CS 365
Introduction to Scientific Modeling
Lecture 1: Modeling

Stephanie Forrest
Dept. of Computer Science
Univ. of New Mexico
Fall Semester, 2014

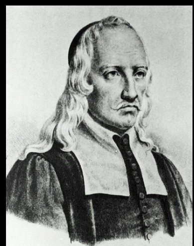
THE UNIVERSITY *of* NEW MEXICO

Introduction

- Three legs of science
  - Experiment
  - Mathematical theory
  - Simulation and modeling
- What are models good for?
  - Tools for analyzing data
  - Methods for discovering new knowledge (3rd leg)
  - Understanding nature as an information-processing system
  - Explaining how something works---mechanisms

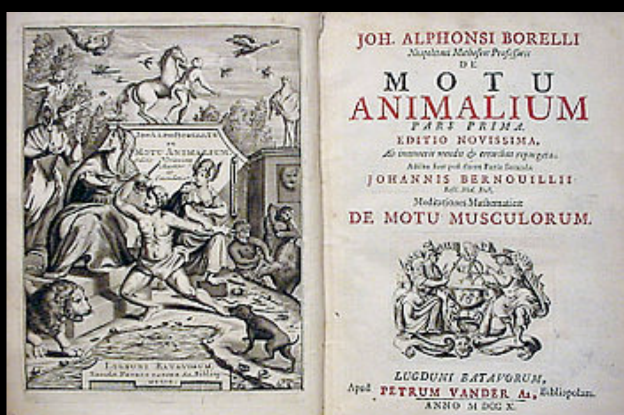
What is a model?

- A hypothetical description, often based on an analogy, used in analyzing something.
- A representation of something in which we are interested.
- The representation is reduced---what to throw away?
- Examples?



Giovanni Alfonso Borelli (1608-1679)

- Father of biomechanics and an early modeler
- Related animals to machines and used mathematics to prove his theories
- *Likened* the action of the heart to that of a piston and reasoned that arteries must be elastic



JOH. ALPHONSI BORELLI

DE

MOTU ANIMALIUM

PARS PRIMA.

EDITIO NOVISSIMA.

JOHANNIS BERNOUILLII

DE MOTU MUSCULORUM.

LUGDUNI BATAVORUM,

Apud PETRUM VANDER AA, Bibliopolam.

ANNO M DCC X.

Representation: Models are expressed in different ways

- Verbal descriptions, e.g.,
  - The "invisible hand" in economics
  - Clonal selection theory in immunology
  - Bohr's model of the atom
- Pictures
- Mathematical equations
- Computer programs (model vs. simulation)

EXAMPLES ??

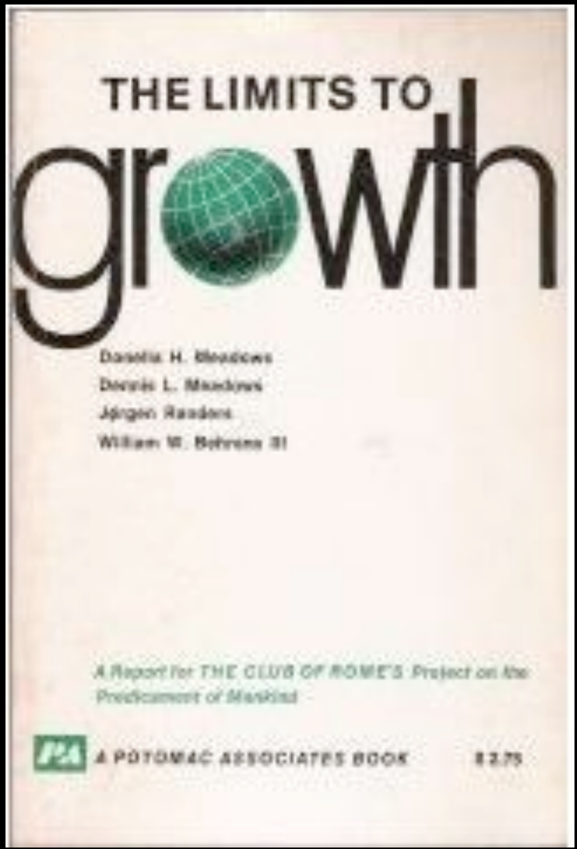
Examples

- A genetic algorithm is a model of Darwinian evolution
- Stock market simulation
- Crash test dummies (model humans)
- Social network models of social relationships
- The carbon cycle: a compartment model (air, ocean, terrestrial ecology, fossil fuels)

