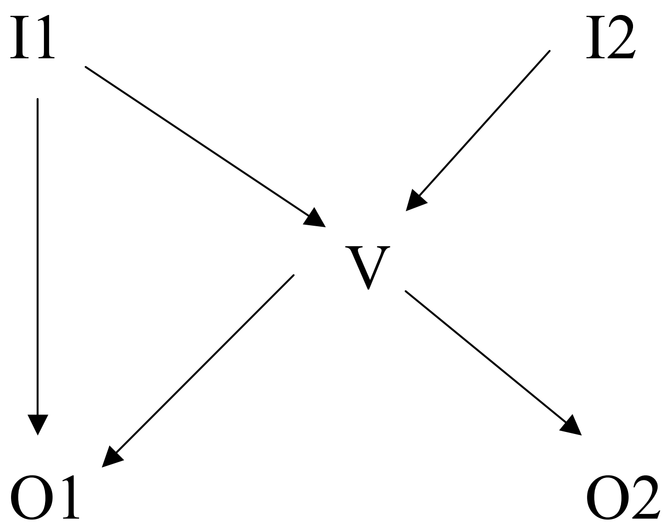
{I1, O1,O2; I2, O1,O2} {I1, O1,O2; I2, O1,O2}

{I1, O1,O2; I2, O1,O2} {I1, O1,O2; I2, O1,O2}

Normal Functioning



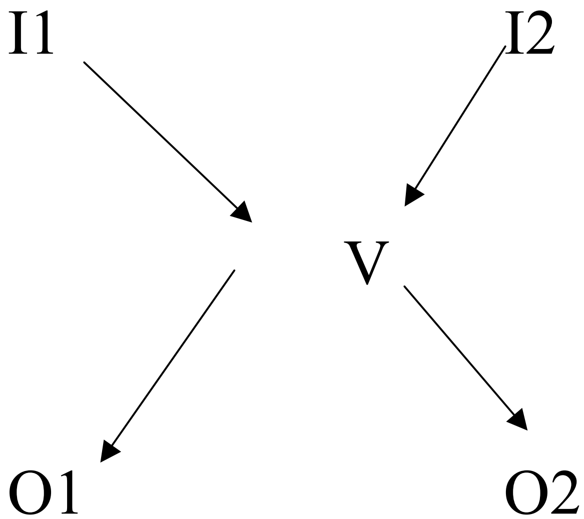
Impossible Deficits to Normal Functioning



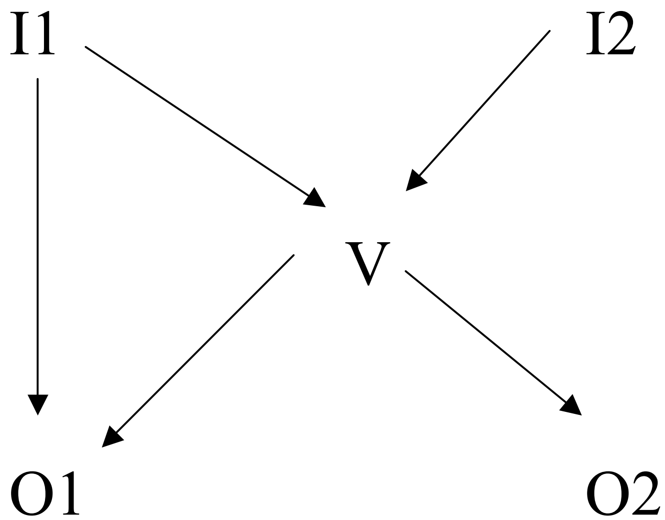
~~{I1, O1,O2; I2, O1,O2}~~ {I1, O1,O2; I2, O1,O2}

{I1, O1,O2; I2, O1,O2} {I1, O1,O2; I2, O1,O2}

What Does the Occurrence of Deficit {I1, O1,O2; I2, O1,O2} Tell Us?

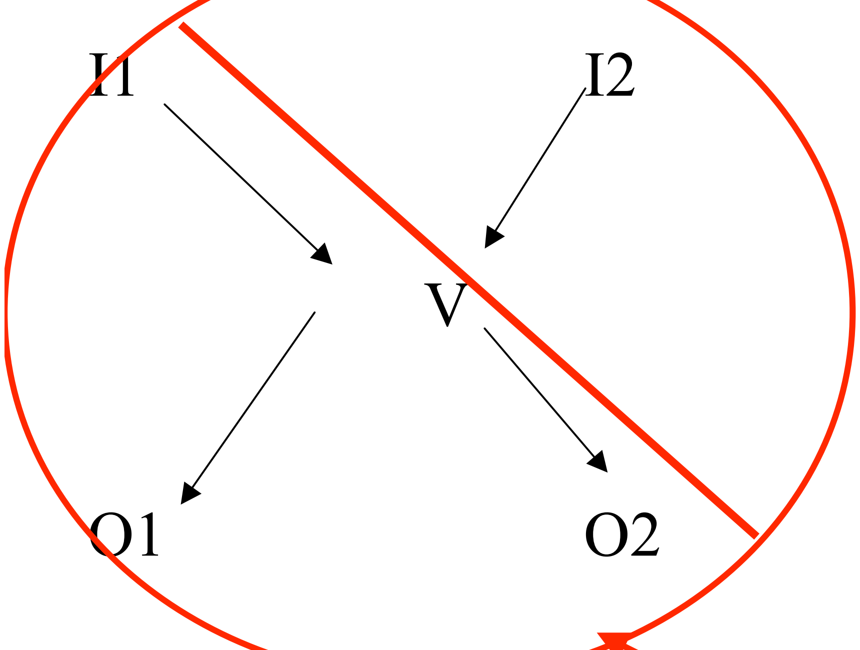


Structure 2

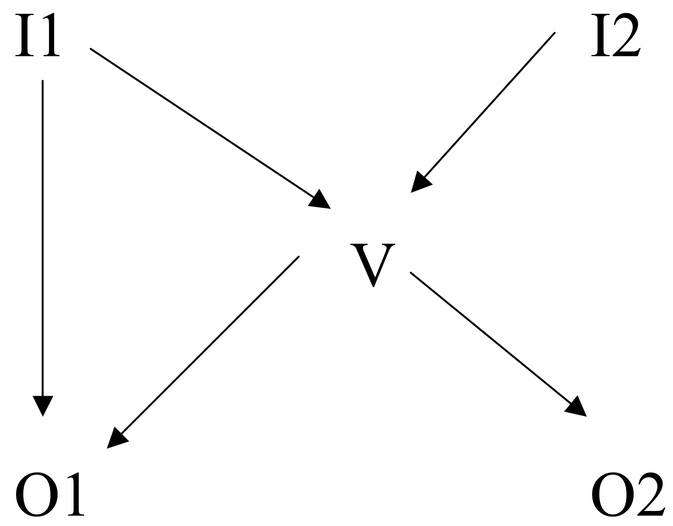


Structure 3

The Occurrence of Deficit {I1, O1,O2; I2, O1,O2} Tells Us that the Normal Structure is not Number 2



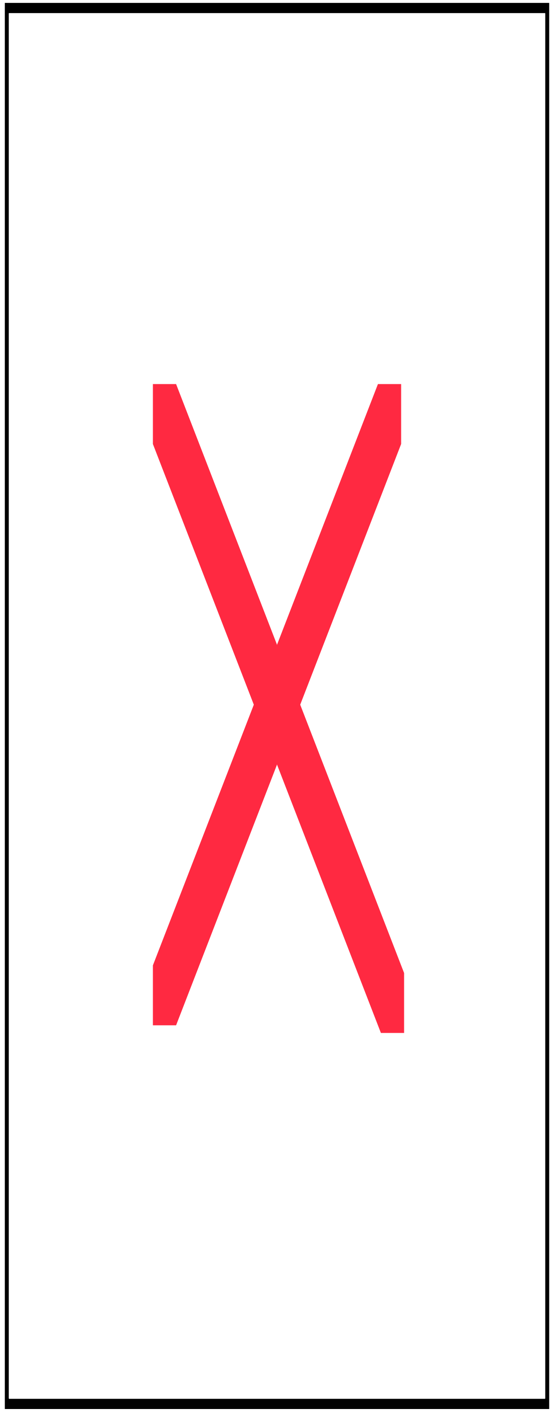
Structure 2



Structure 3

{I1, O1,O2; I2, O1,O2} {I1, O1,O2; I2, O1,O2}

{I1, O1,O2; I2, O1,O2} {I1, O1,O2; I2, O1,O2}



Question: Assuming that every deficit consistent with the true normal structure will eventually be observed, what inference procedure is reliable for the hypothesis space of the six structures?

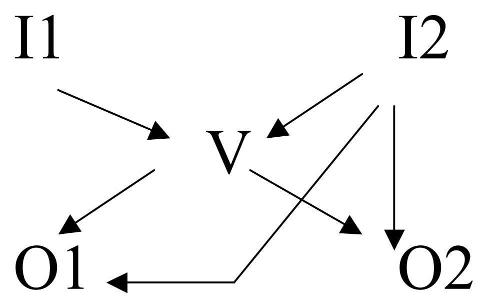
Answer: *Conjecture the normal structure that allows the smallest superset of the observed deficits.*

The procedure conjectures Structure 2 until and if a deficit is observed that is inconsistent with Structure 2; then it conjectures whichever of Structures 3, 4, 5, 6 allow the new deficit, until (and if) another deficit is observed, upon which it conjectures Structure 1.

