Intelligent Control
and Cognitive Systems

Simulations, Agents and Science

Joanna J. Bryson
University of Bath, United Kingdom

Outline

- The simulation controversy in AI
- Snarky introduction to agents
- Science & modelling
- Examples of agent-based models
- Friday: case study of scientific replication

Simulation Controversy

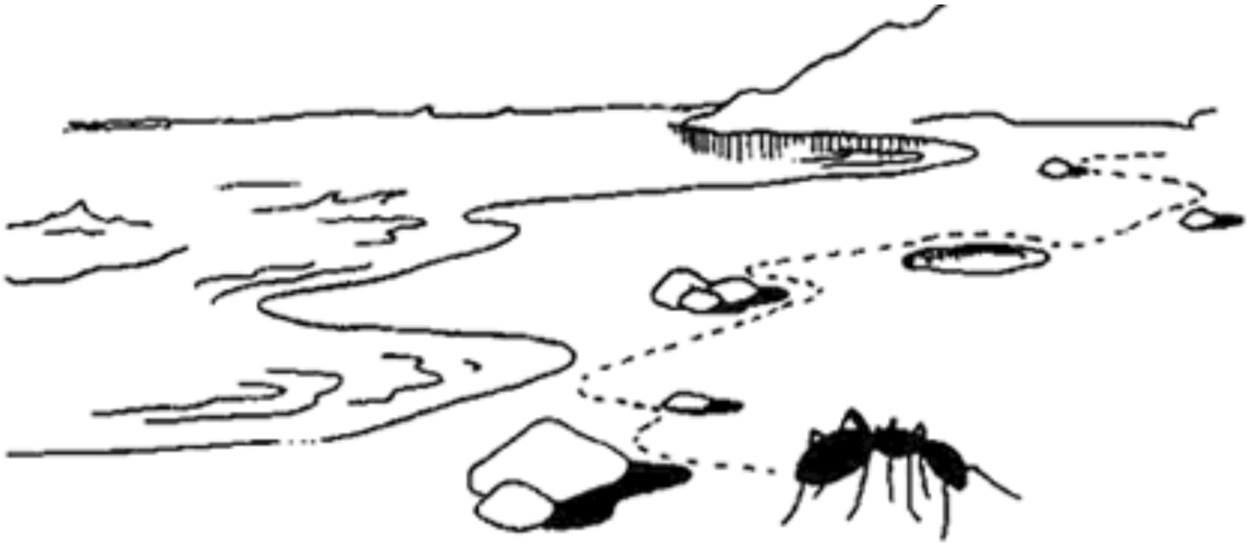
- Premise: AI has failed (so far).
- Cause: Solving the wrong problems.
- Facilitator: Simulation

(Brooks 1986, 1991)

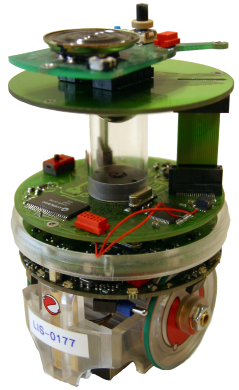
What's Wrong with Simulation?

- Simulations describe the problem.
  - If you really understand the problem, the solution is a SMOP.
  - + No AI ⇒ Getting the problem wrong.
- Simulations simpler than the real world.
  - Apparent complexity of intelligence is just a reflection of complexity of the world. Emergence from interaction.

Ubiquitous Herb Simon Ant Slide



The complexity of an ant's path on a beach is due to the beach more than the ant.



Robots

Living organisms

Partial Response to the Simulation Critique

- Simulations no longer bespoke ∴ harder to cheat.
- Robots also have orders of magnitude less input & output mechanisms than NI.
- Simulated environment should be viewed & reviewed as part of the theory.

Simulation Controversy

- Premise: AI has failed (so far). Really??
- Cause: Solving the wrong problems.
- Facilitator: Simulation

(Brooks 1986, 1991)

Recent AI Success

- Google search vs the Turing Test (David Willshaw & Bob French examples)
- Google cars (sensing, reaction, planning)
- Siri (speech and plan recognition, Internet Actions)
- Watson (learning from texts, understanding queries)

Simulation Controversy

- Premise: AI has failed (so far).
- Cause: Solving the wrong problems.
- Facilitator: Simulation

Maybe!!

(Brooks 1986, 1991)

Outline

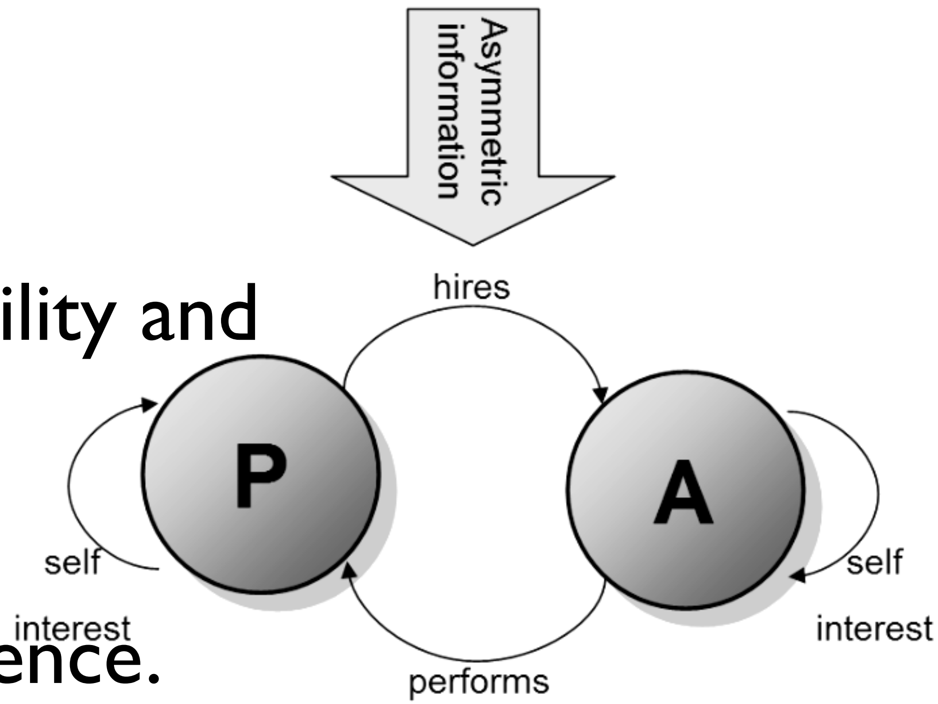
- The simulation controversy in AI
- (Somewhat) snarky introduction to agents
- Science & modelling
- Examples of agent-based models
- Friday: case study of scientific replication

What is an Agent?