Example: Limits to growth model

Meadows et al. 1972

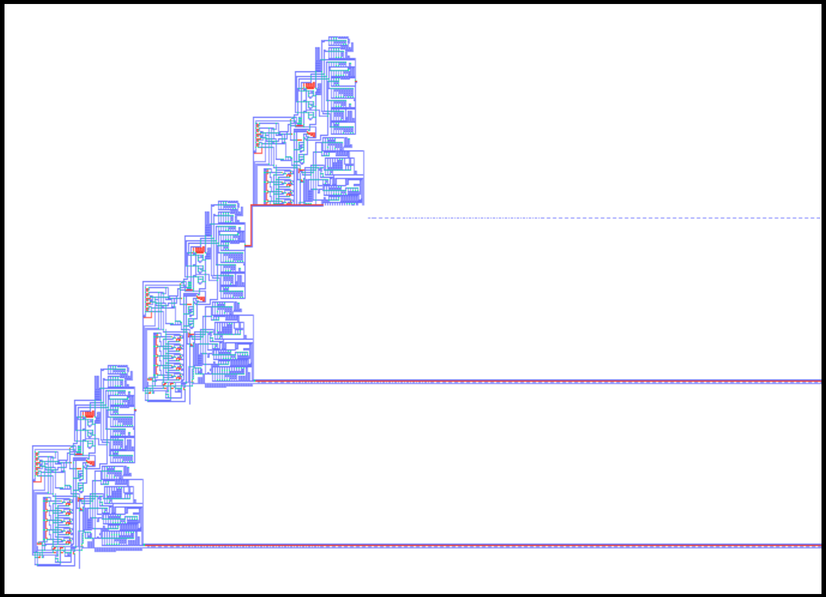
- Computer model:
  - 5 variables: populartion, industrialization, pollution, food production, resource depletion
- Goal: Not to make specific predictions, but to explore how exponential growth interacts with finite resources. Because the size of resources is not known, only general behavior can be explored.
- This process of determining behavior modes is "prediction" only in the most limited sense of the word. … These graphs are *not* exact predictions of the values of the variables at any particular year in the future. They are indications of the system's behavioral tendencies only.



How do we use models?

- Making predictions (conventional use)
  - Quantitative (most analytical models)
  - Qualitative, e.g., critical parameters, regions of stability or instability
  - Validation: Accurate predictions

How do we use models?



ftp://ftp.ira.uka.de/pub/cellular-automata/jvn

- Existence proofs
  - Demonstrate that something is possible (e.g., von Neumann's self-reproducing automaton, patents)
  - Validation: it works

How do we use models?

- Building intuitions about how complex systems work (exploratory):
  - Examples: Flight simulators, Sim City.
  - Study how patterns of behavior, how resources flow, how co-operation arises, arms races.
  - Sensitivity analysis.
  - Discovering lever points for intervention, e.g., vaccines
  - Validation? (Trust, allows a pilot to land safely).

"*The contemplation in natural science of a wider domain than the actual leads to a far better understanding of the actual.* " A.S. Eddington. (first test of Einstein's Theory of Relativity experimentally)

Not everyone agrees

- Def. “Modeling is the application of methods to analyze complex, real-world problems in order to make predictions about what might happen with various actions.” [Shiflet and Shiflet, 2006]

- “Models are metaphors that explain the world we don’t understand in terms of worlds we do. They are merely analogies, provide partial insight, stand on someone else’s feet. Theories stand on their own feet, and rely on no analogies.” [Emanuel Derman, 2012]

How do we evaluate a model?

