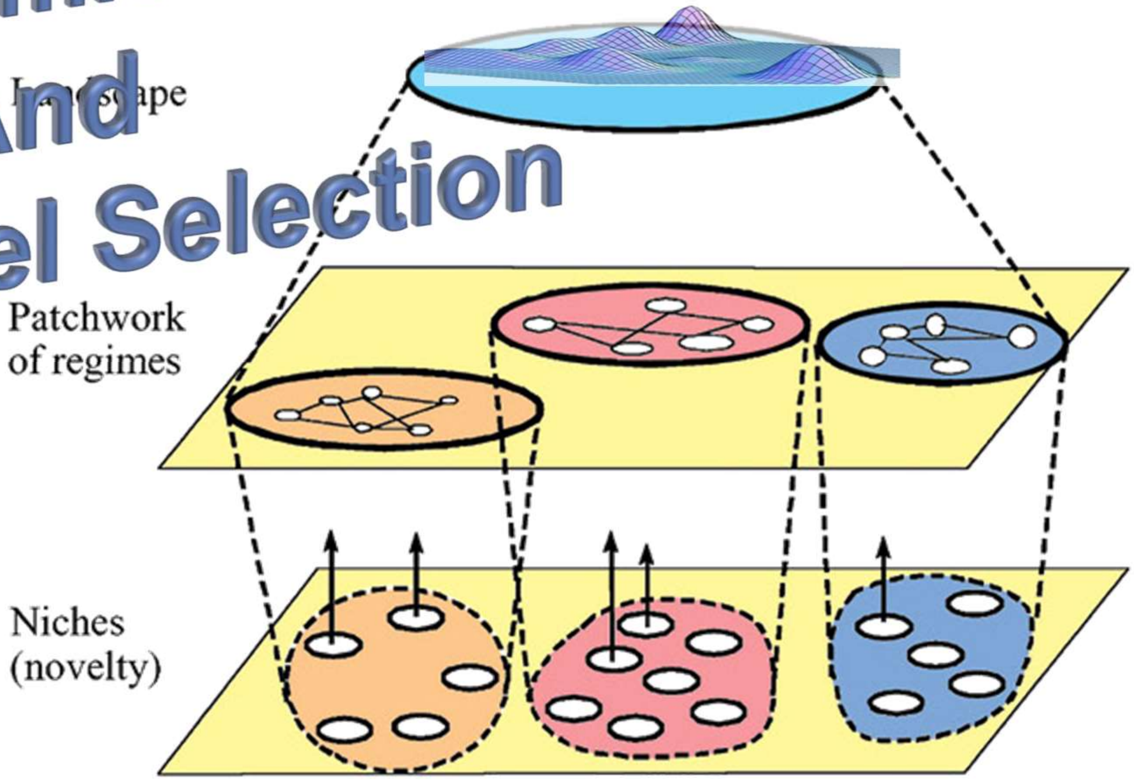
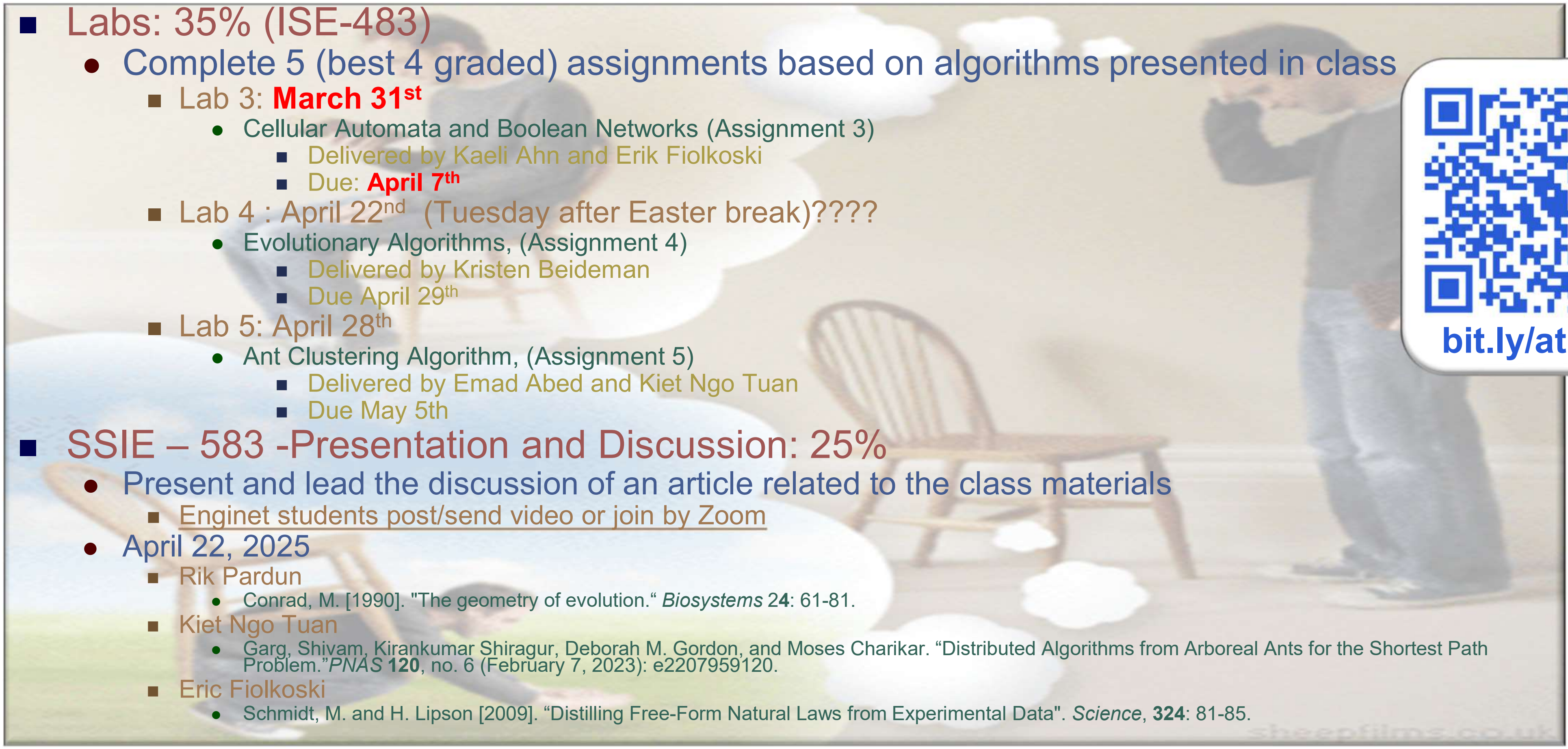
Adding Genomic Complexity And Multi-level Selection



bit.ly/atBIC

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

key events coming up



- Labs: 35% (ISE-483)
  - Complete 5 (best 4 graded) assignments based on algorithms presented in class
    - Lab 3: **March 31st**
      - Cellular Automata and Boolean Networks (Assignment 3)
        - Delivered by Kaeli Ahn and Erik Fiolkoski
        - Due: **April 7th**
    - Lab 4 : April 22nd (Tuesday after Easter break)????
      - Evolutionary Algorithms, (Assignment 4)
        - Delivered by Kristen Beideman
        - Due April 29th
    - Lab 5: April 28th
      - Ant Clustering Algorithm, (Assignment 5)
        - Delivered by Emad Abed and Kiet Ngo Tuan
        - Due May 5th
- SSIE – 583 -Presentation and Discussion: 25%
  - Present and lead the discussion of an article related to the class materials
    - Enginet students post/send video or join by Zoom
  - April 22, 2025
    - Rik Pardun
      - Conrad, M. [1990]. "The geometry of evolution." *Biosystems* **24**: 61-81.
    - Kiet Ngo Tuan
      - Garg, Shivam, Kirankumar Shiragur, Deborah M. Gordon, and Moses Charikar. "Distributed Algorithms from Arboreal Ants for the Shortest Path Problem." *PNAS* **120**, no. 6 (February 7, 2023): e2207959120.
    - Eric Fiolkoski
      - Schmidt, M. and H. Lipson [2009]. "Distilling Free-Form Natural Laws from Experimental Data". *Science*, **324**: 81-85.

bitly

bit.ly/atBIC

final project schedule

- Projects
  - **Due by May 7th** in Brightspace, “Final Project 483/583” assignment
    - ALIFE 2025
      - Not necessarily to submit to actual conference due date
        - May 4 full paper, July 4, abstract
      - https://2025.alife.org/
      - Max 8 pages, author guidelines:
      - https://2025.alife.org/calls#paper-call
      - MS Word, PDF, and Latex/Overleaf templates.
  - Individual or group
    - With very definite tasks assigned per member of group

ALIFE 2025

Tackle a real problem using bio-inspired algorithms, such as those used in the labs.

bit.ly/atBIC

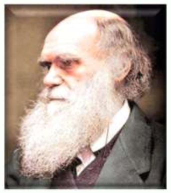
Reusing and expanding labs is highly encouraged.

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK

rocha@indiana.edu
casci.binghamton.edu/academics/i-bic

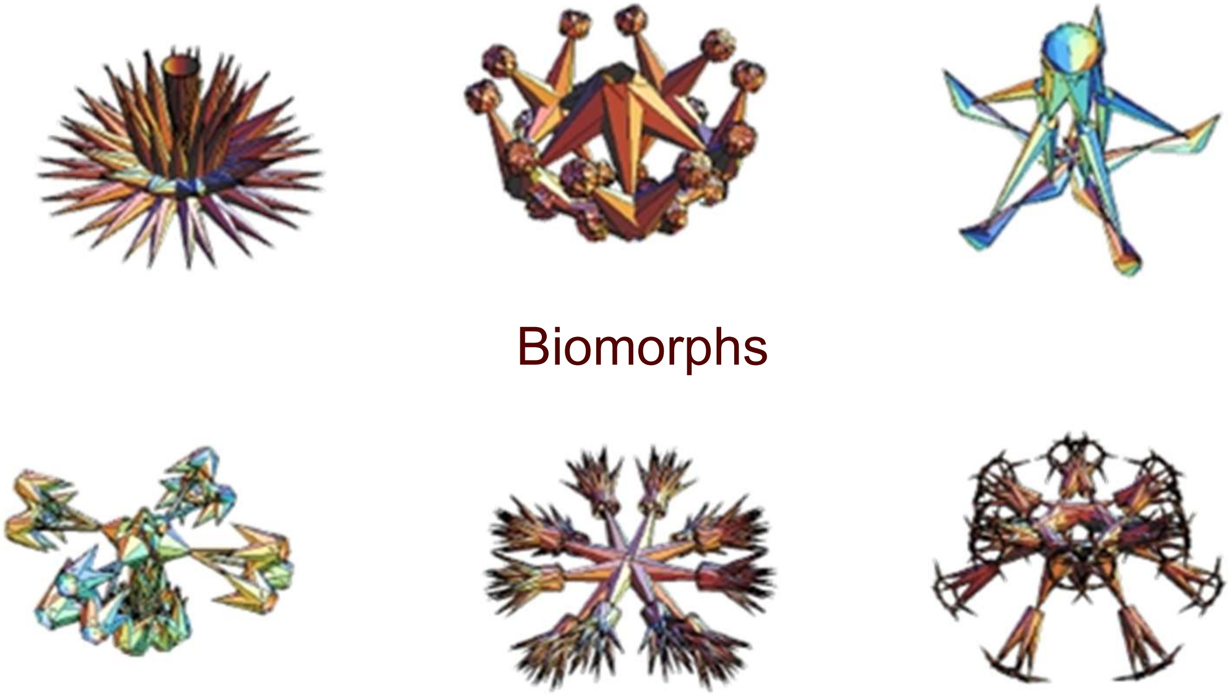
until now

- Class Book
  - Floreano, D. and C. Mattiussi [2008]. *Bio-Inspired Artificial Intelligence: Theories, Methods, and Technologies*. MIT Press.
    - Chapters 1, 4, **and 7**
- Lecture notes
  - Chapter 1: What is Life?
  - Chapter 2: The logical Mechanisms of Life
  - Chapter 3: Formalizing and Modeling the World
  - Chapter 4: Self-Organization and Emergent Complex Behavior
  - Chapter 5: Reality is Stranger than Fiction
  - Chapter 6: Von Neumann and Natural Selection
  - Chapter 7: Modeling Evolutionary Systems
    - posted online @ http://informatics.indiana.edu/rocha/i-bic
- Papers and other materials
  - Optional
    - Nunes de Castro, Leandro [2006]. *Fundamentals of Natural Computing: Basic Concepts, Algorithms, and Applications*. Chapman & Hall.
      - Chapter 2, 7, 8
      - **Chapter 3, sections 3.1 to 3.5**