Generalizing the Example

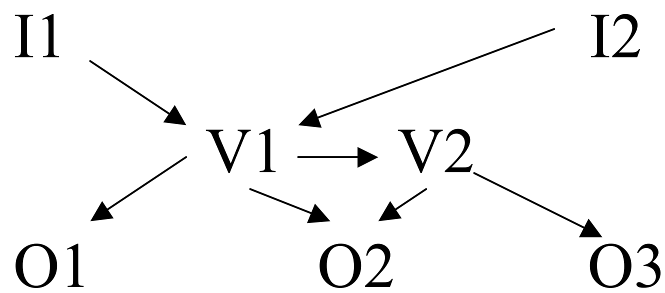
1. Expand the hypothesis space

-For example, allow structures such as



-Or allow internal nodes not directly connected to response nodes

-Or allow compound structures

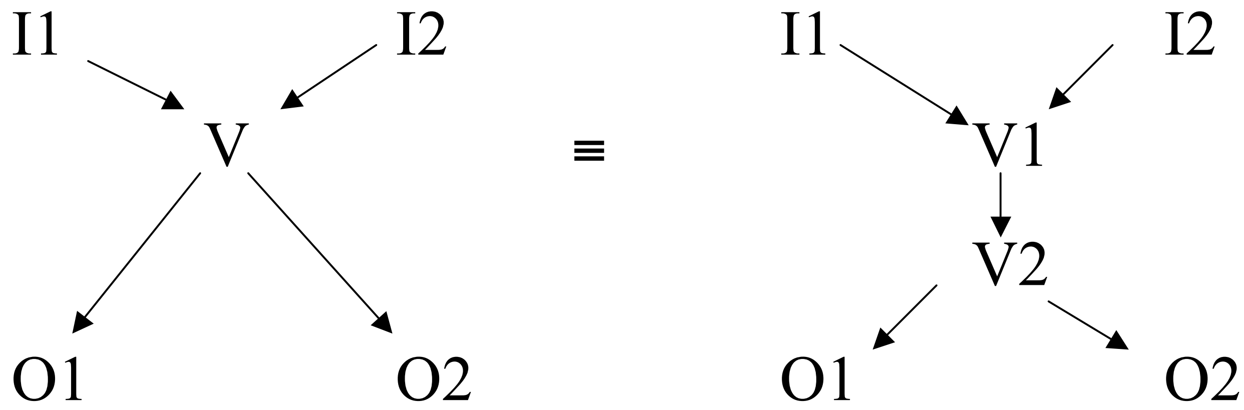


If the hypothesis space is expanded, what changes about the optimal inference procedure?

Only this:

Some structures may be indistinguishable—they imply the same set of possible deficits.

For example: Any structure with two adjacent internal vertices V1, V2 such that every input to V2 is through V1 and every output of V1 is through V2 is indistinguishable from a structure in which V1, V2 are replaced by a single node, V, with the input of V1 and the output of V2



Questions I haven't thought about:

Is there a polynomial (in the number of vertices) time algorithm for deciding whether two structures are indistinguishable?

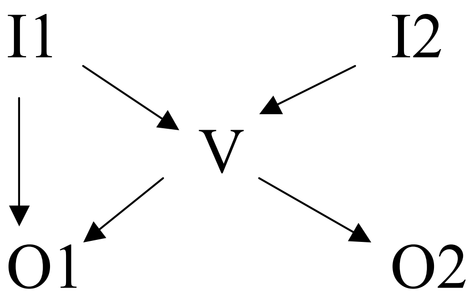
Is there a "nice" graphical characterization of indistinguishability?

Altering the hypothesis space

Jeff Bub suggests the assumption that

A normal response occurs if *at least one* pathway from input to output is intact.

In other words, the output has a disjunctive rather than a conjunctive gating.



Disjunctive gating

implies the deficit pattern

I1, O1,O2; I2, O1,O2

can occur

Conjunctive gating

implies the deficit pattern

I1, O1,O2; I2, O1,O2

cannot occur

We can consider the discovery problem when the structure is unknown and when the gating (“and” or “or”) is also unknown.

With at most one internal node, there are then 24 different graphs and 18 indistinguishability classes.

But whatever the space of structures, and whatever the dependence of normal output on intact pathways, no inference procedure is more reliable than:

Infer the disjunction of the indistinguishable structures that imply the smallest superset of the observed deficits.

Other problems:

Group data versus individual data
Can neural net models explain every conceivable pattern of deficits?

Can be similarly treated.

C. Glymour, *The Mind's Arrows, 2003.*

Learning Chaos

From Mara Harrell

Quick Outline

- Chaotic Systems in Nature
- Linear Systems
  - The Simple Pendulum
- Non-linear Systems and Chaos
  - The Double Pendulum
  - The Logistic Equation

Quick Outline

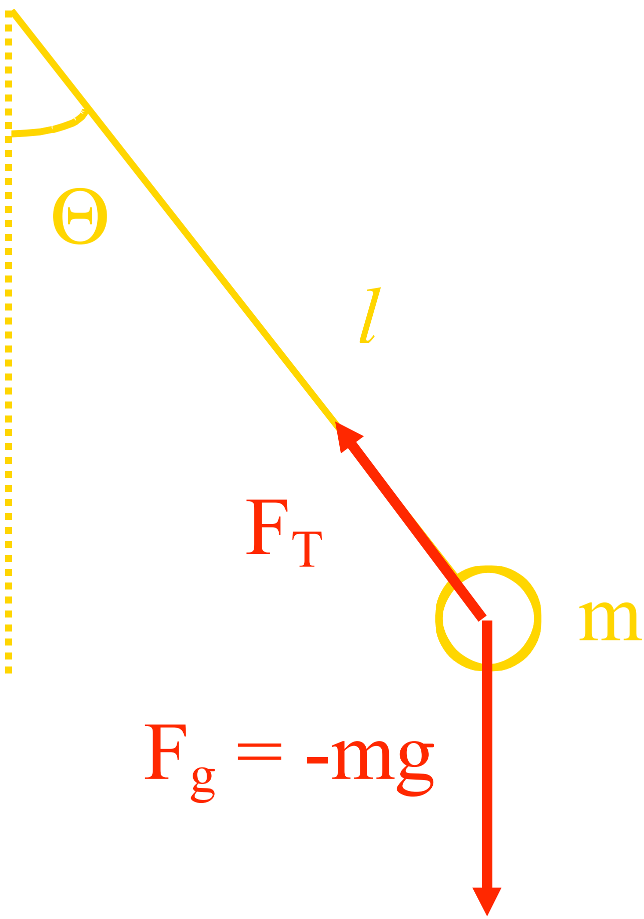
- How Can We Tell?
- The Problem of Induction
- A Modification of the Problem
- Results of the Modified Investigation

Alleged Chaos in Nature

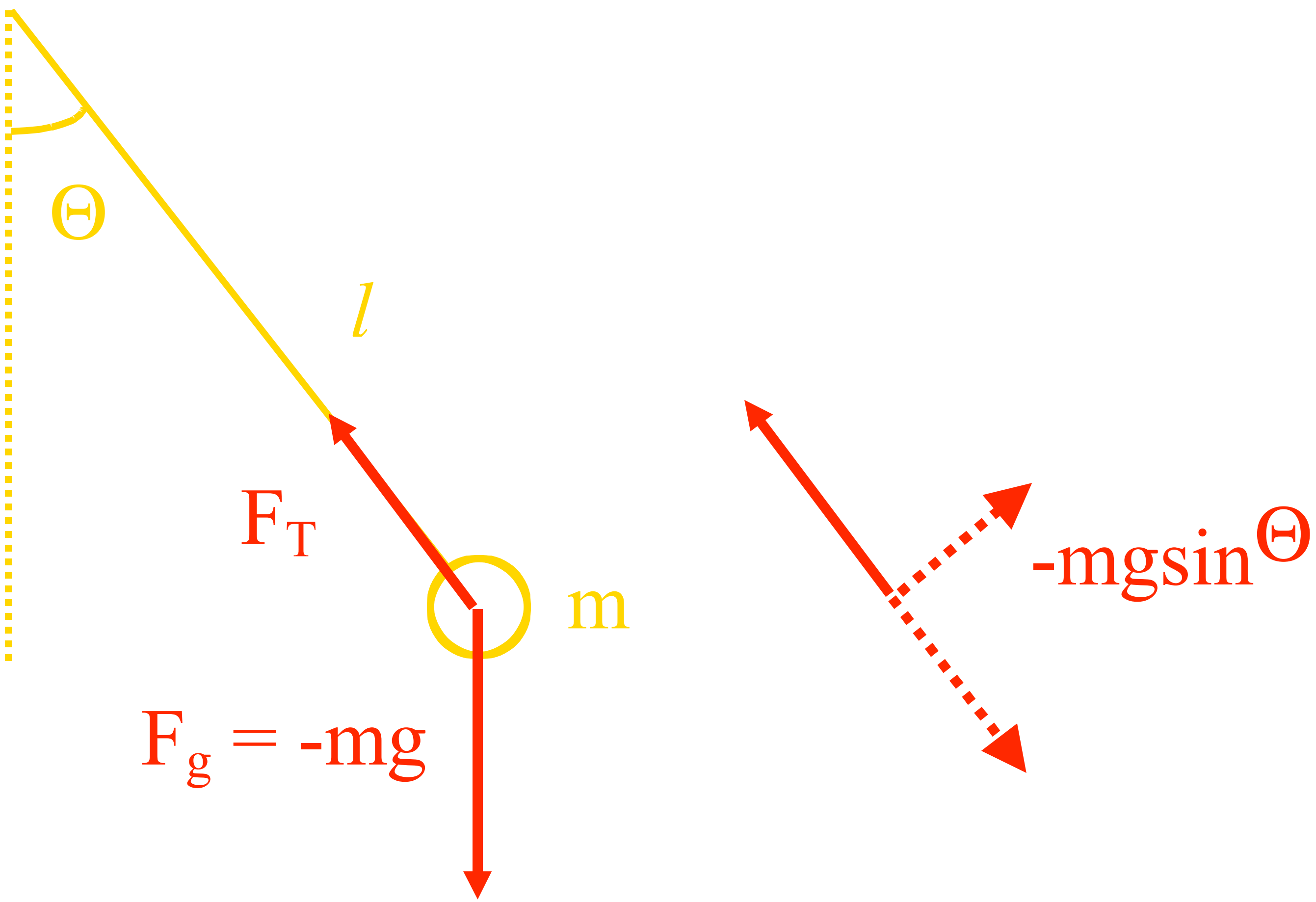
- The Belousov-Zhabotinskii Reaction
- Heartbeats
- The Solar System
- The Brain
  - Epileptic seizures

