

20

THEY OPEN THE'IR HANDS AND LET THE DELICATE WINGS FLAP ONCE.



THE DISTURBANCE RIPPLES OUTWARD, CHANGING THE FLOW OF THE EDDY CURRENTS IN THE UPPER ATMOSPHERE.

THESE CAUSE MOMENTARY POCKETS

OF HIGHER-PRESSURE AIR TO FORM,

WHICH ACT AS LENSES THAT DEFLECT INCOMING COSMIC RAYS, FOCUSING THEM TO STRIKE THE DRIVE PLATTER AND FLIP THE DESIRED BIT.

⇔



xkcd.com

Logistics

- We plan to return graded projects on Wednesday, Thursday.
- I will hold office hours all day Thursday so you can pick up and discuss your graded project.
- Project 2 has been posted.
- Groups have been assigned.

Logistics

- We plan to return graded projects on Wednesday, Thursday.
- I will hold office hours all day Thursday so you can pick up and discuss your graded project.
- Project 2 has been posted.
- Groups have been assigned.

¿Project Questions?

Genetic Algorithms Continued

Lecture 10

Terminology

On chromosomes:

Amino acids biochemical properties nonpolar polar

basic acidic Termination: stop codon

On chromosomes:	1st 2nd base						3rd			
	base		т		С		Α		G	base
4 Bases {T,C,A,G}		ттт	(Pho/E) Phonylalapipo	тст		TAT	(Tur(M) Turoping	TGT		т
	т	ттс	(FIE/F) FIEllylaidillie	тсс	(Ser/S) Serine	TAC	(Tyl/T) Tyrosine	TGC	(Cys/C) Cysteme	С
Codona (coguonaca of 2		TTA		TCA		TAA ^[B]	Stop (Ochre)	TGA ^[B]	Stop (Opal)	Α
Codons (sequences of 3		TTG		TCG	TAG ^[B]	Stop (Amber)	TGG	(Trp/W) Tryptophan	G	
bases) code for each		СТТ	(Leu/L) Leucine	ССТ	(Pro/P) Proline	CAT	(His/H) Histidine	CGT	(Arg/R) Arginine	т
amino acid.	с	стс		ссс		CAC		CGC		С
		СТА		CCA		CAA	(Gln/Q) Glutamine	CGA		Α
Genes (variable length sequences of codons) code for complete proteins.		CTG		CCG		CAG		CGG		G
	A	ATT		ACT	(Thr/T) Threonine	AAT	(Asn/N) Asparagine AG	AGT	(Ser/S) Serine	т
		ATC	(Ile/I) Isoleucine	ACC		AAC		AGC		С
		ATA	-	ACA		AAA		AGA	(Arg/D) Argining	A
		ATG ^[A]	(Met/M) Methionine	ACG		AAG	(Lys/K) Lysine	AGG		G
		GTT		GCT	CT CC CA (Ala/A) Alanine	GAT	(Asp/D) Aspartic acid GGC	GGT		т
	G	GTC		GCC		GAC		GGC		С
Proteins form the		GTA	(vai/v) valine	GCA		GAA	GGA	GGA	(Gly/G) Glycine	A
Phonotypo		GTG	GCG		GAG	AG	GGG	G	G	

Standard genetic code

Proteins form the Phenotype.

Wikipedia, 2-22-2017

Chromosomal location of the TCR α/δ , β and γ chain loci in man



TCR: T cell Receptor

These loci are highly variable in their alleles.

Terminology: Protein Phenotype

Proteins form the phenotype.

The conformation of a protein is how it twists itself.

This conformation process is dynamical.

Proteins vibrate into low energy conformations.



Robot path planning techniques for protein folding. (Lydia Tapia, UNM)

Xinyu Tang, Shawna Thomas, Lydia Tapia, David P. Giedroc, Nancy M. Amato, "Simulating RNA Folding Kinetics on Approximated Energy Landscapes," *Journal of Molecular Biology*, 3811(4):1055-1067, 2007

- Terminology: Protein Phenotype
- Proteins form the phenotype.
- The conformation of a protein is how it twists itself.
- This conformation process is dynamical.
- Proteins vibrate into low energy conformations.
- These states are defined by stable and unstable fixed points and saddle points.

