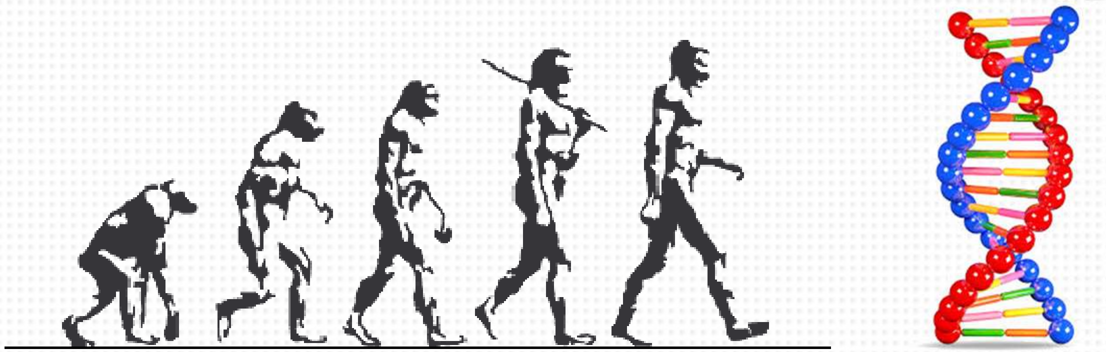
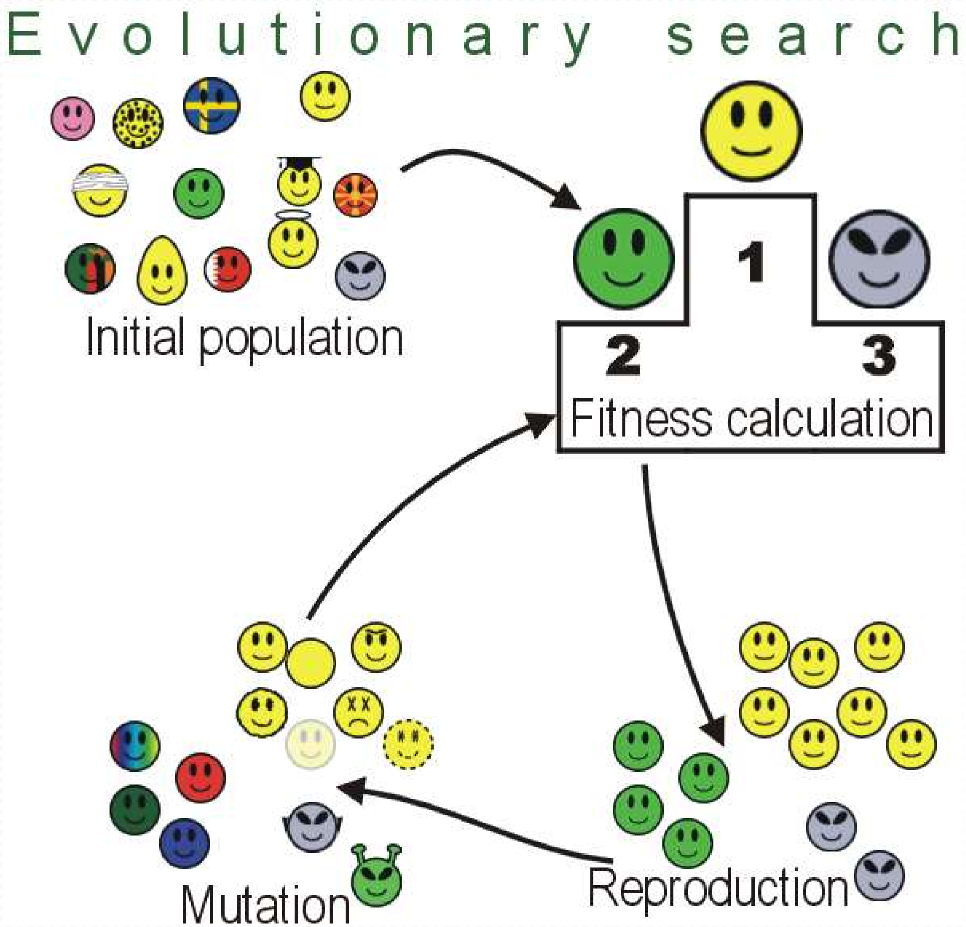
Evolutionary Algorithms

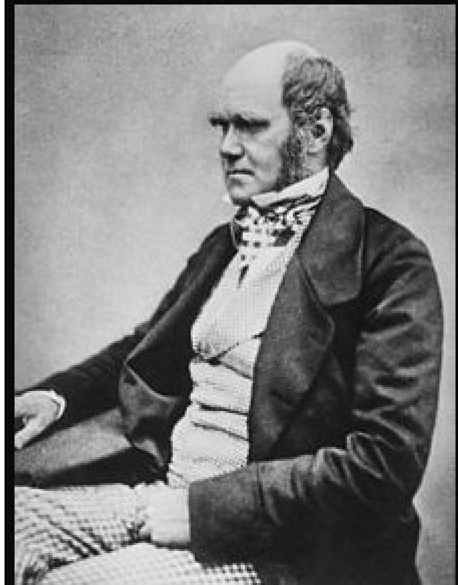
Dr. Bo Yuan



Overview

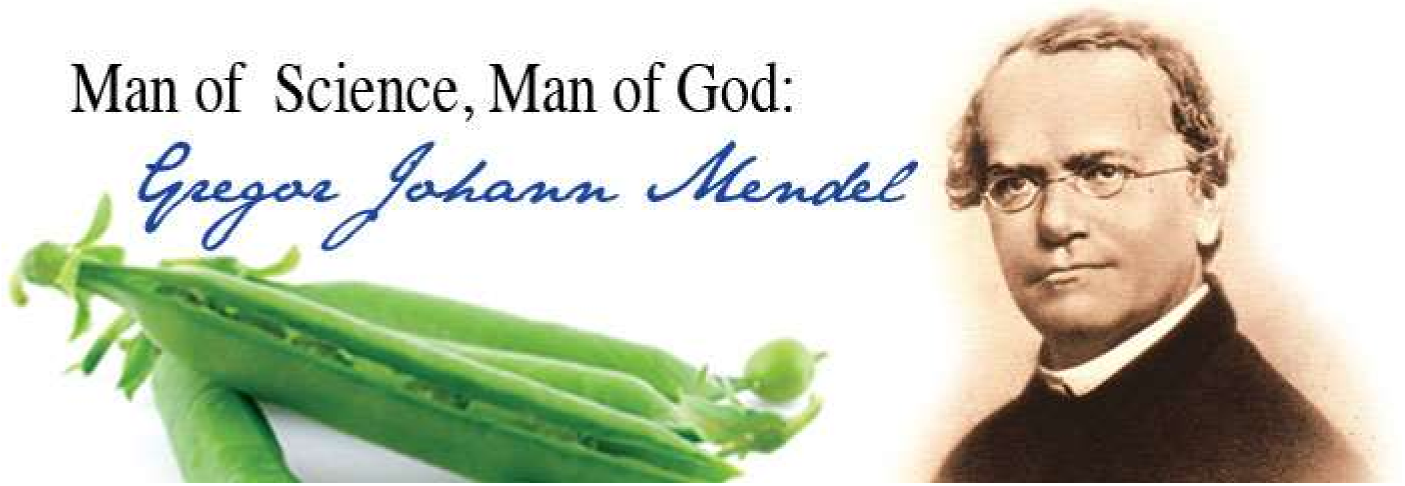
- Global Optimization
- Genetic Algorithms
- Genetic Programming
- Evolvable Things

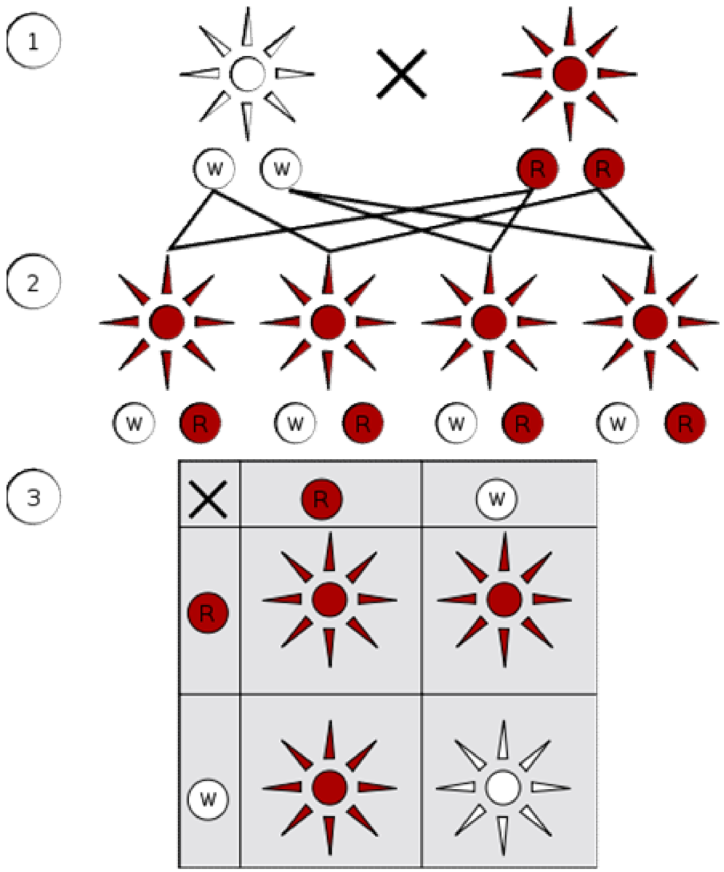


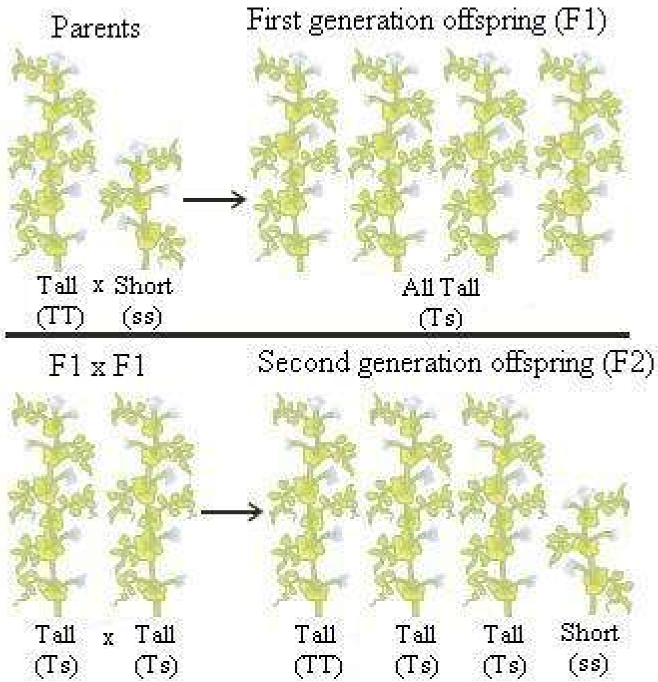


It is not the strongest of the species that survives, nor the most intelligent that survives. It is the one that is the most adaptable to change.

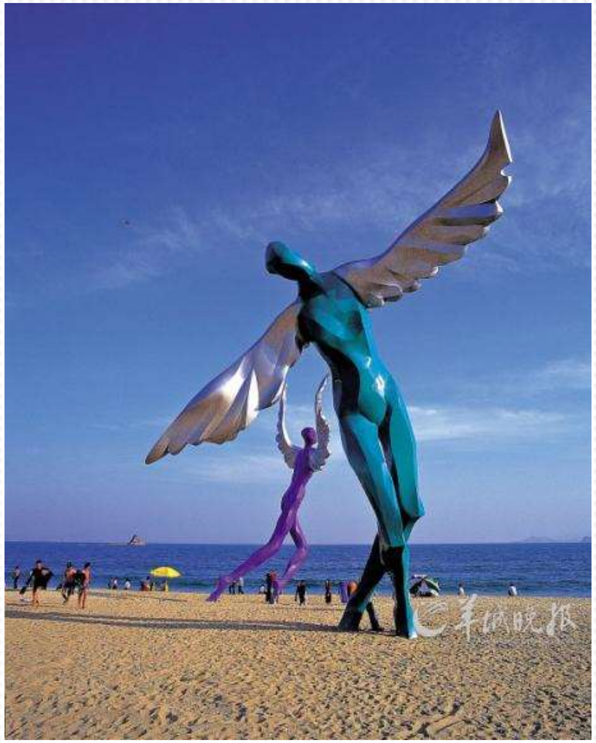
(Charles Darwin)


Man of Science, Man of God:
Gregor Johann Mendel


1
2
3
W
R
X


Parents
First generation offspring (F1)
Tall x Short
(TT) (ss)
All Tall
(Ts)
F1 x F1
Second generation offspring (F2)
Tall x Tall
(Ts) (Ts)
Tall
(TT)
Tall
(Ts)
Tall
(Ts)
Short
(ss)

Learning from Nature



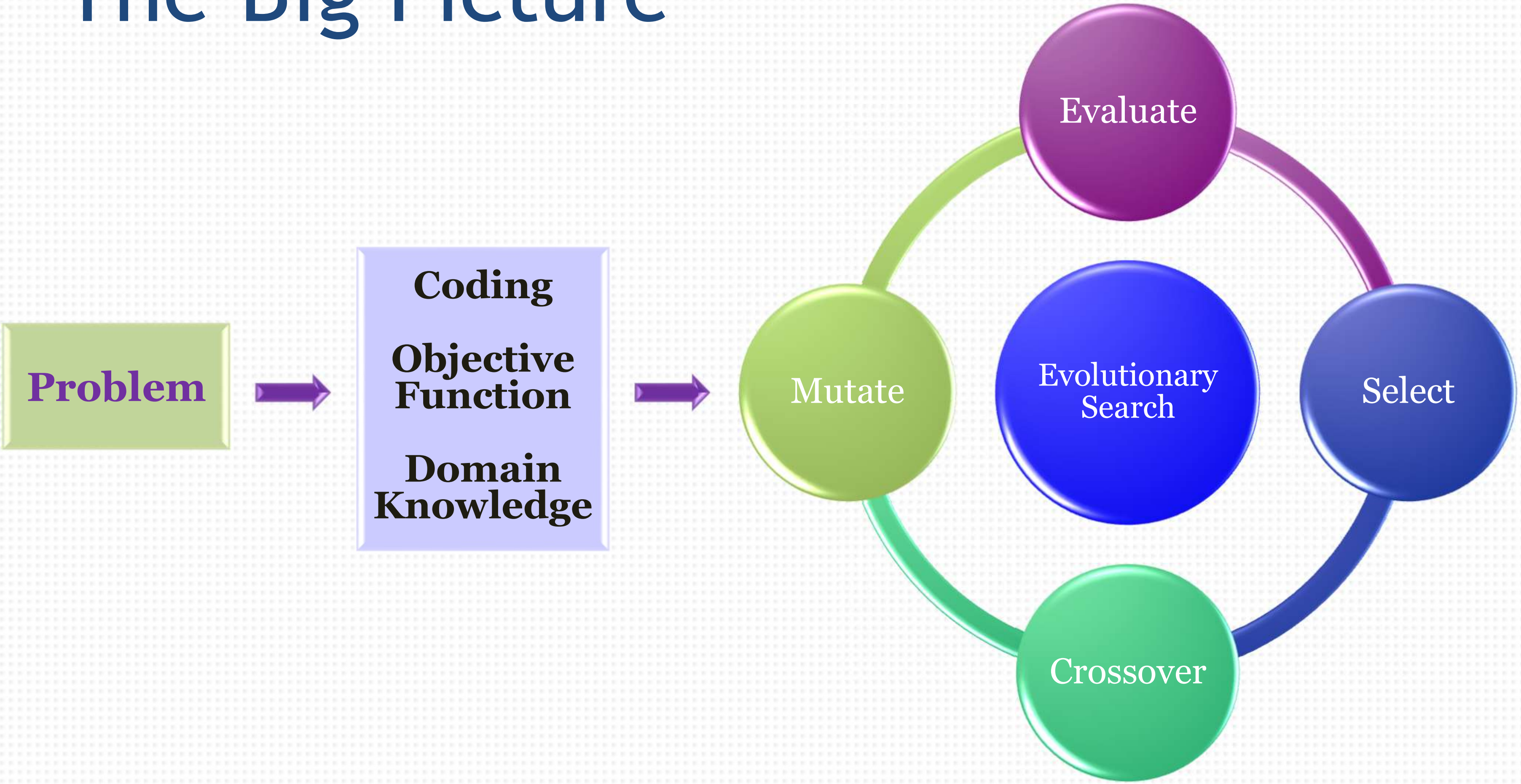
Motivation of EAs

- What can EAs do for us?
  - Optimization
  - Help people understand the evolution in nature.
- What is optimization?
  - The process of searching for the optimal solution from a set of candidates to the problem of interest based on certain ***performance criteria***.
  - Accomplish a predefined task to the highest standard.
    - Job Shop Problem
  - Produce maximum yields given limited resources.
    - Investment Strategy

Key Concepts

- Population-Based Stochastic Optimization Methods
- Inherently Parallel
- A Good Example of Bionics in Engineering
- Survival of the Fittest
- Chromosome, Crossover, Mutation
- Metaheuristics
- Bio-/Nature Inspired Computing

