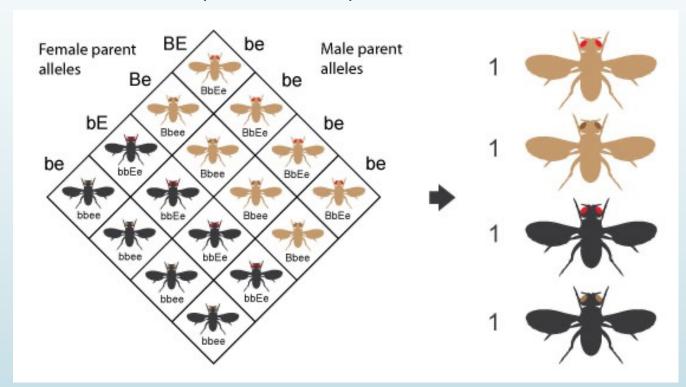
Introduction to Artificial Intelligence

Week 10

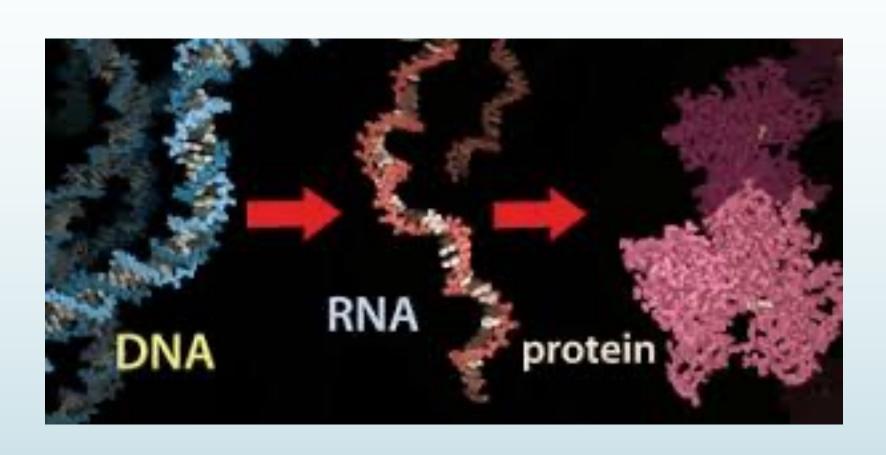
Evolutionary Algorithms (Part II)

Direct v. Indirect Representations

- Genetics Genotype v. Phenotype
- We may have a translation step between our internal representation in the system v. the traits perceived by someone

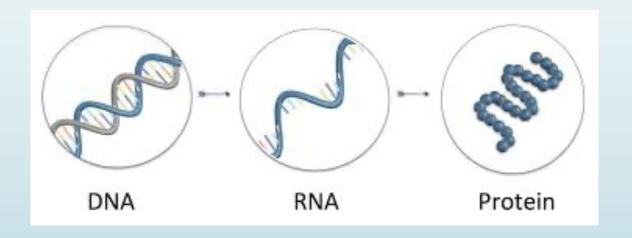


Central Dogma of Molecular Biology



Dogma

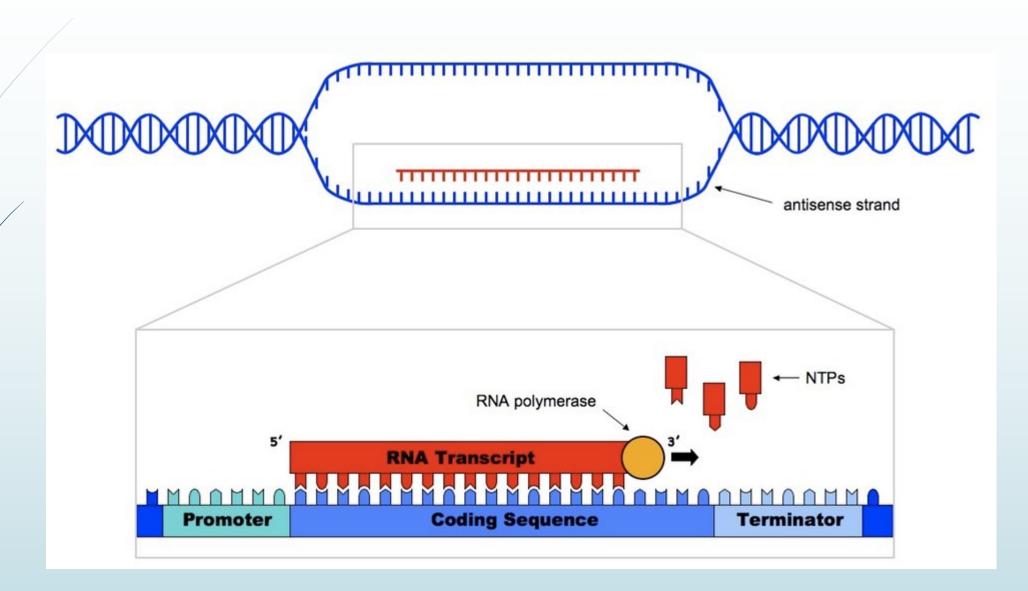
- DNA genes
 - Transcription into
- RNA messages
 - Translation into
- Proteins molecular machines