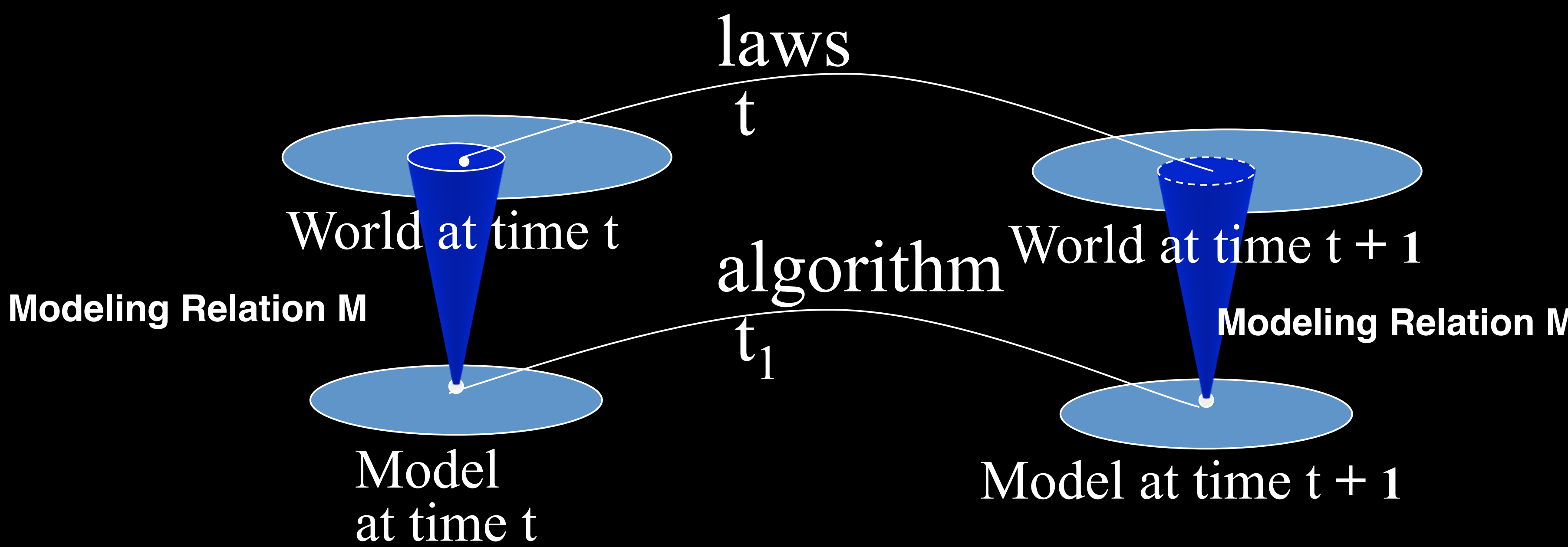
- Parsimony and simplicity
  - Occam's Razor (select the competing hypothesis that makes the fewest new assumptions, when the hypotheses are equal in other respects)
- Accuracy of predictions
  - R^2 and other statistical tests
- Dynamical model works as claimed:
  - Run it. E.g., patented devices.
- Cogency and relevance of ideas that they produce.
- Falsifiability.
- Consistency---formalize notion of model as a *homomorphic map.*

A Framework for Modeling
based on homomorphic maps

- From <u>An Introduction to Cybernetics</u> by Ross Ashby (1957)
- The black box (from electrical engineering)
  - Given a sealed box that has terminals for input and terminals for output
  - Deterine the contents (internal rules) of the box
  - Applies to any system whose internal mechanisms are not fully open to inspection, e.g., the brain
  - How should the experimenter proceed to understand the black box?
  - What properties of the box are discoverable and which are not?
  - What is the most efficient way to investigate the properties of the box?
- Scientists study only a small fraction of a complex system
  - To what extent can systems justifiably be simplified?
  - How to distinguish good simplifications (models) from bad ones?