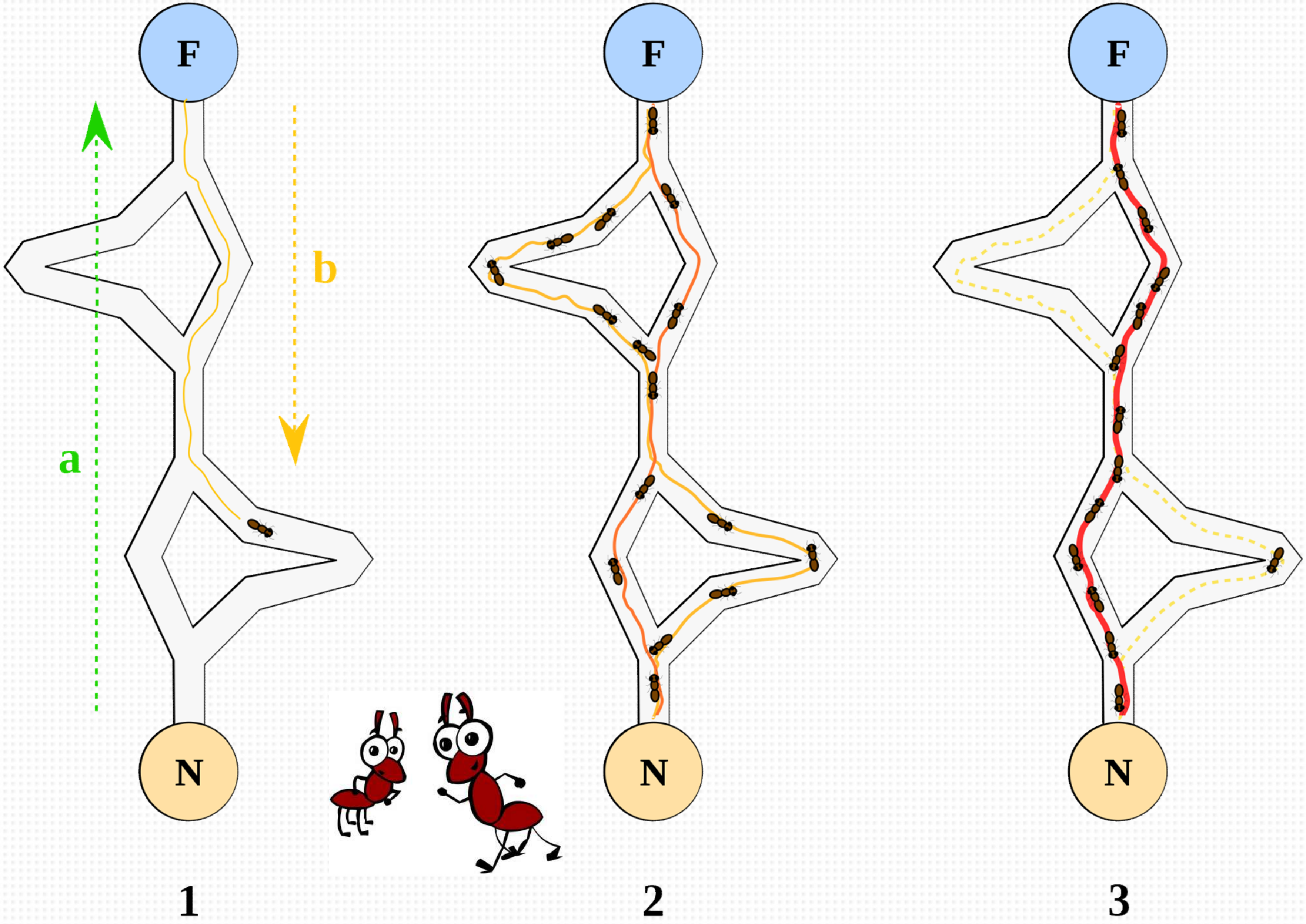
The Big Picture

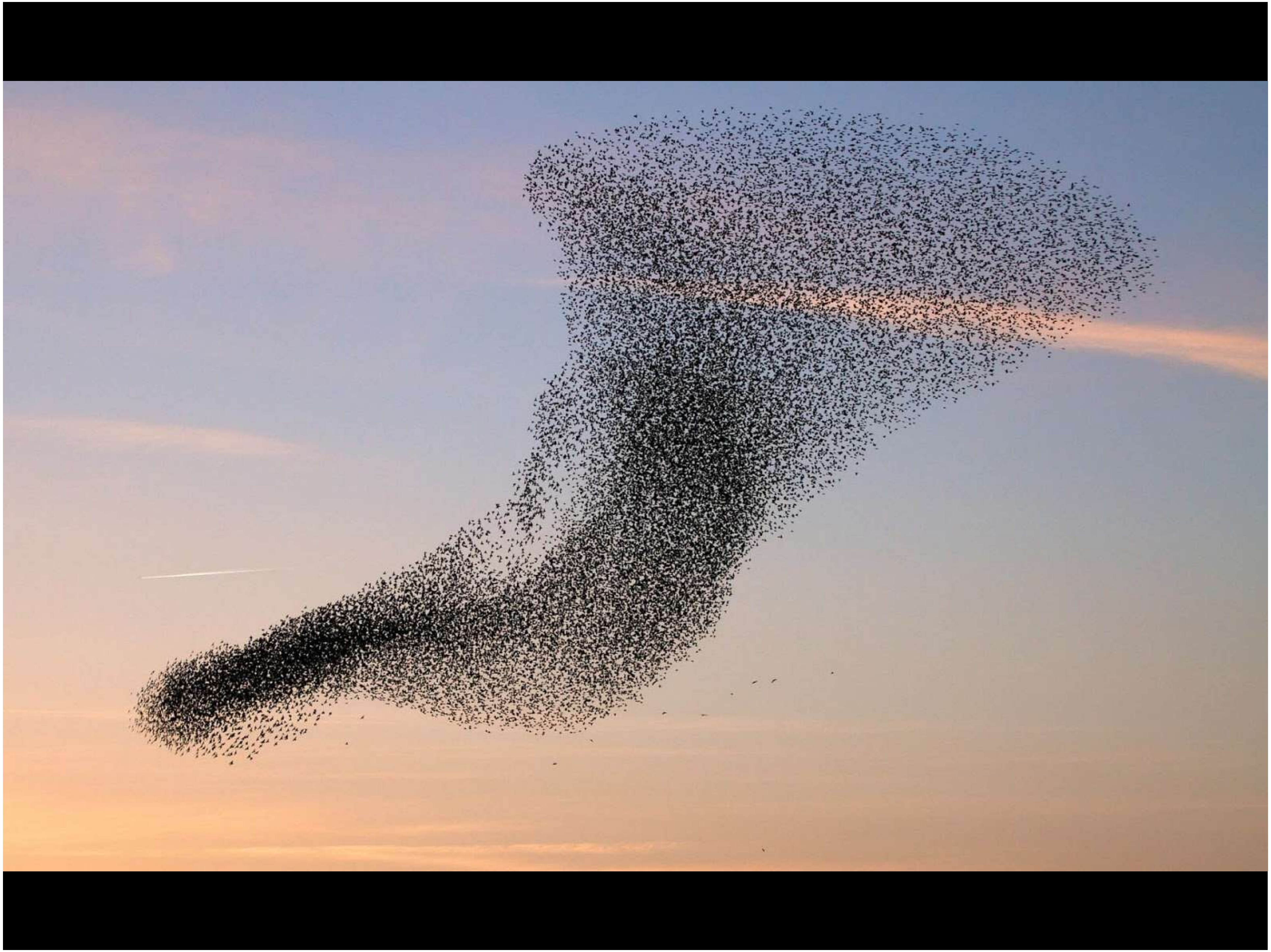


EA Family

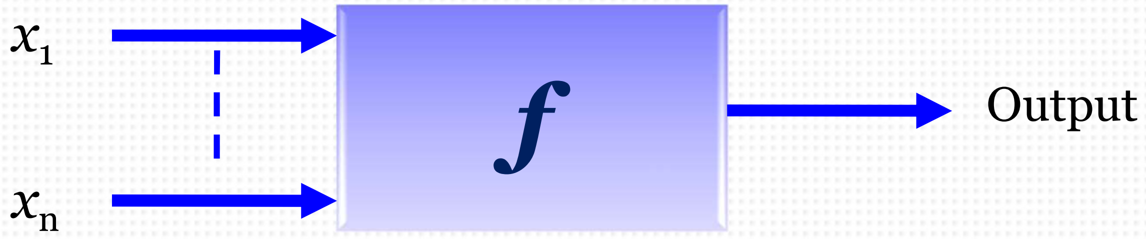
- GA: Genetic Algorithm
- GP: Genetic Programming
- ES: Evolution Strategies
- EP: Evolutionary Programming

- EDA: Estimation of Distribution Algorithm
- PSO: Particle Swarm Optimization
- ACO: Ant Colony Optimization
- DE: Differential Evolution


F
F
F
a
b
N
N
N
1
2
3



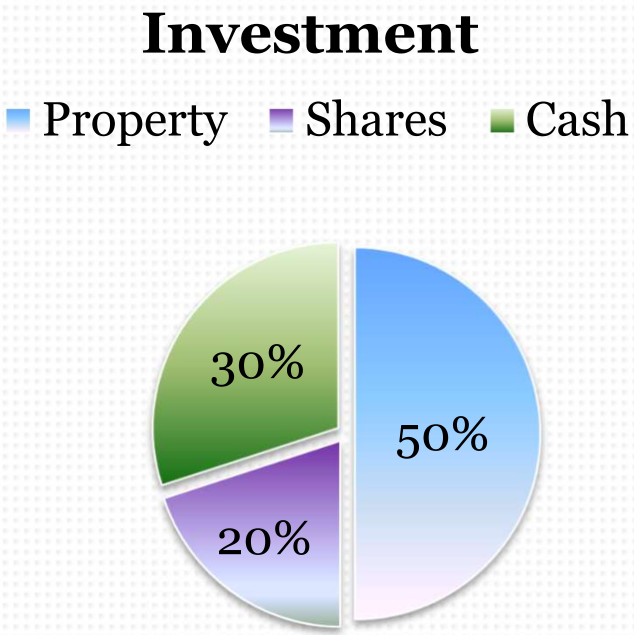
Objective Function

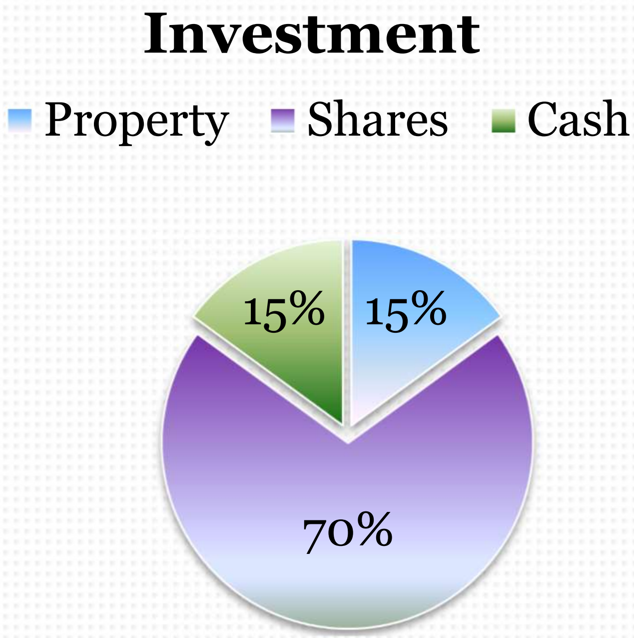


$$f(X) = \sum_{i=1}^{N} x_i^{\;2}$$

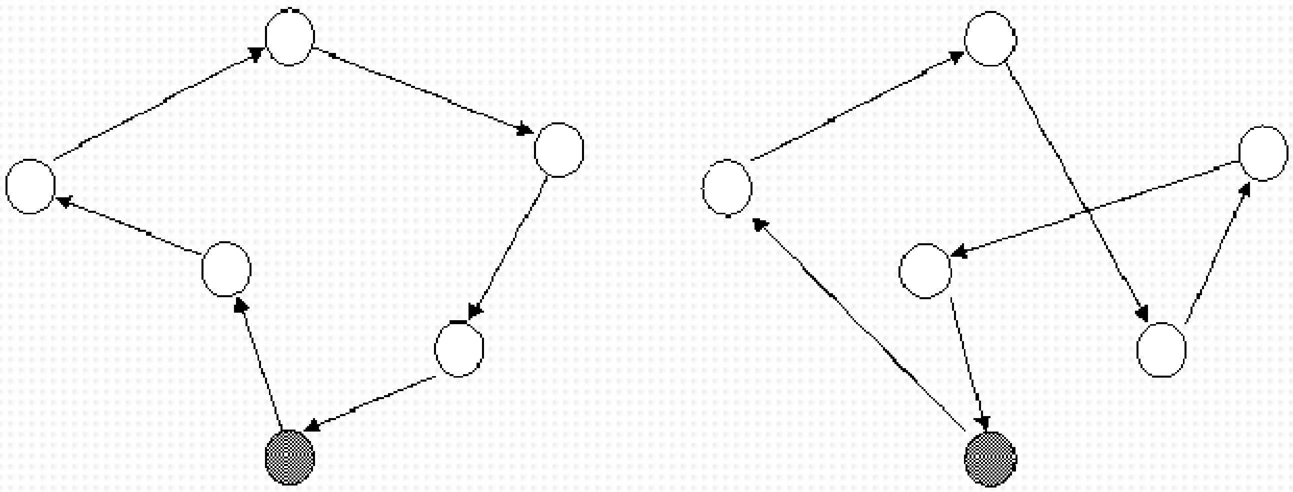
$$f(X) = \sum_{i=1}^{N-1} \left[(1 - x_i)^2 + 100(x_{i+1} - x_i^{\;2})^2 \right]$$

Portfolio Optimization





Travelling Salesman Problem

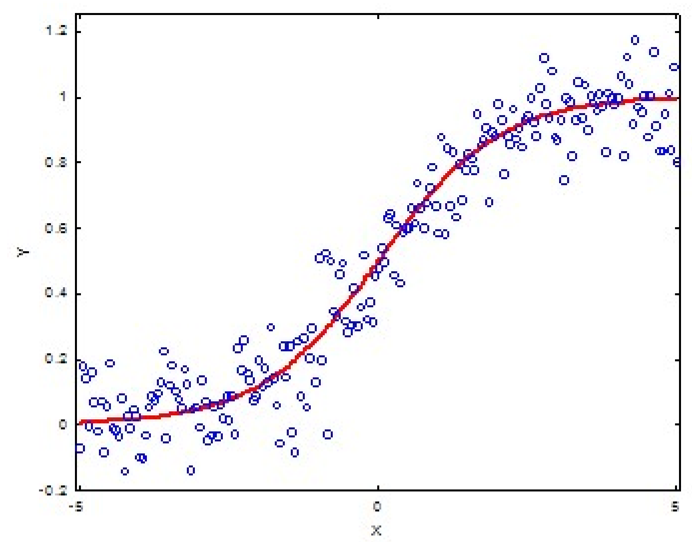


Totally 20!=2,432,902,008,176,640,000 different routes for a 20-city problem.