Transcription

- RNA is a copy of a single gene
- DNA strand unwinds in the presence of RNA polymerase
- Polymerase 'reads' the DNA out and copies a section to a mRNA strand
- Photocopying an Instruction from a Book

RNA Transcription



Translation

- Blocks of 3 bases Codons
- Maps onto 20 amino acids
- Forms a chain of amino acids
- Chains then will fold into the structure
- Protein Structure Defines Function

Known Translation Table

					Seconed	Positi	on			
		U		С		А		G		
		code	Amino Acid	code	Amino Acid	code	Amino Acid	code	Amino Acid	
1		UUU	phe	UCU		UAU	tyr	UGU	cys -	U
/	U	UUC	prie	UCC	ser	UAC		UGC		С
	U	UUA	leu	UCA	361	UAA	STOP	UGA	STOP	A
		UUG	ieu	UCG		UAG	STOP	UGG	trp	G
		CUU	leu	CCU		CAU	his	CGU	arg	U
	С	CUC		CCC	pro	CAC		CGC		С
		CUA	100	CCA		CAA		CGA		A
		CUG		CCG		CAG	giii	CGG		G
		AUU		ACU		AAU	asn	AGU	ser	U
1	A	AUC	ile	ACC	:A thr	AAC		AGC		С
1	^	AUA		ACA		AAA		AGA		A
1		AUG	met	ACG		AAG	lys	AGG		G
		GUU		GCU		GAU	aen	GGU		U
	•	GUC	val	GCC	ala	GAC	GGC	alv	С	
	G	GUA	vai	GCA	aid	GAA	al.	GGA	gly	A
V		GUG		GCG		GAG	glu	GGG		G

Gödel, Esher, Bach

Biology

- DNA
- Transcription
- RNA
- Translation
- Protein
- Codon
- Genotype
- Phenotype

Computer

- Set of all programs
- File System
- A program
- Program Compiler
- Process
- Machine Code Statement
- Source Code
- Running Program

Direct Representation in Problems

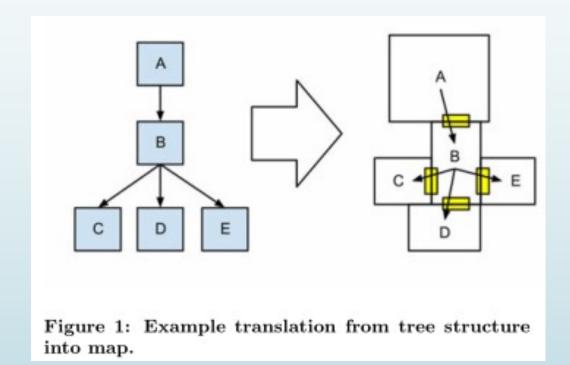
- Direct Representation
 - When the genotype and phenotype are the same
- Examples
 - Settings on a machine
 - Creation of a FSM or other controller
- Majority of optimizations have a direct search for the outcome

Indirect Representation in Problems

- Indirect Representation
 - When the evolved values allow for the construction of a solution to the problem
- Examples
 - Evolving a program for the solution
- This is an interesting emerging method for looking at evolution of structures which lead to solutions or for the entire system

Dungeon Levels

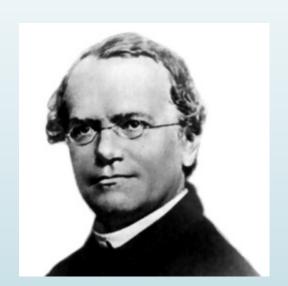
- Genotype a tree of tuples
- Phenotype a board with placement of rooms
- Translation step



Mendelian Genetics

- Discovered by Gregor Johann Mendel in 1865, but not popularized until Thomas Hunt Morgan in 1915
- Ratios of Pea Pods always in divisions of 1/4s
- Dominant and recessive traits
- Punnett Squares:

	R	R
r	Rr	Rr
r	Rr	Rr







What doesn't kill you makes for an interesting feature

Sickle-cell diseases lead to a resistance to Malaria (Nature vol. 515 no. 7526)

A moving target

Migration has led to an increase in the occurrence of sickle-cell disease in countries with previously low incidence of the disorder.

BY SIMON PLEASANTS

cickle-cell disease is a heritable blood disorder that can have devastating effects. It is the result of a minor genetic change — a single error in a single gene - yet it is one of the most common genetic causes of sickness and death1. It is also on the rise. According to a the gene. study published last year, approximately 305,800 children were born with sickle-cell disease in 2010, about two thirds of them in Africa2. By 2050, the study authors predict, those numbers will increase by more than 25%, to more than 400,000 babies born annually with the disorder.

The disease-causing mutation occurs in a gene responsible for the production of haemoglobin, the protein in red blood cells that carries oxygen to tissues throughout the body. Individuals who inherit only one mutant gene typically experience no ill effects, but those who have two faulty copies develop the full-blown disease. The mutation itself is tiny: it replaces one amino acid (called glutamic acid) with another (valine) at one location in the haemoglobin protein. This small change has a major consequence - it turns the flexible, disc-shaped red blood cells rigid, giving them a distorted, sickle-like appearance.