The Simple Pendulum

The Simple Pendulum



The Simple Pendulum

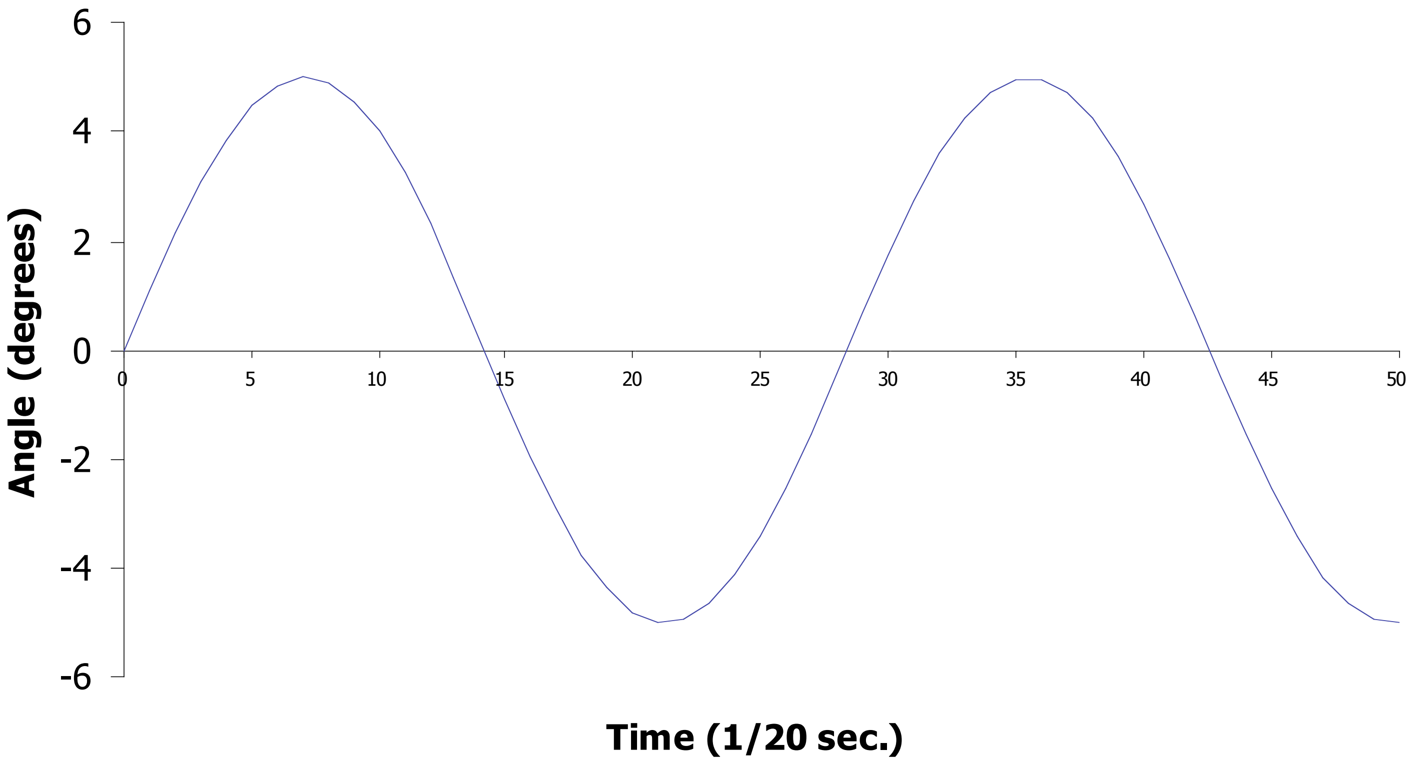


The Simple Pendulum

- $\sin\Theta = \Theta$
- $s = l*\Theta$
- $a = d^2s/dt^2 = l*d^2\Theta/dt^2$
- $ma = mg\sin\Theta$
- $d^2\Theta/dt^2 = g*\Theta/l$
- Solution: $\Theta = A\sin(\omega t)$

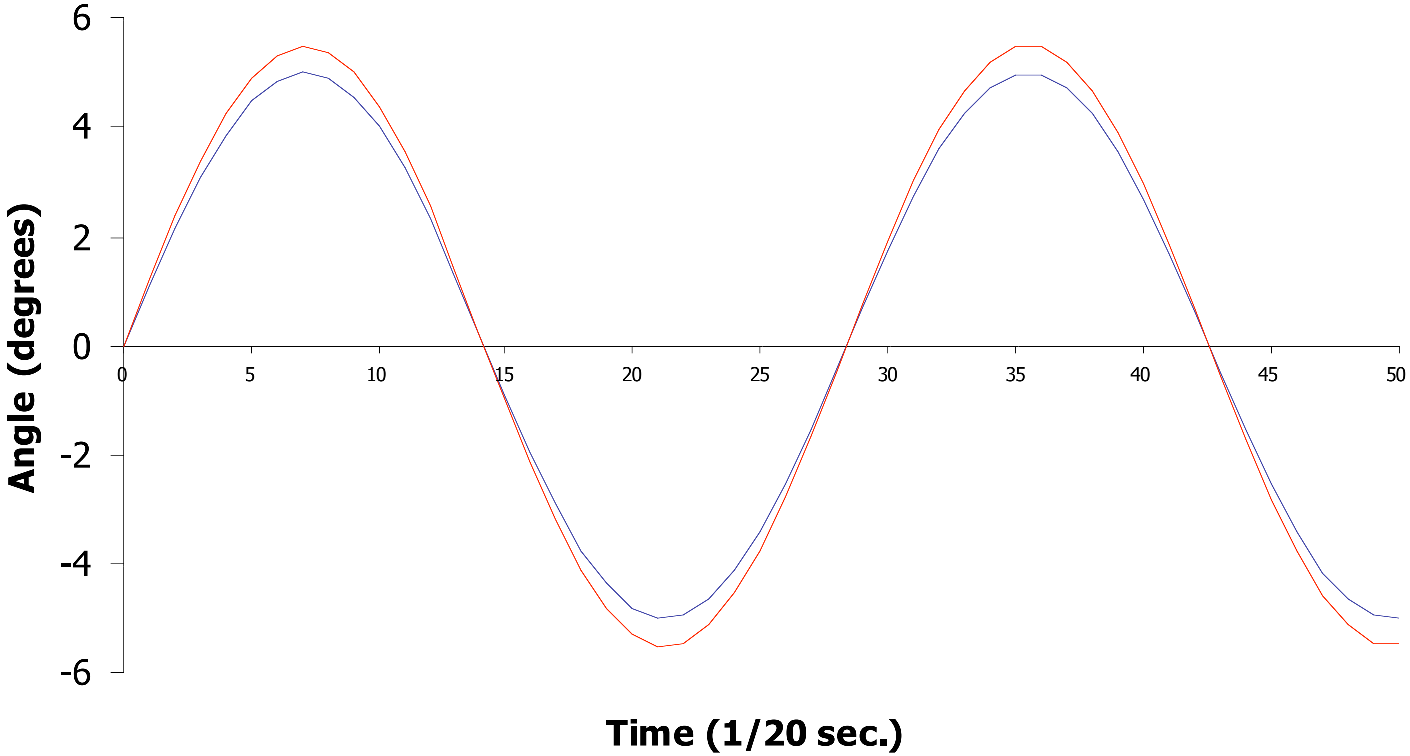
The Simple Pendulum

$\Theta = A\sin(\omega t)$



The Simple Pendulum

$\Theta = A\sin(\omega t)$



The Double Pendulum

The Double Pendulum

Where:

The Double Pendulum

For small oscillations:

[redacted]

The Double Pendulum

- Exhibits Chaos
  - Deterministic Equations
  - Sensitive Dependence on Initial Conditions
  - Exponentially Decreasing Predictability

The Logistic Equation

$N_{t+1} = (R - bN_t) N_t$

- N is the number of organisms
- R is a parameter that governs the growth rate when N is small
- b is a parameter that governs how the growth rate decreases as N increases