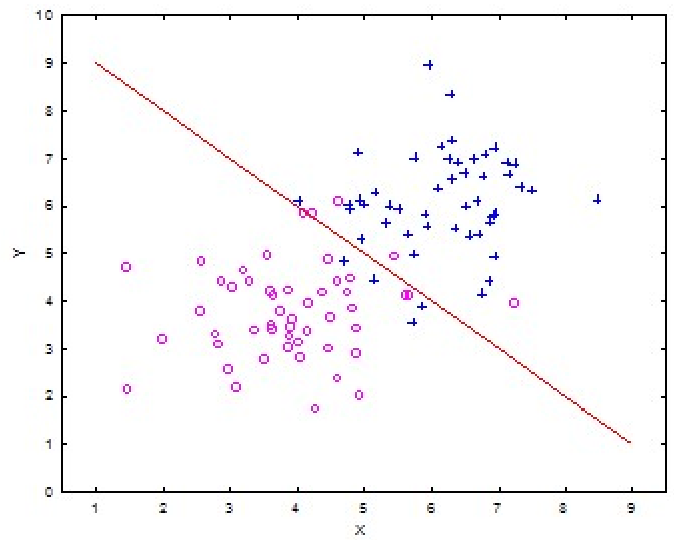
Knapsack Problem

Bin Packing Problem

Machine Learning Problems

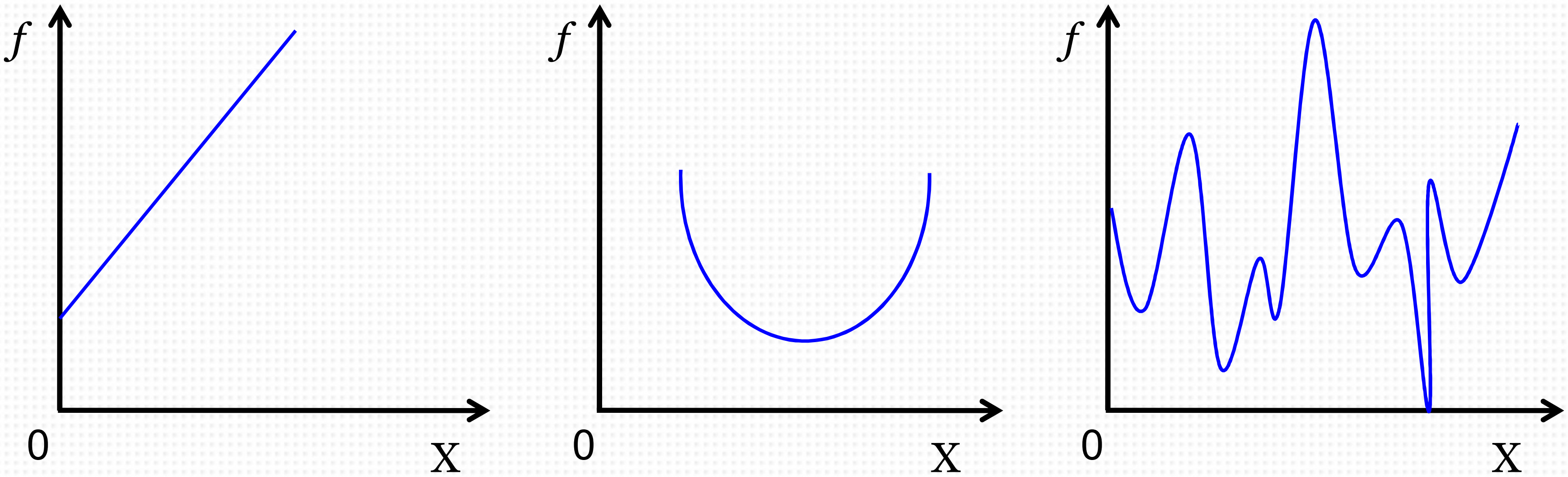


Regression

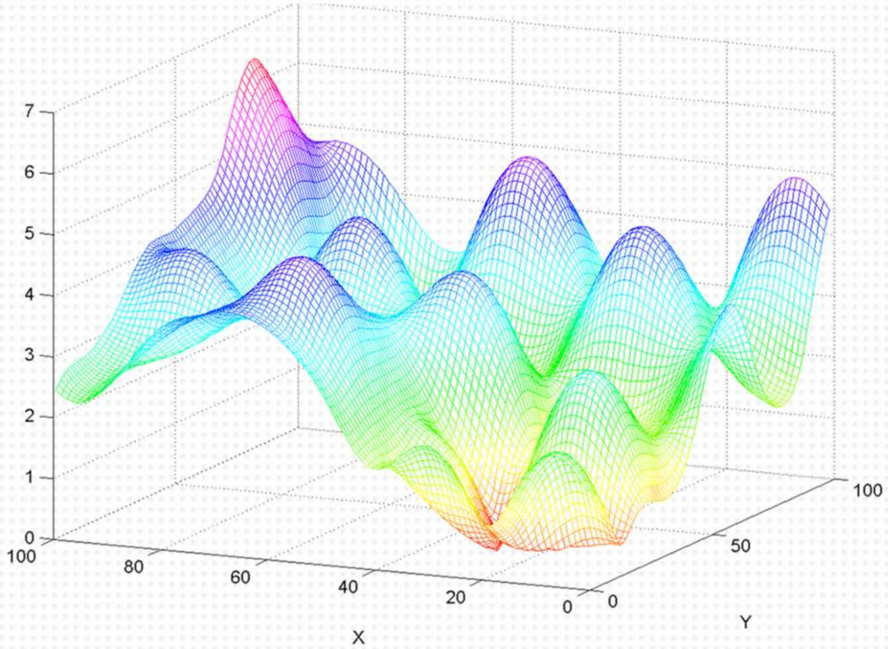


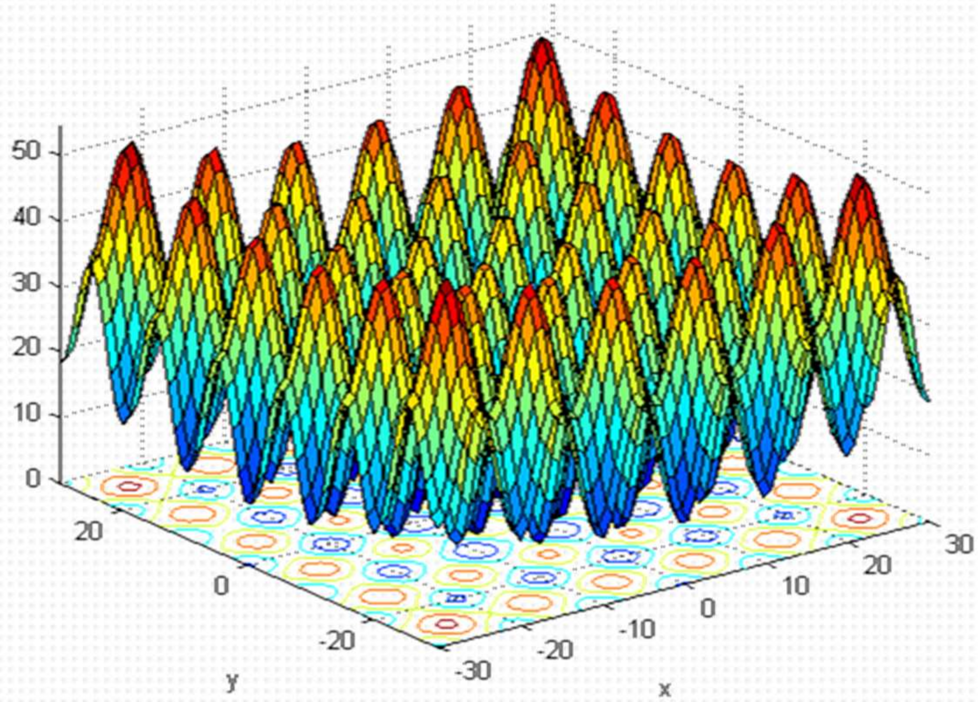
Classification

Local Optima



Local Optima + Dimensionality





The size of the search space grows exponentially!

The Bad News

- Many interesting optimization problems are not trivial.
  - The optimal solution cannot always be found in polynomial time.
- Feature Selection Problem
  - Start with N features.
    - Redundant/Irrelevant/Noisy
  - Select a subset of features.
    - Objective: Highest Accuracy
  - How challenging is this task?

The Bad News

- Given 32 features (N=32)
- 1: selected feature; 0: unselected feature
- The optimal solution may look like:
  - 0010 1100 0000 0100 0000 0000 0000 0001
  - {3, 5, 6, 14, 32}
- However, there are $\underline{2^{32}-1}$ combinations in total.
  - 1 in 4,294,967,295 chance
  - Needle-in-a-haystack situation
  - Unless you are extremely lucky....
- The Limitation of Computation

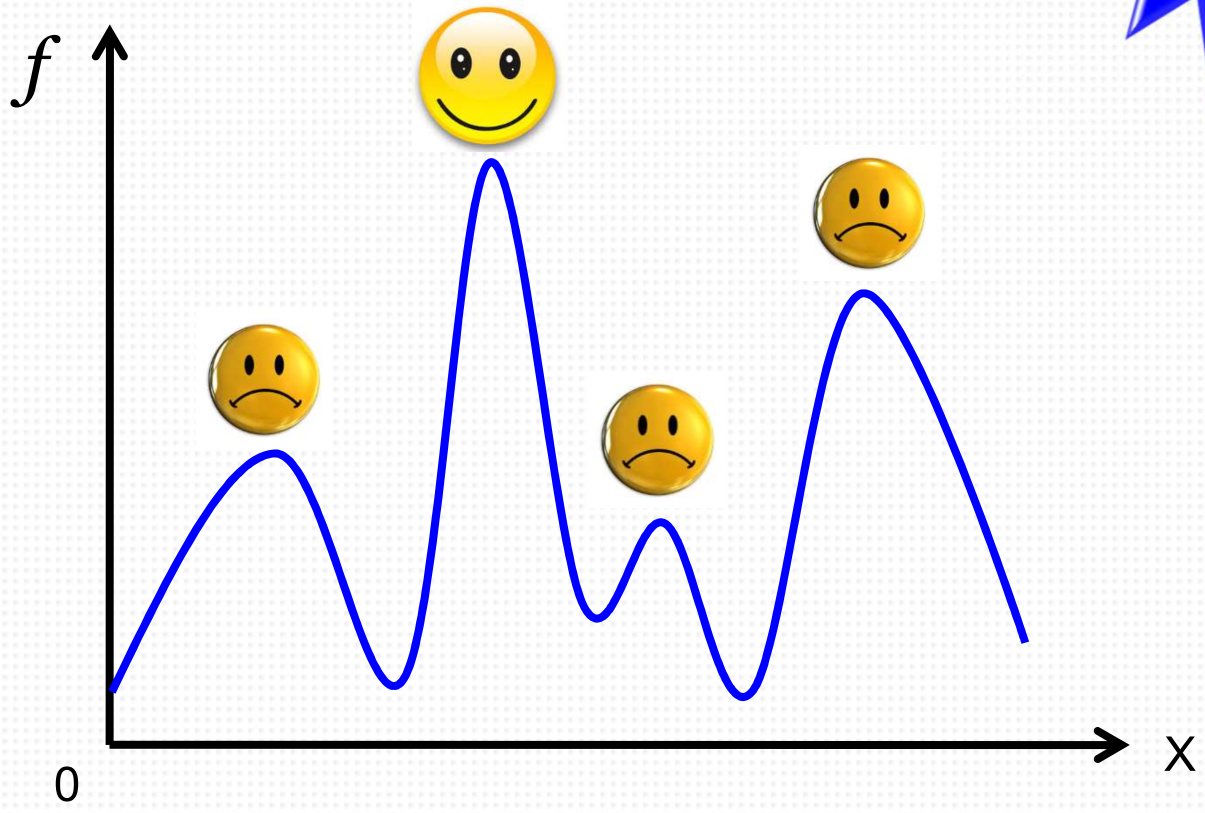
FLOPS

~6×10^{24} KG

- FLoating point OPerations per Second
  - GFLOPS (gigaFLOPS): 10^9 (每秒10亿次运算)
  - TFLOPS (teraFLOPS): 10^{12} (每秒1万亿次运算)
  - PFLOPS (petaFLOPS): 10^{15} (每秒1千万亿次运算)
- Intel Core i7 980 XE: ~100 GFLOPS
- Bremermann's Limit
  - $c^2/h = 9\times10^{16}/6.62606896\times10^{-34} \approx 1.36\times10^{50}$ bits·s^{-1}·kg^{-1}
  - Is it fast enough?

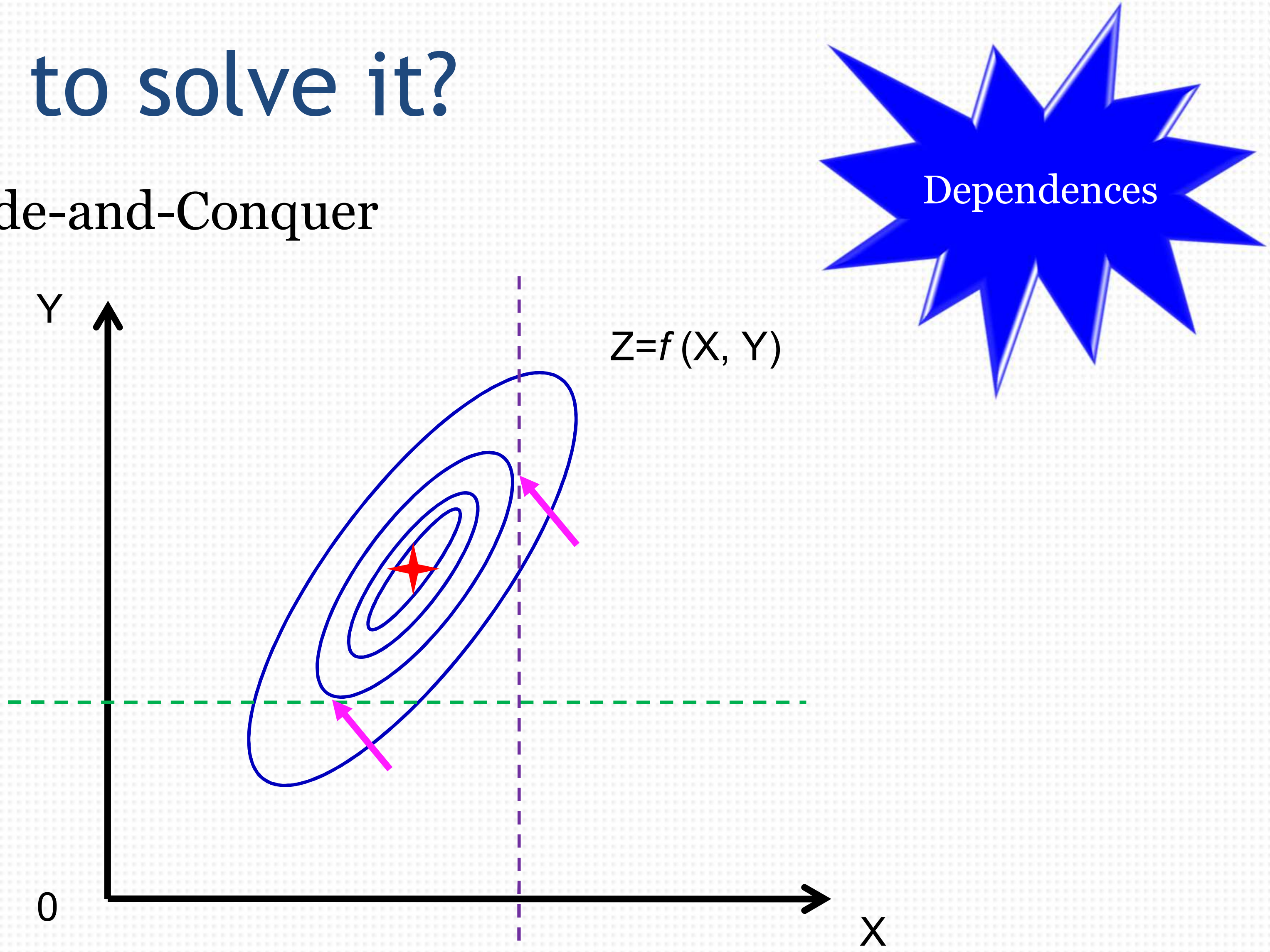
How to solve it?

- Local Search



How to solve it?

- Divide-and-Conquer



Solution: Parallel Search

- Conduct searching in different areas simultaneously.
  - Population Based
  - Avoid unfortunate starting positions.
- Employ heuristic methods to effectively explore the space.
  - Focus on promising areas.
  - Also keep an eye on other regions.
  - More than random restart strategies.
- This is where EAs come into play!

Publications

Top Journals:

- IEEE Transactions on Evolutionary Computation
- Evolutionary Computation Journal (MIT Press)

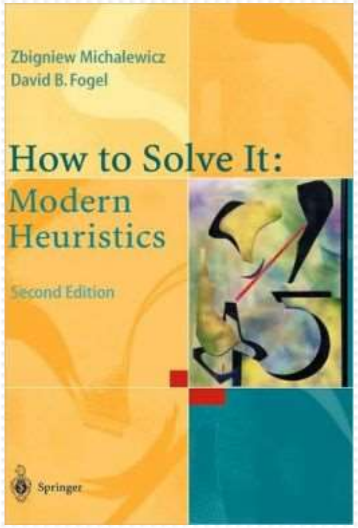
Major Conferences:

- IEEE Congress on Evolutionary Computation (CEC)
- Genetic and Evolutionary Computation Conference (GECCO)
- Parallel Problem Solving from Nature (PPSN)

People

- Prof. Xin Yao
  - The University of Birmingham
  - www.cercia.ac.uk
  - Former EiC: IEEE Transactions on Evolutionary Computation
  - Former President of the IEEE Computational Intelligence Society

- Dr. David Fogel
  - Natural Selection Inc.
  - www.natural-selection.com
  - Blondie24: Playing at the Edge of AI



- Prof. Zbigniew Michalewicz
  - University of Adelaide
  - cs.adelaide.edu.au/~zbyszek
  - How to Solve It: Modern Heuristics

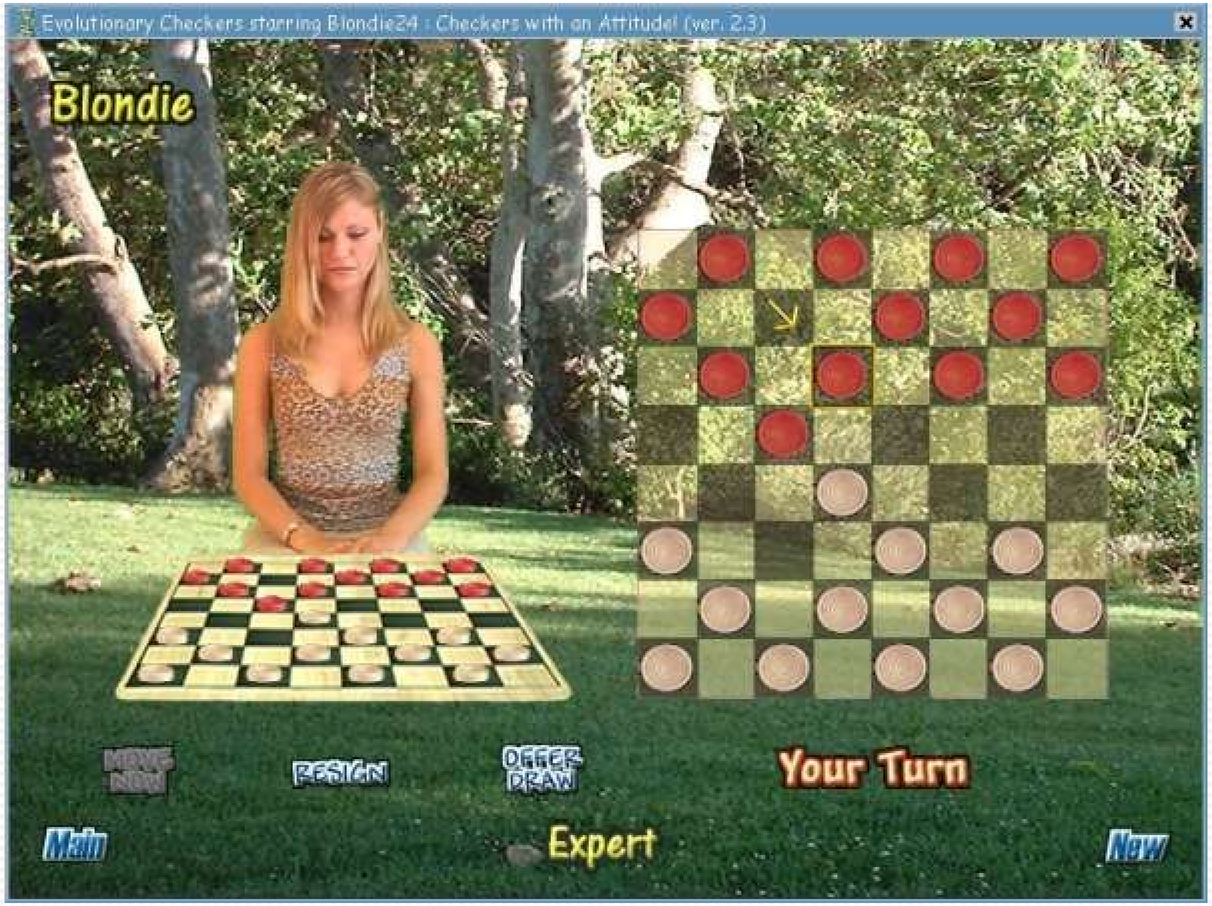
People

- Prof. David E. Goldberg
  - University of Illinois at Urbana-Champaign
  - scholar.google.com/citations?user=BUzKxsoAAAAJ
  - Genetic Algorithms in Search, Optimization and Machine Learning (1989)