objective function may be subjective

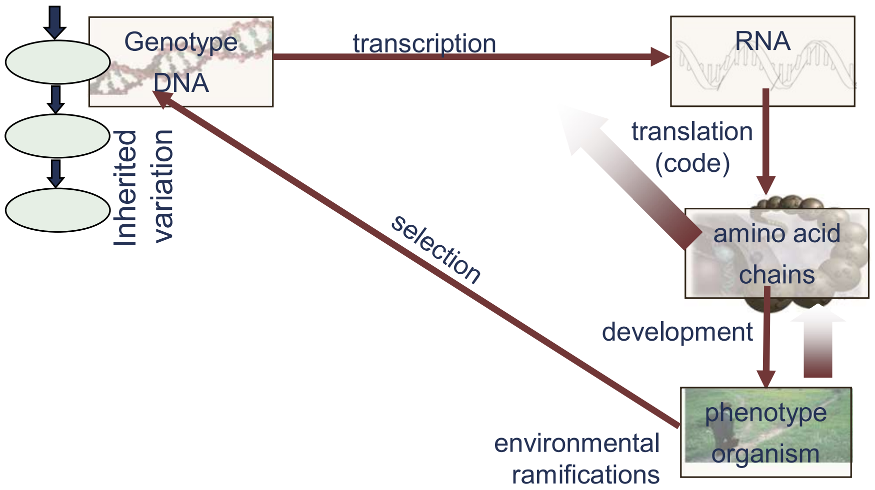
> "Once a Darwinian process gets going in a world, it has an open-ended power to generate surprising consequences: us, for example" Richard Dawkins



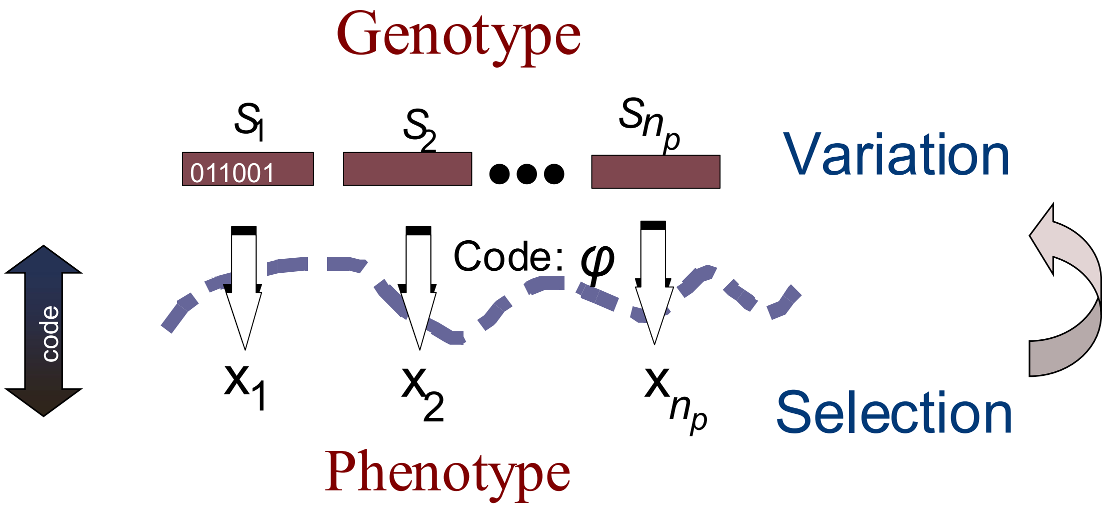
Biomorphs

artificial genotype/phenotype mapping

- Search algorithms based on the mechanics of Natural Selection
  - Based on distinction between a machine and a description of a machine
  - Solution alternatives for optimization problems

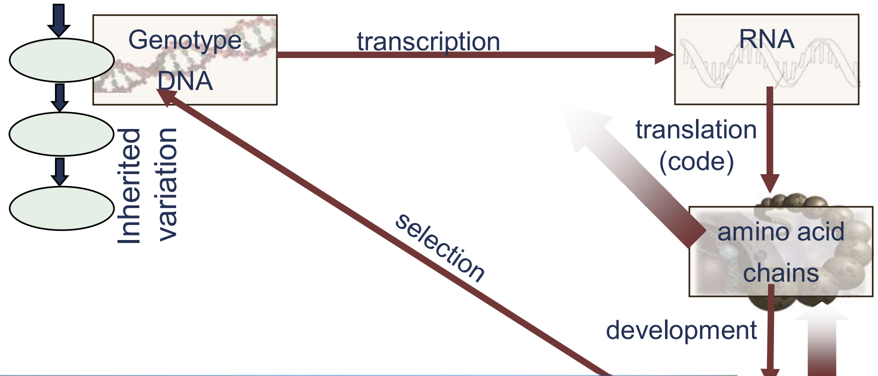


Traditional Genetic Algorithm

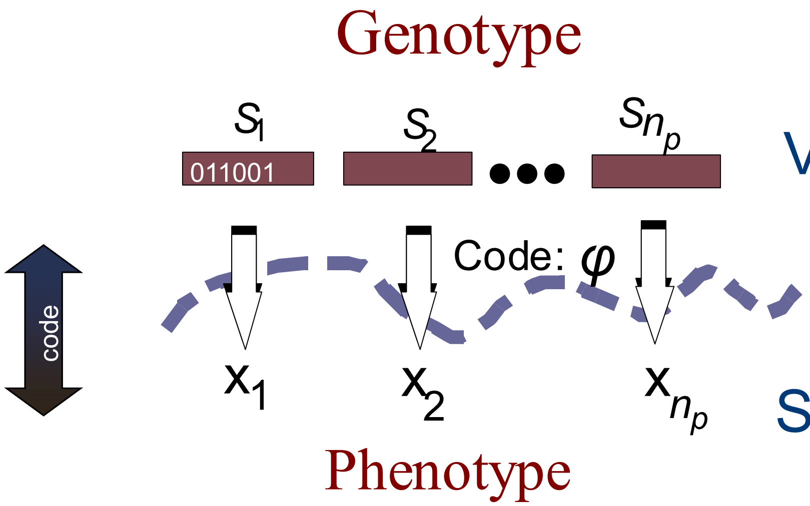


artificial genotype/phenotype mapping

- Search algorithms based on the mechanics of Natural Selection
  - Based on distinction between a machine and a description of a machine
  - Solution alternatives for optimization problems



Traditional Genetic Algorithm



Variat

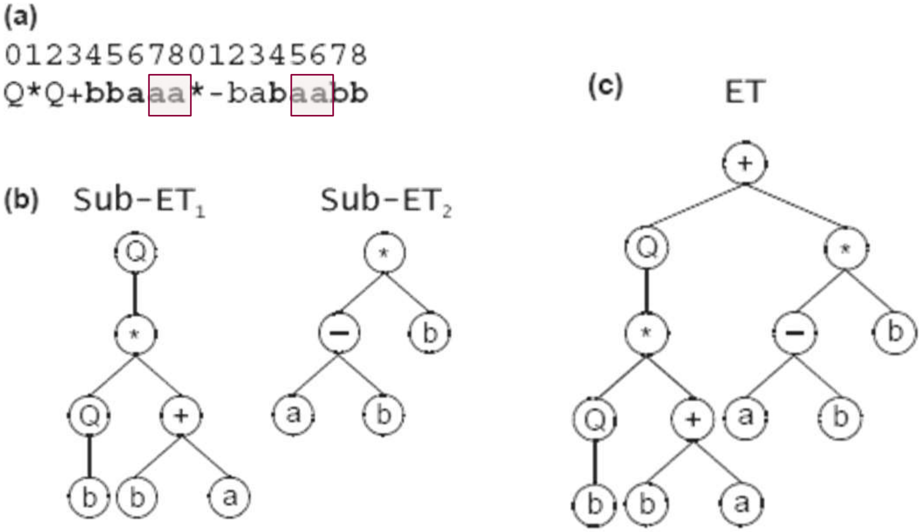
Select

Including a genotype/phenotype map in GP

- Proposed by Candida Ferreira
- Program trees are encoded in fixed-length linear genotypes
- Genotypes
    - Open-reading frame architecture
        - Stop signal not necessarily at end of genotype
    - Non-coding genes are possible
        - Can include genetic operators
    - Genes contain two types of symbols
        - Functions (only at the head) and terminals
    - Multigenic solutions
        - Assembled from non-coding operations between various open-reading frames

012345678012345678012345678
-b*babbab*Qb+abbba-*Qabbaba

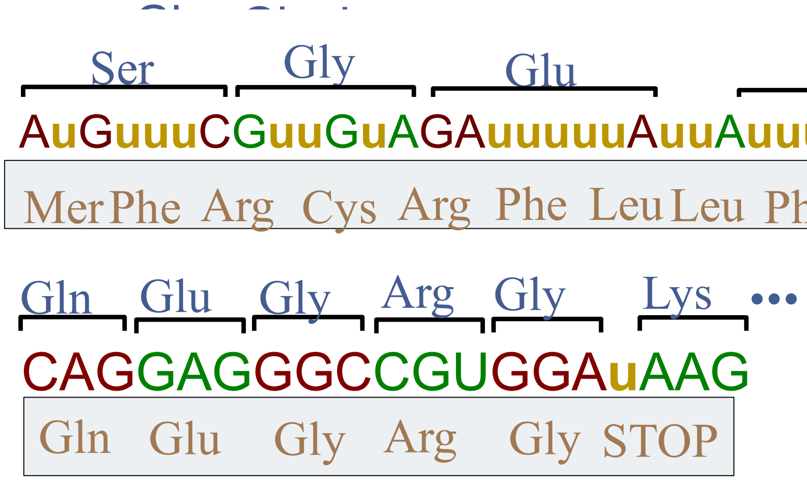
C. FERREIRA [2001]. Gene Expression Programming: A New Adaptive Algorithm for Solving Problems. *Complex Systems*, **13** (2): 87-129.



BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@indiana.edu
casci.binghamton.edu/academics/i-bic

RNA Editing

Including transcription regulation and translation



Ser Gly Glu Lys

AuGuuuCGuuGuAGAuuuuuAuuAuuuuuuuuAuuA

Mer Phe Arg Cys Arg Phe Leu Leu Phe Phe Leu Leu

Gln Glu Gly Arg Gly Lys •••

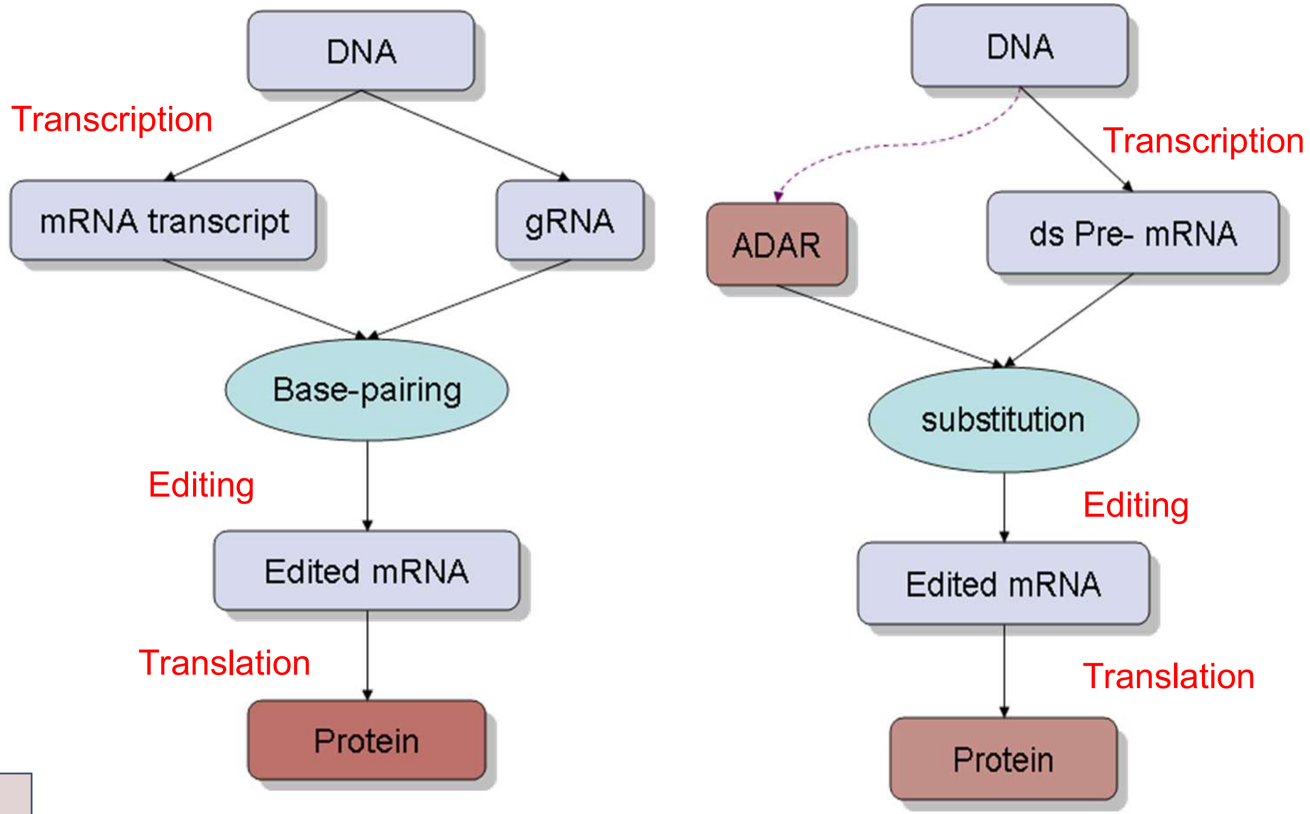
CAGGAGGGCCGUGGAuAAG

Gln Glu Gly Arg Gly STOP

Bass, B.L. (Ed.) (2001). *RNA Editing*.
Benne, R. (Ed.) (1993). *RNA Editing*

- RNA Editing: post-transcriptional alteration of genetic information
  - can be performed by ncRNA structures and proteins (i.e enzyme cascades).
- U-Insertion/deletion RNA Editing (mitochondria of kinetoplastid protozoa)
  - Involve small guide RNAs (gRNA) complementary to the target mRNA, and editosome (multi-protein complex)
    - gRNA is a template for editing
  - insertion/deletion of Uracil (U) residues, usually in coding regions of mRNA transcripts
    - e.g. creation of open reading frames

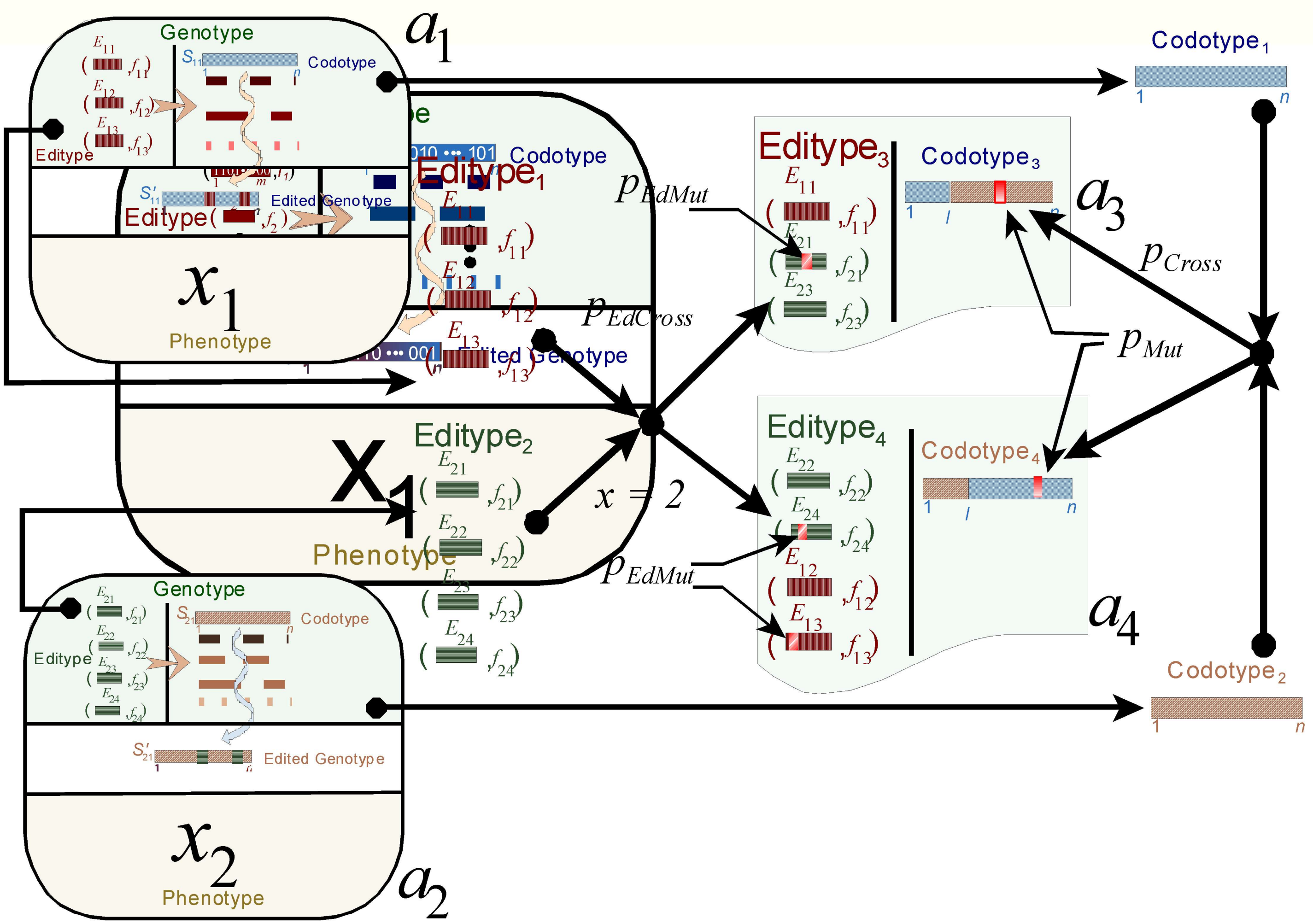
U-insertion and A-to-I subtitution

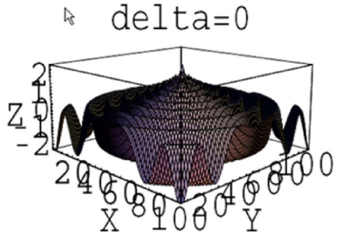


is there a **general principle** at play?

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

RNA Editing





Huang, C-F, et al [2007]. *Evolutionary Computation*. **15**(3).

rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

Simple fitness function

$$\begin{aligned}
s_1 &= 11111{*};\ c_1 = 10\\
s_2 &= {*}{*}{*}{*}{*}11111{*};\ c_2 = 10\\
s_3 &= {*}{*}{*}{*}{*}{*}{*}{*}{*}{*}11111{*};\ c_3 = 10\\
s_4 &= {*}{*}{*}{*}{*}{*}{*}{*}{*}{*}{*}{*}{*}{*}{*}11111{*};\ c_4 = 10\\
s_5 &= {*}11111{*}{*}{*}{*}{*}{*}{*}{*}{*}{*}{*}{*}{*}{*}{*};\ c_5 = 10\\
s_6 &= {*}11111{*}{*}{*}{*}{*}{*}{*}{*}{*}{*};\ c_6 = 10\\
s_7 &= {*}11111{*}{*}{*}{*}{*};\ c_7 = 10\\
s_8 &= {*}11111;\ c_8 = 10
\end{aligned}$$