Xia, Kelin, and Guo-Wei Wei. "Molecular nonlinear dynamics and protein thermal uncertainty quantification." *Chaos: An Interdisciplinary Journal of Nonlinear Science* 24.1, 013103, 2014.





Transition from chaos to periodicity in the chaotic dynamics model (CDM) of bacteriocin AS-48 (PDB ID 1e68). (a) The butterfly wing pattern for one of 70 chaotic oscillators. (b) The solution of original 70 chaotic oscillators. (c) The periodic orbit of the ILDM for bacteriocin AS-48. (d) Bacteriocin AS-48 induced Hopf bifurcation from chaos. All of 70 nonlinear oscillators are in one lag synchronized periodic orbit.

Terminology

- Gene A variable in the genome
- Genotype A string of symbols in the genome
- Phenotype The decoding of a genome
- Locus A position in the genome (*i*th position in a string)
- Allele A value the variable (gene) can take on.
- Epistasis interdependence of genes (nonlinearity)
- Encoding A mapping of a set of features into the genome.
- Biological genes can be independent of loci.



Schemata

Developing a theory of Genetic Algorithms



John Holland, Inventor of the GA.

The following slides are based on material from Goldberg, D. *Genetic Algorithms*, 1989 and Holland J, *Adaptation in Natural and Artificial Systems*, 1993

Consider a sample problem:

Encoding: integers as bitstrings. Integers here are the phenotype and bitstrings are the genotype. We want to maximise the bitstrings according to the fitness function:

$$F(x) = x^2$$

GAs: Instructive Example

$$F(x) = x^2$$

i	Initial Pop.	Phenotype
1	01101	13
2	11000	24
3	01000	8
4	10011	19

GAs: Instructive Example

$$F(x) = x^2$$

i	Initial Pop.	Phenotype	Fitness (f)
1	01101	13	169
2	11000	24	576
3	01000	8	64
4	10011	19	361

Roulette Selection



i	Initial Gen.	Phenotype	Fitness (f)	P(i)
1	01101	13	169	0.14
2	11000	24	576	0.49
3	01000	8	64	0.06
4	10011	19	361	0.31

After selection

i	2 nd Gen.	Phenotype	Fitness (f)
1	01101	13	169
2	11000	24	576
3	11000	24	576
4	10011	19	361

After selection, 1-point crossover

i	2 nd Gen.			Phenotype	Fitness (f)
1	01	10	1	13	169
2	110	00	0	24	576
3	11	00	0	24	576
4	10	01	.1	19	361

After selection, crossover (P = 1.0),

and mutation
$$\left(P = \frac{1}{1000}\right)$$
 No mutation in this case.

i	2 nd Gen	•	Phenotype	Fitness (f)
1	0110	0	12	144
2	1100	1	25	625
3	11 01	11	27	729
4	10 00	00	16	256

GAs: Schemata What information about the search space does this table contain?

Bitstring	Fitness
01101	169
11000	576
01000	64
10011	361

GAs: Schemata What information about the search space does this table contain?

Notice that bitstrings starting with 1s have higher fitness.

Bitstring	Fitness
01101	169
11000	576
01000	64
10011	361

- 1. Similarity among strings in the population.
- A causal relationship between the strings and the fitness function.

Schemata capture this.

Bitstring	Fitness
01101	169
11000	576
01000	64
10011	361

- 1. Similarity among strings in the population.
- A causal relationship between the strings and the fitness function.

Schemata capture this.

Bitstring	Fitness
01101	169
11000	576
01000	64
10011	361

Schemata define functional equivalence classes.

Schemata are defined over the string alphabet plus a metasymbol '*'

", is just a wildcard (not a Kleene star if you are in CS500).

Schemata (or similarity templates)

define equivalence classes.

10*01

Schemata (or similarity templates) define equivalence classes.

 $10 * 01 = \{10101, 10001\}$

Schemata (or similarity templates) define equivalence classes.

$10 * 01 = \{10101, 10001\}$ *000*

Schemata (or similarity templates) define equivalence classes.

 $10 * 01 = \{10101, 10001\}$ $*000 * = \{00000, 00001, 10000, 10001\}$

Schemata (or similarity templates) define equivalence classes.

 $10 * 01 = \{10101, 10001\}$ $*000* = \{00000, 00001, 10000, 10001\}$ 0 * 1 * *

Schemata (or similarity templates) define equivalence classes.