Sickled cells are not only less pliable than healthy red blood cells, they're also stickier. This causes them to adhere to each other and plug up narrow blood vessels, reducing blood flow and preventing adequate oxygen from being delivered to the tissues that need it. This in turn leads to many complications, some of which are life threatening. People with sickle-cell disease commonly experience episodes of acute pain, known as crises, that can last from a few hours to a few days. Other complications include anaemia, leg ulcers, jaundice, kidney damage, high blood pressure, gallstones, increased susceptibility to infections and stroke. These problems arise from a variety of complex mechanisms, such as altered pain receptors (see page S8) and the decreased lifespan of red blood cells1.

In high-income countries, improved management of the disease and medication, including strong painkillers, has meant that patients with sickle-cell disease can expect a better quality of life. In low-income nations, however, those with sickle cell are not as fortunate: in Africa, at least 50% of children born with the condition die before their fifth birthdav3.

Sickle-cell disease is most prevalent in

subcontinent and the Caribbean are all affected. There is evidence to suggest that the geographic concentration of people with sickle-cell disease arose because people with sickle-cell trait those who carry the sickle-cell gene but are gen erally healthy - have enhanced resistance to malaria4, resulting in preferential selection of

The malaria-causing parasite typically invades red blood cells and multiplies inside them. But although sickle-cell carriers can be infected, the presence of sickle haemoglobin confers resistance to malaria. There are a number of different theories, one of which is that subtle biological changes in the blood - pos-

tions have increased disease prevalence in areas "In Africa, at not previously associleast 50% of ated with the disorder. children born The Sickle Cell Disease with sickle-Association of America cell disease estimates that there are die before about 100,000 people their fifth living with the disorder birthday." in the country. The US Department of Health

of Africa - a crippling health burden in a region

that is ill-equipped to bear it, and one that lacks

the financial and health-care resources present

Increased population movement associated

with globalization is bringing sickle cell to

regions that previously had negligibly low inci-

dences of the disease. Migration of people with

the sickle-cell trait or sickle-cell disease caused

the faulty gene to flow from high-frequency

sickle-cell gene areas, such as Africa and India. to western Europe, the eastern coast of South

America, and North America7. These migra-

in wealthy countries.

and Human Service's National Heart, Lung and Blood Institute suggests that sickle-cell disease affects about 1 in 500 African American babies, a number about ten times greater the general population.

Ironically, improved medical treatment leading to longer life expectancy is contributing to the increased prevalence of sickle-cell disease. Penicillin, taken daily, helps to prevent and treat infection in young children, folic acid stimulates the production of red blood cells, the drug hydroxycarbamide (known as hydroxyurea in the United States) reduces the number and severity of painful crises, and blood transfusions help to temper acute cases of the disease. With so many interventions, more people will survive to pass on sickle-cell genes to their children. ■

- Makani, J. et al. Sci. World J. http://dx.doi. org/10.1155/2013/193252 (2013).
- Piel, F. B. et al. PLoS Med. http://dx.doi. org/10.1371/journal.pmed.1001484 (2013
- 3. Grosse, S. D. et al. Am. J. Prev. Med. 41, S398-S405
- Allison, A. C. Br. Med. J. 1, 290–294 (1954). Ferreira, A. et al. Cell 145, 398-409 (2011).
- Cyrklaff, M. et al. Science 334, 1283-1286 (2011). 7. Piel, F. B. et al. Lancet 381, 142-151 (2013).

CARRIED IN THE BLOOD

sickle-cell disease is often fatal in low-income nations. But researchers are beginning munderstand more about the disease and how to treat it. By Simon Pleasants.

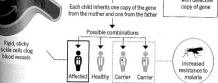
t is no coincidence that sickle-cell disease and malaria share similar global distributions. The gene responsible for sickle haemoglobin confers ased resistance to the malaria parasite when inherited from only one parent. Such a trait allowed carriers of the gene (HbS distribution, on the map below in red)1 to survive and reproduce in areas where malaria is present (green)2

Athough malaria is indigenous in many areas, the sickle haemoglobin gene has a more narrow distribution than malaria because the mutation rose in Africa. Limited migration from Africa to certain regions (for example, China) has restricted its spread.

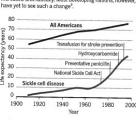
frequency (%)







As a result of more effective treatments, the life expectancy of people with sickle-cell disease in the United States has increased dramatically. Most developing nations, however



UNDERSTANDING THE DISEASE

The knowledge of and available treatments for kle-cell disease have evolved over the past 175 years.

Hydroxycarbamide, a cancer therapy, found to reduce the frequent, painful complications of sickle-cell disease

Haemoglobin protein

disorder is caused by a

single amino acid change.

sequence shows the

transplants sickle-cell disease in 9 of 10 severely affected adults.

Stem-cell

Low-oxygen conditions found to be First apparent case of the cause of sickled red blood cells in sickle cell reported. people with the disease

Linus Pauling and maps human colleagues show that haemoglobin sickle-cell haemoglobin genes, differs in structure from | pinpointing the | 90% in children with healthy haemoglobin.