- Prof. Kenneth A. De Jong
  - George Mason University
  - www.cs.gmu.edu/~eclab
  - De Jong Test Suite (PhD Thesis, 1975)
  - Evolutionary Computation: A Unified Approach (MIT Press, 2006)

- Prof. Melanie Mitchell
  - Portland State University
  - web.cecs.pdx.edu/~mm
  - An Introduction to Genetic Algorithms (MIT Press, 1996)

Blondie24

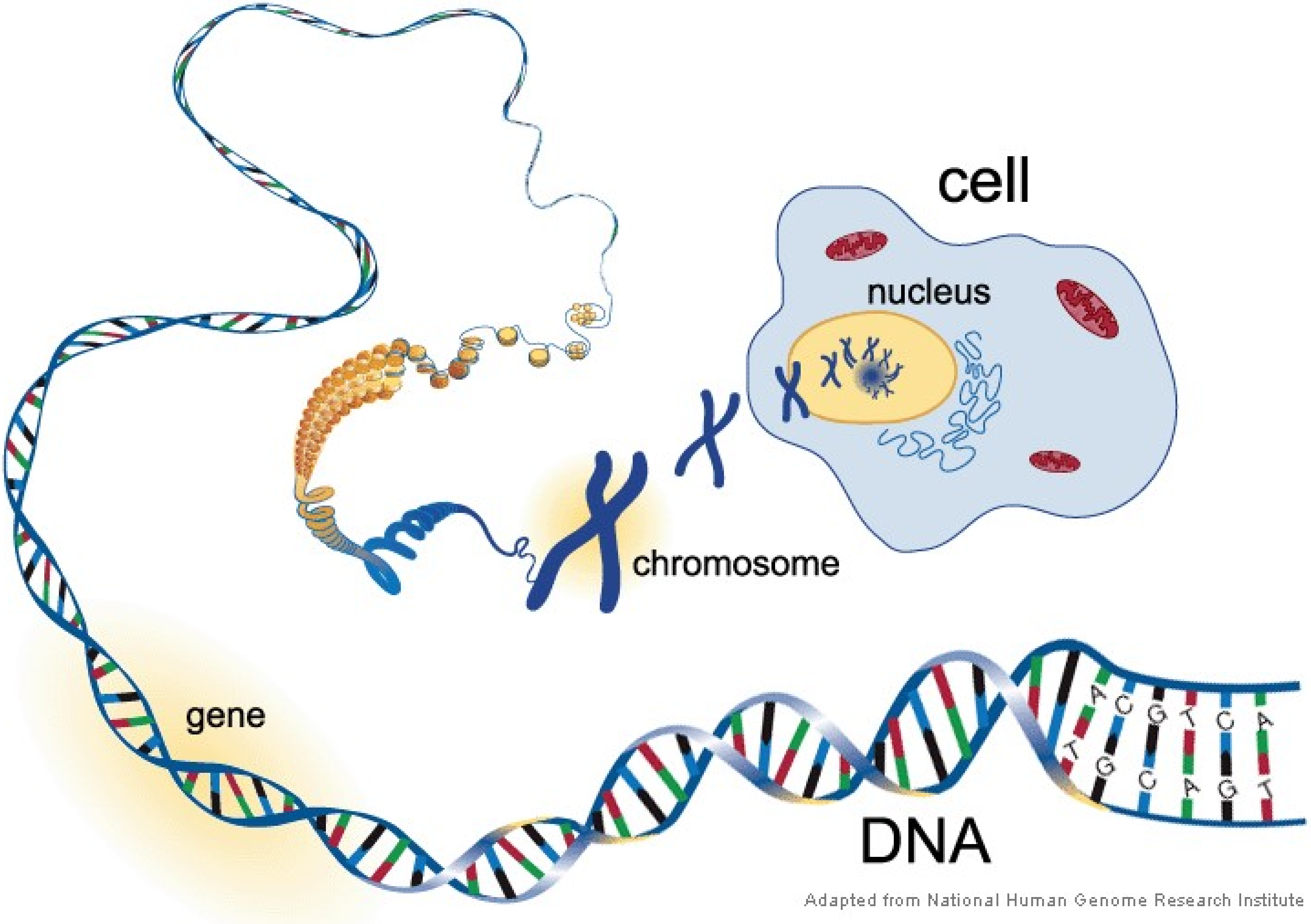

Evolutionary Checkers starring Blondie24 : Checkers with an Attitude! (ver. 2.3)
Blondie
RESIGN
OFFER DRAW
Your Turn
Main
Expert
New

Review

- Evolutionary Algorithms
  - Key Concepts
  - EA Family
  - Where to find related information?
- What is optimization?
- Typical Optimization Problems
- What makes a problem difficult to optimize?

A Gentle Introduction to Genetic Algorithms

Biology Background



Biology Background

- Gene
  - A working subunit of DNA
- Gene Trait
  - For example: colour of eyes
- Allele
  - Possible settings for a trait (e.g., Blown, Gray and Blue)
- Genotype
  - The actual genes carried by an individual
- Phenotype
  - The physical characteristics into which genes are translated

Genetic Algorithms

- John Holland
  - Adaptation in Natural and Artificial Systems, 1975
- Inspired by and (loosely) based on Darwin's Theory
  - Chromosome
  - Crossover
  - Mutation
  - Selection (Survival of the Fittest)
- Basic Ideas
  - Each solution to the problem is represented as a chromosome.
  - The initial solutions may be randomly generated.
  - Solutions are evolved during generations.
  - Improved gradually based on the principle of natural evolution.

Basic Components

- Representation
  - How to encode the parameters of the problem?
  - Binary Problems
    - 10001 00111 11001 .....
  - Continuous Problems
    - 0.8 1.2 -0.3 2.1 .....

- Genetic Operators
  - Crossover
    - Exchange genetic materials between two chromosomes.
  - Mutation:
    - Randomly modify gene values at selected locations.

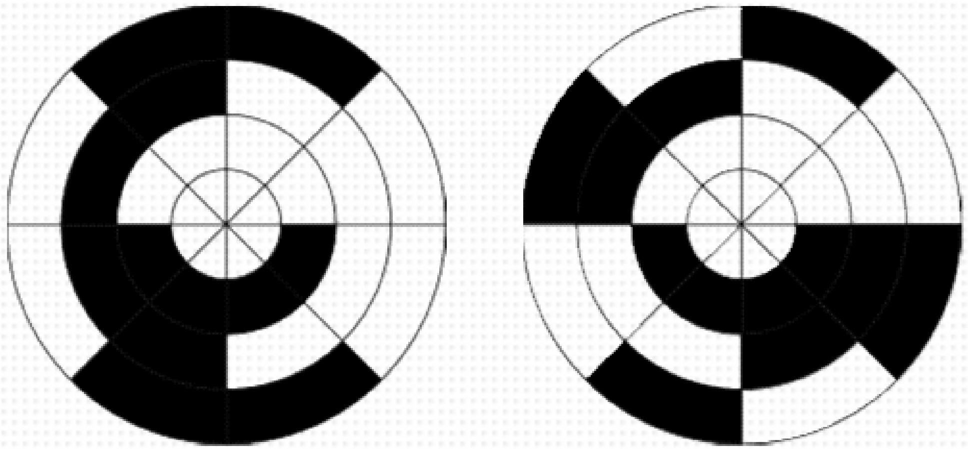
- Selection Strategy
  - Which chromosomes should be involved in reproduction?
  - Which offspring should be able to survive?

Representation

- Individual (Chromosome)
  - A vector that represents a specific solution to the problem.
  - Each element on the vector corresponds to a certain variable/parameter.
- Population
  - A set of individuals
  - GAs maintain and evolve a population of individuals.
  - Parallel Search → Global Optimization
- Offspring
  - New individuals generated via genetic operators
  - Hopefully contain better solutions.
- <u>Encoding</u>
  - Binary vs. Gray
  - How to encode TSP problems?

Binary vs. Gray

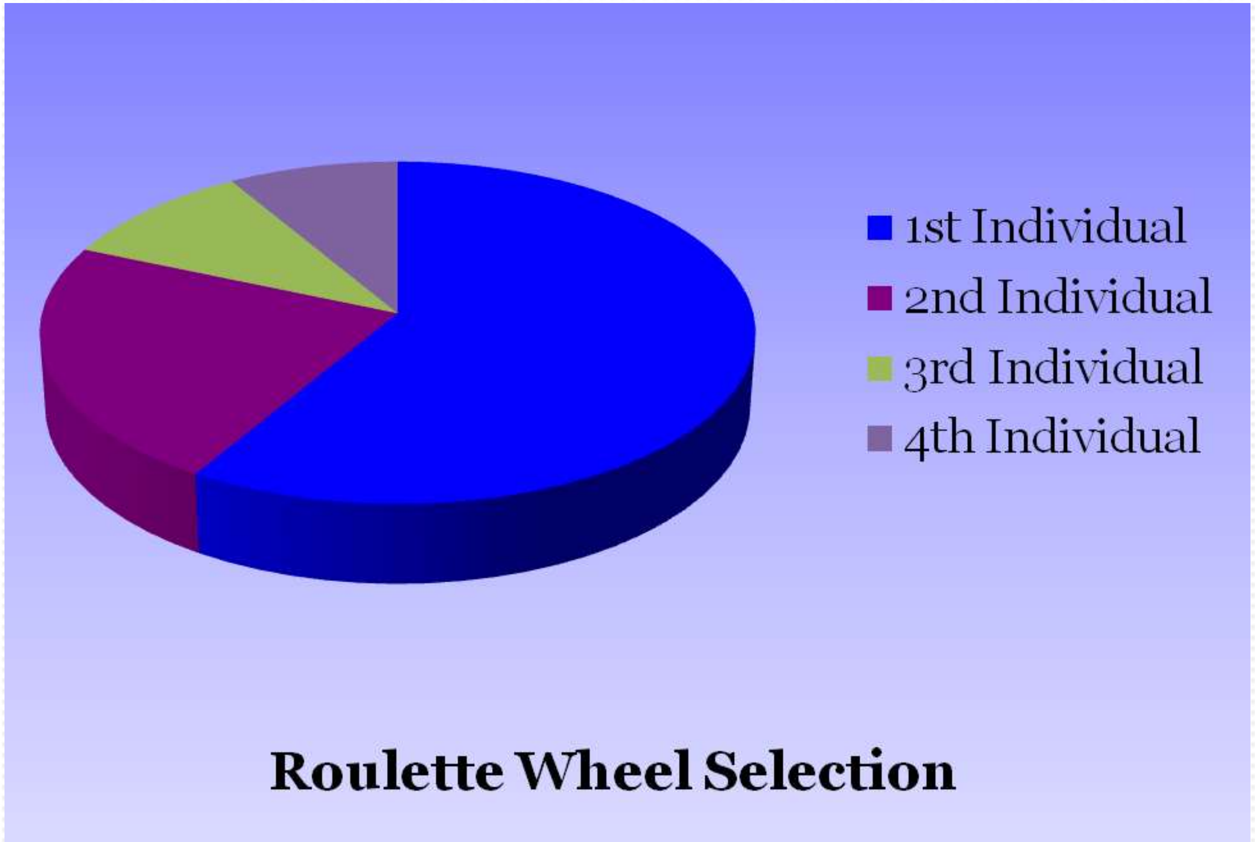
Decimal	Binary	Gray
0	000	000
1	001	001
2	010	011
3	011	010
4	100	110
5	101	111
6	110	101
7	111	100



Rotary Encoder

Selection I

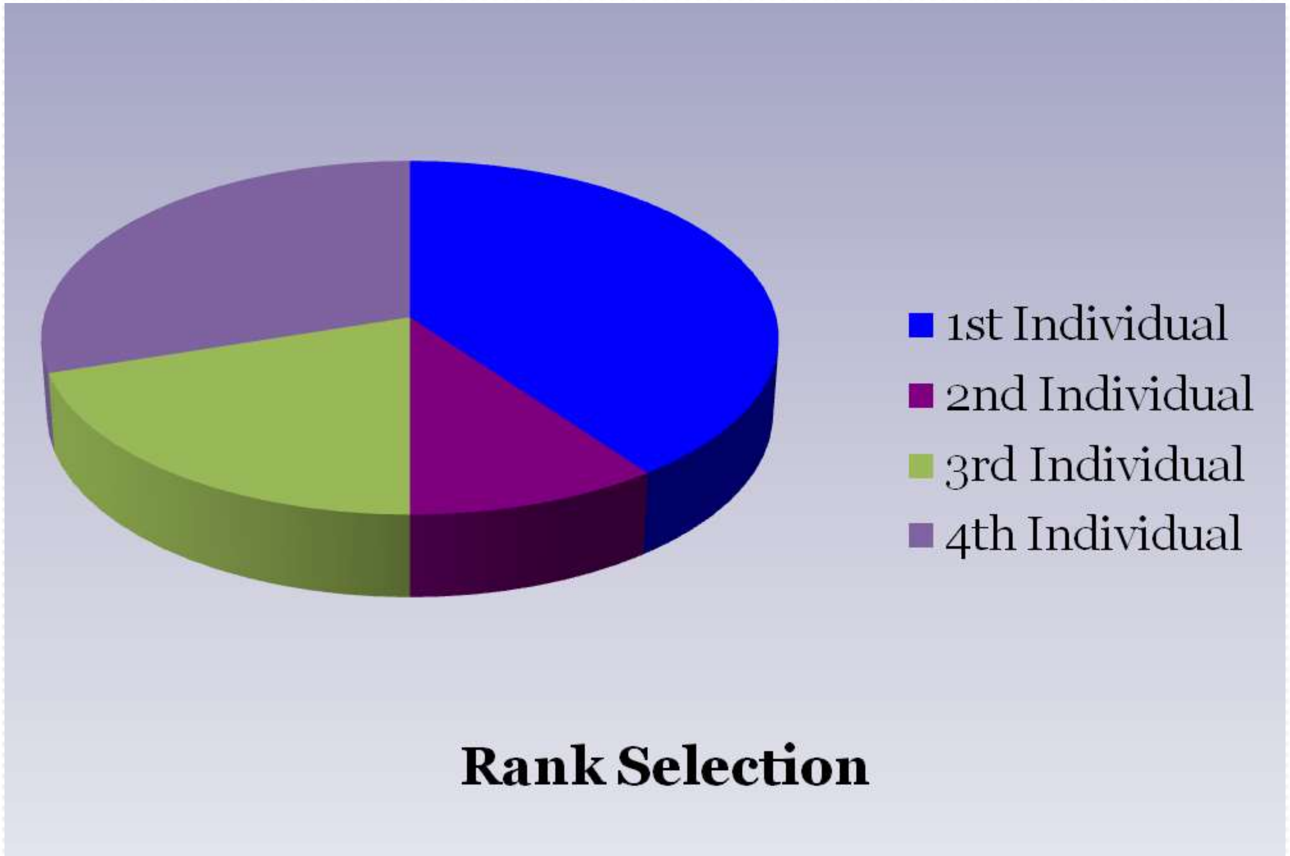
	Values
1	8.2
2	3.2
3	1.4
4	1.2



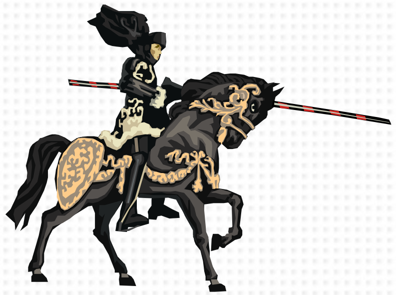
Negative Values?

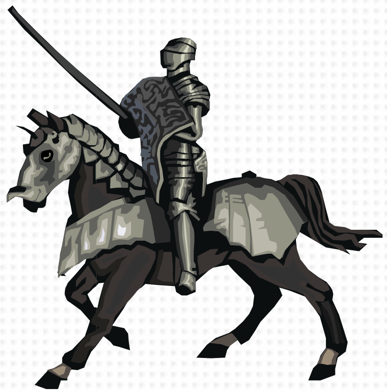
Selection II

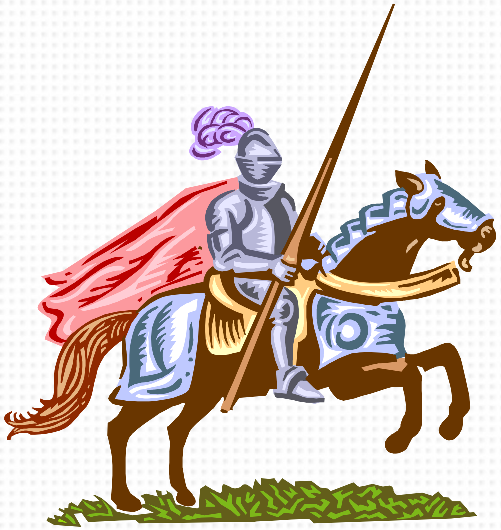
	Values	Rank
1	8.2	4
2	-3.2	1
3	1.4	2
4	-1.2	3



Selection III







FIGHT!