The Logistic Equation

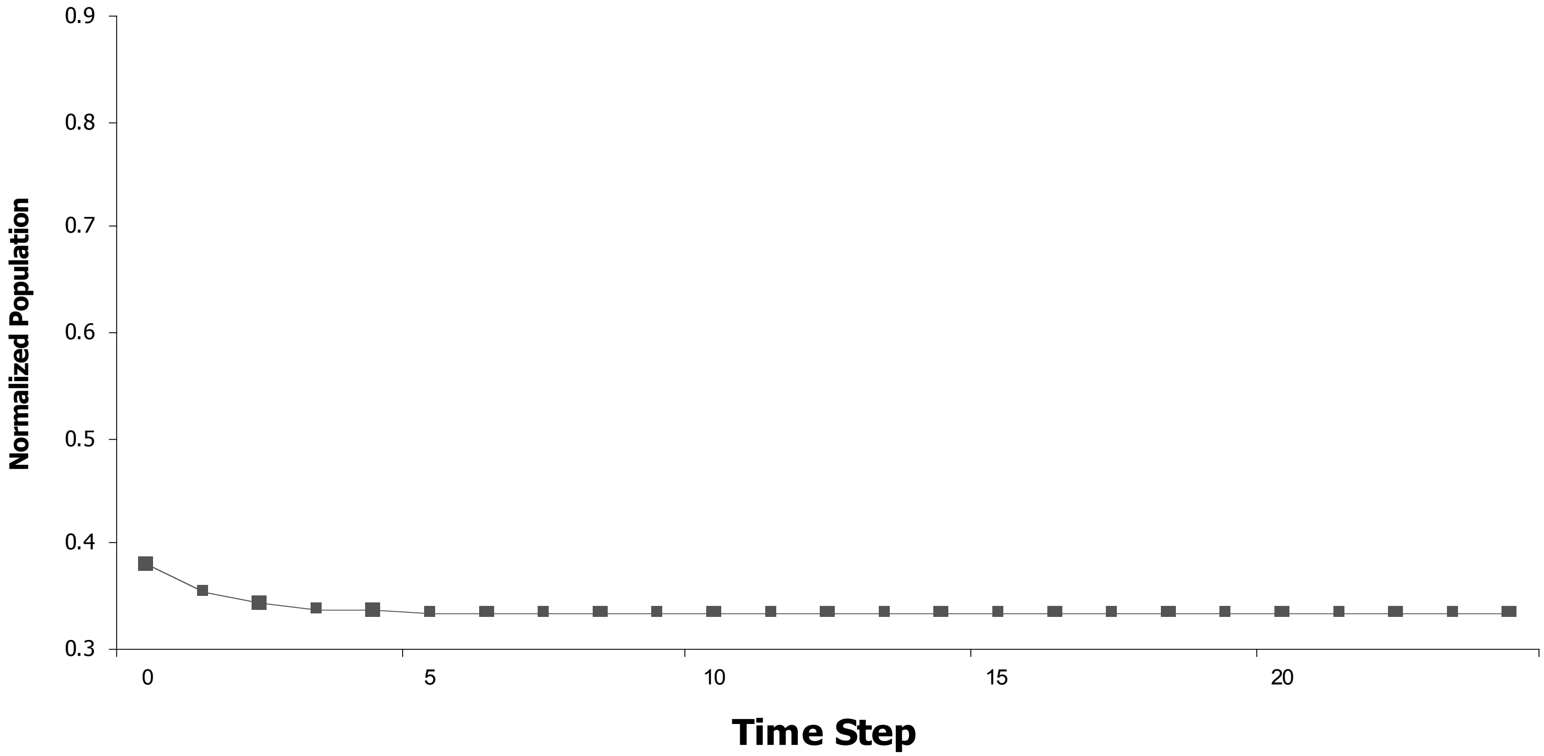
$N_{t+1} = (R - bN_t)\, N_t$

$x_t = bN_t/R$

$x_{t+1} = Rx_t(1 - x_t)$

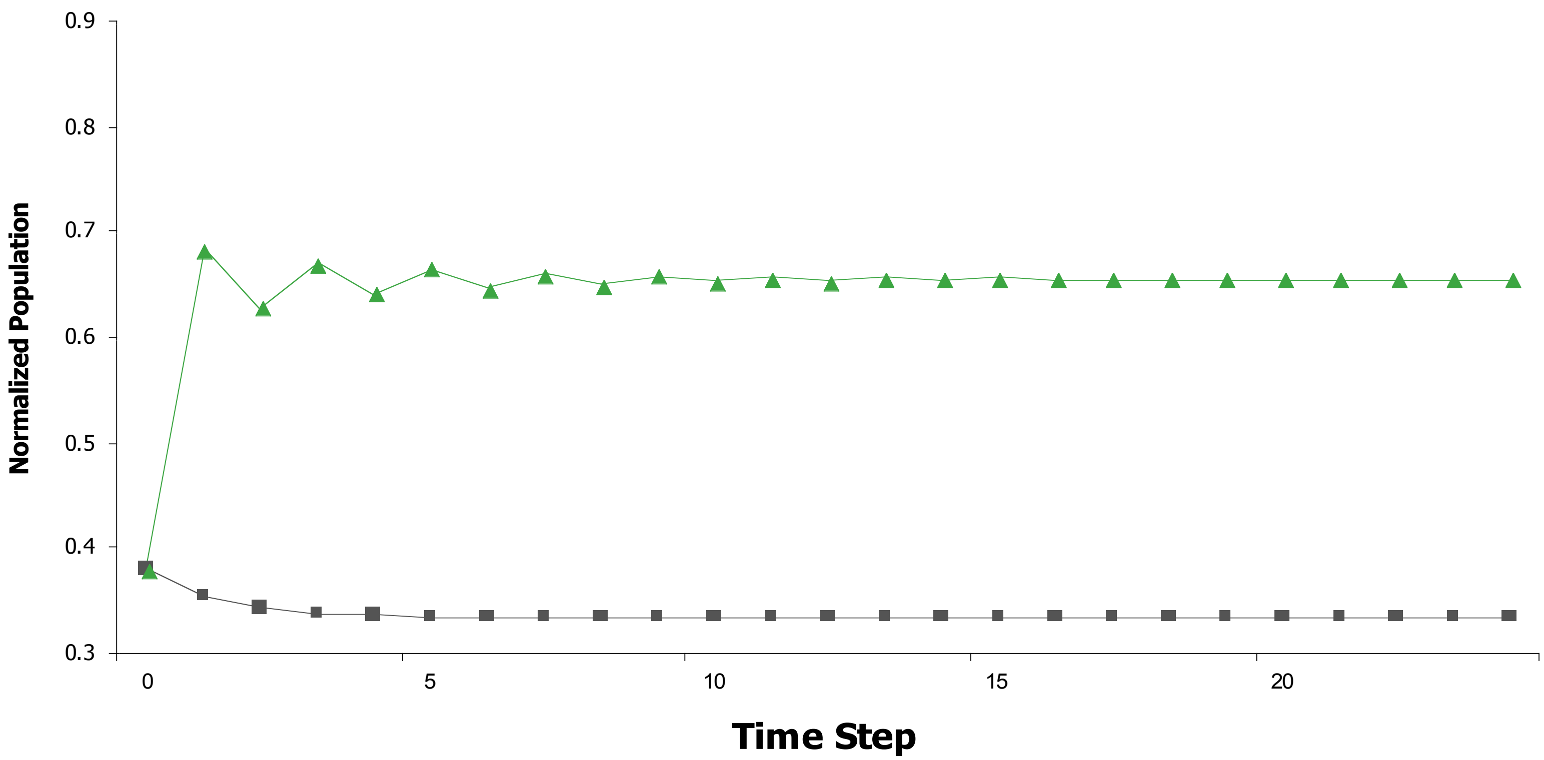
The Logistic Equation

$x_{t+1} = Rx_t(1 - x_t)$ $R = 1.5$



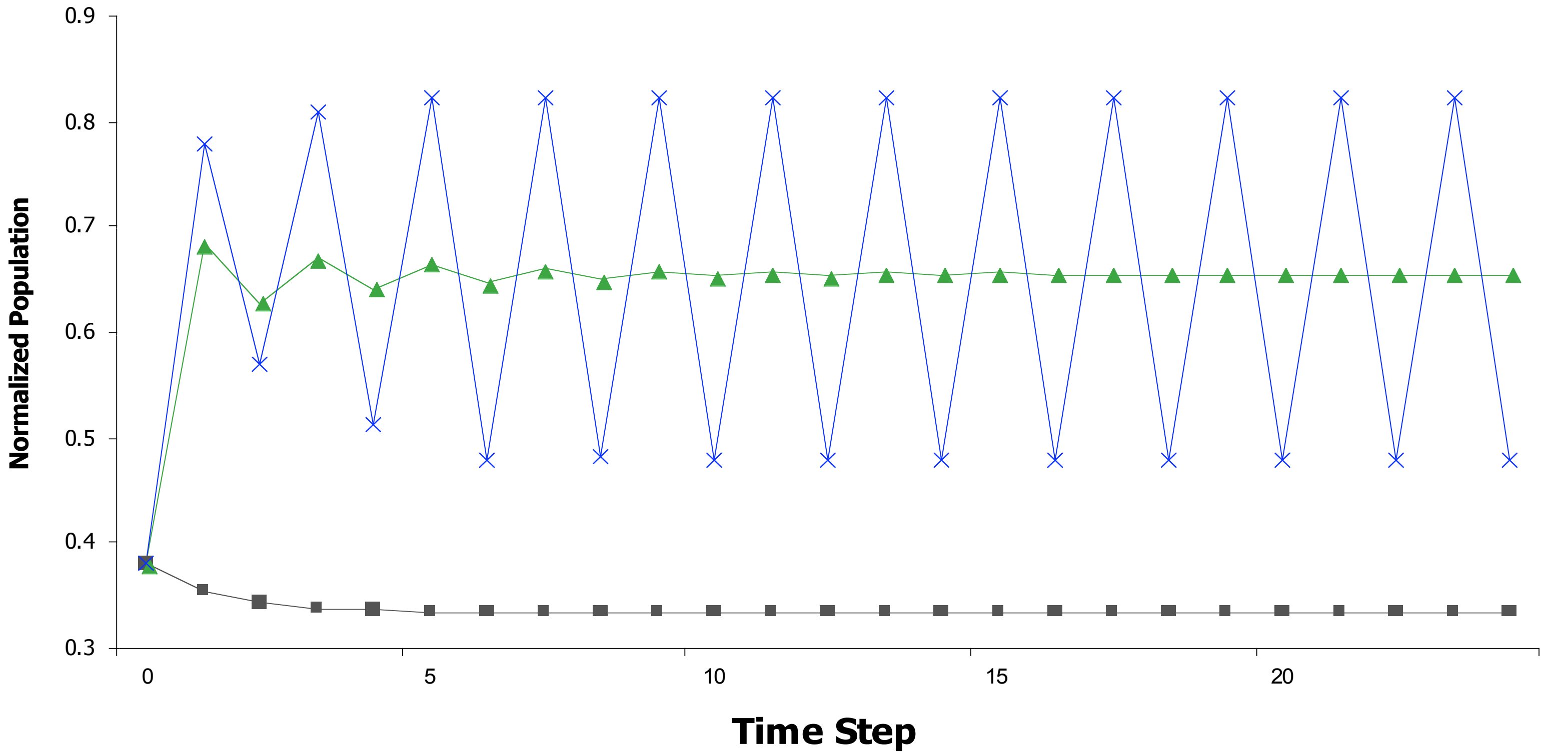
The Logistic Equation

$x_{t+1} = Rx_t(1 - x_t)$ R = 2.9



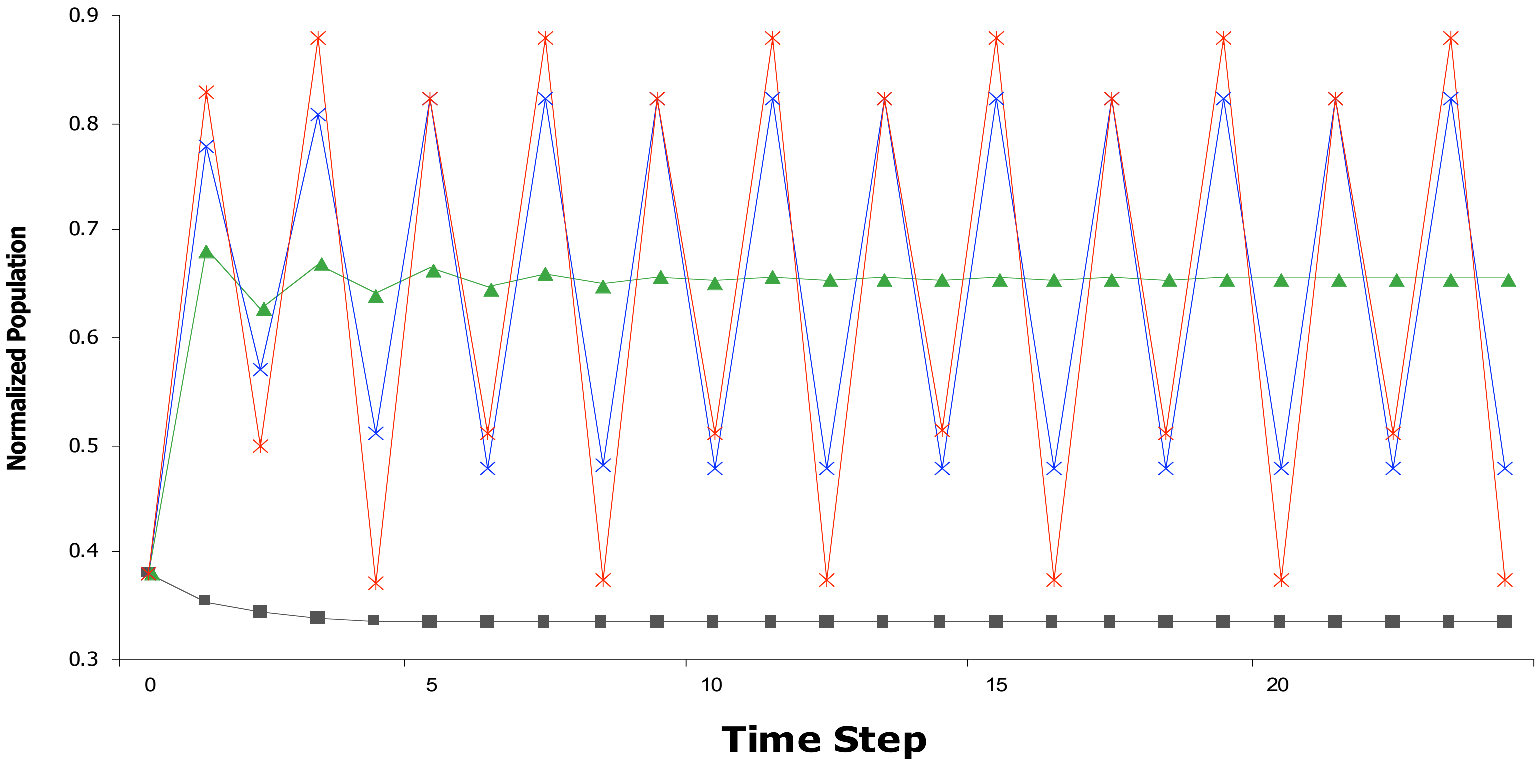