- $10 * 01 = \{10101, 10001\}$
- $*000* = \{00000, 00001, 10000, 10001\}$

0 * 1 * * All strings of length 5 with a 0 in the first position and a 1 in the third.

How many schemata are there for an alphabet with cardinality kand genome length N?

How many schemata are there for an alphabet with cardinality kand genome length N?

 $(k+1)^{N}$

We can bound the number of schemata in a population.

We can bound the number of schemata in a population. In our example:

The number of schemata for an individual genome is 2^5 . because it can take on its actual value or the wildcard. (2 values)

We can bound the number of schemata in a population. In our example:

The number of schemata for an individual genome is 2^5 . because it can take on its actual value or the wildcard.

If the population has n individuals there will be at most $n2^2$ schemata.

So what? How does this help us?

So what? How does this help us?

Can think in terms of useful classes of genomes, since some of the variation won't matter.

Holland makes an optimality argument for GAs using schemata.

So what? How does this help us?

Can think in terms of useful classes of genomes, since some of the variation won't matter.
Schemata

Take a couple of minutes to discuss with your neighbour the effect on

Schemata in the current population of:

- Reproduction
- Crossover
- Mutation



First we consider reproduction:

First we consider reproduction: Schemata with higher fitness tends to increase each generation. No new schemata appear.

First we consider reproduction: Schemata with higher fitness tends to increase each generation. No new schemata appear. Formally, for a schema S in population A, The number of individuals with schema S at time t + 1, given the number at time t is

$$m_S(t+1) = m_S(t) \frac{f(S)}{\operatorname{mean}(f(A))}$$

Consider a schema with fitness $c \times mean(A)$, and c > 1

Consider a schema with fitness $c \times mean(A)$, and c > 1

We can rewrite the increase in representation for this schema, S, as: $m_S(t+1) = m_S(t) \left[\frac{c \times \text{mean}(A)}{\text{mean}(A)} \right]$

We can rewrite the increase in representation for this schema, S, as:

$$m_S(t+1) = m_S(t) \left[\frac{c \times \operatorname{mean}(A)}{\operatorname{mean}(A)} \right]$$

 $m_S(t+1) = c \times m_S(t)$

We can rewrite the increase in representation for this schema, S, as:

$$m_S(t+1) = m_S(t) \left[\frac{c \times \operatorname{mean}(A)}{\operatorname{mean}(A)} \right]$$

 $m_S(t+1) = c \times m_S(t)$

$$m_S(t) = c \times c \times \cdots \times c \times m_S(0)$$

We can rewrite the increase in representation for this schema, S, as:

$$m_S(t+1) = m_S(t) \left[\frac{c \times \operatorname{mean}(A)}{\operatorname{mean}(A)} \right]$$

 $m_S(t+1) = c \times m_S(t)$

 $m_S(t) = c \times c \times \cdots \times c \times m_S(0) = c^t$ Exponential

Now consider the effect of 1-point crossover on schemata.

Now consider the effect of 1-point crossover on schemata.

Schemata survive if they are not cut by the crossover

For example: more 1 * * 0 is less likely than * * 10*to be destroyed.

Formally:

Probability of disruption: $P_D = P_{\times} \frac{D(S)}{L_S - 1}$ where, 1 fewer crossover P_{\times} is the crossover probability, sites than genes D(S) is the defining length of schema S, L is the string/genome length.

Defining length: distance between the first and last non-wildcard symbols.

How about mutation?

Mutation is more likely to destroy "higher order" schemata. The order is the number of fixed (non-wildcard) symbols.

Mutation is more likely to destroy "higher order" schemata. The order is the number of fixed (non-wildcard) symbols.

More formally,

- Probability of disruption: $P_D = \left(1 (1 P_{\text{mut}})^{O(S)}\right)$ where,
- $P_{\rm mut}$ is the mutation probability,
- O(S) is the order of the schema S.

Also called the fundamental Theorem of Genetic Algorithms

The schema theorem:

Fit schema with lower defining length

and lower order increase exponentially in the population over time.

These are called "building blocks"

Also called the fundamental Theorem of Genetic Algorithms

Schema Theorem

More formally,

$$m_S(t+1) \ge m_S(t) \frac{f(S)}{\operatorname{mean}(S)} \left[1 - P_{\times} \frac{D(S)}{L-1} - O(S) P_{\operatorname{mut}} \right]$$

Building block hypothesis: A genetic algorithm seeks optimal performance through the juxtaposition of short, low-order, high- performance schemata, called the building blocks.