<u>Tournament Size?</u>

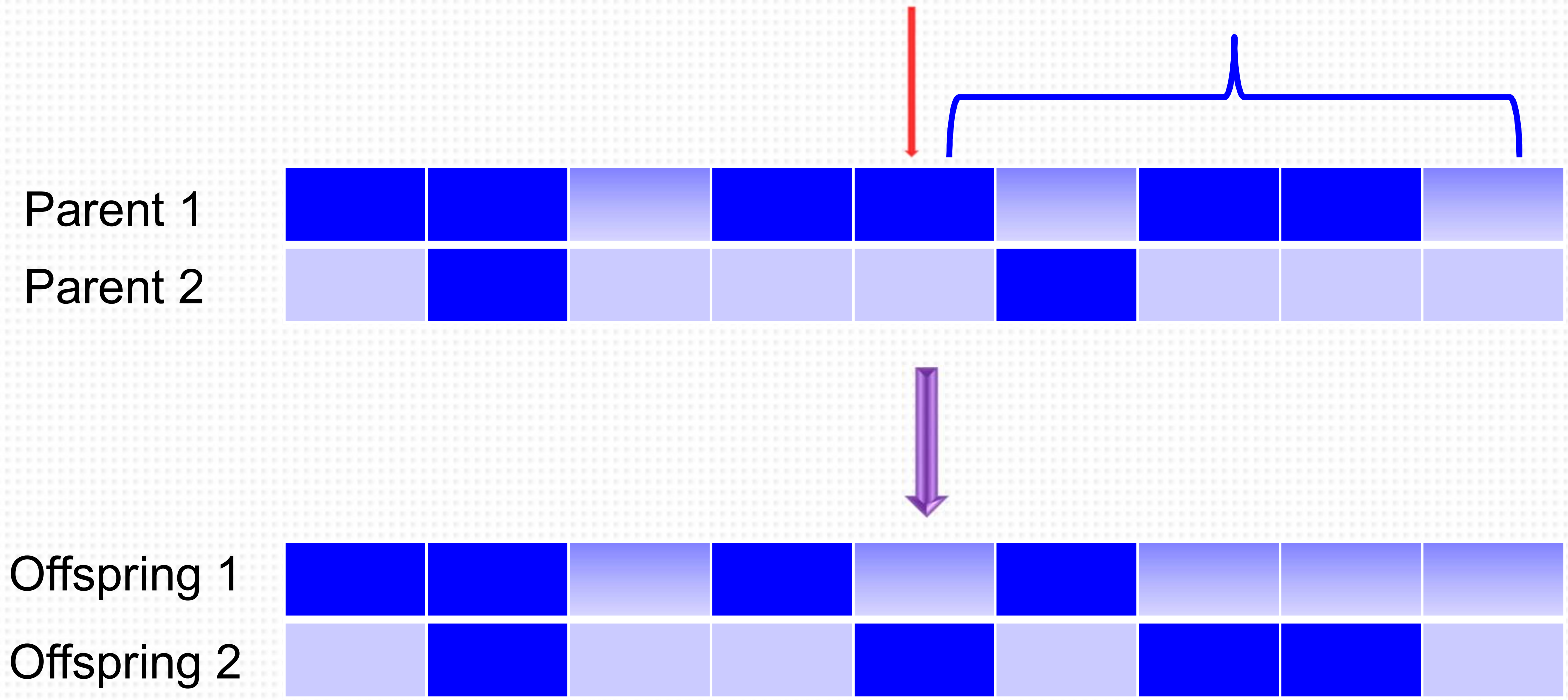
Tournament Selection

Selection IV

- Elitism
  - Offspring are not necessarily better than their parents.
  - The best individual in the current population may be lost.
    - Destroyed by crossover & mutation
  - Copy the best individual to the next generation.
  - Improve the stability and performance of GAs.

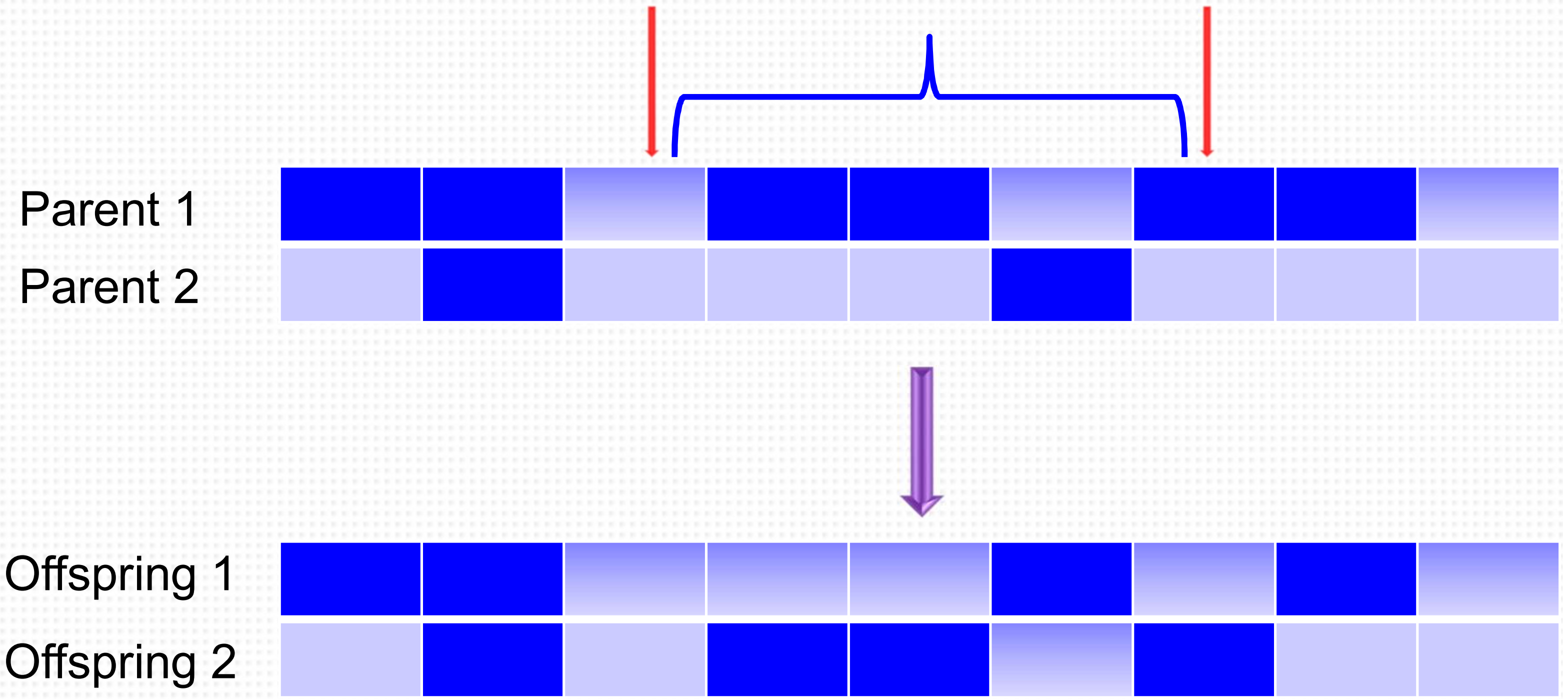
- Offspring Selection
  - Usually the old population is replaced by the offspring.
  - Are their any other options?
    - (μ, λ) Strategy
    - $(\mu+\lambda)$ Strategy

Crossover I



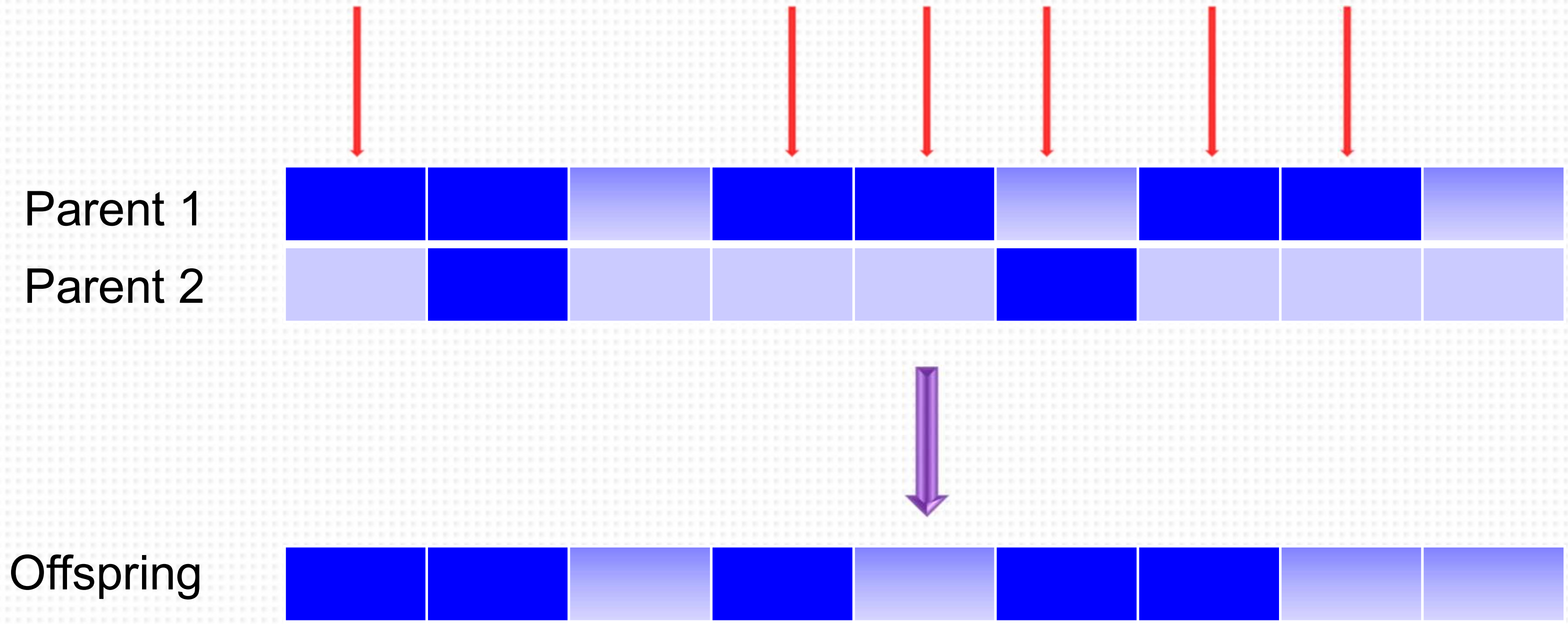
One Point Crossover

Crossover II



Two Point Crossover

Crossover III



Uniform Crossover

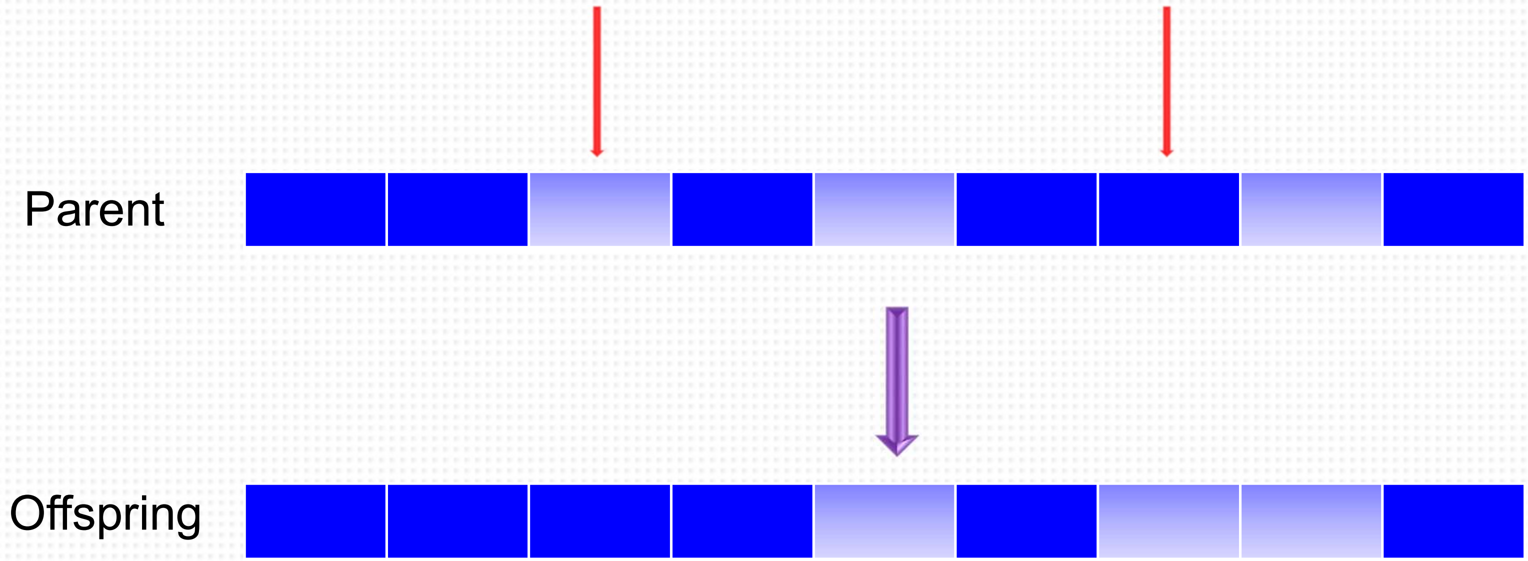
Is It Always Easy?

Parent 1	1	2	3	4	5	6	7	8	9
Parent 2	3	1	2	8	4	9	5	7	6

Offspring 1	1	2	3	4	4	9	5	7	6
Offspring 2	3	1	2	8	5	6	7	8	9

Crossover of Two Individuals for TSP

Mutation



Mutation vs. Crossover

Mutation is mainly used to maintain the genetic diversity.

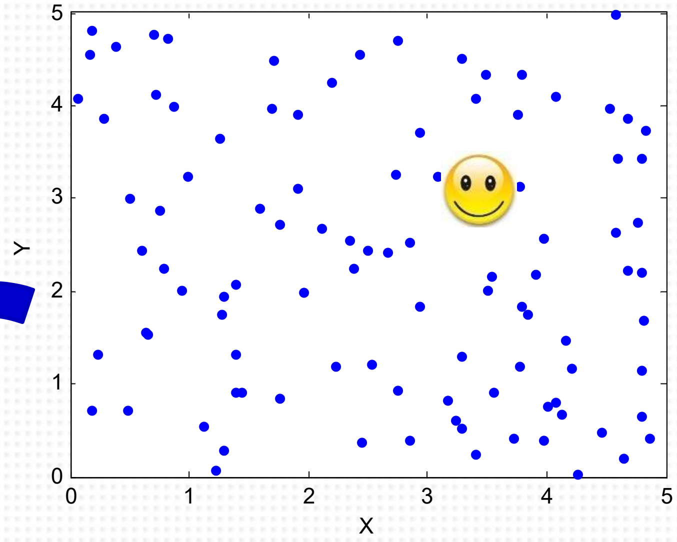
Population Diversity

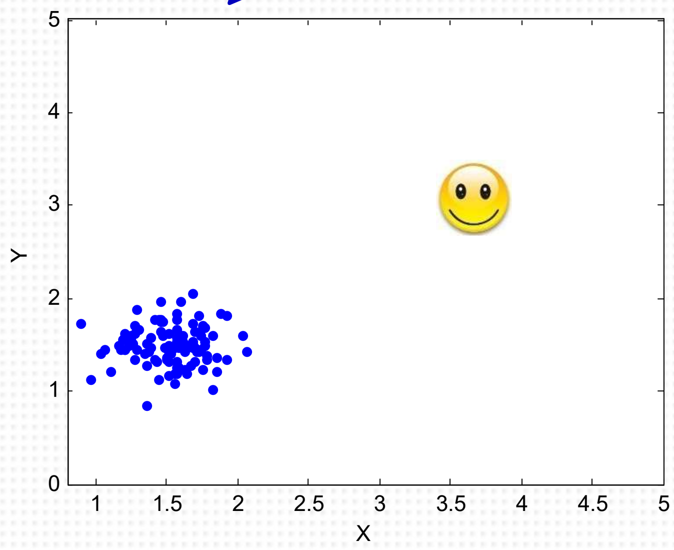
Population

	0	
	1	
	0	
	1	
	1	
	0	
	0	

1
1
1
1
1
1
1

kth dimension



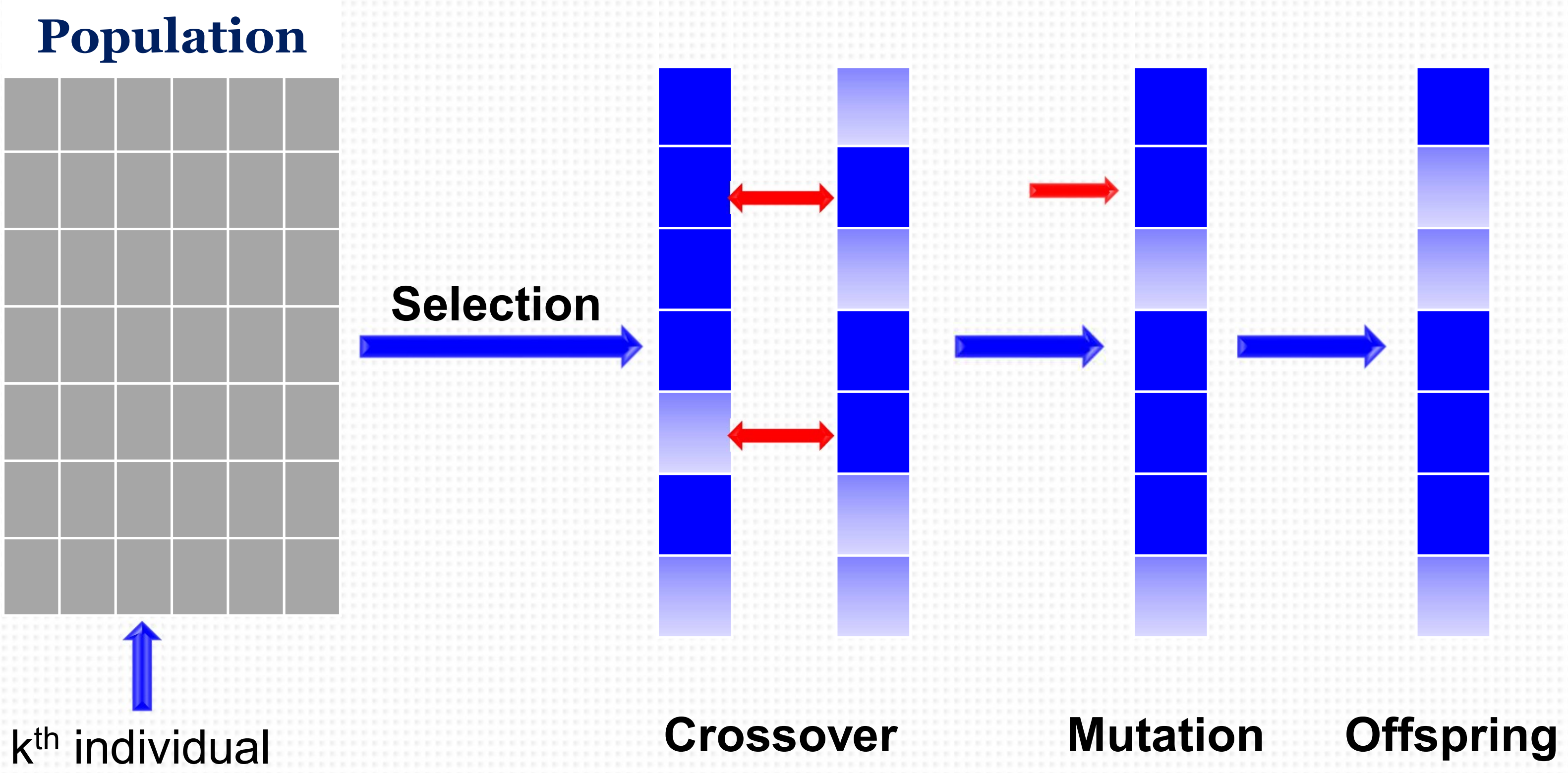


Loss of genetic diversity → Premature Convergence

Selection vs. Crossover vs. Mutation

- Selection
  - Bias the search effort towards promising individuals.
  - Loss of genetic diversity
- Crossover
  - Create better individuals by combining genes from good individuals.
  - Building Block Hypothesis
  - Major search power of GAs
  - No effect on genetic diversity
- Mutation
  - Increase genetic diversity.
  - Force the algorithm to search areas other than the current focus.
- Exploration vs. Exploitation