Richard Flavell Periodic blood transfusions shown to reduce risk of stroke by

ces: 1. Piel, F. B. et al. Lancet 381, 142-151 (2013); 2. Piel, F. B. et al. Nature Comms. 1, 104 (2010); 3. US National Institutes of Health: National Heart, Lung and Blood Institute,

James Herrick provides

sickle-cell anaemia.

first formal description of

NORMAL CELLS

SICKLE CELLS

Sickle cells contain a defective form of haemoglobin that, when not bound to oxygen, rescent shape. The sticky, rigid cells block small

capillaries, which results in pain and other serious

complications. Sickle cells have a lifespan of

sibly higher carbon monoxide levels5 or the interaction between oxygen and sickle haemoglobin6 — makes people with the sickle-cell Simon Pleasants is a copy editor at Nature. trait less of a hospitable host to the parasite and usually prevents the disease from taking hold.

About 70% of global cases of sickle-cell disease are in Africa. The World Health Organization estimates that in some African countries, including the Democratic Republic of Congo, Ghana and Nigeria, prevalence of sickle-cell trait is between 10% and 30%, resulting in an estiareas with hot climates - Africa, the Indian mated disease prevalence of at least 2% in most

EA Representation

- A data structure and its defined variation operators
- Choice of representation defines the search space
 - Defines the possible solutions
 - Data structure choice
 - Defines the connection
 - Variation Operation
- How we represent a problem is a key to a good EA
- Good representations allow smaller search spaces and faster searches through the space

Self Avoiding Walks

■ Take an NxN grid, a Self avoiding walk of size N-1 is a walk with NxN steps visiting every square once

V	>	V	>	V	X
٧	٨	V	٨	٧	٨
٧	^	V	^	V	٨
٧	٨	V	^	V	٨
٧	٨	V	^	V	٨
>	٨	>	^	>	٨

How Can We Represent an Agent's Movement?

- Directionally we could put each of the movement directions into a string
 - UP, DOWN, LEFT, RIGHT
- The previous example would be:
- The size of the search space is 4 directions raised to N-1 steps

Can We Do Better?

D	R	D	R	D	X
D	U	D	U	D	U
D	U	D	U	D	U
D	U	D	U	D	U
D	U	D	U	D	U
R	U	R	U	R	U

Agent with Direction Saved

- Lets have an agent which saves its current direction it faces in
- Now movements are Forward (F), Turn Left + Forward (L), Turn Right + Forward (R)
- The solution string from before is:
 - FFFFFRRFFFRRFFFRRFFFRRFFF
- Only 3 symbols raised to the N-1
- Large reduction in the search space

What Do We Lose

F	R	R	R	R	X
F	F	F	F	F	F
F	F	F	F	F	F
F	F	F	F	F	F
F	F	F	F	F	F
R	R	R	R	R	R

Search Space has Been Reduced

- All legal walks still retained in the Saved Direction
- Walks where we backtrack are removed remember this is not a movement we would be interested in
- The only walks we have reduced from the solution space where those who would never provide a correct solution reduction of the space without consequence

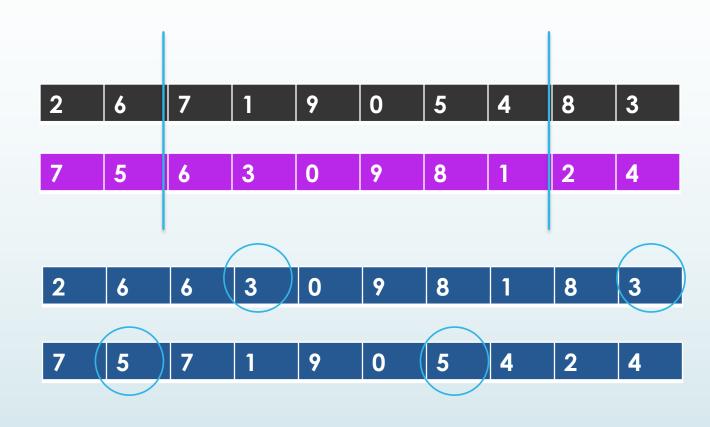
Epistasis

- Genetics Epistatic gene, in genetics, a gene that determines whether or not a trait will be expressed
- EAs the actions made in earlier parts of the chromosome having larger impacts on the fitness value of the chromosome
- Self Avoiding Walk actions made early in the string will have large impacts on the fitness
- For example, changing the first move to a R will run the walk off the stage, giving it 0 fitness rather than 35, whereas changing the last symbol to a R will only reduce the fitness by 1