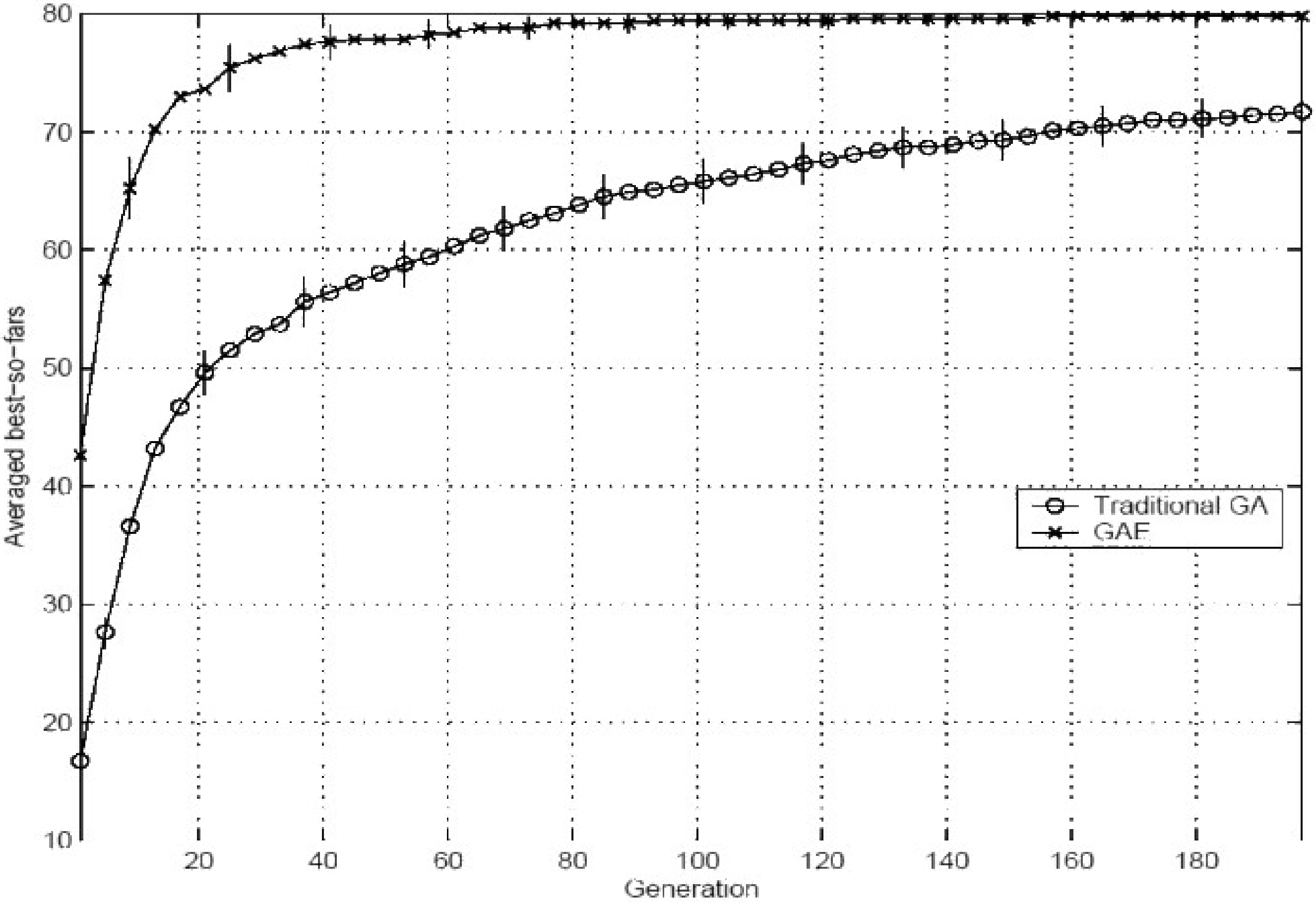
Royal road

$$F(x) = \sum_{s_i \in S} c_i \sigma_{s_i}(x)$$

- Miniature of "Royal Road" function (Forrest and Mitchell, 1993)
    - Schemata $S = (s_1, \ldots, s_8)$
    - c_i is a value assigned to each schema s_i
    - $\sigma_{si}(x)$ = 1 if x is an instance of s_i and 0 otherwise
    - Fitness of the global optimum string (40 1's) is 10*8 = 80

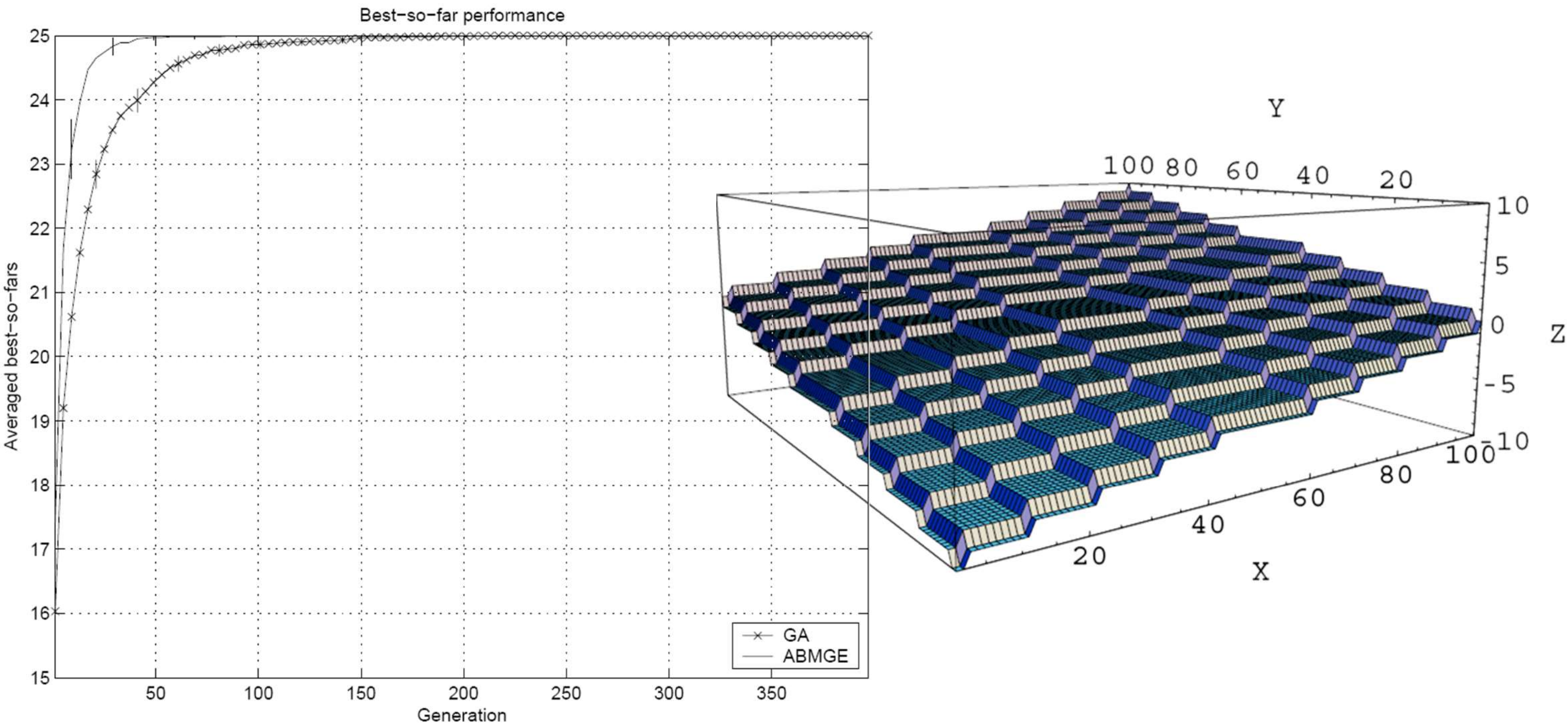
Simple example

50 runs for small royal road testbed

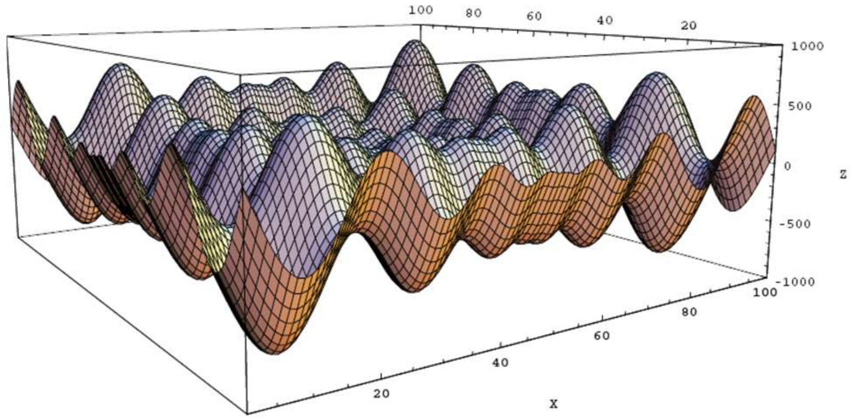


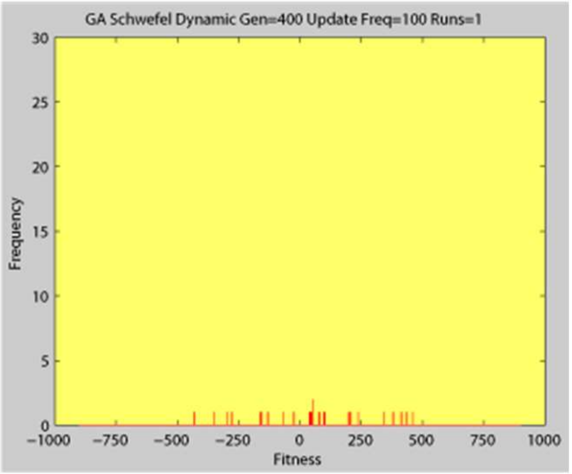
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