Premature convergence is the main challenge. Building blocks can cause premature convergence.



A Royal Road Problem: Testing the Building Block Hypothesis



Figure 1: Example Royal Road Function. $F(x) = \sum_{s \in S} c_s \sigma_s(x)$, where x is a bit string, c_s is a value assigned to the schema s (here, $c_s = order(s)$), and $\sigma_s(x) = \begin{cases} 1 & \text{if } x \text{ is an instance of } s \\ 0 & \text{otherwise.} \end{cases}$

Mitchell, Melanie, Stephanie Forrest, and John H. Holland. "The royal road for genetic algorithms: Fitness landscapes and GA performance." *Proceedings of the first european conference on artificial life*. 1992.



WE'VE DECIDED TO DROP THE CS DEPARTMENT FROM OUR WEEKLY DINNER PARTY HOSTING ROTATION.

https://xkcd.com

Logistics

- Midterm exam March 10th.
- Midterm review moved to March 6th.
- Project 2 is due on March 10th.
- Project reviews are due on March 20th.
- The in class competition will be on March 20th.
- Transitioning from Genetic Algorithms to Cellular Automata
- We will be using genetic algorithms in the next project to explore cellular automata and game theory.

Theory of Self-Reproducing Automata - John von Neumann

- David Shubsda
- Joshua Ridens
- Turn in your review forms in to Bianca at the end of the presentation.
- I will put David and Joshua's slides on the course website.

Suppose you can play k slot machines. The i^{th} machine pays jackpot following a Gaussian probability distribution with mean $= \mu_i$, and variance $= \sigma_i^2$



Suppose you can play k slot machines. The i^{th} machine pays jackpot following a Gaussian probability distribution with mean $= \mu_i$, and variance $= \sigma_i^2$



The challenge is to win the most money possible (or lose the least) over time.

Reason about the best trade-off



Reason about the best trade-off



A strategy:

- Given a maximum of N pulls at the arm we could try:
- Allocating an equal number, n, of pulls, where (kn < N), to each arm,
- then use all the remaining time to pull on the arm with
- the most payouts.

Reason about the best trade-off



Given N, μ_i , and σ_i^2 , we can define a loss function, L(N, n).

GAs: Optimality of Trial Allocation

- There is a trade-off between sampling near the best observed and exploring the fitness landscape.
- Loss due to searching near the currently known optimum is due to *sampling error*.
- Loss due to choosing areas that are not known to be good we might call *performance loss*.

GAs: Optimality of Trial Allocation

• The presence of a trade-off suggests an optimisation problem.

Given N, μ_i , and σ_i^2 ,

Reason about the best trade-off



we can define a loss function, L(N, n).

$$L(N,n) = |\mu_1 - \mu_2| \left[(N-n)q(n) + n(1-q(n)) \right]$$

The 2-arm case generalises

De Jong, K. A. An analysis of the behaviour of a class of genetic algorithms. Diss. PhD thesis, University of Michigan, 1975.

GAs: k-armed bandits $L(N,n) = |\mu_1 - \mu_2| \left[(N-n)q(n) + n(1-q(n)) \right]$

Where q(n) is the probability that the worst arm is the best arm after n pulls. In other words q(n) is the probability that you got unlucky as a function of the explore/exploit trade-off.

De Jong, K. A. An analysis of the behaviour of a class of genetic algorithms. Diss. PhD thesis, University of Michigan, 1975.

q(n) decreases exponentially with n. To minimise q(n)we allocate $O(e^n)$ trials to the known best.

"The ratio of trials of the observed best to the second best grows exponentially."



Holland J, Adaptation in Natural and Artificial Systems, 1993

Holland showed that allocating an exponential number of pulls to the best performing slot machine minimises the loss function L(N, n).

This exponential allocation to exploitation best solves the problem.

Holland, J., "Genetic algorithms and the optimal allocation of trials." SIAM Journal on Computing 2.2 (1973): 88-105.

Holland showed that allocating an exponential number of pulls to the best performing slot machine minimises the loss function L(N, n).

This exponential allocation to exploitation best solves the problem.