Models as Homomorphic Maps

Commutativity of the Diagram



M is an equivalence relation

Model M is valid iff this is a homomorphic map:

$M(t(x)) = t_1(M(x))$, for all x

Remarks

- What if the inside of the box is not deterministic?
  - Search set of inputs/outputs that take more variables into account (refine the model) and see if the results is deterministic. OR
  - Look for statistical determinancy, i.e., determinancy in the averages, as in Markov chains.
- Models themselves are seldom regarded in all their detail
  - Only aspect of the model is related to system of study
  - E.g., a tin mouse may be a satisfactory model of a living system, if we ignore the "tinniness" of the model and the "proteinness" of the living mouse
  - Thus, two systems (the system of study and the model) are often related such that a homomorphism of one is isomorphic with a homomorphism of the other.

Common Modeling Assumptions

- Homogeneity (all agents are identical / stateless)
- Equilibrium (no or very simple dynamics)
- Random mixing
- No feedback (learning)
- Deterministic
- No connection between micro and macro phenomena

Models with these assumptions can produce some interesting features, e.g., tipping points (R_0).

Features of Complex Systems

- Heterogeneous agents
- Non-equilbrium (non-linear dynamics)
- Contact structure (networks, nonrandom mixing)
- Learning / Feedback (agents can change behavior)
- Stochastic behavior (interesting behavior in the tails)
- Emergence (multi-scale phenomena)

Modeling Complex Systems is Difficult

- Closed form solutions rarely exist:
  - Features from previous slide
- Detailed simulations are problematic:
  - Can never hope to get all the details correct.
  - Because systems are nonlinear, small errors can have large consequences.
- Evolution is key:
  - Basic components change over time.
  - Individual variants matter (hard to do theory).
- Discreteness (e.g., time, state spaces, and internal variable values)
  - Techniques developed to study nonlinear systems are not always directly applicable
- Spatial heterogeneity

Scientific Models

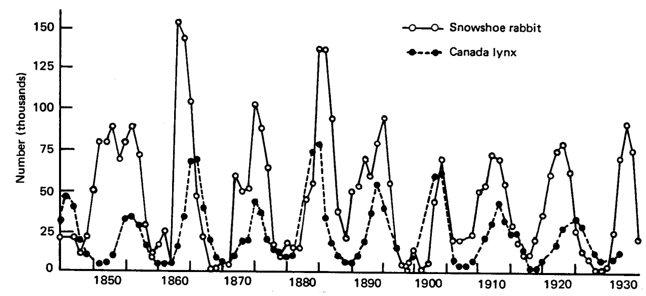
- Continuous vs. Discrete
  - E.g., Differential equation vs. Cellular automaton
- Deterministic vs. Probabilistic
  - Dynamical system vs. Markov chain
  - Cellular automaton vs. genetic algorithm
- Spatial vs. nonspatial
- Data-driven vs. theory-driven
  - Bayesian networks vs. expert system

Classical ODE
(ordinary differential equation)

- Represent how a process changes through time as a differential (difference) equation
  - Time is continuous (discrete)
  - Model components are continuous (density)
  - Deterministic
  - Nonspatial (in standard case)

$$\frac{dx}{dt} = Ax - Bxy$$

$$\frac{dy}{dt} = -Cy + Dxy$$

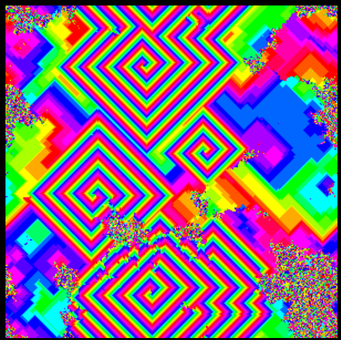


ODE Characteristics

- Describes the global behavior of a system
- Averages out individual differences (stateless)
- Assumes infinite size populations of model components
  - E.g., all possible genotypes always present in population
- Spatial homogeneity
- Easier to do theory and make quantitative predictions
- Examples
  - Maxwell's equations
  - Mackey-Glass systems
  - Lotka-Volterra systems