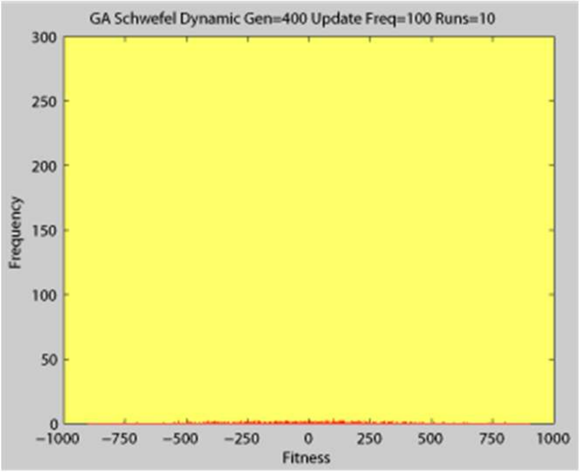
De Jong F3 function



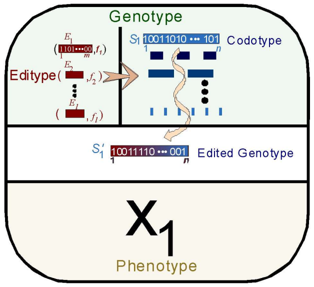
example run: Schwefel function



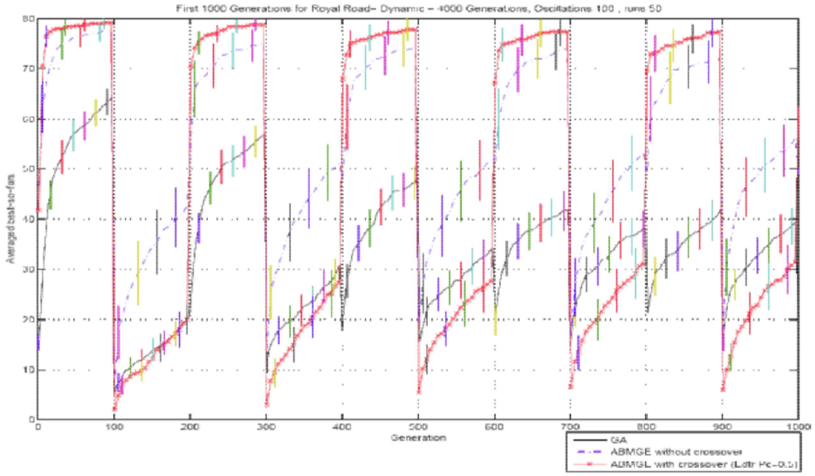

GA Schwefel Dynamic Gen=400 Update Freq=100 Runs=1


GA Schwefel Dynamic Gen=400 Update Freq=100 Runs=10

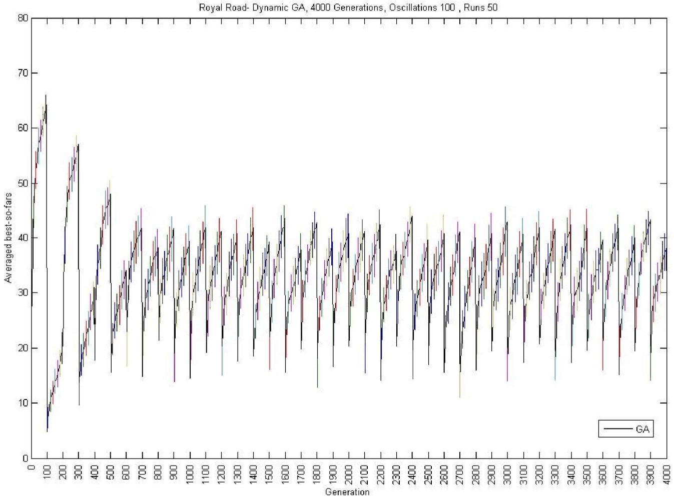
ABMGE on oscillatory fitness landscapes (Royal Road)



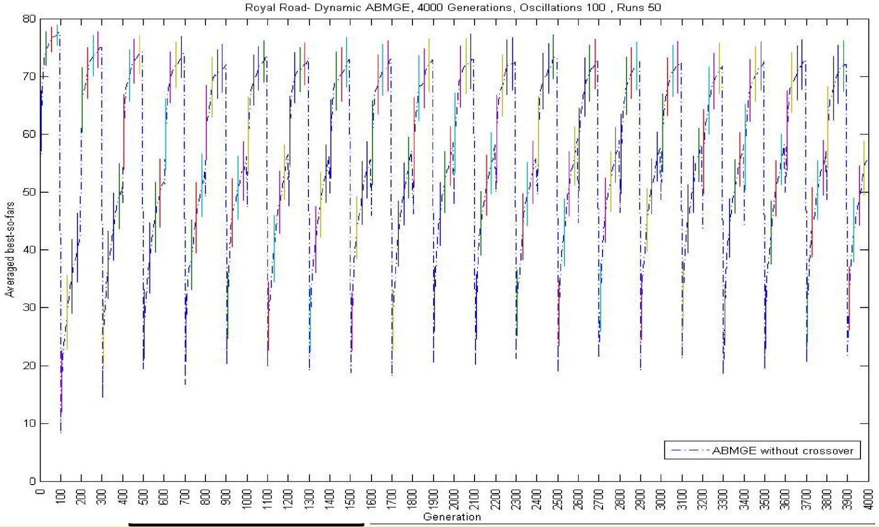
$$
\begin{aligned}
s_1 &= 11111***********************************;\ c_1 = 10\\
s_2 &= *****11111******************************;\ c_2 = 10\\
s_3 &= **********11111*************************;\ c_3 = 10\\
s_4 &= ***************11111********************;\ c_4 = 10\\
s_5 &= ********************11111***************;\ c_5 = 10\\
s_6 &= *************************11111**********;\ c_6 = 10\\
s_7 &= ******************************11111*****;\ c_7 = 10\\
s_8 &= ***********************************11111;\ c_8 = 10
\end{aligned}
$$



GA

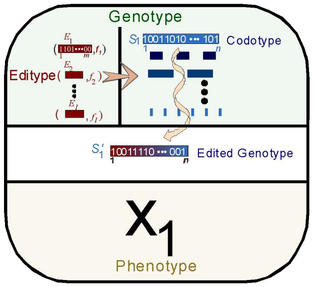


ABMGE

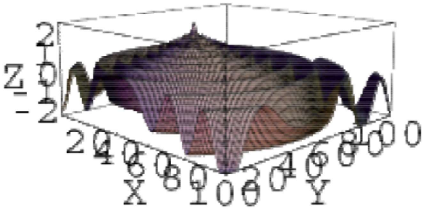


Huang, C-F, et al [2007]. *Evolutionary Computation.* **15**(3).

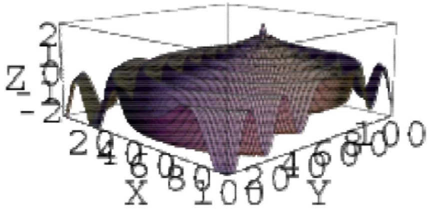
ABMGE on dynamical environments

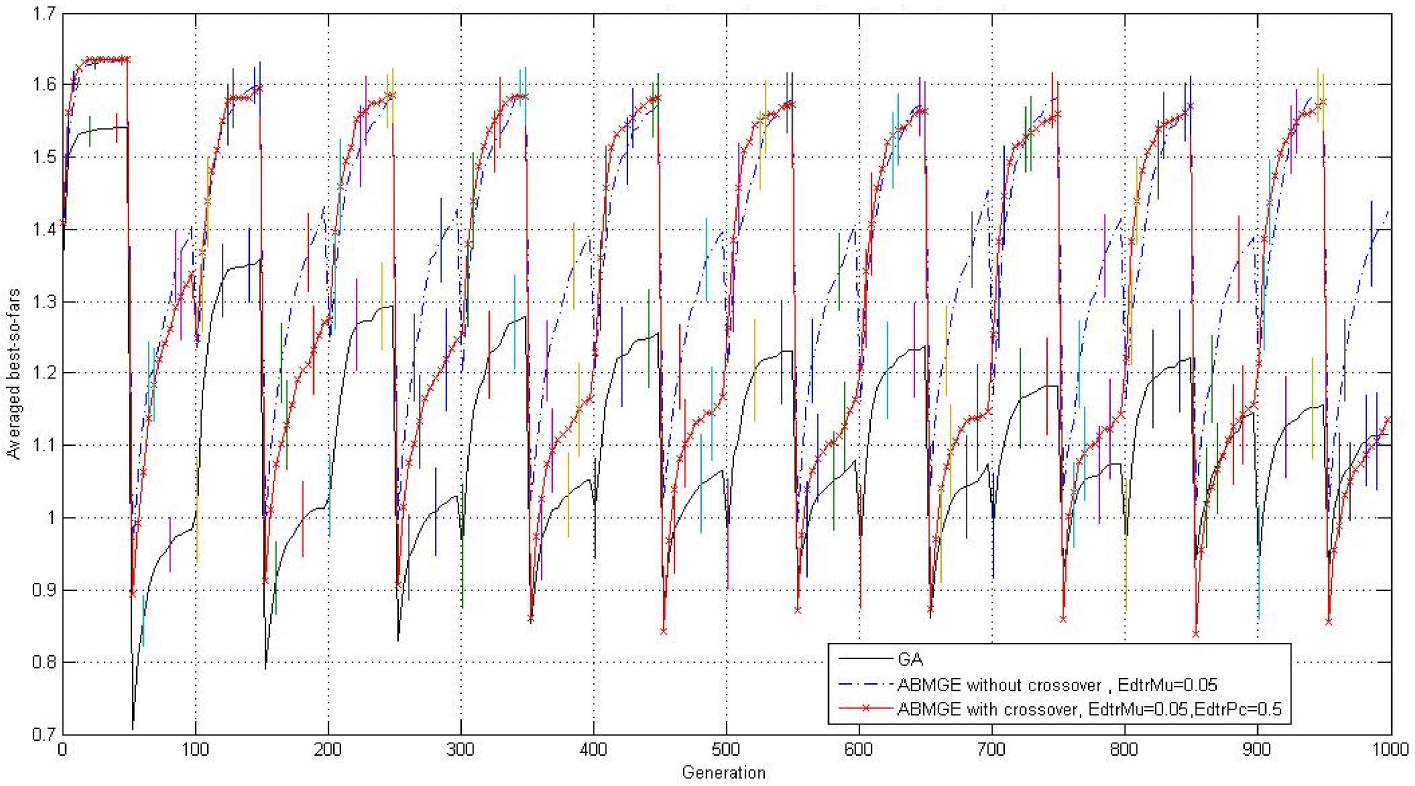


delta=1



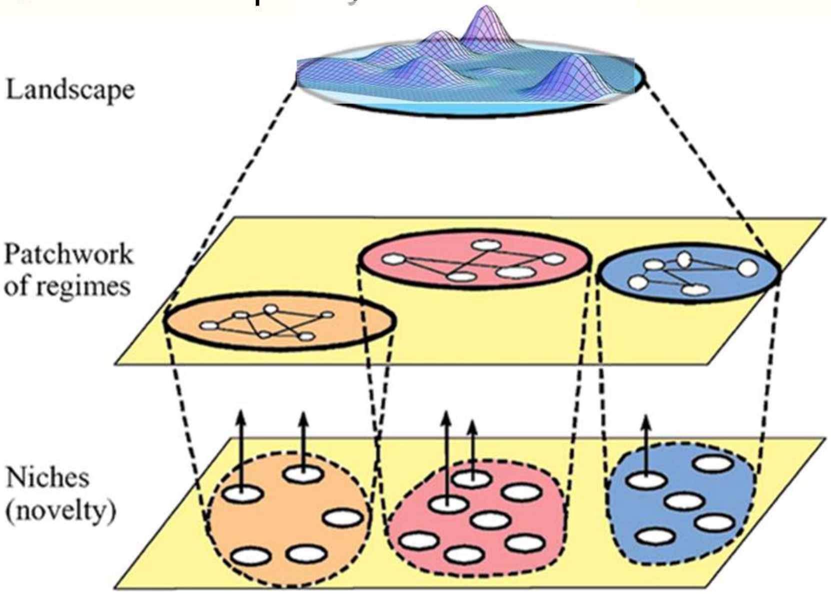
delta=-1

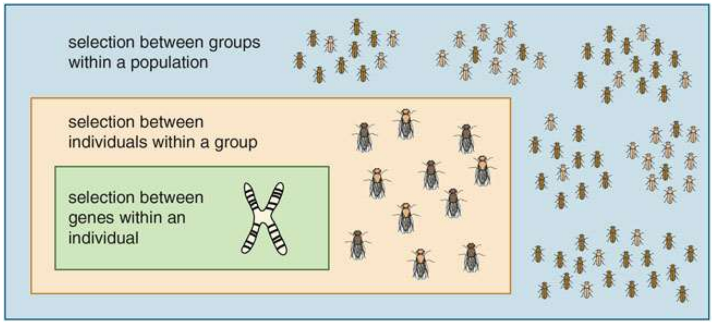


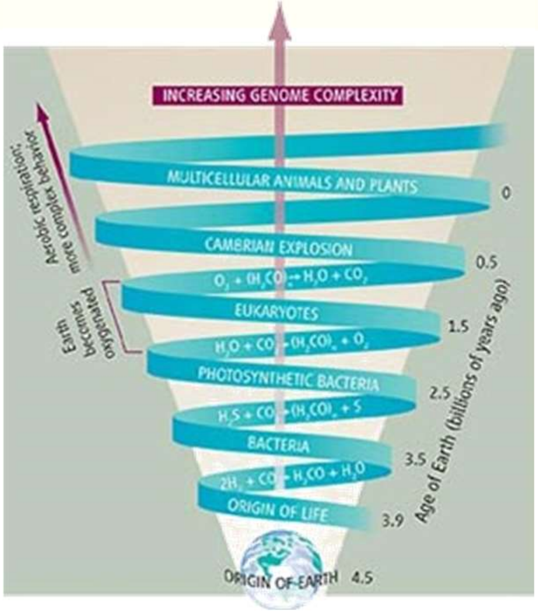


Huang, C-F, et al [2007]. *Evolutionary Computation*. **15**(3).