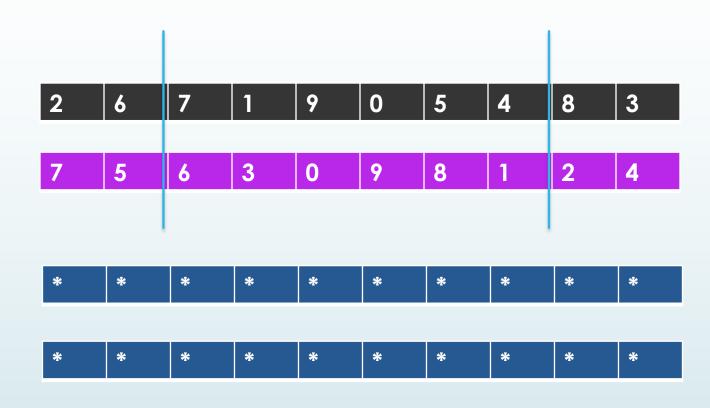
TSP-Ordered Gene

- GAs representation is the tour of cities
- Natural idea of ordering in gene

Two Point Crossover



Partial Mapped Crossover

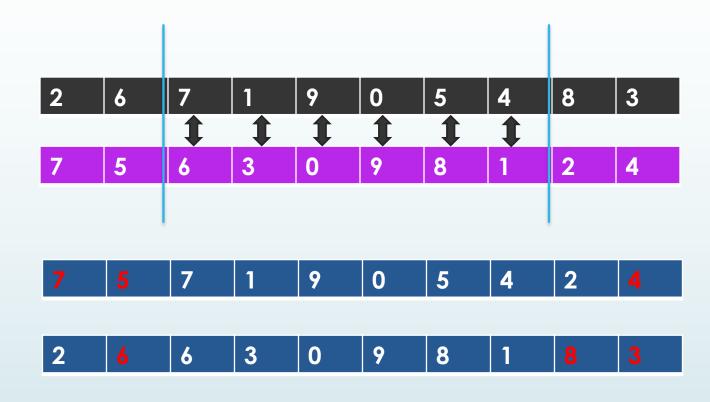


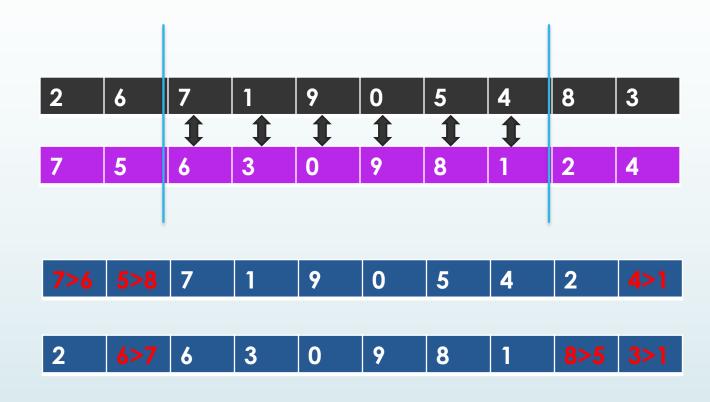
Partial Mapped Crossover

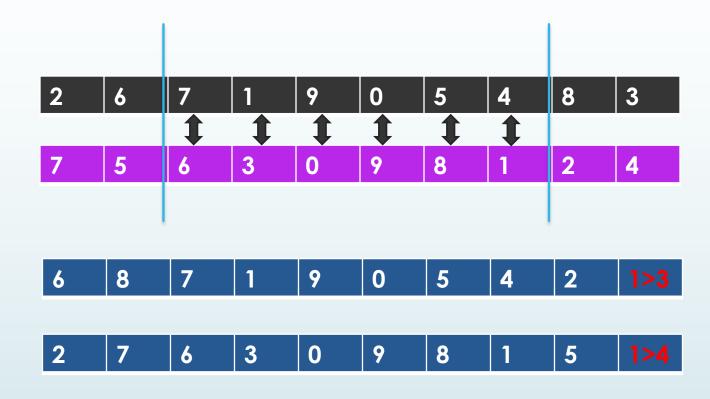


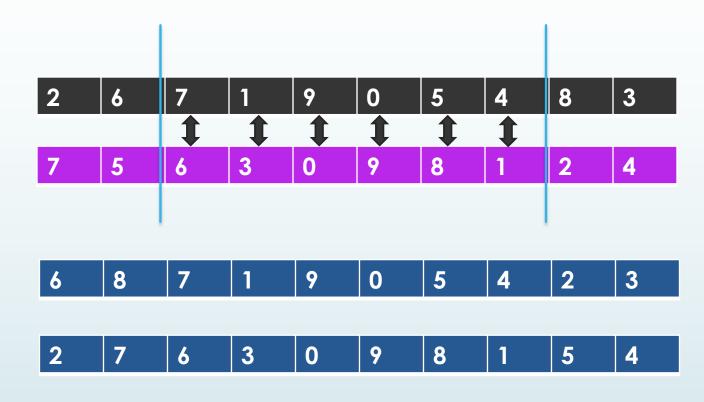
Partial Mapped Crossover











Indirect Real Numbered Representation

- Move from direct to indirect representation
- Real valued string
- Crossover as per one and two point
- Mutation is now as per ES add a random number from a distribution
- Mapping from
- Example

0	1	2	3	4	5	6	7	8	9
34.3	53.1	22.1	65.5	34.5	12.4	0.7	66.3	96.2	0.12

9, 6, 5, 2, 0, 4, 1, 3, 7, 8.