Cellular Automata (CA)

- Basic model
  - Regular arrangement (lattice) of cells
  - Cells are in one of a finite number of states each time step
  - All cells have the same synchronous update rule
  - Cells have a local interaction neighborhood

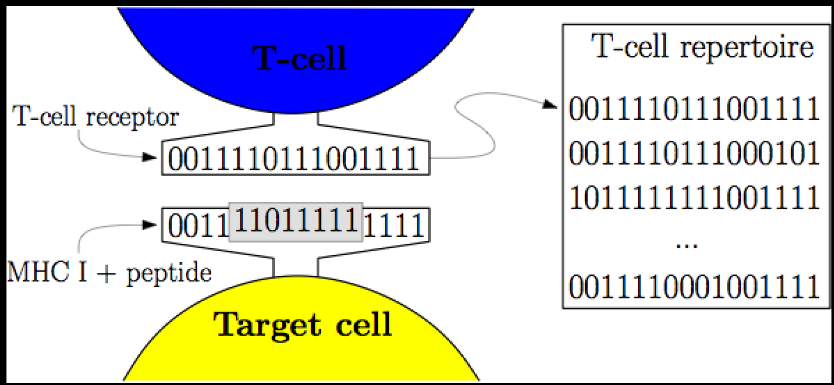


CA Characteristics

- Naturally spatial and discrete
- Individuals have state
  - But all cells have the same rule set
  - Cleverness required to design the uniform rule set
  - Correspondence: What does each cell represent?
- Mathematical analysis is challenging
- Examples
  - Hydrodynamics, fluid dynamics
  - Forest fire simulation
  - Ferro magnetic modeling (ISING models)

Agent-based Models (ABM)

- A computational artifact that captures essential components and interactions (I.e. a computer program)
  - Each component represented explicitly
  - Aids visualization
- Study the behavior of the artifact, using theory and simulation:
  - To understand its intrinsic properties, and wrt modeled system.
- CAs are a kind of minimal ABM



ABM Characteristics

- Encodes a theory about relevant *mechanisms*:
  - Mechanisms give rise to macro-properties without being build in from the beginning
  - Very different kind of explanation than simply predicting what will happen next
    - Example Cooperation emerges from iterated PD model
  - Simulation as a basic tool
  - The mechanistic theory cannot always be stated cleanly
- Many contingent behaviors
- Limited scalability (time, error rates, population sizes)
- No single answer, how to interpret results?
  - Observe distribution of outcomes
- Examples:
  - Cellular automata
  - Genetic algorithms
  - Digital immune systems
  - Sugarscape
  - Stock market models

Studying Complex Systems with ABM *(taken from Axelrod, 2005)*

- Can address problems that are fundamental to many disciplines:
  - Path dependency
  - Effects of adaptive vs. rational behavior
  - Effects of network structure
  - Cooperation among egoists
  - Diffusion of innovation
- Facilitate interdisciplinary collaboration:
  - "A prosthesis for interaction"
- Useful tool when closed form mathematical analyses are intractable
  - E.g., the evolution of sex
- Can reveal unity across disciplines
- Can be a "hard sell"
  - Reality vs. clarity

How do we build models?

- Three elements in a computational model:
  - The system of interest
  - The model
  - The computer
- Modeling: The relationship between the real system and the model
- Simulation: The relationship between the computer and the model

Basic Modeling Approaches

- Top-down
  - Analyze problem / system of interest
    - Collect data: Live with your data!
  - Formulate a model
  - Make simplifying assumptions
  - Determine variables and model relationships between them, and submodels
  - Solve/run model
  - Compare to data
- Bottom-up
  - Collect data (training data)
  - Formulate a model to account for the data
  - Test model on additional data for goodness of fit