Genomic complexity and multi-level selection







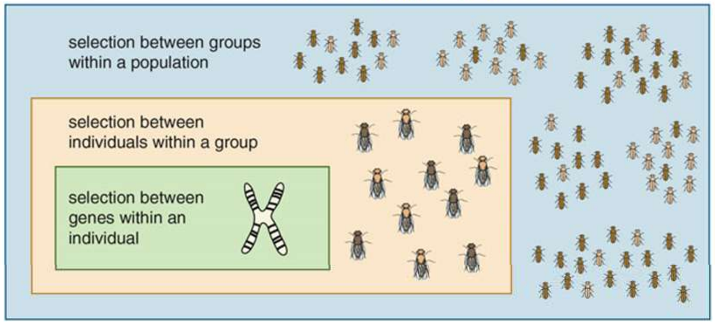
Wilson, David Sloan, and Edward O. Wilson. "Evolution ‘for the Good of the Group’ ”
American Scientist **96**.5 (2008): 380-389.

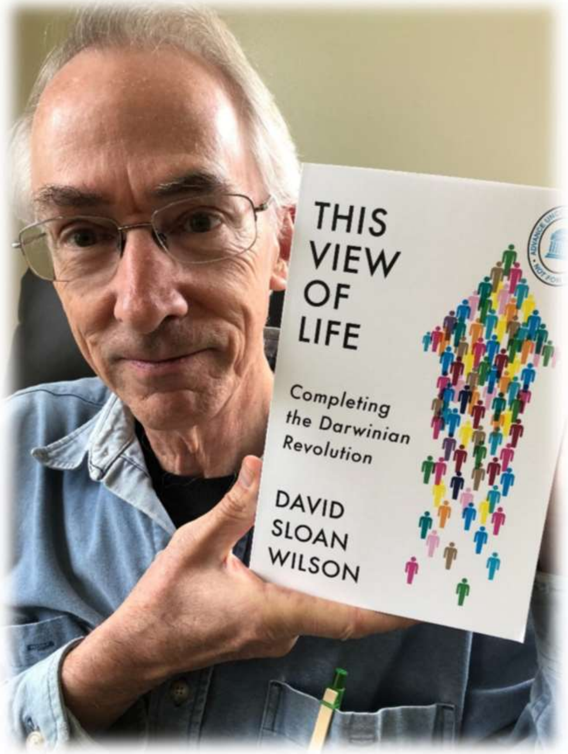
Genomic complexity and multi-level selection

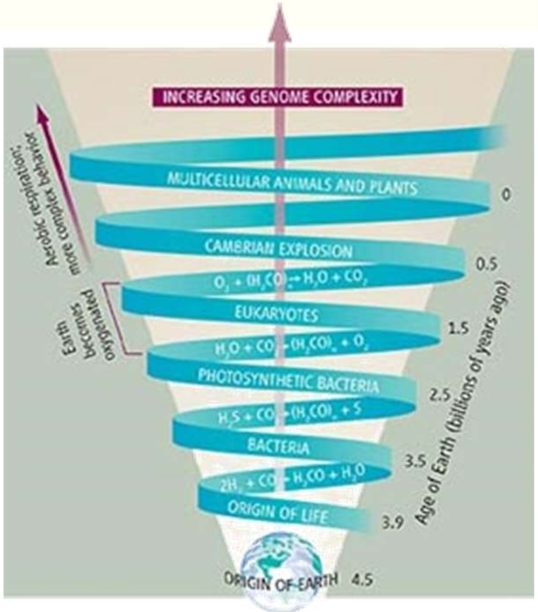
AMERICAN Scientist

Evolution "for the Good of the Group"

BY EDWARD O. WILSON, DAVID SLOAN WILSON







Wilson, David Sloan, and Edward O. Wilson. "Evolution ‘for the Good of the Group’ " *American Scientist* **96**.5 (2008): 380-389.

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK

rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

units of selection

moral men might not do any better than immoral men but tribes of moral men would certainly "have an immense advantage" over fractious bands of pirates. (Charles Darwin)

- Multilevel selection theory
  - Selection occurs in multiple levels simultaneously
  - No general-case scenario, each organism on a case-by-case basis
    - David Wilson and E.O. Wilson
- Experiments with *Pseudonas fluorescens*
  - Oxygen-exhausting bacteria in liquid
  - Groups with enough altruists survive
- Kin-selection as special case of group selection
  - Leading to various, diverse (selectable) groups with high genetic similarity
- Sociobiology
  - Selfishness beats altruism within groups. Altruistic groups beat selfish groups.

"Morality is herd instinct in the individual".
(Friedrich Nietzsche)

INDIVIDUAL SELECTION

"Altruist" *Pseudomonas fluorescens* bacteria (*below left*) carry a gene for secreting a polymer that enables mats of bacteria to float and thus access oxygen easily; "freeloaders" (*below right*) lack the gene. Producing the polymer costs extra energy, so freeloaders reproduce faster than altruists. Natural selection acting on individuals alone would drive the altruists to extinction.

GROUP SELECTION

But group selection appears to operate, too—at least in the laboratory—favoring mats of *P. fluorescens* bacteria in which some altruists persist. Only mats that include enough polymer-secreting altruists will float and thus survive to reproduce themselves, altruists included (*below left*). Mats in which individual selection leads to too many freeloaders will sink, drowning the entire bacterial colony (*below right*). Such mats leave no progeny.

Sci. American, Jan 2009 (Steve Mirsky)

altruism and selection

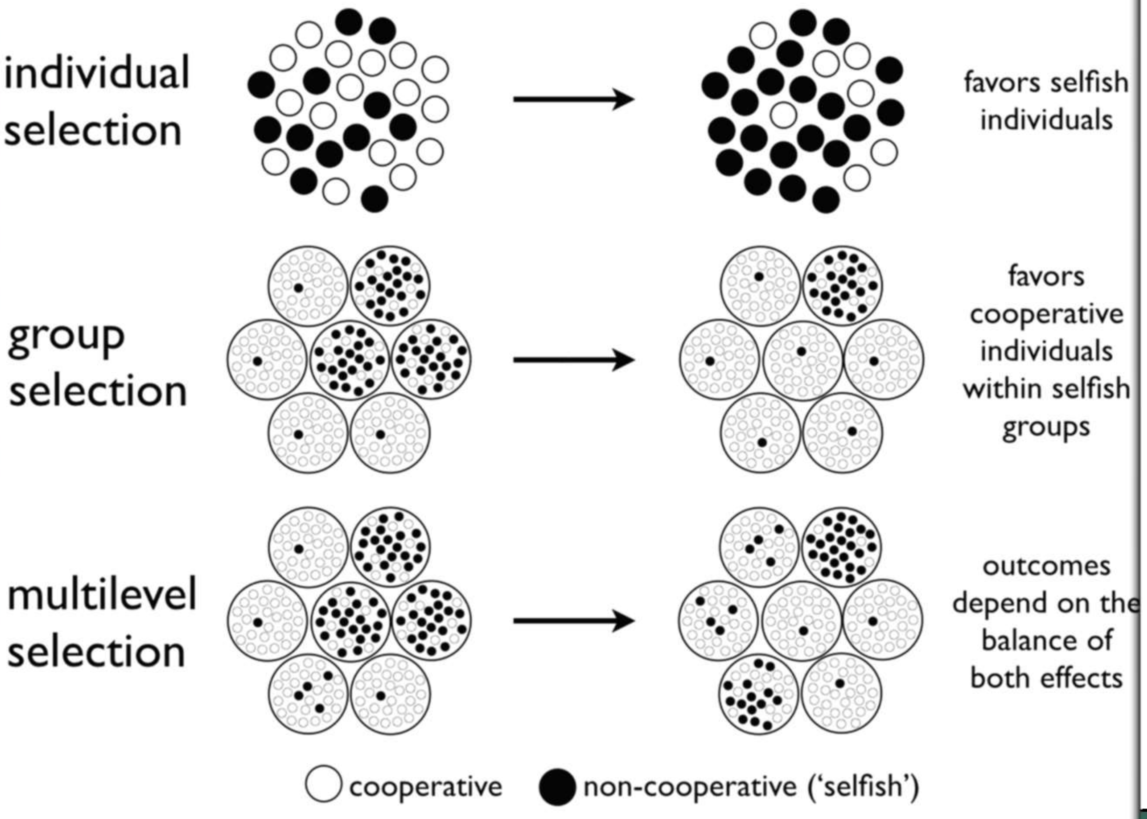
units of select

moral men might
certainly “have ar

- Multile
  - Sele
  - No g
    case
    - I
- Experi
  - Oxy
  - Gro
- Kin-sel
  - Lea
    gen
- Sociob
  - Self
    bea

“Morality
(Friedrich Nietzsche)

Waring, Timothy M., Michelle Kline Ann, Jeremy S. Brooks, Sandra H. Goff, John Gowdy, Marco A. Janssen, Paul E. Smaldino, and Jennifer Jacquet. "A multilevel evolutionary framework for sustainability analysis." *Ecology and Society* 20, no. 2 (2015).



INDIVIDUAL SELECTION

“Altruist” *Pseudomonas fluorescens* bacteria (*below left*) carry a gene for secreting a polymer that enables mats of bacteria to float and thus access oxygen easily; “freeloaders” (*below right*) lack the gene. Producing the polymer costs extra energy, so freeloaders reproduce faster than altruists. Natural selection acting on individuals alone would drive the altruists to extinction.