Indirect Real Numbered Representation

- Direct to Indirect method
- No longer need to worry about ordering
 - Ordering is now taken care of in the list
 - Sorting algorithms are fast
- Evolutionary Changes are less likely to numbers which are closer together
 - Evolution also has a direct control on the heritability

Selection of Representation

- Has a large impact on the outcome of the evolution
- Defines how your space will look the search space in terms of the connectivity on the space
- Genetic operations are defined by how they act upon the representation of the problem

Iterated Prisoner's Dilemma

- Two player game
- Story of two convicts who are placed in interrogation and played against each other
- Exploitable player allows a score >3 against it

		Player 1	
		C	D
Player 2	C	(3,3)	(5,0)
	D	(0,5)	(1,1)

Iterated Prisoner's Dilemma

- Game playing agents
- Representations
 - Function Stack Machines FST
 - Cellular Encoded FSM CAT
 - ► Finite State Machines AUT
 - Boolean Formulas TRE
 - Markov Chains MKV
 - Look-up Tables LKT
 - IF-ACTION-SKIP Lists ISC
 - Boolean Formula with Delays DEL

Co-evolution

- The fitness of the population is based not upon the target, but the other members within the population
- Used to create game player agents
- For IPD
 - Play each of the other players in the population once
 - Score is what you make off of all the other members in the population

Cooperative Ability

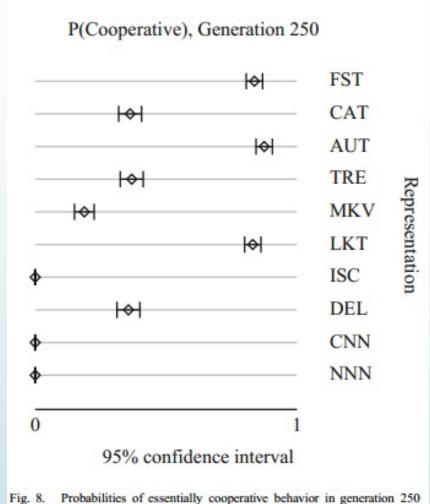
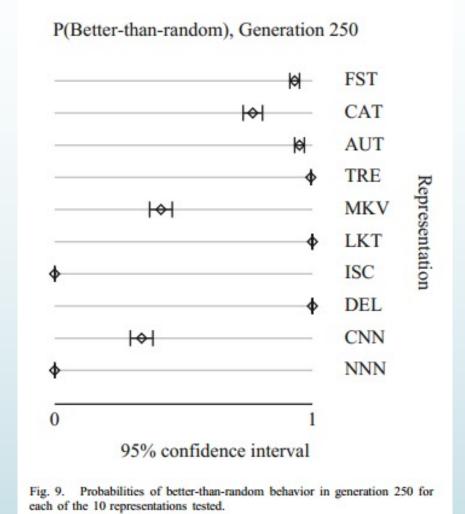


Fig. 8. Probabilities of essentially cooperative behavior in generation 250 for each of the 10 representations tested.

Play in the Population v. Random player



Mhys

1. Initial Population

Random initialization of all of these techniques has some differences in the ability to create an early cooperator

 Variation Operators with Different Probabilities of Creating "Stupid" Offspring

May allow for offspring which are easy to exploit

3. Fidelity of Reproduction

Differences in the ability to represent ones parents. Some reproduction rules lead to drastically different children than parents

4. Granularity of Reproduction

Cooperation must be valued to be sustained, if newer models are allowed in which can be exploited it allows for exploiters to overrun the population

Epistasis

- Genetics Epistatic gene, in genetics, a gene that determines whether or not a trait will be expressed
- EAs the actions made in earlier parts of the chromosome having larger impacts on the fitness value of the chromosome

String Players

- String of actions
- Actions made in a step-wise order or are representations of the player
- Have been used for a number of simulation ideas:
 - Sunburn Space Ships
 - VIPs Virtual Politician Simulator
 - Lessons to a young man courting Dating Simulator
 - Dawn of the Dead Government Responses to Zombies
- Co-evolution and Evolution to a target
 - Sunburn co-evolved
 - VIPs co-evolved
 - Lessons Based on a model of the potential date
 - Dawn of the Dead based on an epidemic model

Sunburn

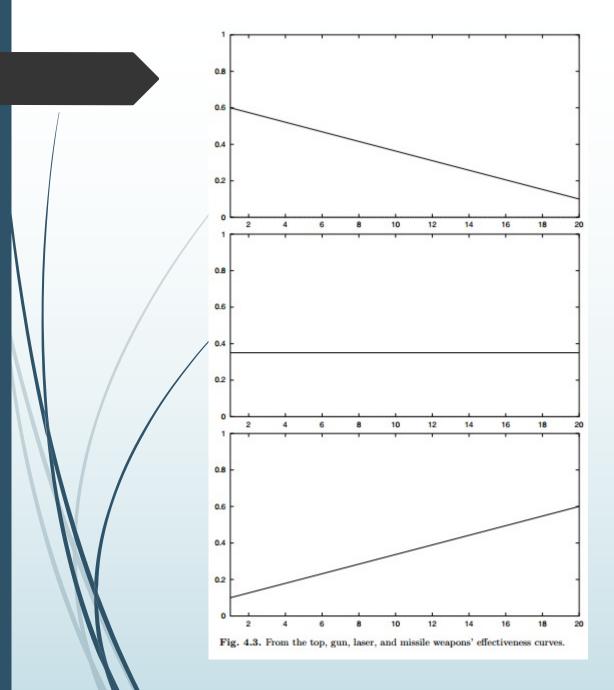
- Space Warfare Simulation
- Student Project by John Walker
- Ships are represented by a string which is an indirect method of building a starships
- Ships compete in a gladiatorial tournament selection method
- ► Ships have a distance measure from each other 1 dimension