- An animal or an animat (New AI), *or*
- A module of a program, treated anthropomorphically (e.g. only communicates to other parts through language, has beliefs, goals) for software engineering reasons (MAS), *or*
- A simple entity representing an individual (ABM).

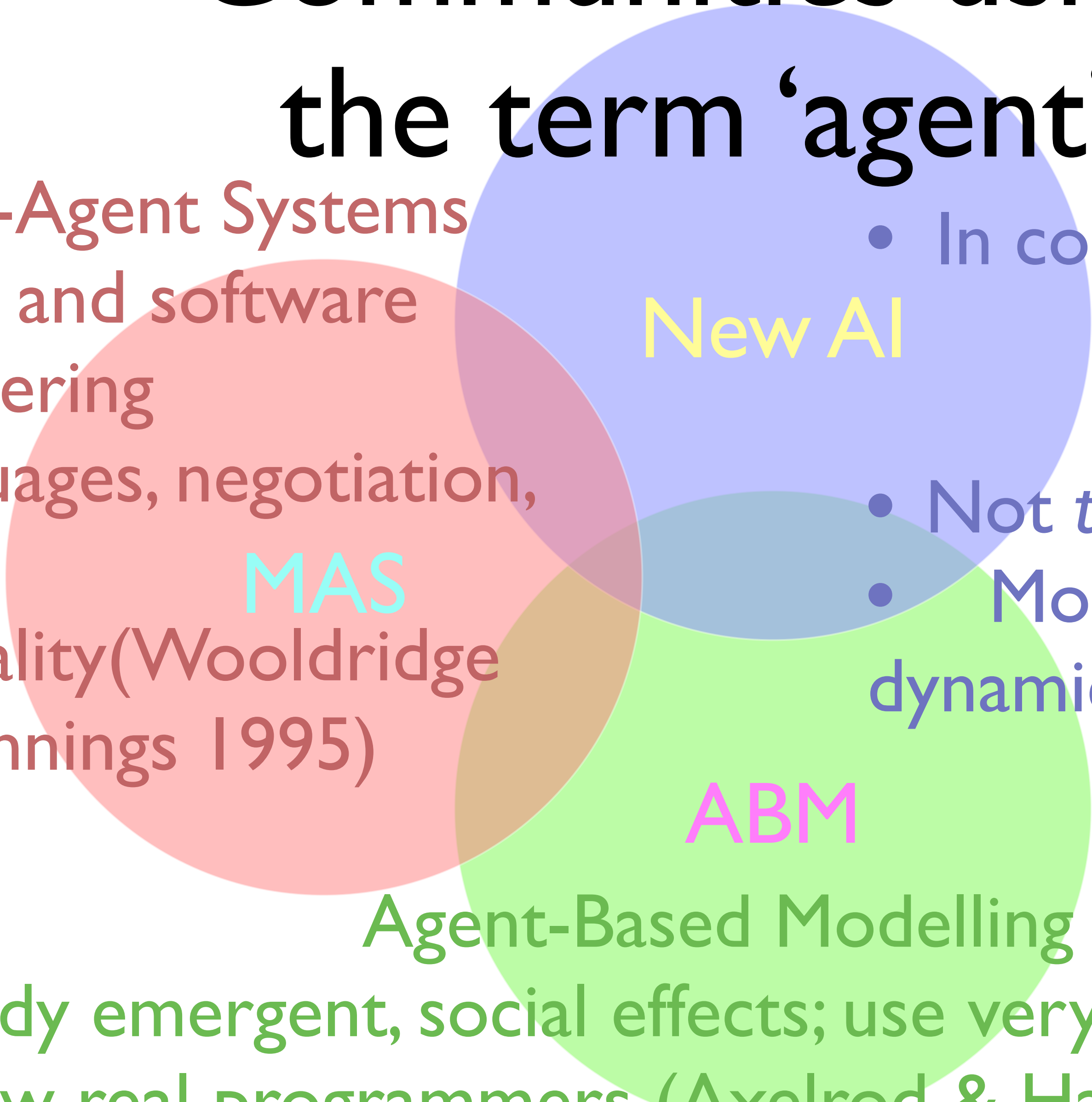
What is an Agent really?

- Philosophy defines an agent as an actor in the world, something that facilitates change,
  - e.g. chemical agents.
- Agency implies responsibility and intentionality,
  - e.g. the Principle Agent Problem in Political Science.



Mister XOOO, wikipedia

Communities using the term 'agent'