GROUP SELECTION

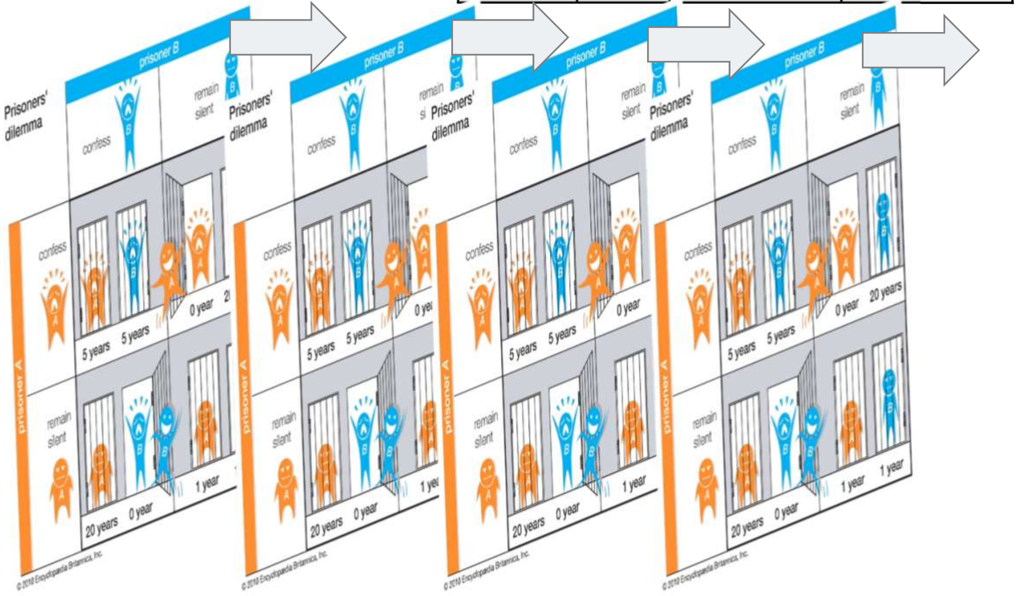
But group selection appears to operate, too—at least in the laboratory—favoring mats of *P. fluorescens* bacteria in which some altruists persist. Only mats that include enough polymer-secreting altruists will float and thus survive to reproduce themselves, altruists included (*below left*). Mats in which individual selection leads to too many freeloaders will sink, drowning the entire bacterial colony (*below right*). Such mats leave no progeny.

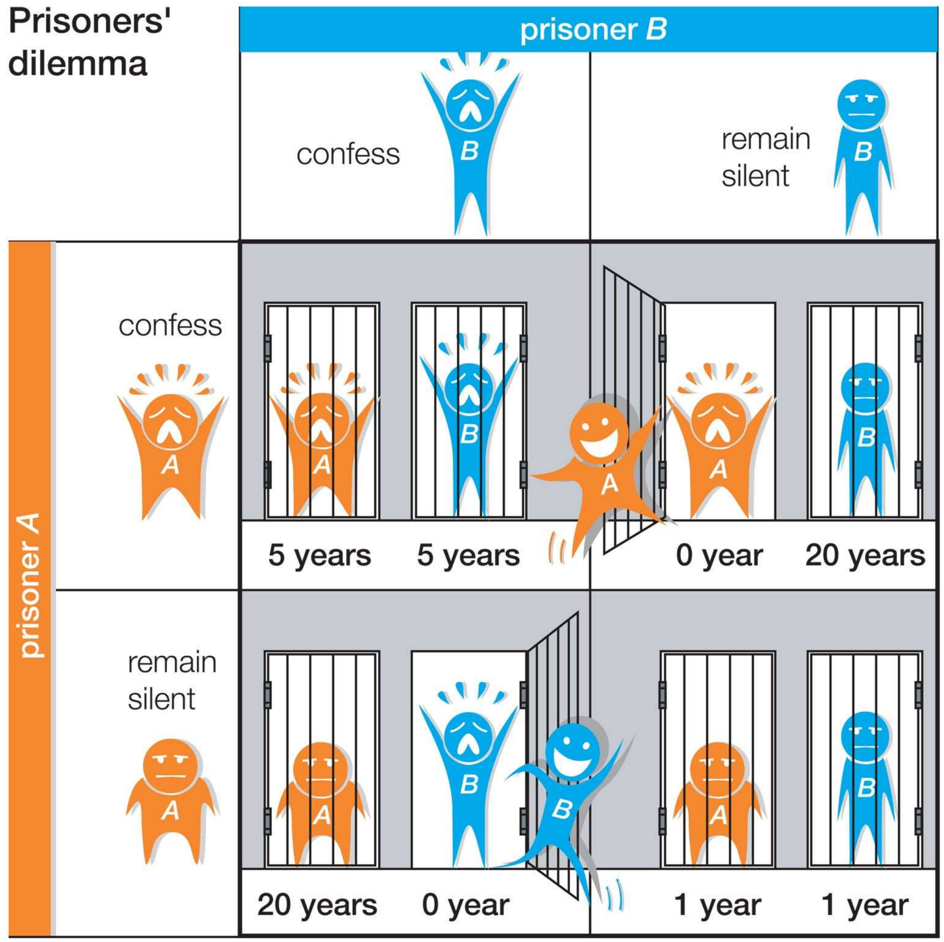
Sci. American, Jan 2009 (Steve Mirsky)

iterated prisoner's dilemma

encoding

		P2	
		C	D
P1	C	(3,3)	(0,5)
	D	(5,0)	(1,1)

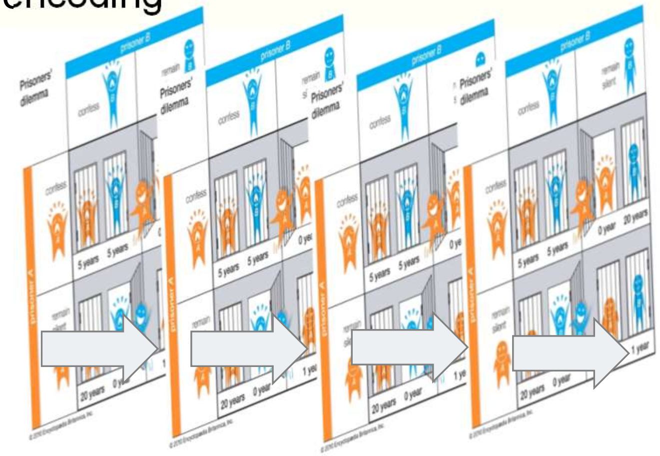




© 2010 Encyclopædia Britannica, Inc.

iterated prisoner's dilemma

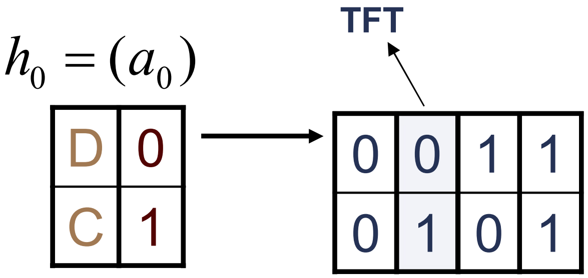
encoding



		P2	
		C	D
P1	C	(3,3)	(0,5)
	D	(5,0)	(1,1)

$$h_m = (a_{m-1},\ldots,\overset{\text{agent}}{a_1},\overset{\text{opponent}}{a_0})$$

Lindgren's iterated game for agents with memory



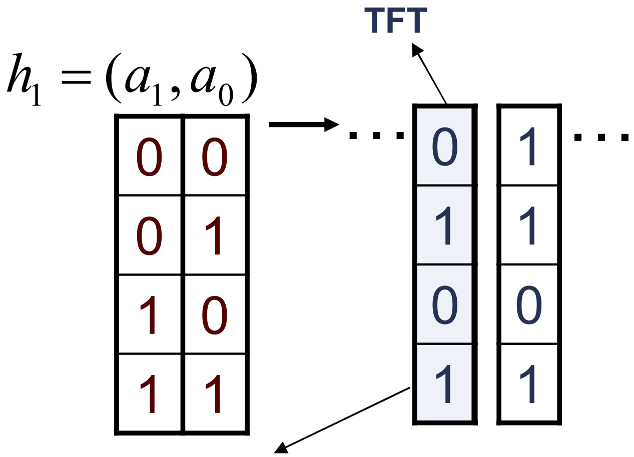
$$h_0 = (a_0)$$

D	0
C	1

→

	TFT		
0	0	1	1
0	1	0	1

4 possible strategies (genotype=2 bits)



$$h_1 = (a_1, a_0)$$

0	0
0	1
1	0
1	1

→ . . .

TFT	
0	1
1	1
0	0
1	1

. . .

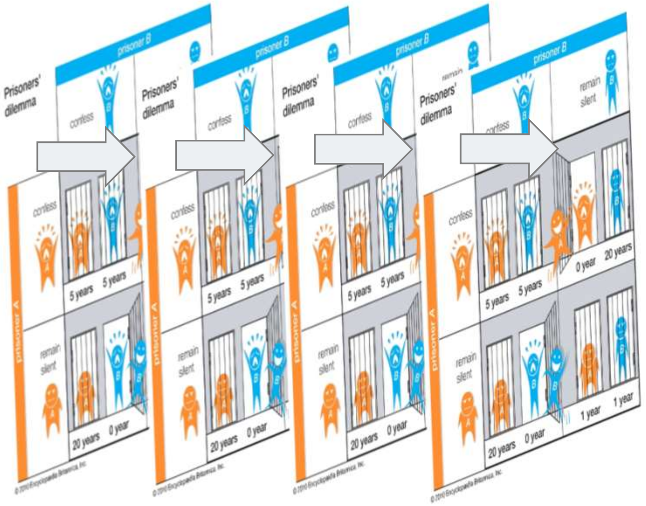
Used in the evolutionary search by GA (tournament selection)

16 possible strategies (genotype=4 bits)

Lindgren, Kristian. "Evolutionary phenomena in simple dynamics." *Artificial life II* (1991): 295-312.

Iterated prisoner's dilemma

memory 0 strategies



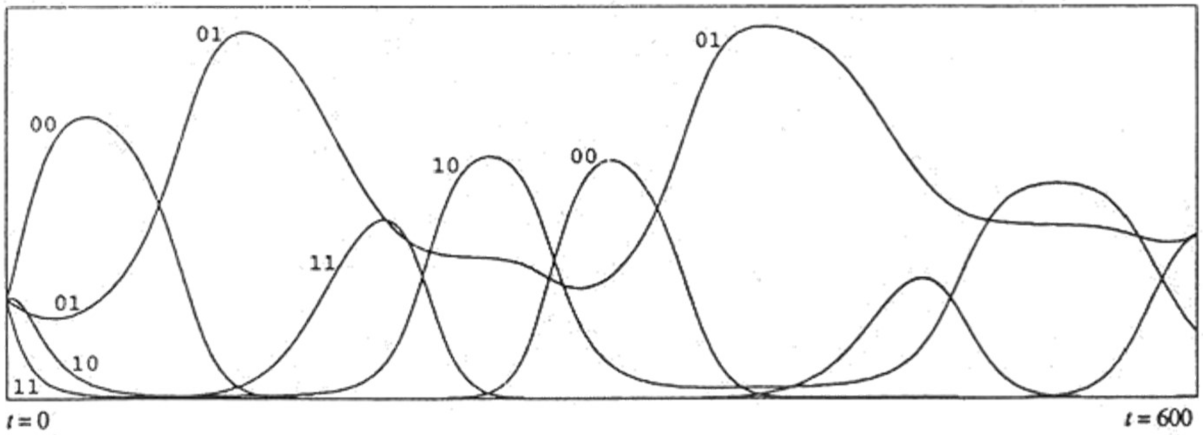


FIGURE 1 The evolution of a population of strategies starting with equal fractions of the memory one strategies [00], [01], [10], and [11] is shown for the first 600 generations. The fractions of different strategies are shown as functions of time (generation).

$$h_0 = (a_0)$$

D	0
C	1

→

TFT

0	0	1	1
0	1	0	1

4 possible strategies

Lindgren, Kristian. "Evolutionary phenomena in simple dynamics." *Artificial life II* (1991): 295-312.