The Logistic Equation

$x_{t+1} = Rx_t(1 - x_t)$ R = 3.3



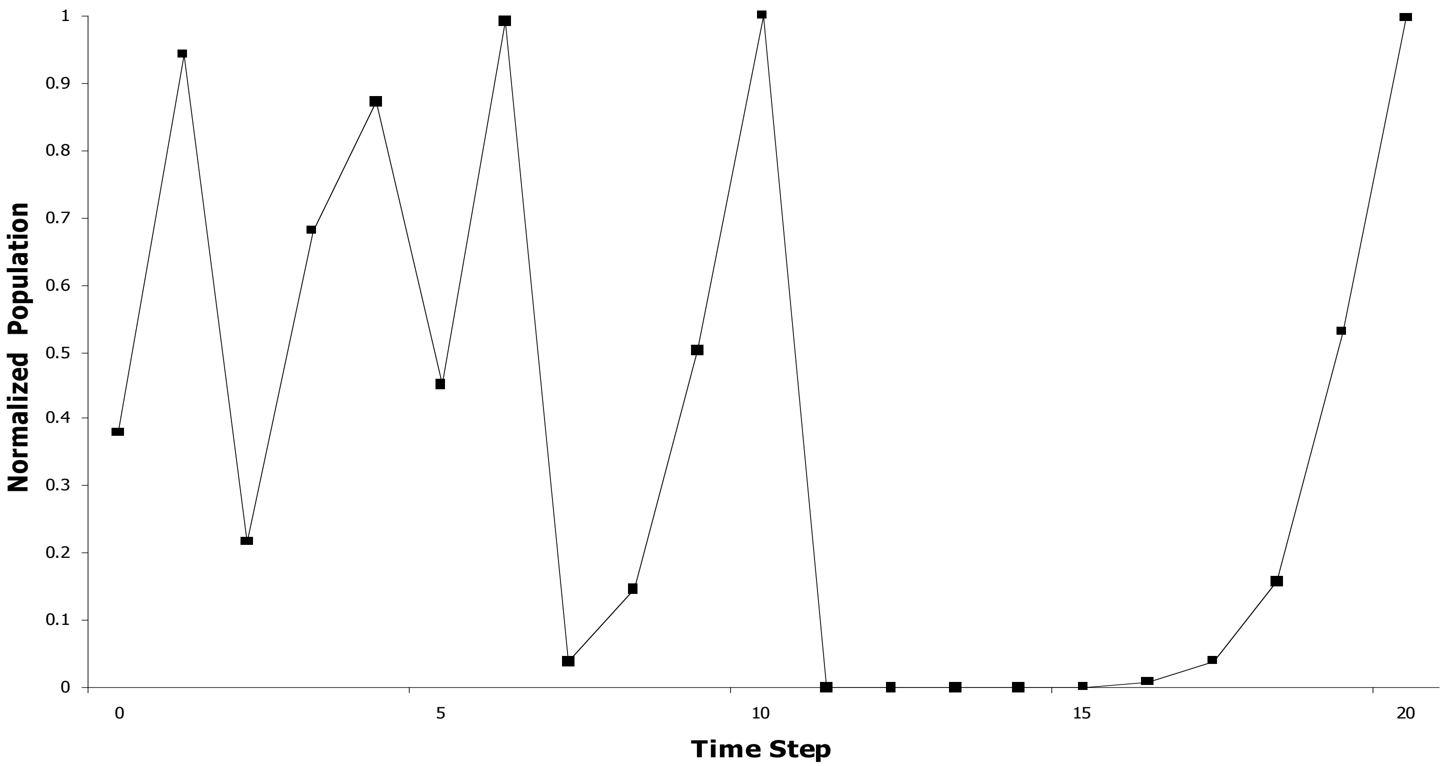
The Logistic Equation

$x_{t+1} = Rx_t(1 - x_t)$ R = 3.52



The Logistic Equation

$x_{t+1} = Rx_t(1 - x_t)$ $R = 4$

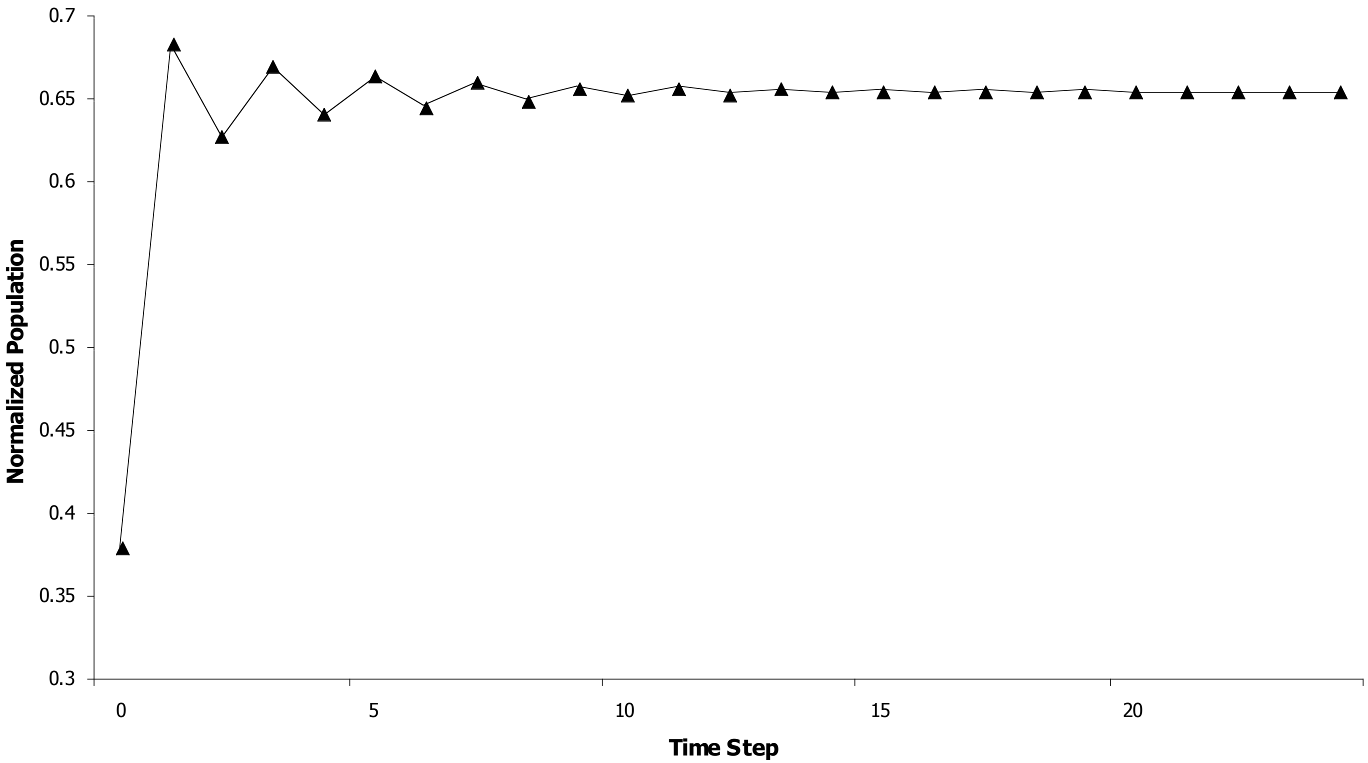


The Logistic Equation

- Aperiodicity
- Apparent randomness
- Sensitive Dependence on Initial Conditions