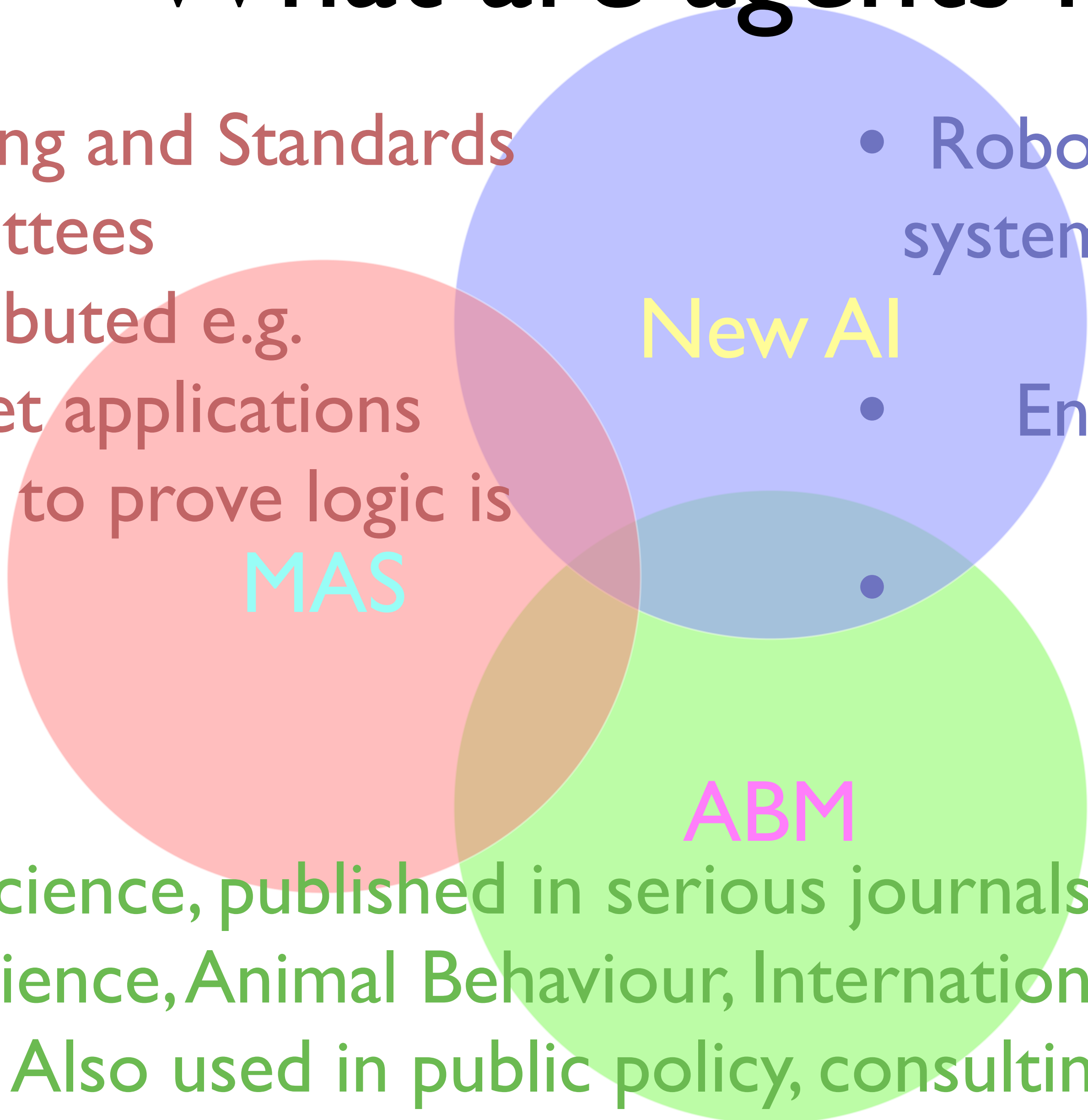
- Multi-Agent Systems
- Logic and software engineering
- Languages, negotiation, voting, optimality(Wooldridge and Jennings 1995)

- In contrast to "Good Old-Fashioned AI" (GOFAI)
- Not *that* new (1985-)
- Modular, embodied, dynamic (Brooks 1986)

- Agent-Based Modelling
- Study emergent, social effects; use very simple agents
- Few real programmers (Axelrod & Hamilton 1981)

What are agents for?

- Funding and Standards committees
- Distributed e.g. Internet applications
- Want to prove logic is useful.



- Robots and cognitive systems (e.g. Roomba, Aibo, iCub)
- Entertainment, VR, games
- Want to create human-level AI.

- Science, published in serious journals, e.g. Nature, Science, Animal Behaviour, International Relations.
- Also used in public policy, consulting, logistics.
- Want to be the next Operations Research (OR).

Outline

- The simulation controversy in AI
- Snarky introduction to agents
- Science & modelling
- Examples of agent-based models
- Friday: case study of scientific replication

Simulations as Science

- A simulation is a hypothesis like any other.
  - Thesis / model specified so completely it can be run on a computer.
- Consequences of model assessed by sampling.
  - Model behaviour compared to target system's using standard hypothesis testing.

Simulations as Good Science

- The output of a model is not data about the world!
    - Data about the hypothesis.
    - Predictions of the hypothesis.
- Simulations are new, some people make mistakes here (e.g. Hemelrijk et al, *Behaviour* 2005; cf. de Vries, *Behaviour*, 2009).

Modelling as Science

- Simulations are one form of modelling.
  - Other forms of modelling have been around longer, e.g. differential equations.
- Excellent text on modelling: Kokko (2007), Modelling for Field Biologists, CUP.
  - "We use models because our brains aren't big enough to understand all the consequences of our theories."

The Map of Germany Problem

- People (not just Brooks) often complain that a model leaves out a salient detail.
- A map of Germany that leaves out no details is the same size as Germany.
- Akin to overfitting–utility requires generality.
- Need to know a model's purpose.

Agent-Based Modelling

- Describe essential features of the environment.
- Specify the behavior of individuals.
- See if the consequences of individuals acting in an environment are what you predicted.
- (Examples soon.)

Science with ABM

- As with any theory, be as general as you can be and still get the behaviour you are trying to explain.
- If two models both predict data equally well, the simplest model wins.

Occam's Razor

Science Is Never That Easy!

- "Be as general as you can be and still get the behavior you are trying to explain."
  - In fact, may start at level of intuition, then simplify.
- "If two models both predict data equally well, the simplest model wins."
  - Simplicity/accuracy tradeoff can be tricky.

True of All Science, Not Just ABM

- "Be as general as you can be and still get the behavior you are trying to explain."
  - In fact, may start at level of intuition, then simplify.
- "If two models both predict data equally well, the simplest model wins."
  - Simplicity/accuracy tradeoff can be tricky.

Science special to ABM

"Brains big enough", e.g. Whitehouse et al. 2012

- Just trying to build the model may make you realize there were things you didn't know about your target system.