The Complete Picture



GA Framework

Initialization: Generate a random population P of M individuals

Evaluation: Evaluate the fitness $f(x)$ of each individual

Repeat until the stopping criteria are met:

- **Reproduction:** Repeat the following steps until all offspring are generated
  - **Parent Selection:** Select two parents from P
  - **Crossover:** Apply crossover on the parents with probability P_c
  - **Mutation:** Apply mutation on offspring with probability P_m
  - **Evaluation:** Evaluate the newly generated offspring
- **Offspring Selection:** Create a new population from offspring and P

Output: Return the best individual found

Parameters

- Population Size
  - Too big: Slow convergence rate
  - Too small: Premature convergence

- Crossover Rate
  - Recommended value: 0.8

- Mutation Rate
  - Recommended value: 1/L
  - Too big: Disrupt the evolution process
  - Too small: Not enough to maintain diversity

- Selection Strategy
  - Tournament Selection
  - Truncation Selection (Select top T individuals)
  - Need to be careful about the selection pressure.

A Few More Words

- No Free Lunch!
  - So called “optimal” parameter values do not exist!
  - Vary from problems to problems.
  - Need some trials to find suitable parameter values.
- Randomness
  - Inherently stochastic algorithms
  - Independent trials are needed for performance evaluation.
- Why does it work?
  - Easy to understand & implement (No maths required!)
  - Very difficult to analyse mathematically.
  - Converge to global optimum with probability 1 (infinite population).
  - Schema Theorem

Feature Selection

- Select a subset of features from the original features.
  - Dimension Reduction
  - Remove irrelevant features

- Motivation
  - Less challenging in training
  - Faster running time
  - Higher accuracy
  - Better understanding

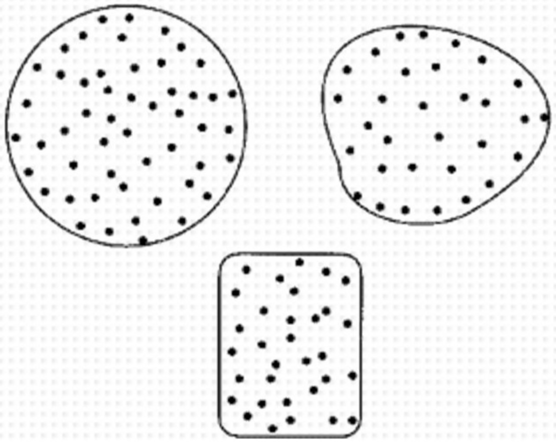
- Approaches
  - Filter Method: Does not consider classification error.
  - Wrapper Method: Does consider classification error.

GAs & Feature Selection

- Representation
  - Binary strings
- Fitness Function
  - Classification errors (KNN, SVM, ANN etc.)
- Evolve a population of candidate subset of features.
- Major Issues
  - The number of candidate features (Curse of Dimensionality)
  - Computational cost
- Reading Materials
  - "A Survey of Genetic Feature Selection in Mining Issues" CEC'99

Clustering

- Definition
  - The process of partitioning a group of data into K groups based on certain similarity metric.
  - Unsupervised Learning (Unlabeled Data)



K-Means

- 1. Choose K points as the initial cluster centres.
- 2. Assign each point to the cluster with the closest centre.
- 3. Recalculate the positions of the K centres.
- 4. Repeat Steps 2 and 3 until the centres no longer move.

Major Issue: Sensitive to the initial K centres!

Could get stuck at suboptimal solutions!

GAs & Clustering

- Use GAs to final an optimal set of cluster centres!

- Representation
  - A string of K centres
  - Length: K·D (D is the data dimensionality)

- Fitness Evaluation
  - Assign each data to the nearest cluster.
  - Calculate the new cluster centres.
  - Update the old cluster centres (individuals).
  - The quality of clustering is measured by:

$$M = \sum_{i=1}^{K} M_i$$

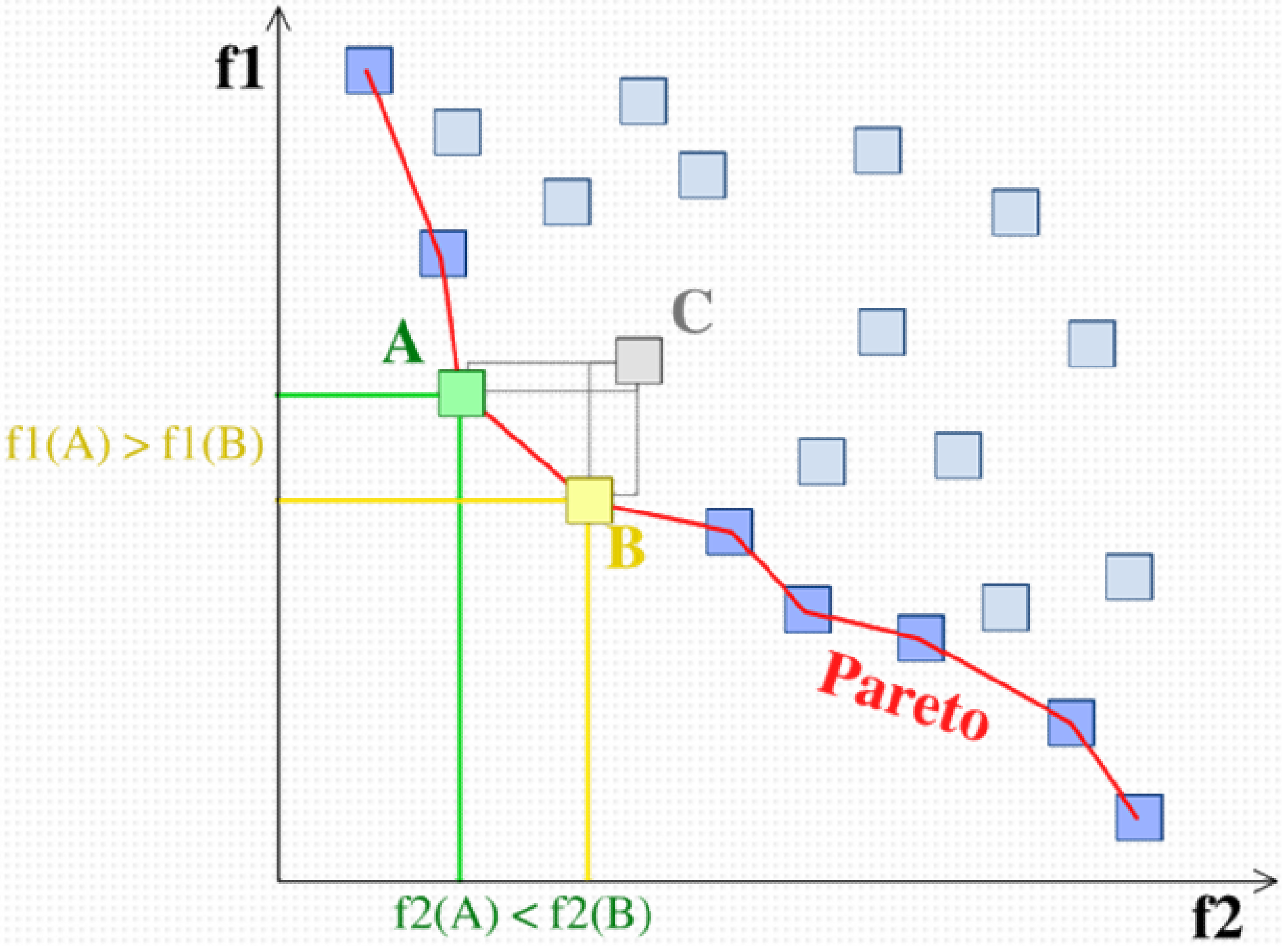
$$M_i = \sum_{x_j \in C_i} \left\| x_j - z_j \right\|$$

- Reading Materials
  - "Genetic Algorithm-Based Clustering Technique"
  - Pattern Recognition , vol. 33(9), 2000

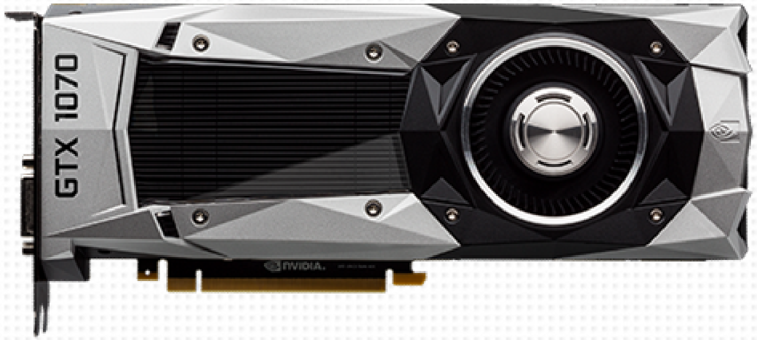
Open Questions

- Parameter Control
  - Parameter Tuning
  - Heuristics for changing parameter values on the run
- Constraint Handling
  - All real-world problems have constraints.
    - Box-Bounded Constraints
    - Linear Constrains
- Multi-Objective Optimization
  - Objectives may be contradicting.
    - Model Complexity vs. Prediction Accuracy
    - Profit vs. Risk

Pareto Front



Parallel Computing



There are many parallel components in EAs …

Review

- Basic Concept of GAs
- Major Components
  - Crossover
  - Mutation
  - Selection
  - Representation
- GA Framework
- Practical Issues
- Applications in Feature Selection & Clustering

EC Digest: https://listserv.gmu.edu/cgi-bin/wa?A0=ec-digest-l

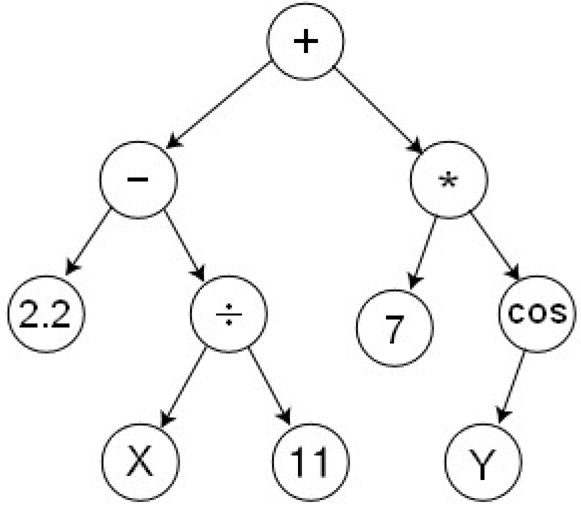
GENETIC PROGRAMMING AND EVOLVABLE THINGS

GA vs. GP

- GP is a branch of GAs.
  - Crossover/Mutation/Selection
- Representation
  - GA: Strings of numbers (0/1)
  - GP: Computer programs in tree structure (LISP)
- Output
  - GA: A set of parameter values optimizing the fitness function
  - GP: A computer program (Yes, a computer program!)
- John Koza
  - *Genetic* Programming: *On the Programming of Computers by Means of Natural Selection*. Cambridge, MIT Press, 1992.
  - www.genetic-programming.org

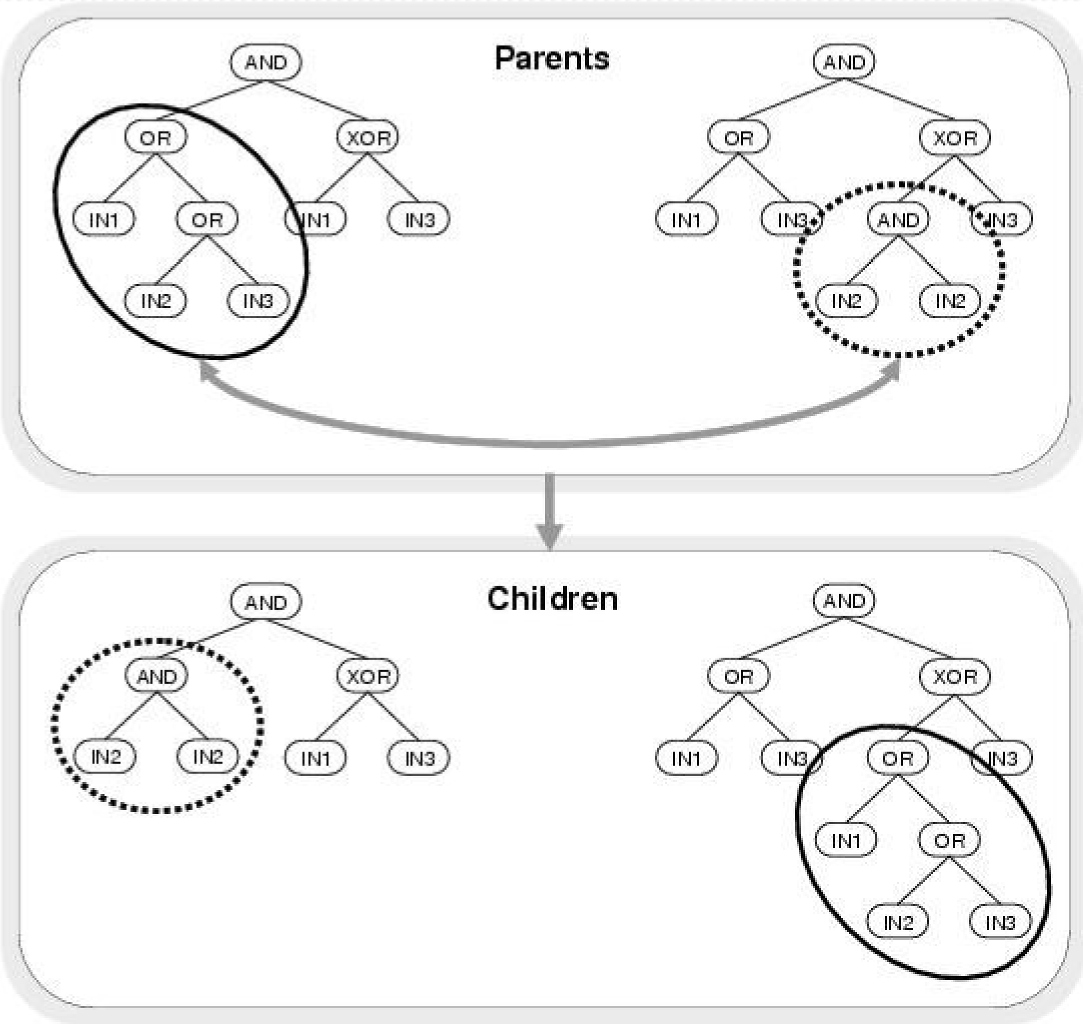
Functions & Terminals

- Important Concepts in GP
- Building Block of Computer Programs
- Terminal Set
  - Variables: x, y, z …
  - Constants: 1, 2, 3,…
- Function Set
  - +, -, *, /
  - Problem specific functions