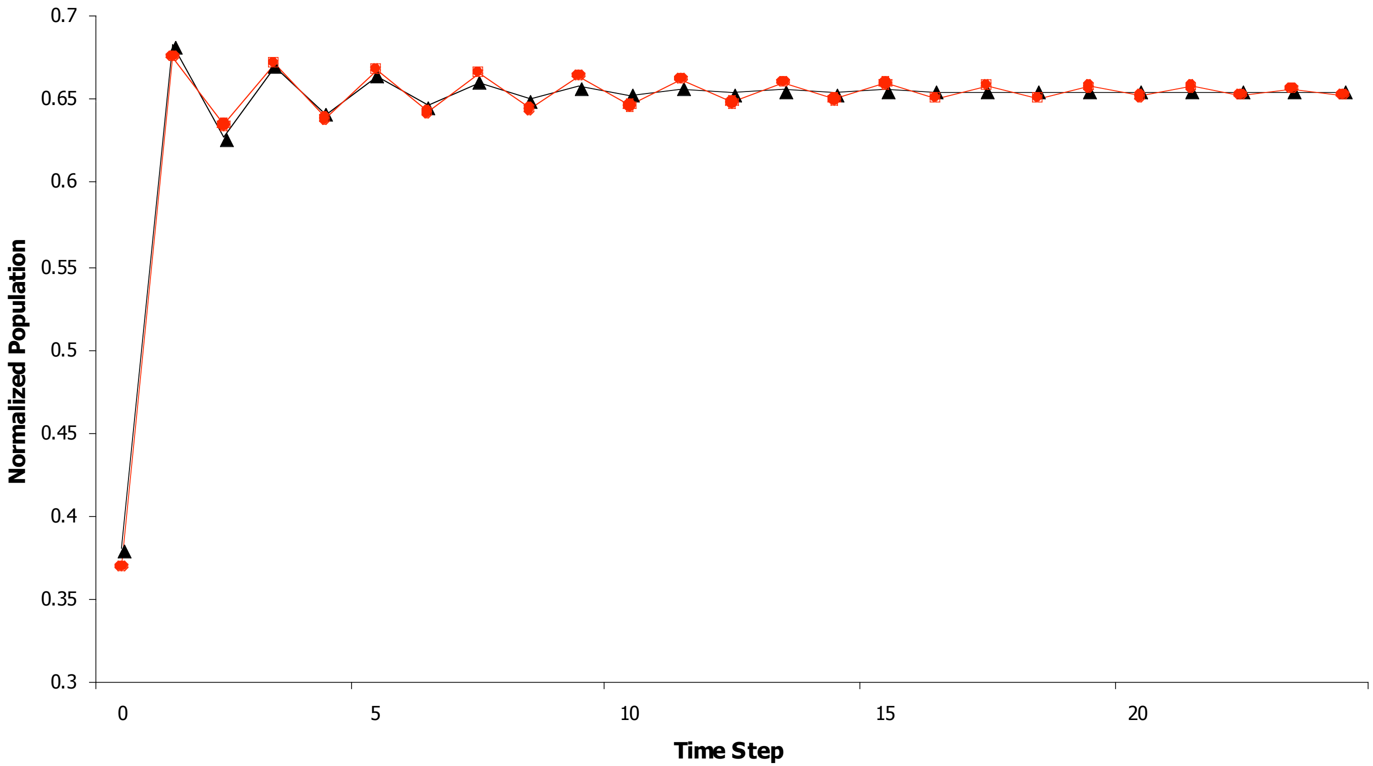
The Logistic Equation

$x_{t+1} = Rx_t(1 - x_t)$ $R = 2.9$



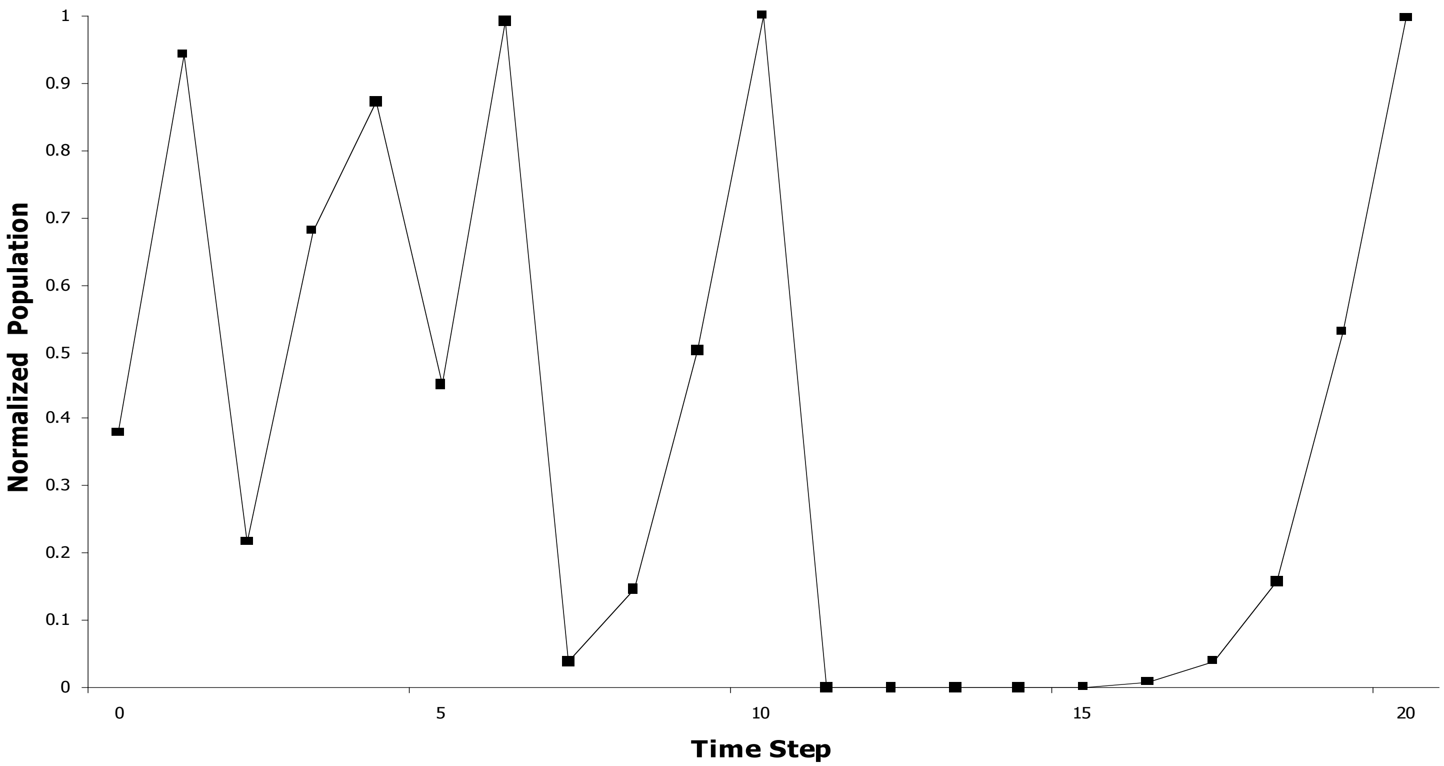
The Logistic Equation

$x_{t+1} = Rx_t(1 - x_t)$ $R = 2.9$



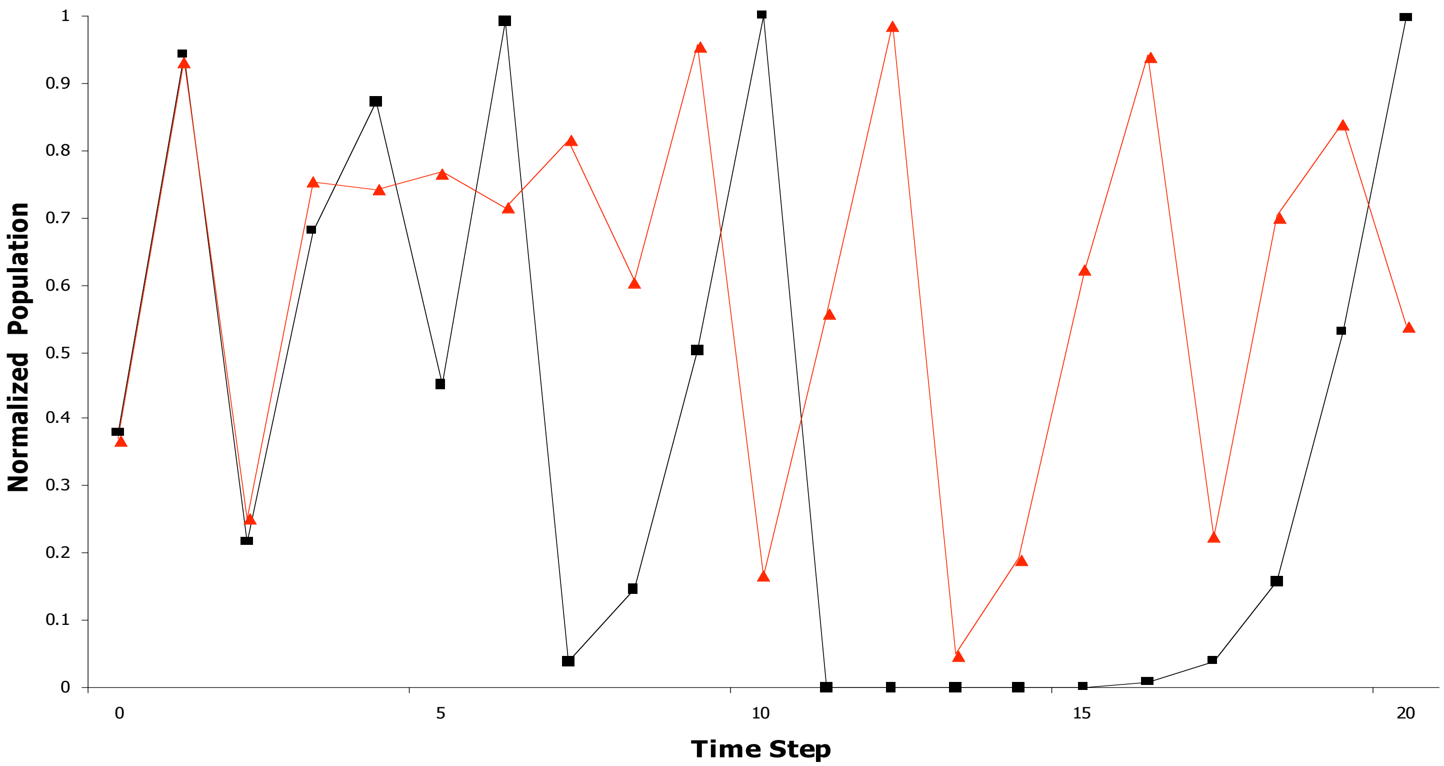
The Logistic Equation

$x_{t+1} = Rx_t(1 - x_t)$ $\qquad$ $R = 4$



The Logistic Equation

$$x_{t+1} = Rx_t(1 - x_t) \qquad R = 4$$



Formal Definition of Chaos

A chaotic trajectory is one that is:

- bounded,
- aperiodic,
- deterministic,
- and displays sensitive dependence on initial conditions.