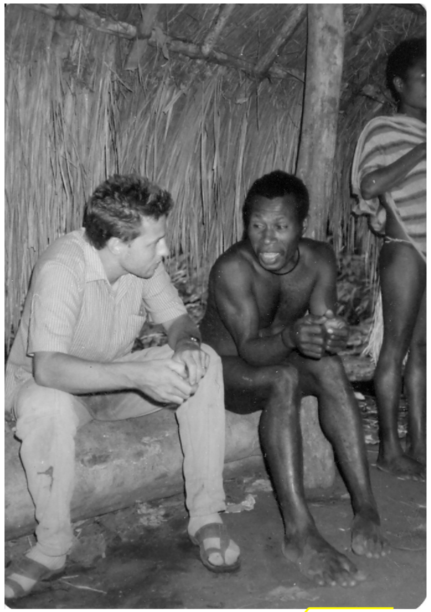
Science special to ABM

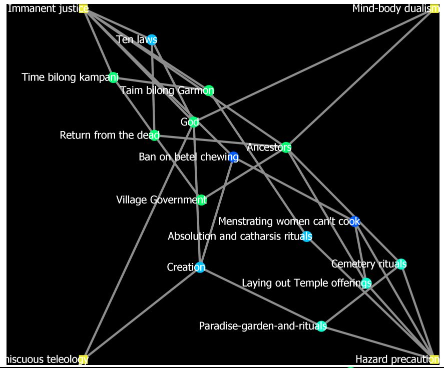
"Brains big enough", e.g. Whitehouse et al. 2012

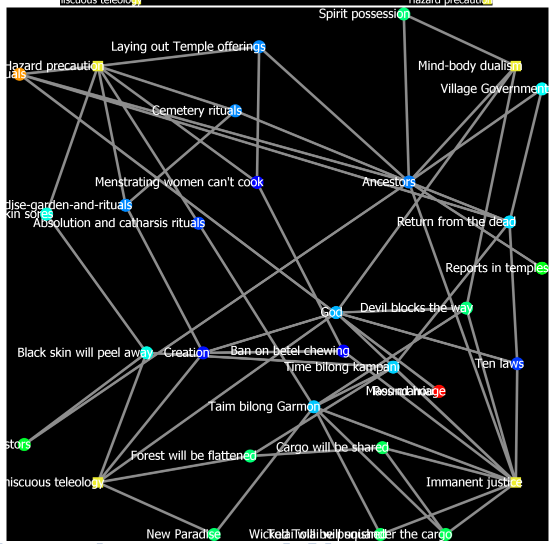
- Just trying to build the model may make you realize there were things you didn't know about your target system.
- If you match the world in more ways than you predicted, then this is convergent evidence for your theory.

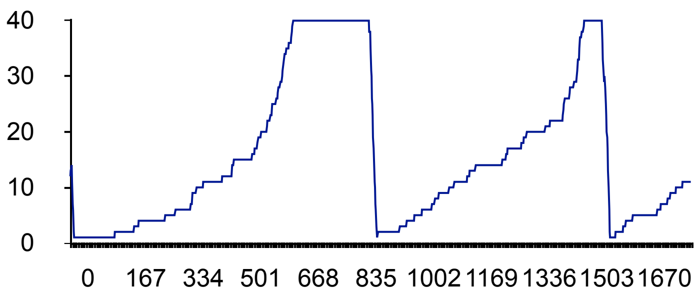
Specification as Theory Building

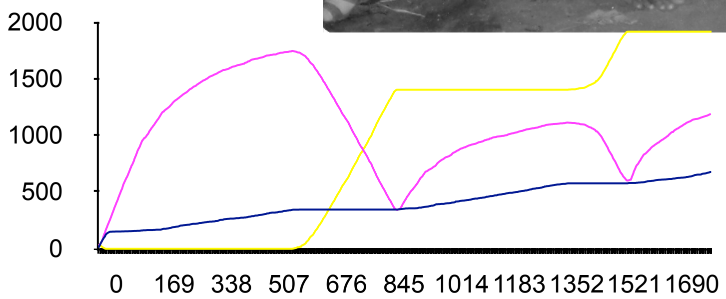
- Theory building is an essential step of science.
- The process of building a simulation may uncover incompleteness or fallacies in a model.