Sunburn Ship

- Length N string of characters
- (S)hields Take a hit of damage before dropping
- Weapons
 - (G)uns Short Range Physical Projectiles which drop in effectiveness as distance increases. Slow moving and able to dodge
 - (L)asers Can be equally effective at any range
 - (M)issiles Guided missiles which take some time in order to ensure a lock. They have a better long range effectiveness as they track their targets
- (D)rives allow a ship to move one space
- String kept in order Shields are moved to front of string and tripled. Each system is worth one point of damage
- Ships have a preferred distance

Weapon Effectiveness

- A weapon hitting does one system worth of damage
- A damaged system cannot be used
- Weapons are probabilistic in their effect
- Hit based upon the weapon effectiveness curve a function based on the current range
 - Arr PG(r) = 0.6 (r 1)/38
 - \rightarrow PL(r) = 0.35
 - \rightarrow PM(r) = 0.1 + (r 1)/38









Example Ship Translation

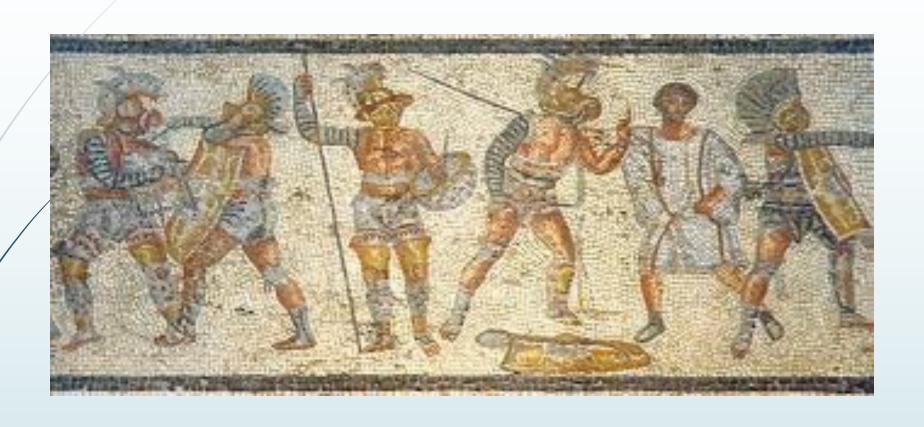
- SGLMSDSMGSLD,16
- Systems with shields in the chromosome
- Translation moves these to the front and triples them
- 4 shield nodes buys you 12 shields
- SSSS SSSS SSSS GLMDMGLD,16

Ship Controllers

- Ships activate their drives in unison one at a time with the goal of moving to their key range
- Ships fire weapons one at a time in unison, based-on their order in the string LLGML

FIRE! - LASER, LASER, GUN, MISSILE, LASER

Gladiatorial Selection



Gladiatorial Fights

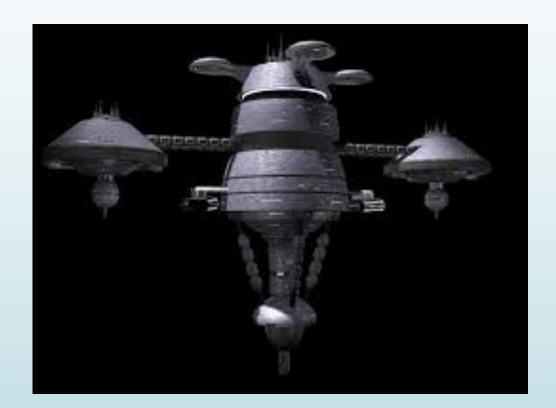
- Steady State Population
- 1. Pull out two ships from the population
- 2. Simulate their battle
- 3. If there is a winner keep for breeding
- 4. Else return to 1
- 5. Pull out two ships from the population
- 6. Simulate their battle
- 7. If there is a winner keep for breeding
- 8. Else return to 5
- 9. Breeding Step

Win Conditions

- A ship without any points remaining is destroyed
- 100 Turns without damage Draw
- Are there any others?

Starbases

- Strip down to one drive in the final position
- Load up on Missiles and Shields



Gunrunners

- Lots of drives
- MORE GUN



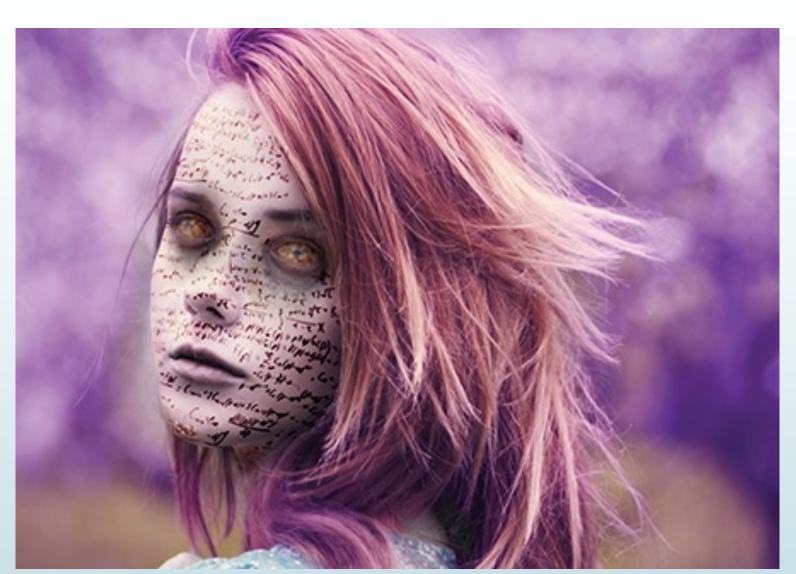


Representational Issues

- Position within the chromosome has a large effect on the range and the damage
- Systems 'go down' in order
 - Expendable systems at front
 - Important systems near end
 - Win condition is a disabling of drives ensure at least one drive is at the end of the string
- Epistatic with importance on the end of the string