Sensitive Dependence

- Measured by the Lyapunov exponent, λ
  - λ measures the divergence of nearby trajectories
  - $\lambda > 0$ implies sensitive dependence on initial conditions
  - There are algorithms for estimating λ given finite data

How Can We Tell?

- Traditionally through complicated data analysis
  - Hard even for a one-dimensional system where we know the equation (Logistic Equation)
- Harder if we don't have the equation
  - Consider knowing it is aperiodic

The Problem of Induction

- Plato
- Sextus Empiricus
- Hume

- For all possible worlds, there is no method that will determine the truth of a universal hypothesis on the basis of a finite amount of singular data.

What Can We Do?

- Introduce background assumptions
- Change the criterion of success

Towards a Solution

- Karl Popper
- Hilary Putnam
- Kevin Kelly

Kelly’s Generalized Framework

- Introduce background assumptions
  - Exclude some possible worlds
- Change the criterion of success
  - Verify, Refute, Decide
  - By a deadline, with certainty, in the limit, gradually

Kelly’s Generalized Framework

- Important criteria for this problem
  - Verification (Refutation) in the limit: stabilize to “true” (“false”) iff
    the hypothesis is true (false)
    - Equivalent to Bayesian convergence given appropriate prior probabilities
  - Gradual verification (refutation): Real-valued conjectures converge to 1.0 (0.0) iff the hypothesis is true (false)

Applications to Real Science

- There aren’t many
  - O. Schulte: Conservation Laws in Particle Physics
  - C. Glymour: Box and Arrow diagrams in cognitive neuropsychology

Framework for Learning Chaos

- Assume the system is a one-dimensional map: $x_{d+1} = f(x_d)$
  - Not assuming the map is deterministic
- Assume the data are a sequence of real numbers
  - Will consider various assumptions about the accuracy and precision of this data

Results for Ideal Data

- Boundedness is verifiable in the limit, but only gradually refutable.
- Aperiodicity is refutable in the limit, but only gradually verifiable.
- Determinism is refutable with certainty, but only verifiable in the limit

Results for Ideal Data

- A positive Lyapunov exponent is verifiable in the limit, but only gradually refutable.
- Chaos is verifiable in the limit, but only gradually refutable.

Results for Non-Ideal Data

- Finitely precise data
  - Boundedness is verifiable in the limit and gradually refutable
  - Aperiodicity is underdetermined
  - Determinism is underdetermined
  - A positive Lyapunov exponent is verifiable in the limit and gradually refutable

Results for Non-Ideal Data

- Noisy data (additive truncated Gaussian)
  - Boundedness is verifiable in the limit and gradually refutable
  - Aperiodicity is underdetermined
  - Determinism is underdetermined
  - A positive Lyapunov exponent is verifiable in the limit and gradually refutable

Further Project Directions

- Future technical research
- Implications for an adequate reliabilist theory of knowledge
- Implications for the practice of scientists
- Account of chaos as an important explanatory tool in science

Other Possible Applications

- Other limiting reliability analyses in philosophy of science
- For example, parameter estimation in modern theories of gravitation (the Parametrized Post-Newtonian Framework)

Future Technical Research

- Results for positive Lyapunov exponent in the case of non-ideal data
- Determine the background conditions necessary for verification and refutation of chaos in the case of non-ideal data
- Results for multi-variable systems

Technical Definitions

- These data are bounded ⇒ There is a y such that for all data x_i from the system, the absolute value of x_i is less than or equal to y ($\exists y\ \forall i\ |x_i| \leq y$)
- These data are aperiodic ⇒ For all T greater than 0, and all N greater than or equal to 0, there is an n greater than N such that $x_{n+\mathrm{T}}$ is not equal to x_n ($\forall \mathrm{T} > 0, \forall \mathrm{N} \geq 0, \exists n > \mathrm{N}\ x_{n+\mathrm{T}} \neq x_n$)

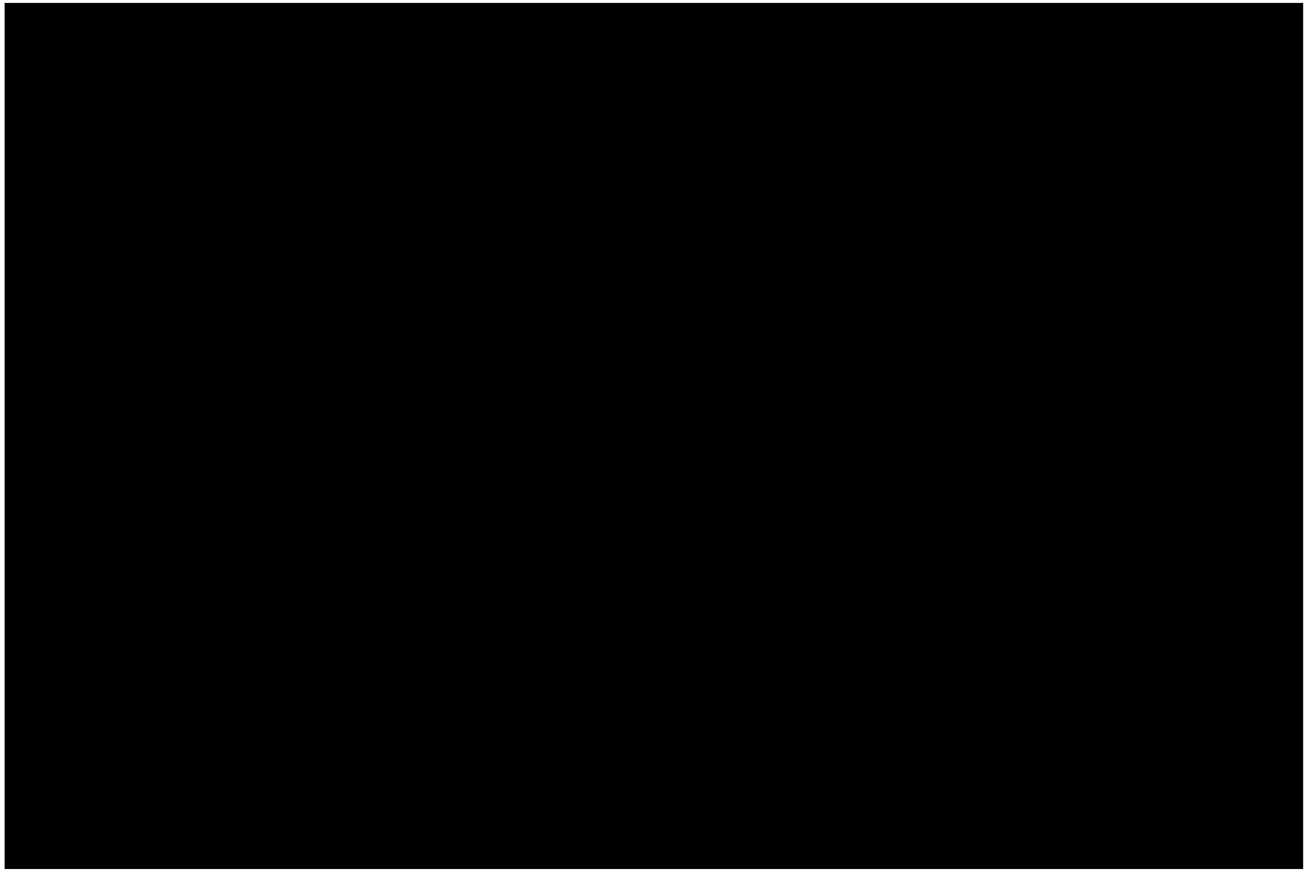
Technical Definitions

- These data are deterministic $\Rightarrow$ For all x_i, all x_j, if $x_i = x_j$, then for all n, $x_{i+n} = x_{j+n}$ ($\forall i \, \forall j \, [(x_i = x_j) \rightarrow \forall n \, (x_{i+n} = x_{j+n})]$)
- These data display sensitive dependence to initial conditions (have a positive Lyapunov exponent) $\Rightarrow$ There is an ε and an M greater than 0 such that for all i greater than M, λ_i is greater than ε ($\exists \varepsilon \, \exists \, M > 0 \, \forall i > M \quad \lambda_i > \varepsilon$)

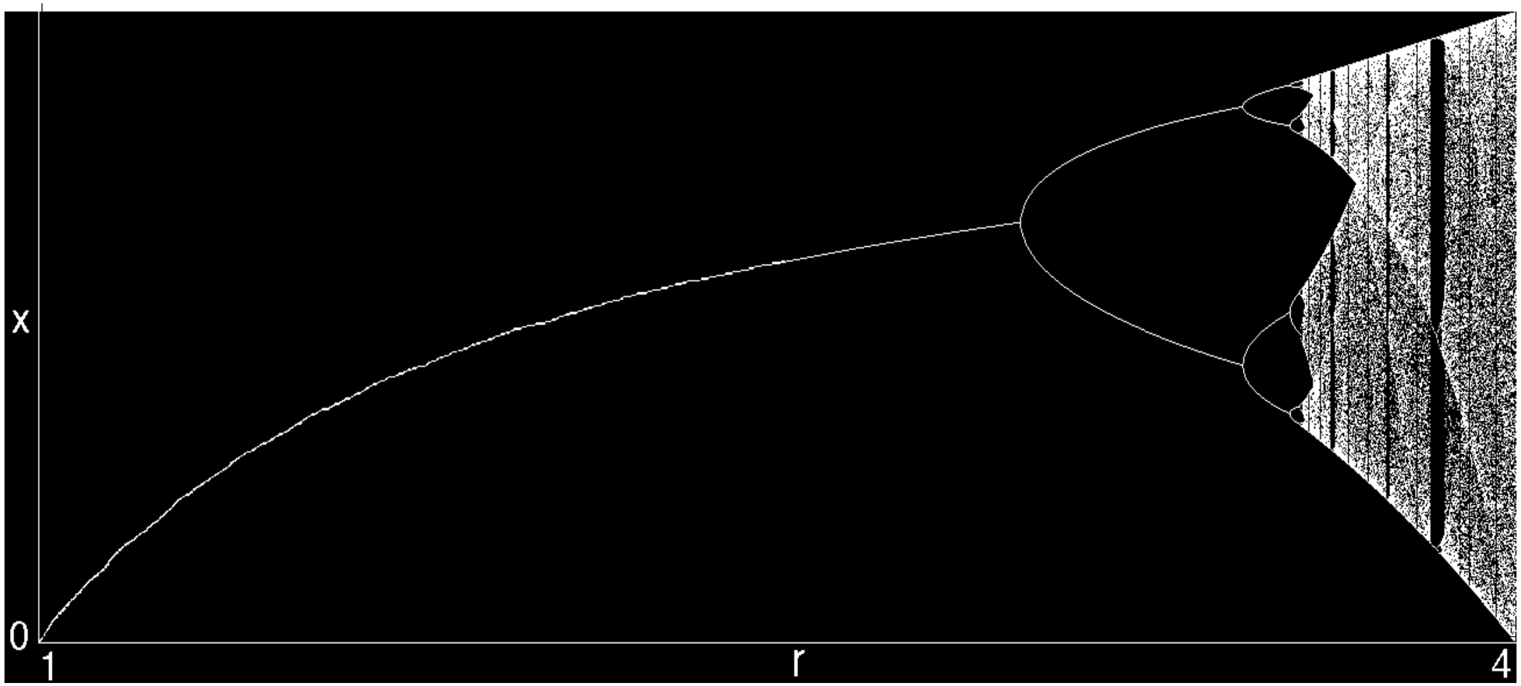
Proof Sketch

- Example: Boundedness is verifiable in the limit
  - Pick a number, q_0
  - Conjecture “true” if the absolute value of the most recent datum is less than q_i
  - Conjecture “false” if the absolute value of the most recent datum is greater than q_i and pick a new $q_{i+1} > q_i$ for comparison