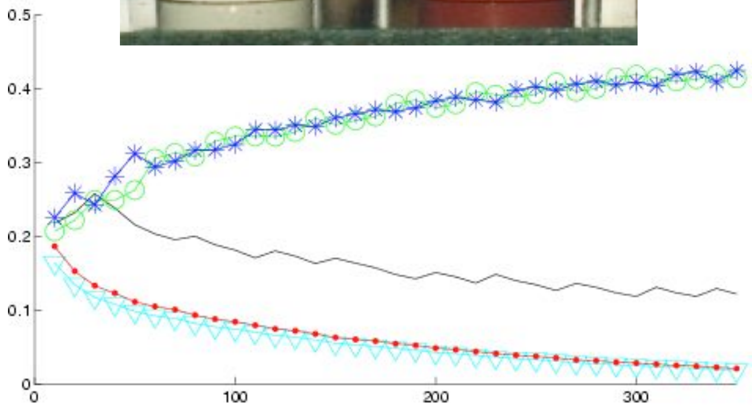


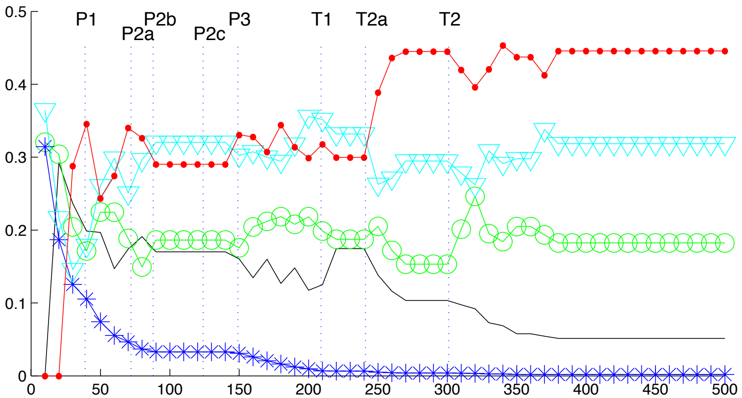


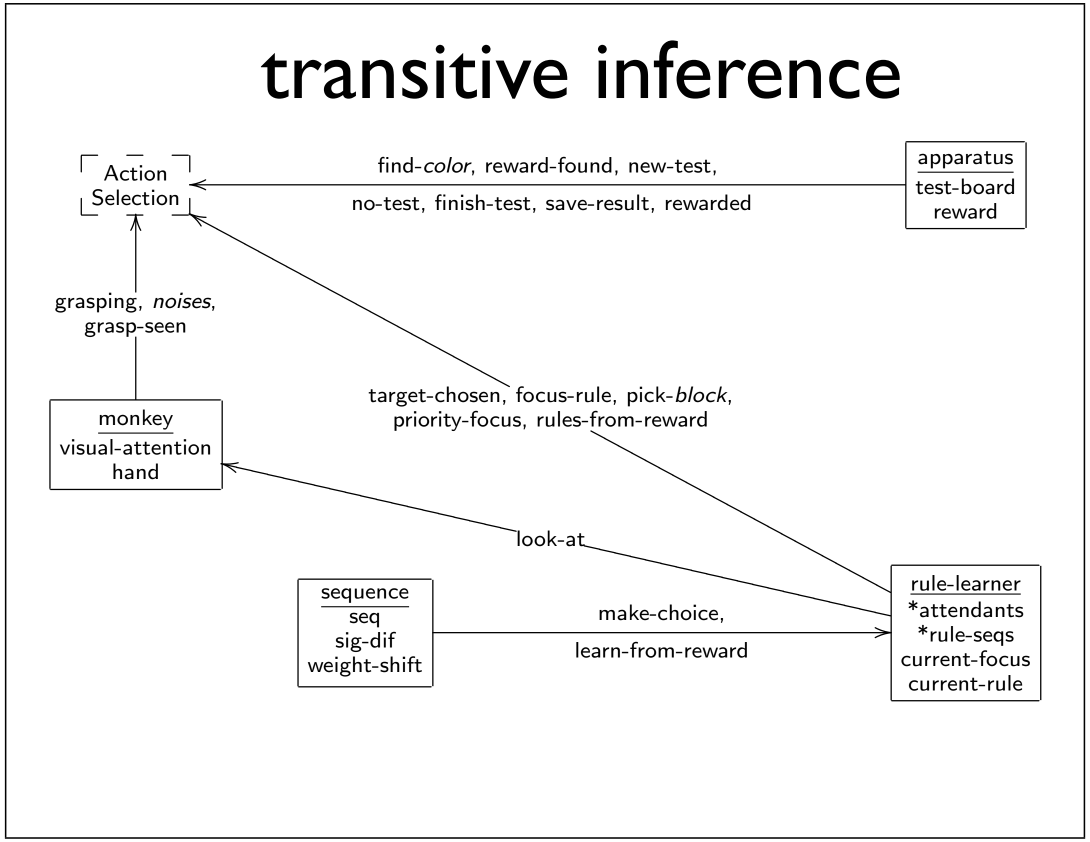


e.g. Whitehouse's Modes Theory of Religiosity (Whitehouse, Khan, Hochberg & Bryson 2012)

“Emergent” Outcomes Add Evidence







(Bryson & Leong, *Animal Cognition* 2007)

Joanna J. Bryson and Jonathan C. S. Leong “Primate Errors in Transitive ‘Inference’” Animal Cognition, **10**(1):1–15, January 2007.

Comparison to Data

Table 2 Production-rule-stack equivalents to solutions by *Saimiri sciureus* subjects (last column) and by two-tier AI subjects undergoing various forms of training[a]

		Regime starting *ED*		Regime starting *AB*		Starting *AB*
	No regime	After training	After testing	After training	After testing	McGonigle and Chalmers (1992)
$s(A)s(B)s(C)$	8	51	41	–	–	–
$s(A)s(B)a(E)$	12	68	26	–	–	–
$s(A)a(E)a(D)$	3	–	1	4	2	2
$s(A)a(E)s(B)$	7	4	16	3	1	2
$a(E)a(D)s(A)$	9	–	1	57	50	–
$a(E)a(D)a(C)$	8	–	–	59	47	1
$a(E)s(A)a(D)$	7	3	–	4	11	–
$a(E)s(A)s(B)$	2	1	13	–	3	–
Total correct	56	127	98	127	114	5
Total	288	144	144	144	144	7

[a]The distribution of solutions for two-tier agents is strongly determined by the order training pairs are presented. The analysis of the live monkeys' correlated stacks reported in the last column was performed by Harris and McGonigle (1994)

(Bryson & Leong, *Animal Cognition* 2007)

Joanna J. Bryson and Jonathan C. S. Leong "Primate Errors in Transitive 'Inference'" Animal Cognition, **10**(1):1–15, January 2007.

You don’t understand it
if you can’t build it.

Josh Epstein, Brookings Institute

Examples

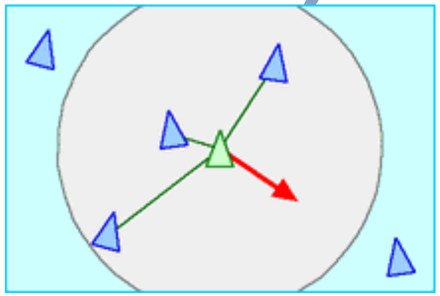
google also “flocking starlings”

Flocking

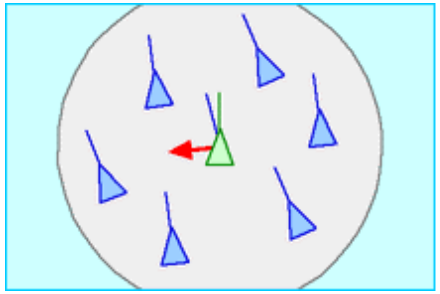
http://researchinprogress.tumblr.com/

Boids (Reynolds 1987)

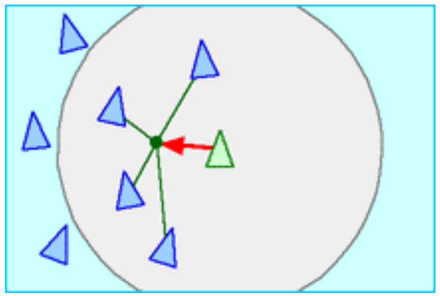
- Separation: avoid crowding local flockmates



- Alignment: steer towards the average heading of local flockmates



- Cohesion: move toward the average position of local flockmates



COURSE: 07

COURSE ORGANIZER: DEMETRI TERZOPOULOS

"BOIDS DEMOS"