Mathematical Modeling of Zombies

(2014)



Zombies

- Based on term project
 - Clinton Innes/Joseph Alexander Brown
- Defines the actions which can be made by a government during an outbreak as a string of actions
 - Based on movies of Romero
- Zombies follow a model of epigenetic networks used in actual diseases, with some modifications to represent the raising of the dead

Model

Extension on the Susceptible, Infected, Removed (SIR) Model to Zombie, Susceptible, Infected, Removed (ZSIR) Model

- Z[t] = Z[t-1] + I[t-d] R[t] // zombies are now the previous zombies, plus those raised from the infected after a delay, minus those removed (i.e. killed)
- S[t] = S[t-1] I[t] // the remaining population has some infected
- I[t] = β * Z[t-1] * S[t-1] // zombies infect at a rate of β based on the current number of zombies and humans who are uninfected
- R[t] = α * Z[t-1] * S[t-1] // zombies are removed at a rate of α based on the current number of zombies and humans who are uninfected

β-value

- Zombie's ability to infect humans
- Fast v. Slow zombies Aggression
- Airborne, bite based, Virus Infectivity
- Vaccinations for the virus Immunity
- lacktriangle Aggression and Infectivity Determine the eta

a-value

- Humans ability to kill zombies
- Weaponry/Training of Humans
- Aggressiveness
- α-value is defined based on their training and their want to hunt zombies

Fitness Evaluation

- Number of survivors at the end of the simulation time
- Trouble at edges
 - If all humans die these situations are equally as bad in this model
 - Allow for a negative fitness one of the actions sends in Army into the area sending in troops who all die is worst than not – changed the definition of fitness to number killed based on original population
- Might be in a losing state at the end with still having a positive population
 - Time till we save the population
 - May want to have a multiple objective model balance between humans and zombies

Government Actions

- Government has a number of actions available to it
- Each time step of the simulation is assumed to be the time needed to coordinate the effort and abilities of the population
- Governments begin with an amount of popularity, and some actions are probabilistic based on the popularity

Actions

- Quarantine Government issues order to population to round up suspected infected and keeps them from the population
- Warn Government provides warnings to the populace
- Kill Government opens local armories to the populace
- ► Fear Monger Government scares population with propaganda over the danger of the zombies
- Army* Government sends in troops
- Science* Government sends in scientists to produce a better vaccination
- Vaccinate* Government issues vaccinations
- *stochastic effect

Actions

- Quarantine Zombies infectiveness down, slight increase in approval of government
- Warn increase in Human effectiveness, approval of government increased slightly
- Kill large increase in Human aggressiveness/effectiveness, chances of zombie interaction increase
- Fear Monger increase in Human effectiveness slightly, approval of government increased
- Army* increase in Human effectiveness if population approves mission, 100 extra humans, government loses approval when zombies outnumber survivors – Sending our sons into harms way needlessly
- Science* increase in vaccine effectiveness if money approved based on governmental popularity
- Vaccinate* if vaccine is effective then reduction in the zombie infectivity. Government loses approval if number of zombies is less than number of humans "I don't see the need."
- *stochastic effect

Early Actions in the Model

- Early on:
 - Vaccination is not very effective
 - Humans are not well trained
 - Have different approvals of their government
- Early actions are Fear Monger, Warn, Quarantine
- Later Actions are Kill and Quarantine
- Very Rarely is Army called in
 - Provides more bodies

Government Approval

- Governments without approval are more likely to Fear Monger
 - I can make them love me
- Governments with approval are more likely to Warn
 - I can get better
- Governments are unlikely to use actions which are stochastic
 - Only a few times will it work, known actions with produced effect are more heritable into later populations
 - Large cost in using them
 - However, both the loved and feared governments both show they would like to have the option – they will work to gain public opinion – vestigial effect? (note that this finding was not significant at the 95% confidence but at 80% - given more time to evolve?)

Other Models

- Lessons to a young man courting
 - Work to gain money
 - Phone call
 - Present
 - Ask out
 - Etc.
- VIPs
 - Raise Funds
 - Adopt popular program
 - Bribe/Pander
 - Do what the other guy did last time
 - Lay low
 - Negative Campaign
 - Scandal

SAT-Problem

- William M. Spears. "Evolutionary Algorithms: The Role of Mutation and Recombination" Springer, 2000
- Looking at producing problems with various properties
- SAT problem
 - Basis of NP-complete
 - Polynomial
- "it turns out that epistatic problems can be easily constructed from conjunctive SAT expressions"

SAT-Problems

- lacktriangle $(a \lor b) \land (c \lor d)$
- Each conjunct can be seen as the allele of an organism
- A conjunct may appear in a number of solutions to the same problem
- *Bloat issue*

Sunburn

- What crossovers would make sense?
- What crossovers would disrupt structure?
- What about the preferred distance?