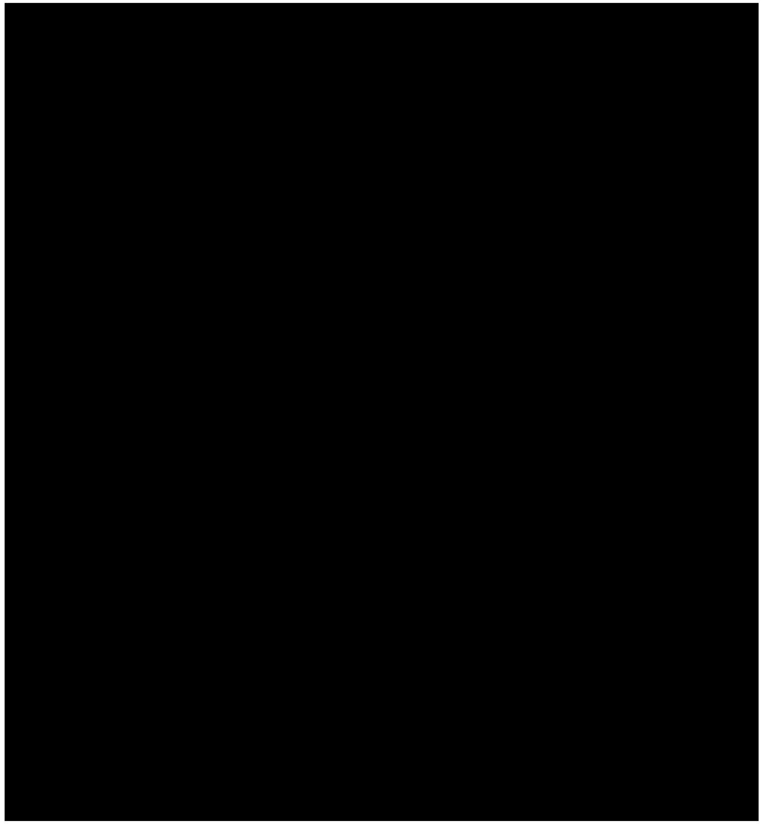
The Lyapunov Exponent



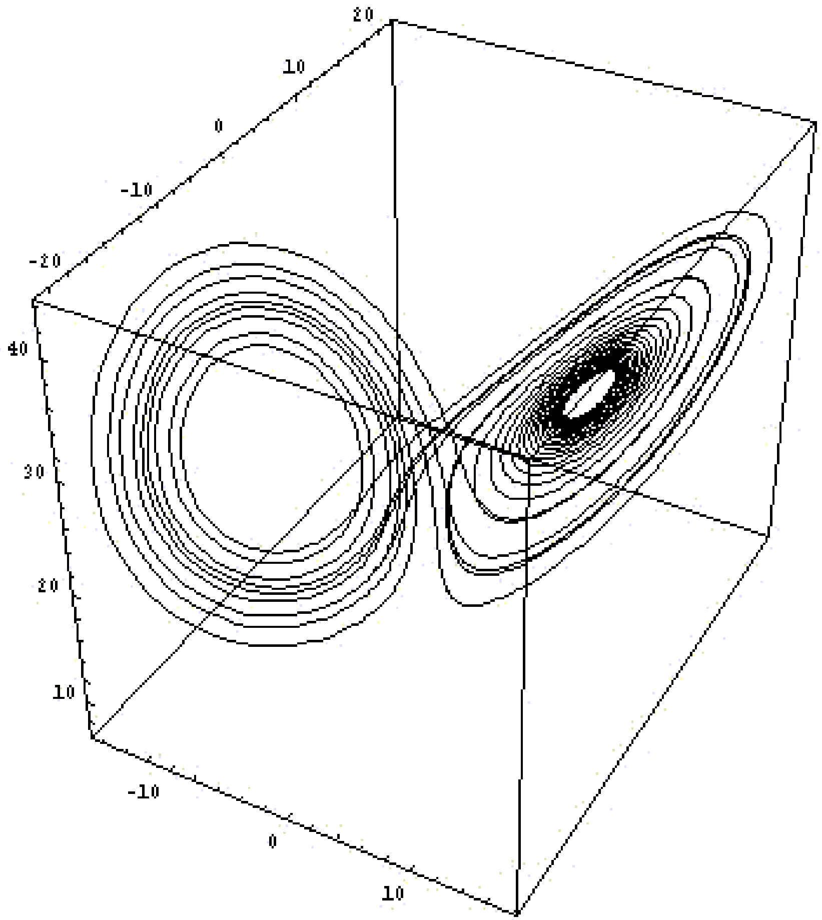
The Logistic Equation



The Lorenz Equations



The Lorenz Attractor



Learning Chaos

Shallice, etc.

- Say damage to modules can be partial
- Say some normal behaviors put smaller computational demands on modules they depend on than do other normal behaviors for the same module

But What Happens if Damage Can Be Partial?

- Suppose output 1 <- B -> output 2 and damage is always total.
- Then whenever Output 1 is abnormal so is output 2 and vice-versa—so one can tell they have a common channel—a common cause.
- But if B can be partially damaged, and some outputs that depend on B can still be normal, output 1 can be abnormal and output 1 normal, or the other way, or both

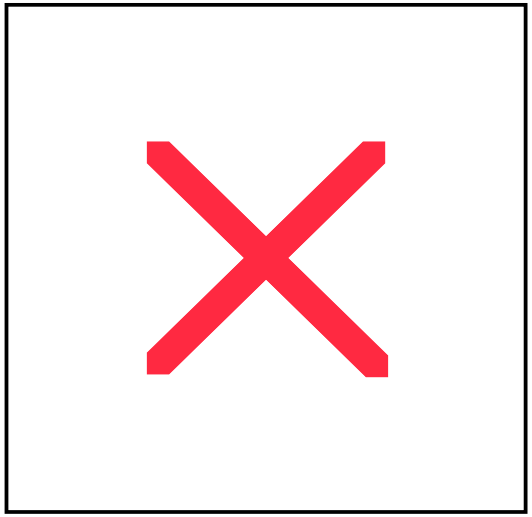
Unless

- One knows a partial ordering of children of a module M—partially ordered by how likely it is that damage to M makes the child module abnormal….
- Nothing can be learned about functional structure from behavioral profiles of brain damaged subjects

“Averaging performance O1* through On* would be justified if we could assume than Ms, Cs and Ls are equivalent in relevant respects for patients P1 through Pn. We cannot....however, for the Ls. ...in our research with brain-damaged patients...we may legitemately average their performance if and only if we have demonstrated empirically that the patients have equivalent functional lesions.”

-Caramazza & McCloskey, 1988, 522-523.

Caramazza and McCloskey Argument about Group Data in Neuropsychology



Pi = ith patient

M = normal cognitive organization

C = cognitive task

Li = lesion of kind i

Oi* = abnormal performance

Or, Farah:

“Traditional neuropsychological group study designs…are not appropriate for answering most questions about cognitive processes: These groups will be heterogeneous with respect to the impairments that are the subject of study, and we therefore risk basing our conclusions on average performance profiles that are artifactual, in that they may not exist in any one case.”

(Visual Agnosia, p. 145)

“we...risk basing our conclusions on average performance profiles that are artifactual, in that they may not exist in any one case.”

But in box and arrow models:

if deficit A can be produced from normal architecture N by some lesion, and if also

deficit B can be produced from normal architecture N by some lesion, then

deficits A and B can be produced in the same subject by a combination of lesions.

So what’s the risk?

Good inferences:

Deficits A, B..K occur in Group, therefore deficits A, B,…K are simultaneously possible.

Bad inferences:

Deficits A, B…K but not M occur in Group, therefore A,B,…K but not M are simultaneously possible.

Deficits A and B are correlated in Group, therefore the normal architecture allows a single lesion that produces A, B.

Ideally, under the same assumptions as for individual data in our examples, for some groupings the rule of inference applied to grouped data is as reliable as any possible procedure.

But,

(1) if a deficit is rare and lesions that produce it are grouped with lesions that do not, the deficit may not be statistically significant.

(2) if combinations of deficits are not recorded (e.g.,

I2, O1, O2 data are omitted) the procedure is not reliable.

(3) if some groups are omitted, the procedure is not reliable.

X

What if we do not use co-occurrences of deficits in a Group, but instead use marginal frequencies of deficits?

For structure 2, for example, every deficit <I1,O1> must be accompanied either by deficit <I1, O2) or by deficit <I2,O1> or both.

Hence if structure 2 is the normal architecture, *necessarily*:

fr(<I1, O1> <= fr(I1,O2> + fr(<I2, O1>)

Structure 2 implies 4 such inequalities in frequencies of deficits.

Structure 1 implies none of these inequalities

Each of the 4 other structures implies a distinct threesome of these inequalities.

Procedure: *conjecture the structure that implies the frequency constraints so far observed.*

(Bates, McWhinney, et al.??)

Procedure is obviously not reliable: any set of frequency constraints can be realized by structure 1.

Procedure: *Eliminate any structure that implies a frequency constraint not satisfied by the data.*

This procedure is correct, but it may not be informative--for example, if all 4 constraints satisfied by structure 2 hold, the procedure yields no information about the true structure.

General Conclusion:

Group data is not useless.

But, issues of individual variation and experimental error aside, it may be less informative than individual profiles, depending on how the data are aggregated.