CRAIG REYNOLDS
SILICON STUDIOS, MS 3L-980
2011 NORTH SHORELINE BLVD.
MOUNTAIN VIEW, CA 94039-7311

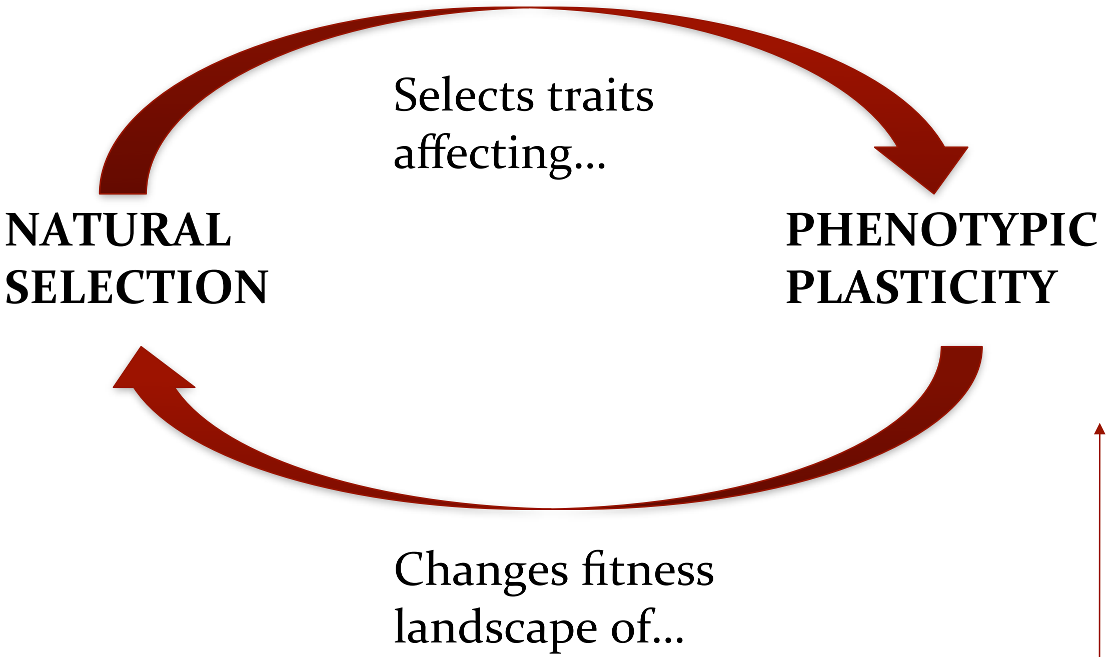
The Baldwin Effect: History

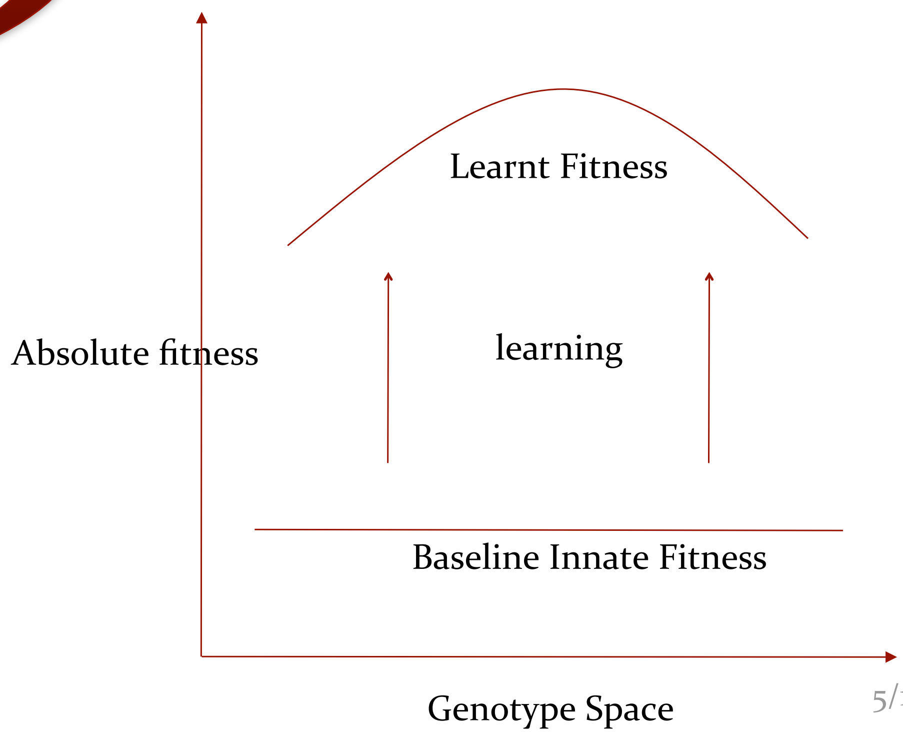
- ‘the effect through which an initially learned response to environmental change evolves a genetic basis’
- Late 1800’s intellectual context:
  - Fossil record shows clear signs of rapid, directed evolution.
  - Natural selection is neither fast, nor directed.
  - Lamarckism has been discredited by ‘Weissman barrier’.
- Baldwin (1896), Morgan (1896) and Osborn (1896) proposed that learning might indirectly support rapid and seemingly directed evolution.
- Controversial, important early application of AI simulation (Hinton & Nowlan 1987; Maynard Smith 1987; Borenstein 2006; Paenke 2008).

The Baldwin Effect: How it Works

- Information from individual learning cannot pass into the genome directly.
- However, it can have impact on the lifetime fitness of an individual.
- Hence, it can increase (or decrease) the fitness difference between genotypes.
- This will accelerate (or decelerate) the rate of genetic change.