What does this have to Do with the Fundamental Theorem?

Holland showed that allocating an exponential number of pulls to the best performing slot machine minimises the loss function L(N, n).

This exponential allocation to exploitation best solves the problem.

What does this have to Do with the Fundamental Theorem?

The fundamental theorem shows that GAs allocate exponentially increasing samples to known best solutions.

Diversity

- A pitfall of GAs is that they may converge too early on a local maxima.
- If the population were infinite this wouldn't be a problem.
- The smaller the population the more likely they are to converge prematurely.
- Building blocks can increase premature convergence for some problems.
Schemata: Implicit Parallelism

Holland also showed in 1985 that if Kis the number of stringsprocessed each generation and N is the number of schemata Then $N \in O(K^3)$.

Diversity: Island GAs





Recall Darwin's famous finches. The diversity is in part due to evolution on islands.

Diversity: Island GAs

What would this look like as an algorithm?

Diversity: Elitism

- Diversity is great for avoiding local minima.
- How do we keep from moving too far from regions we know are good.
- Roulette selection can easily discard the most fit individuals.

Diversity: Elitism

- Diversity is great for avoiding local minima.
- How do we keep the population from forgetting about regions we know are good.
- Roulette selection can easily discard the most fit individuals.
- High mutation rates can move the whole population to a lower local maxima.
- Copying the highest fitness individual into the next generation unchanged is called *elitism*.

Diversity: Elitism

• Elitism guarantees that the GA will converge*

(Applies to simulated annealing and artificial immune system optimization as well)

*Villalobos-Arias et al, "Asymptotic Convergence of Some Metaheuristics Used for Multiobjective Optimization", *Foundations of Genetic Algorithms*, 2005

Selection Pressure: Tournaments

- Selection pressure is a measure of how harsh we make the world.
- At one extreme only the very fittest individual would survive.
- At the other extreme everyone survives.

Selection Pressure: Tournaments

- If selection pressure is too high the GA will converge prematurely on a local minima.
- If selection pressure is too low the GA will not converge at all.

Selection Pressure: Tournaments

• Tournament selection is one way to tune the selection pressure.

Evolving Programs

Project 2

Project 2: Core Wars

- Alexander Dewdney, Mathematician and Computer Science
- Wrote a series of articles called Mathematical Recreation for Scientific American
- One of these was Core Wars
- Inspired by a real life incident (Creeper and Reaper).
- Christopher Langton invited Dewdney to present core wars At the first Alife conference.
- Playing with Core Wars was once considered dangerous.
- Capable of self-mutation...



Battle Info		Max Processes:	
Core Size: 8000	Max Cycles: 80000	Imp	0%
Battle Status: Stopped		Mice	0%
Completion:	0% 0	Midget	
Tournament Type: Round Robin Show Results		Piper	
Round 1 of 1	Matchup 1 of 6		

Some x86 Assembly language

Address Instruction 77E814EE mov 77E814F1 mov 77E814F8 add 77E814FB jmp 77E81500 push 77E81501 xor 77E81503 cmp 77E81505 push 77E81506 push

Register or RAM Address esi,dword ptr [edi+8] dword ptr [ebp+64h], 0Ah esi, 4Ah 77E7E91A ebx ebx,ebx ecx,ebx es1 edi

;redcode ;author: T77 ;name: Example1 ;assert CORESIZE=8000 && MAXLENGTH > 100 MOV 0, 1

- immediate addressing ;redcode ;author: T77 ;name: Example2 ;assert CORESIZE=8000 && MAXLENGTH > 100 MOV 2, @2 JMP -2 DAT #0, #0

;redcode ;author: T77 ;name: Example2 ;assert CORESIZE=8000 && MAXLENGTH > 100 ADD **#4**, 3 MOV 2, @2 — Program counter is here JMP -2 DAT #0, #**4**





This program places dats separated by 4 addresses forever.



;redcode ;author: T77 ;name: Example3 ;assert CORESIZE=8000 && MAXLENGTH > 100 Create new thread spl 0 mov 2, <-1 And continue with next jmp -1, -1 instruction

```
;redcode
;author: T77
;name: Example3
;assert CORESIZE=8000 && MAXLENGTH > 100
spl 0
                    < Indirect with
mov 2, <-1
                    predecrement
jmp -1, -1
```