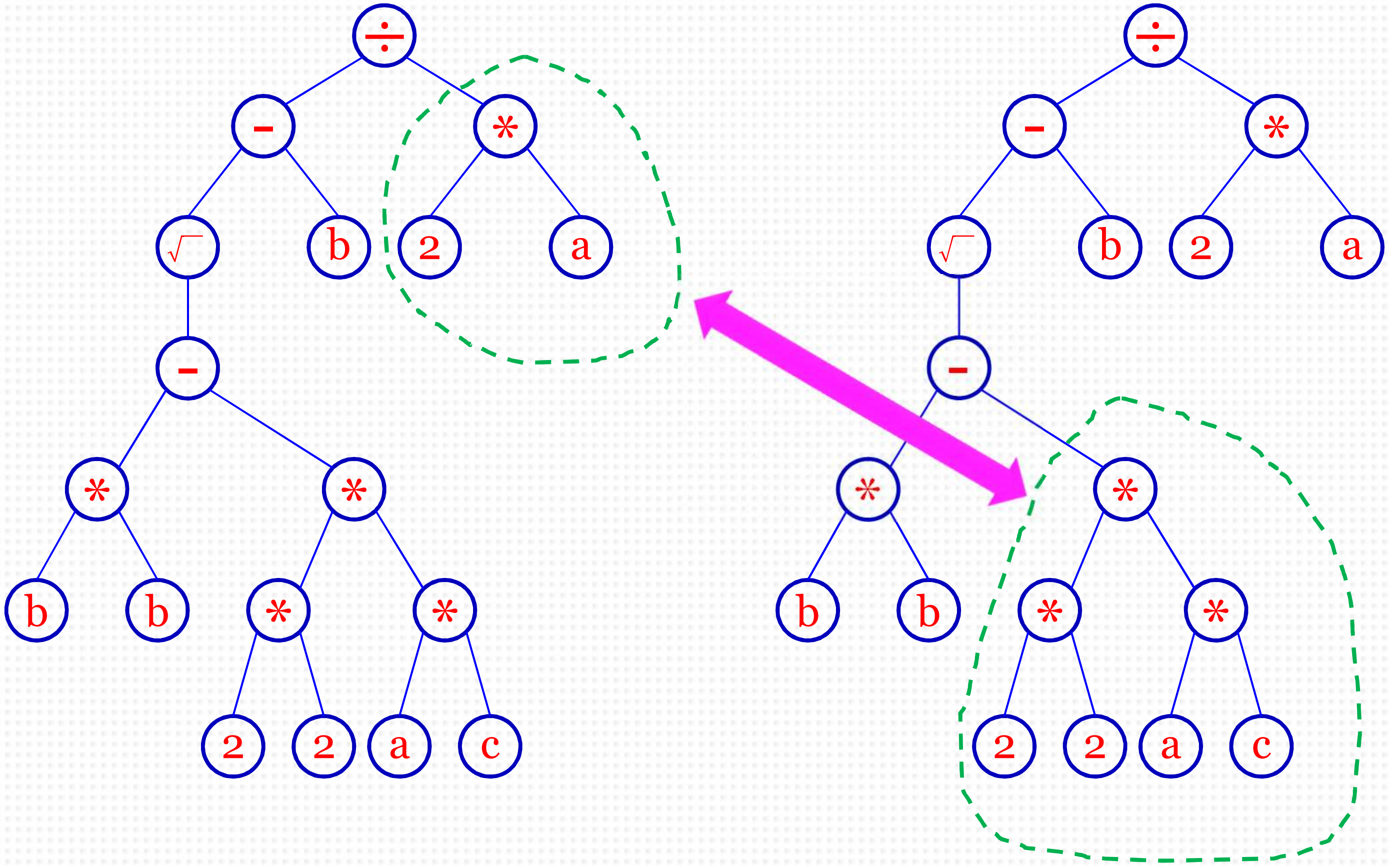


$$\left(2.2 - \left(\frac{X}{11}\right)\right) + \left(7 * \cos(Y)\right)$$

Crossover

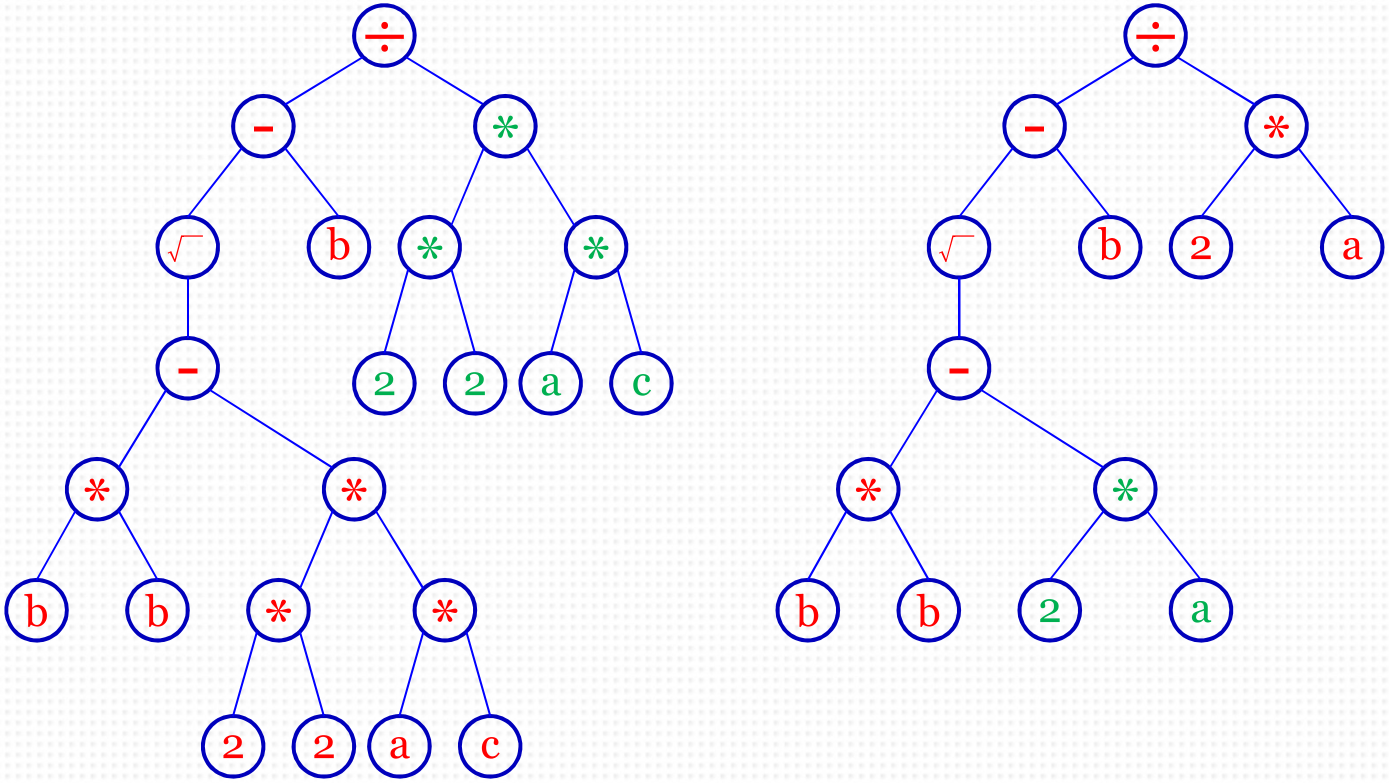

Parents
AND
OR
XOR
IN1
OR
IN1
IN3
IN2
IN3
AND
OR
XOR
IN1
IN3
AND
IN3
IN2
IN2
Children
AND
AND
XOR
IN2
IN2
IN1
IN3
AND
OR
XOR
IN1
IN3
OR
IN3
IN1
OR
IN2
IN3

Crossover



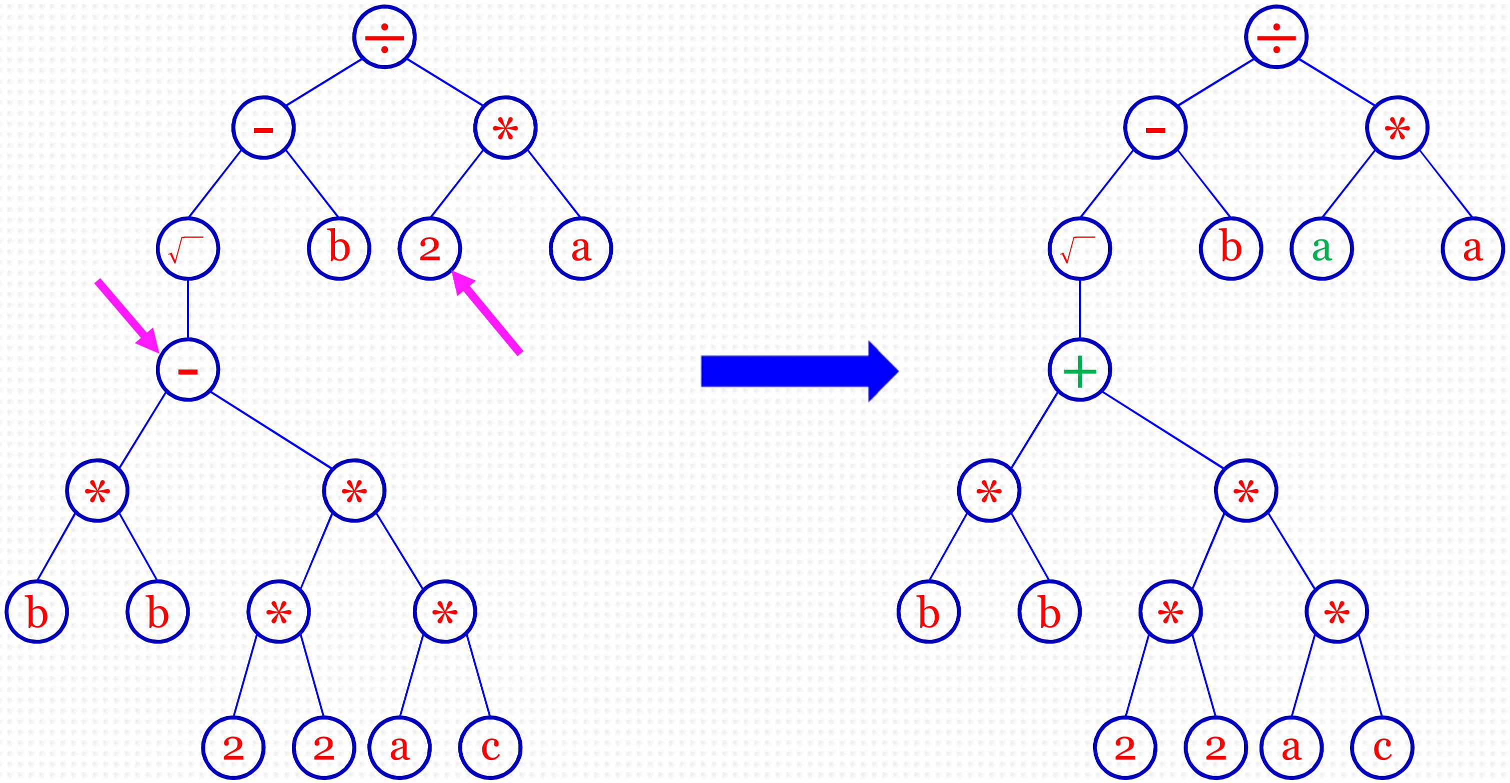
Identical Parents

Crossover

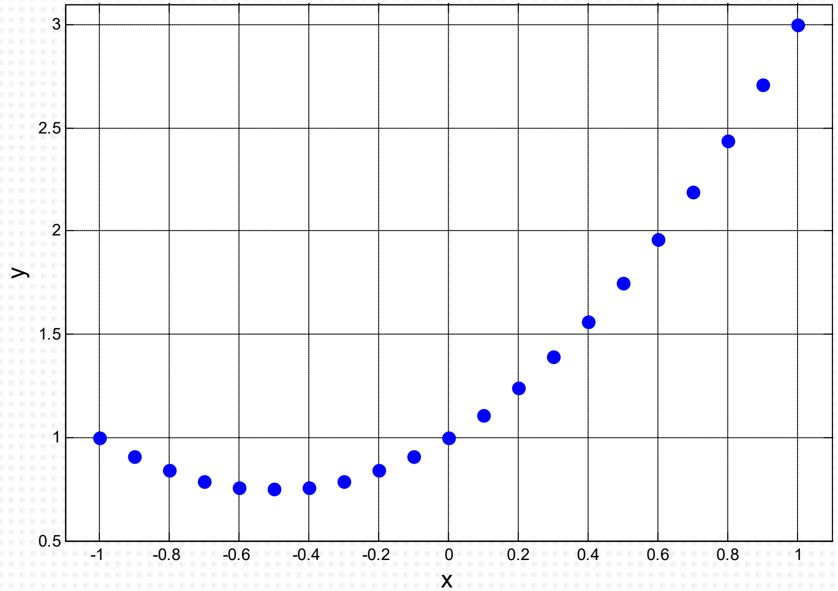


Identical Parents → Different Children

Mutation

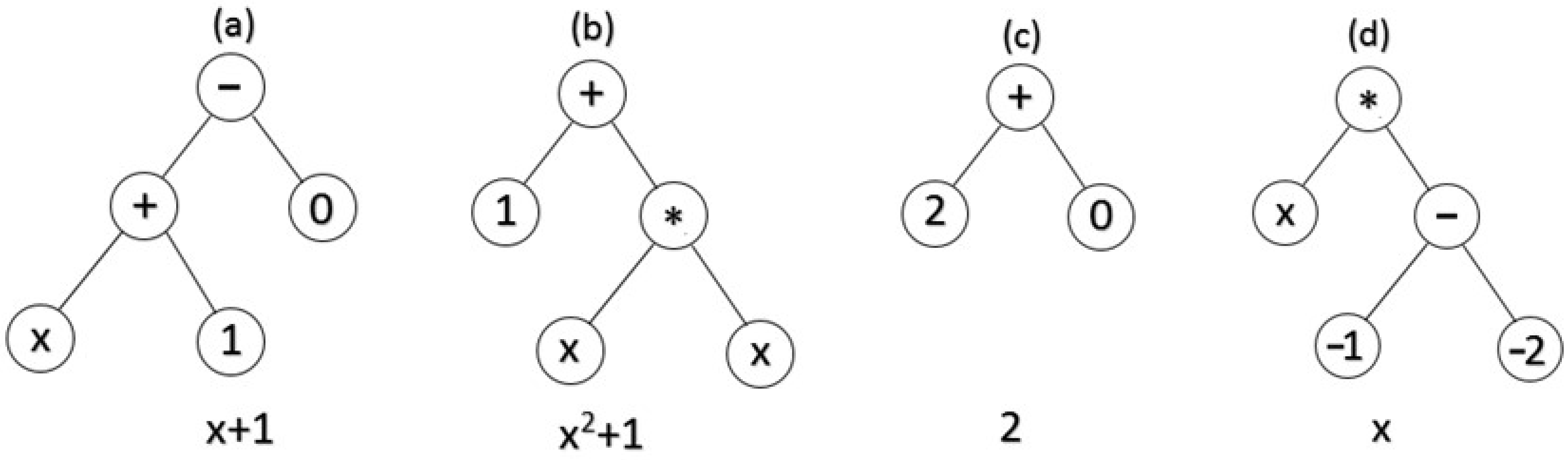


An Example of GP



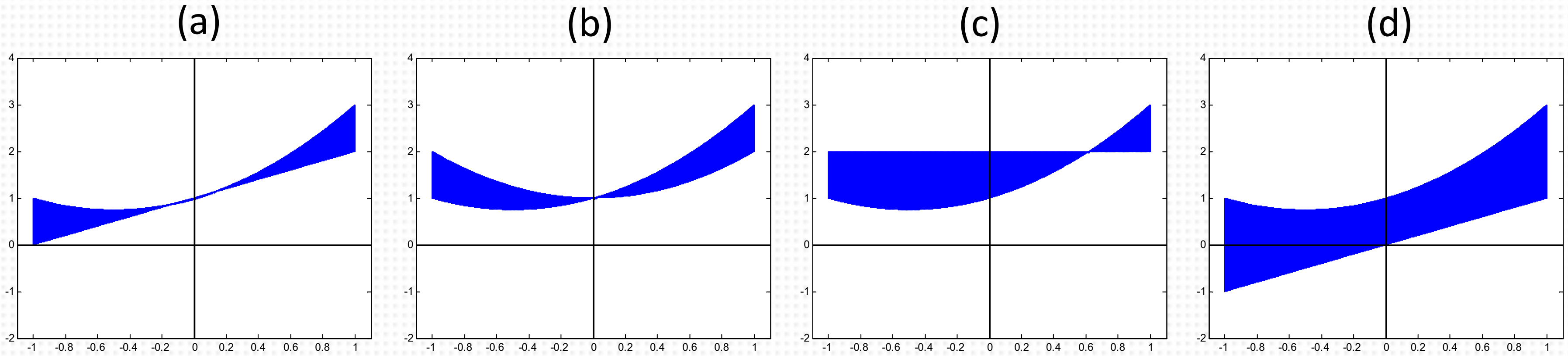
- Symbolic Regression
- Given a set of data points (x, y)
- Evolve a quadratic polynomial
  - Approximation of y given x
  - Find the underlying function $y=f(x)$.
- Fitness Evaluation
  - Absolute values of the differences (errors)
  - $\sum|y-y'|$