The Baldwin Effect Illustrated





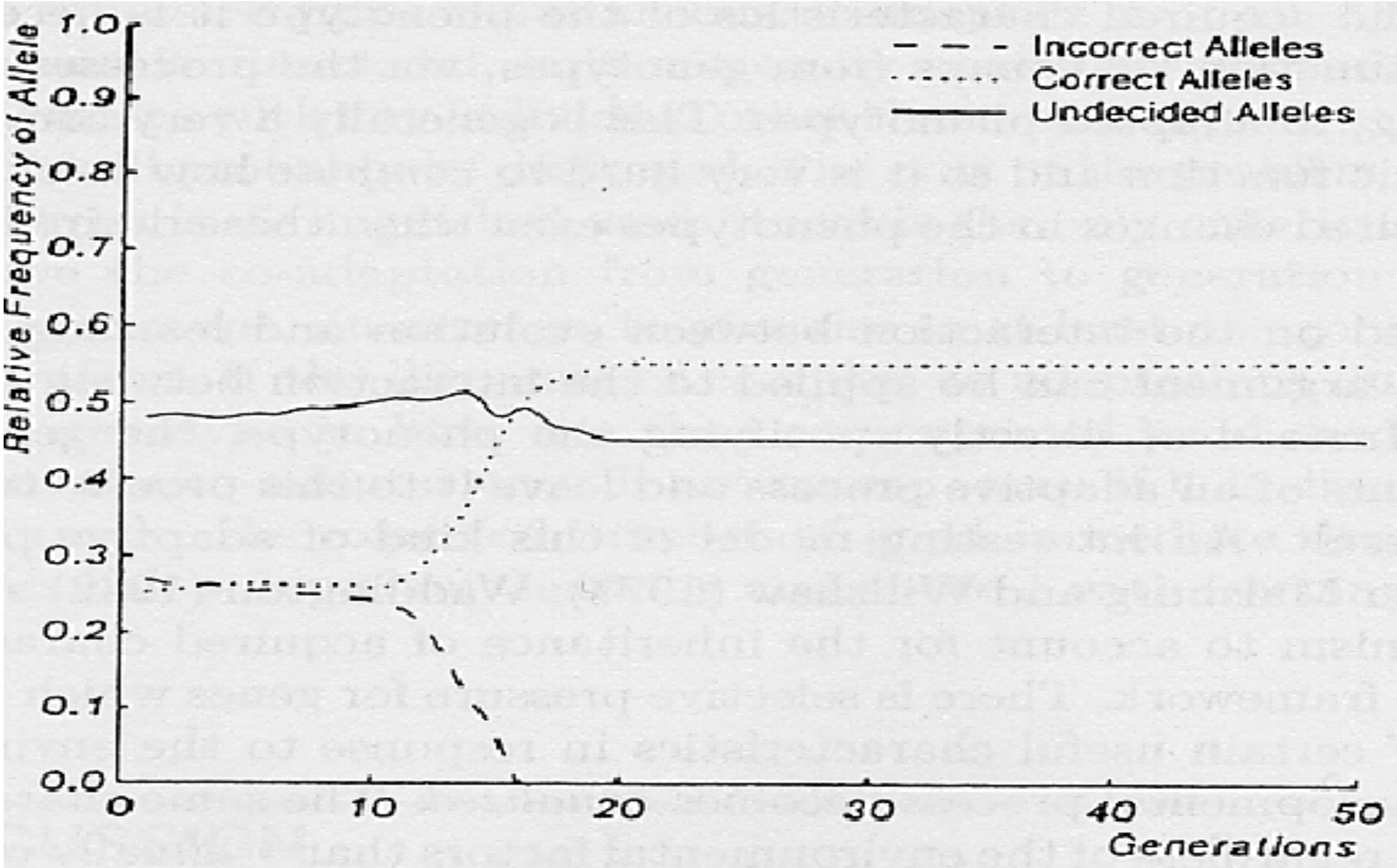
Evidence: Hinton & Nowlan's (1987) Simulation

- Hypothetical organism with 20 two-valued traits, each associated with a gene.
- Fitness improved only if all 20 traits have the advantageous value.
- The genes can have three alleles:
  - Advantageous (represented as 1)
  - Deleterious (represented as 0)
  - Plastic (represented as ?)

Evidence: Hinton & Nowlan's (1987) Simulation

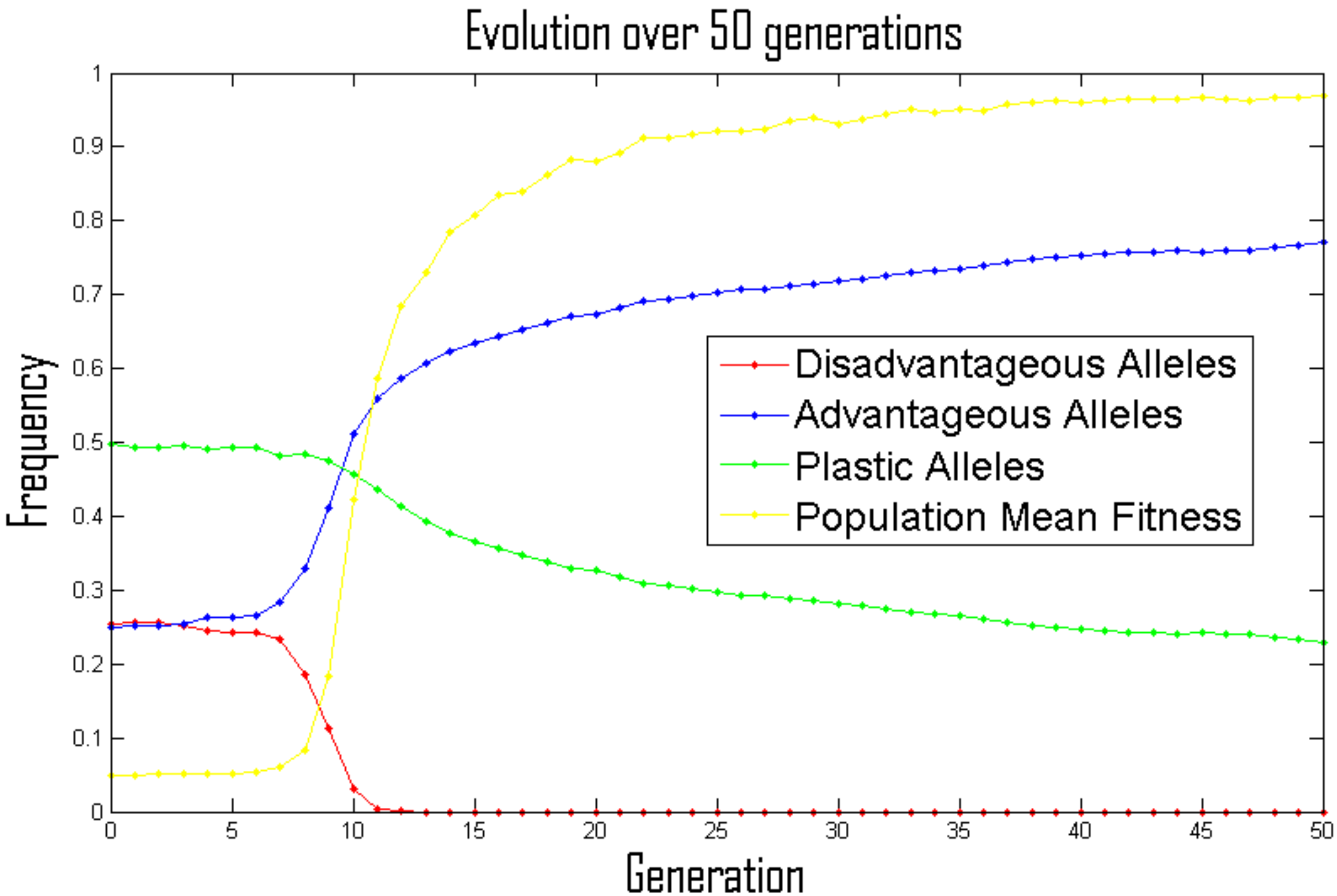
- Learning (within generation)
  - 1,000 learning trials during reproductive lifespan.
  - Each learning trial all plastic loci randomly replaced with a trait equivalent to 1 (adaptive) or 0 (deleterious) until / unless optimum genome found.
- The Organism
  - **Does not know** which trait values are 'advantageous'.
  - **Does know** when it has found the 'adaptive phenotype'.
  - Fitness payoff for learned phenotype proportionate to amount of lifetime remaining after discovery.
- Reproduction: Sexual, Single-point crossover, No mutation.