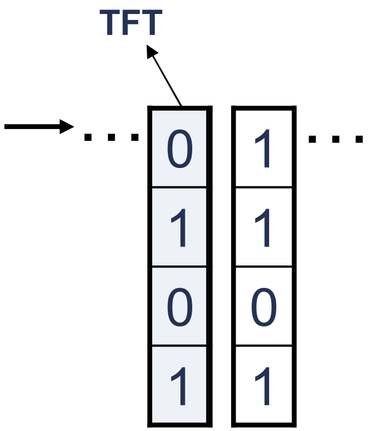
iterated prisoner's dilemma

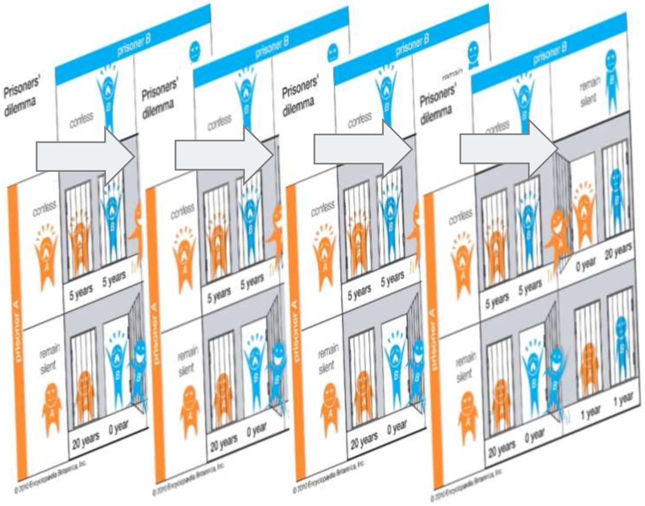
higher memory rules

16 possible strategies (genotype=4 bits)

$$h_1 = (a_1, a_0)$$

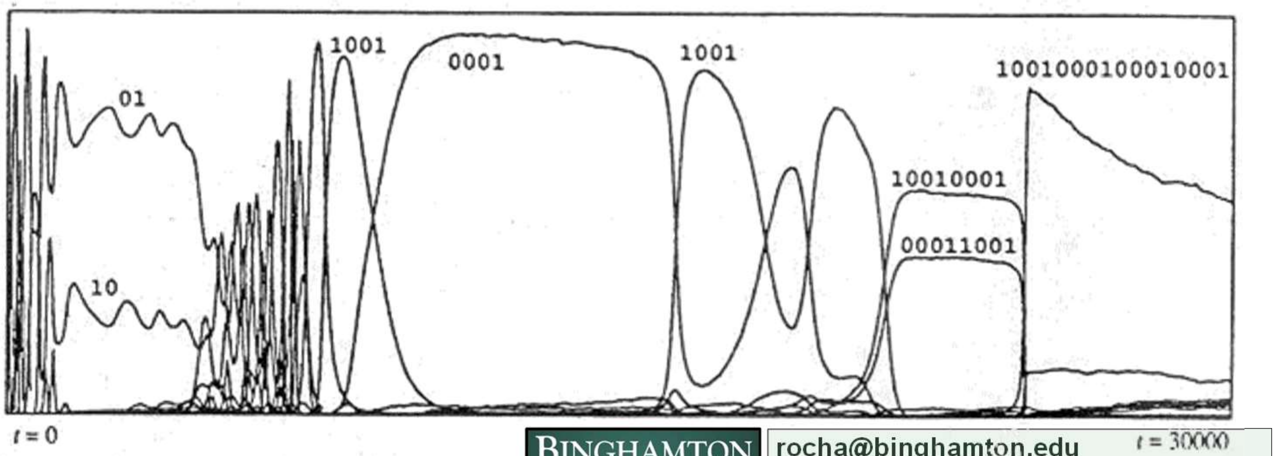
0	0
0	1
1	0
1	1





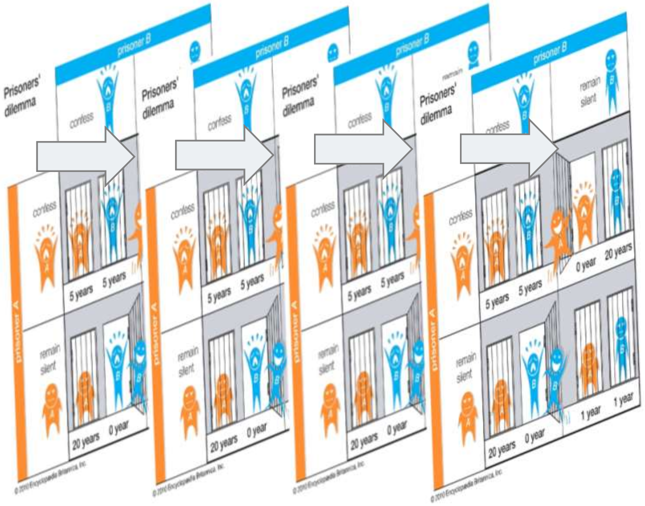
Used in the evolutionary search by GA (tournament selection)

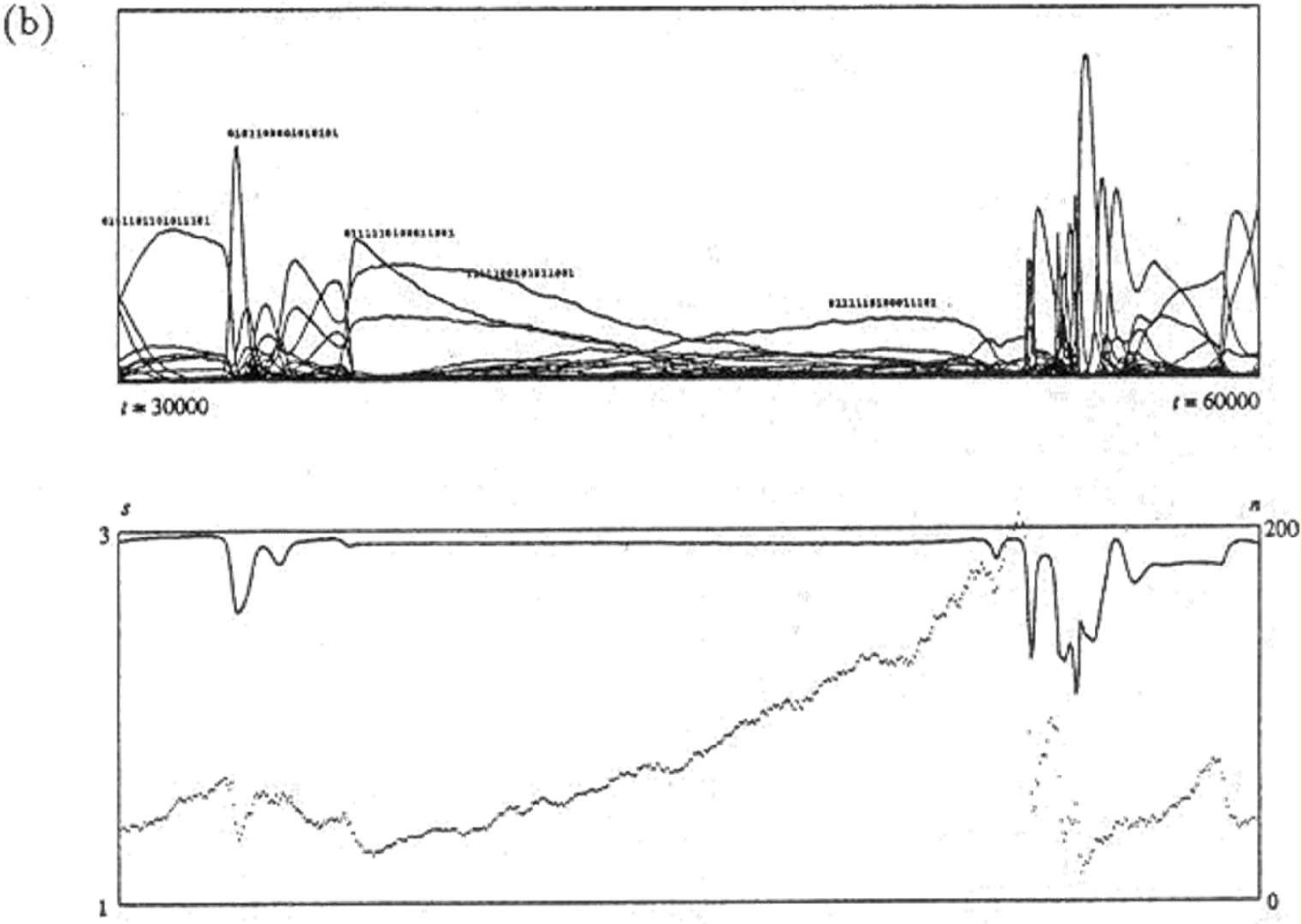
GA uses variable length genotype



Lindgren, Kristian. "Evolutionary phenomena in simple dynamics." *Artificial life II* (1991): 295-312.

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic





Punctuated equilibria and complex evolutionary dynamics

Lindgren, Kristian. "Evolutionary phenomena in simple dynamics." *Artificial life II* (1991): 295-312.

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK

rocha@binghamton.edu
casci.binghamton.edu/academics/i-bic

Next lectures

readings

- Class Book
  - Floreano, D. and C. Mattiussi [2008]. *Bio-Inspired Artificial Intelligence: Theories, Methods, and Technologies*. MIT Press.
    - **Chapter 7**
- Lecture notes
  - Chapter 1: What is Life?
  - Chapter 2: The logical Mechanisms of Life
  - Chapter 3: Formalizing and Modeling the World
  - Chapter 4: Self-Organization and Emergent Complex Behavior
  - Chapter 5: Reality is Stranger than Fiction
  - Chapter 6: Von Neumann and Natural Selection
  - Chapter 7: Modeling Evolutionary Systems
    - posted online @ casci.binghamton.edu/academics/i-bic
- Papers and other materials
  - Optional
    - Nunes de Castro, Leandro [2006]. *Fundamentals of Natural Computing: Basic Concepts, Algorithms, and Applications*. Chapman & Hall.
      - Chapter 5, 7.7, 8.3.1, 8.3.6,