Original Results



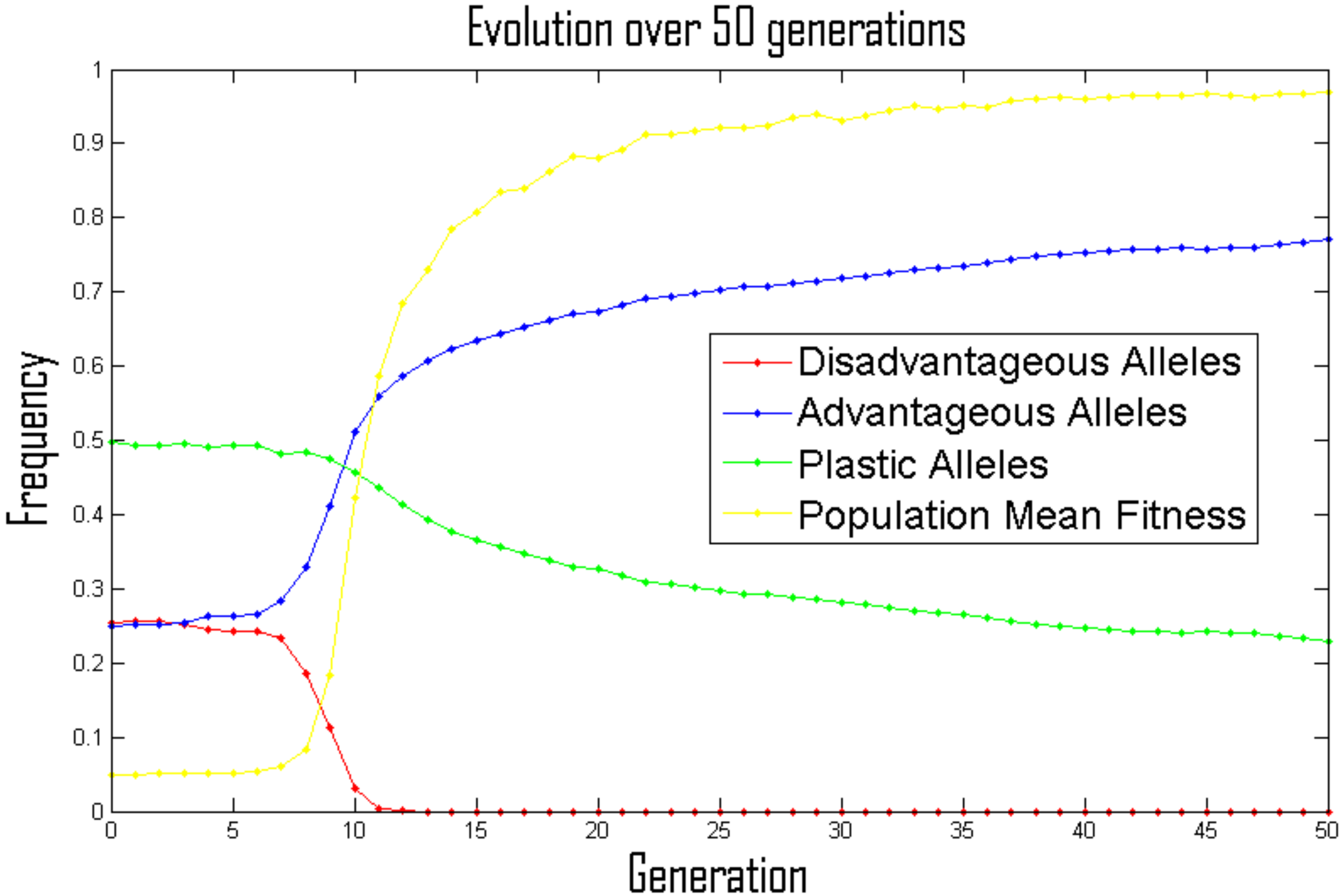
Hinton & Nowlan (1987)

Hinton & Nowlan: Results



Richards MSc (Bath) 2008

Hinton & Nowlan: Results



Richards MSc (Bath) 2008

Hinton & Nowlan: Results

- Learning accelerates evolution.
  - Problem takes 1,000's of generations to solve by genetic evolution alone; would overshoot.
- Evolution selects against learning when learning is costly (less reliable than a genetic solution).
- Learning decelerates evolution when learning is cheap (almost as reliable as a genetic solution), maintains variation.

cf. Maynard Smith *Nature* 1987