Generation 0



T = {X, R}   F = {+, -, *, %}

Generation 0

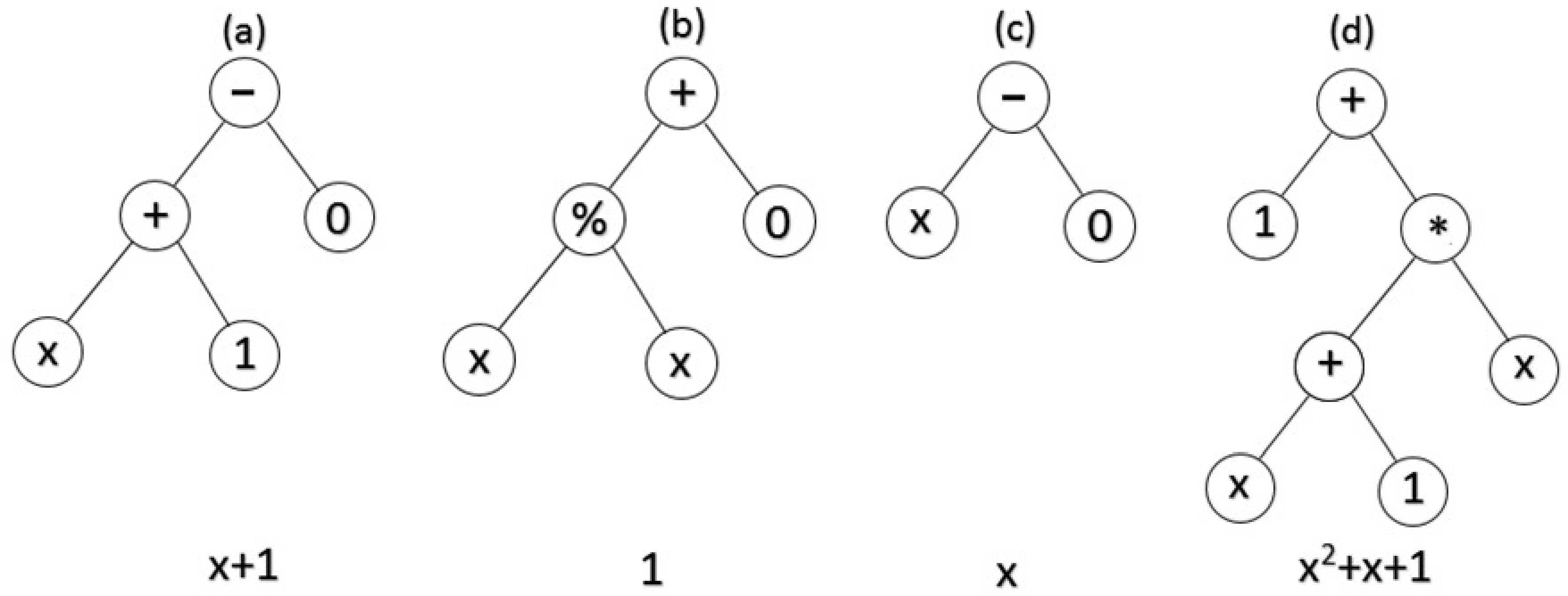


0.67 | 1.0 | 1.70 | 2.67

Fitness Values

$$Ground\ Truth: y = x^2 + x + 1$$

Generation X

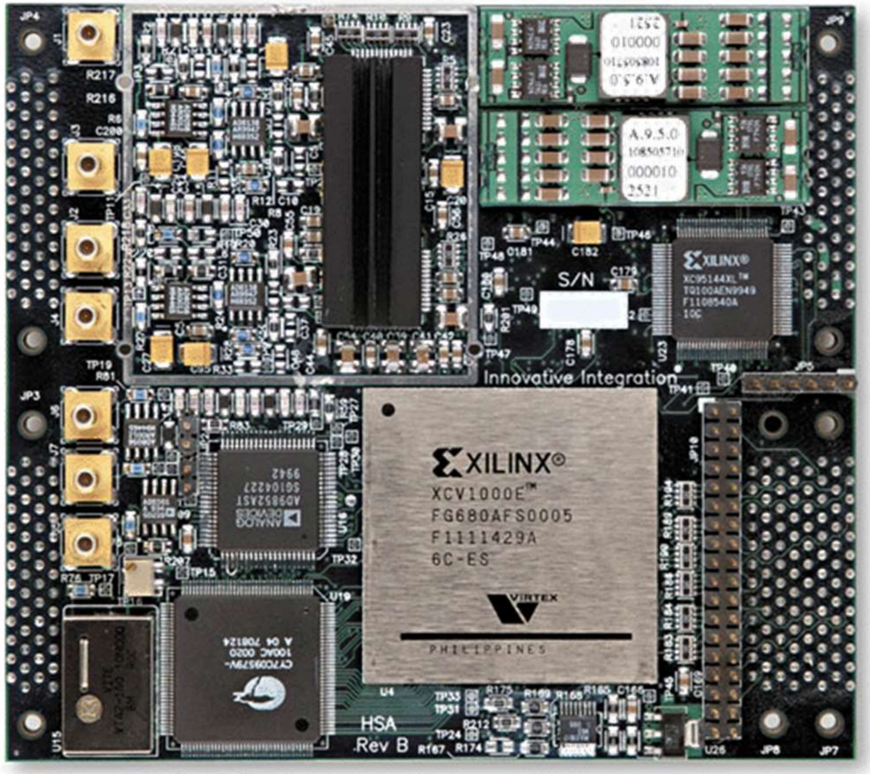


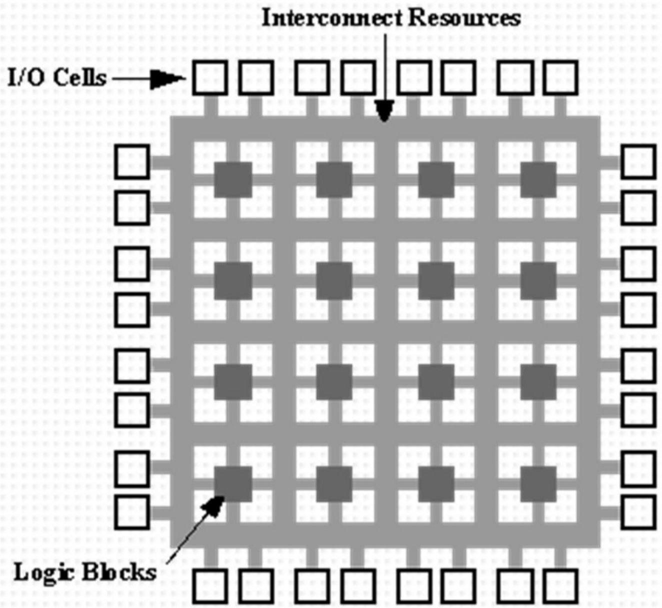
Optimal Solution

Review

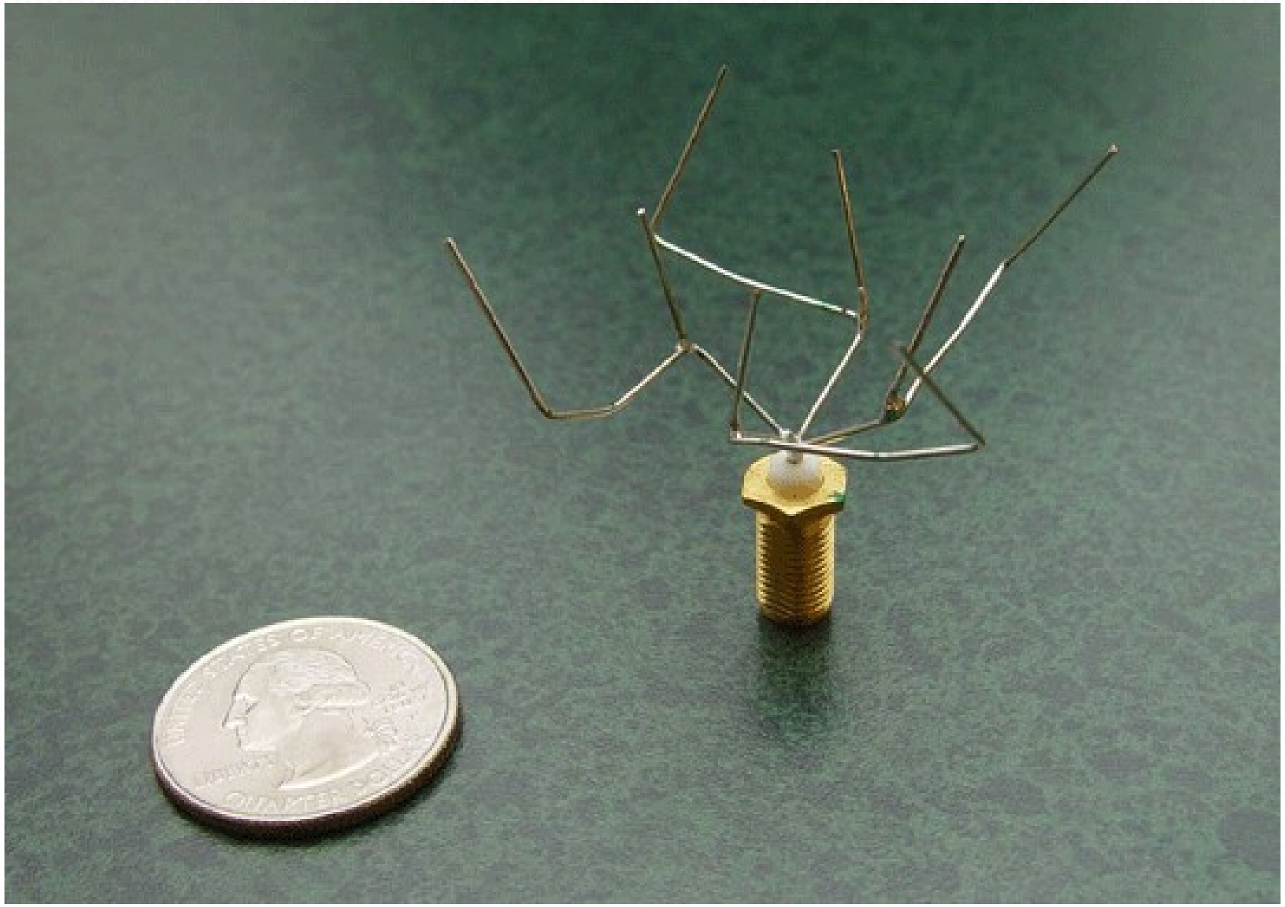
- Tree-Structured Computer Programs
  - Flexible in length
  - Problem specific functions and terminals
- Applications
  - Regression problem
  - Control problem
  - Classification
- Genetic Programming and Evolvable Machines Journal
- Evolvable Hardware
  - Circuits that change their architecture and behavior dynamically and autonomously by interacting with its environment.

Evolvable Circuits

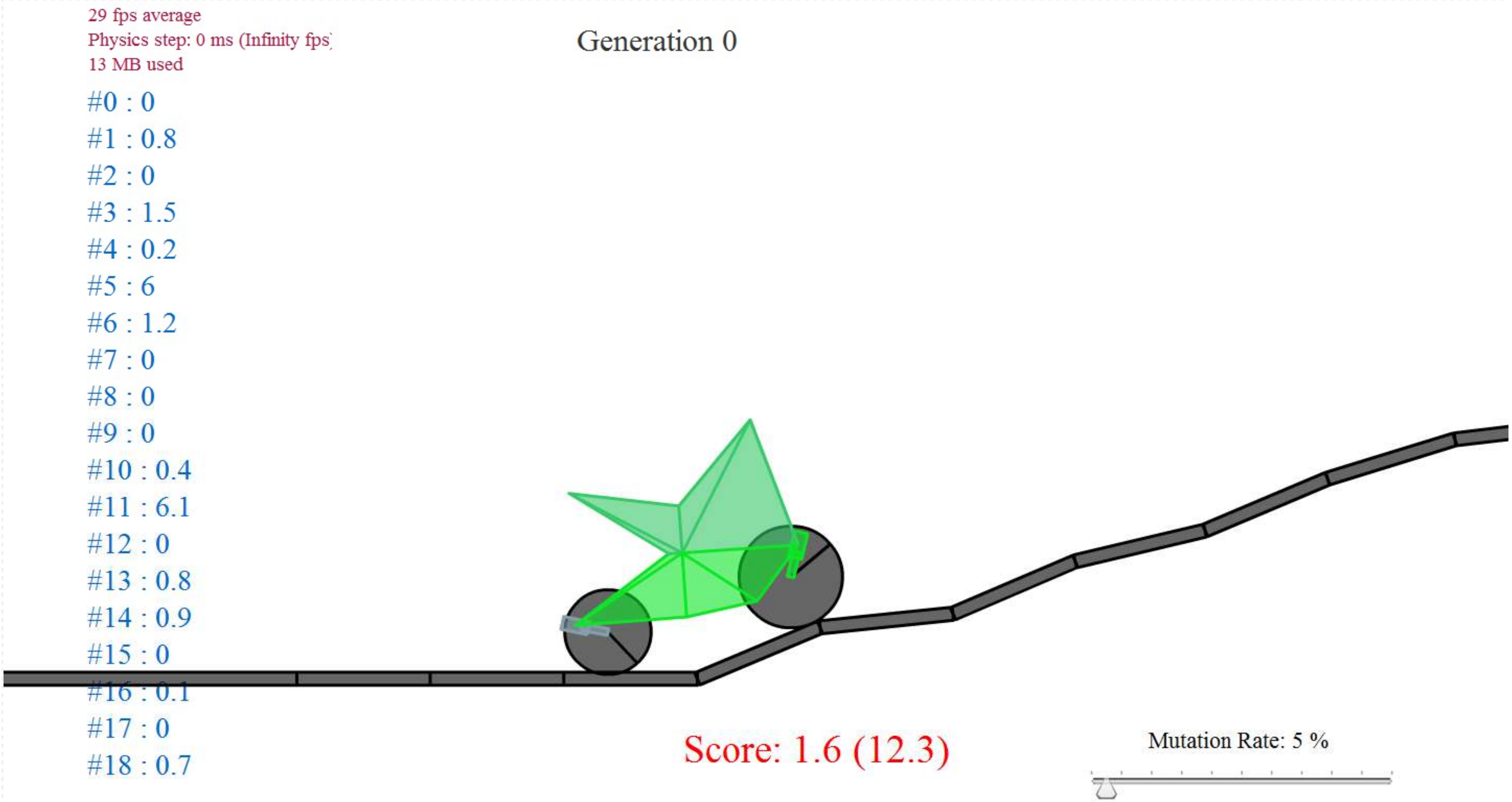




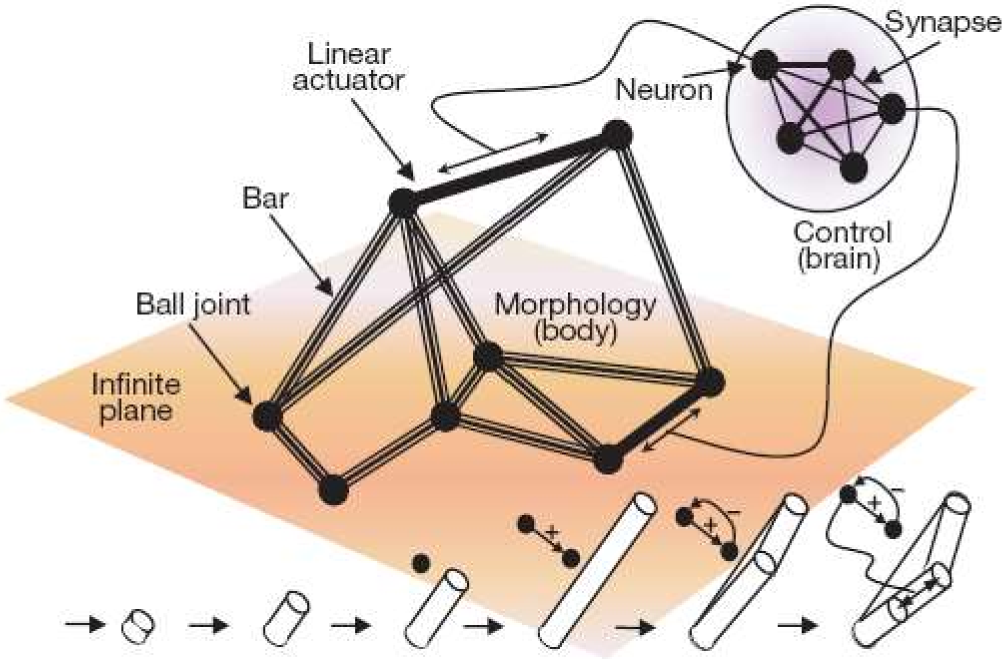
Antenna for NASA

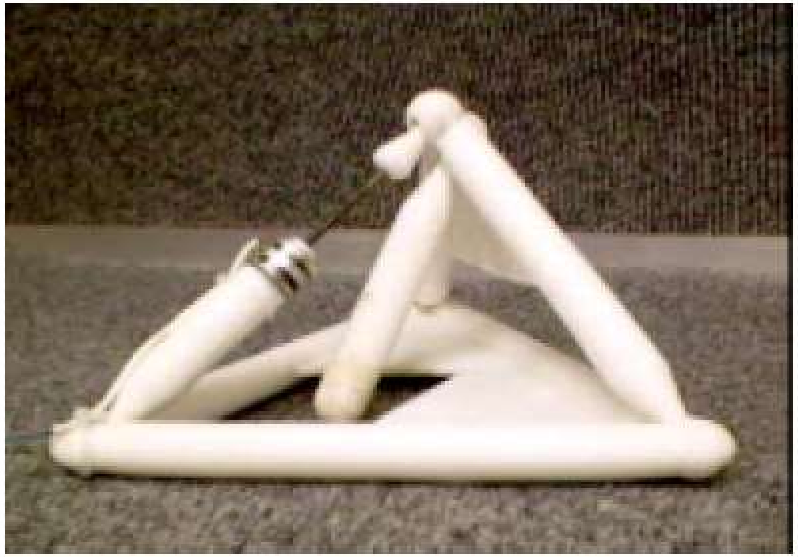


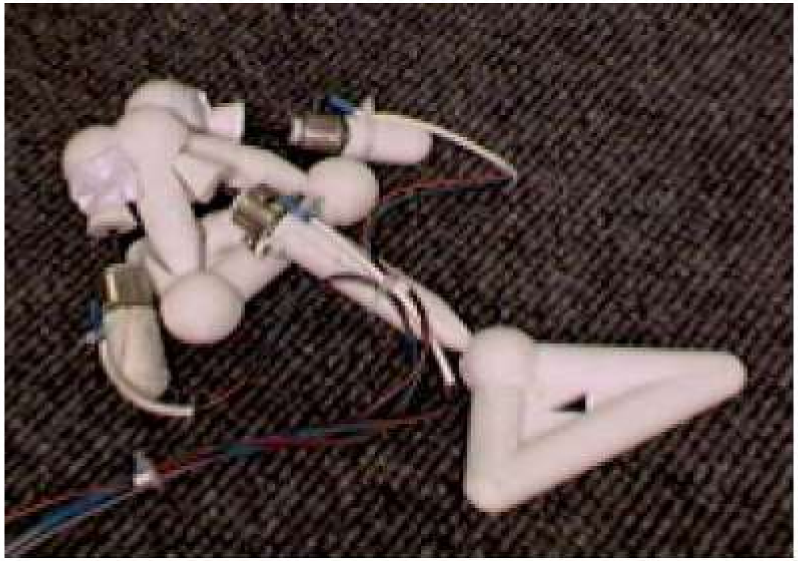
Car Design


29 fps average
Physics step: 0 ms (Infinity fps)
13 MB used
Generation 0
#0 : 0
#1 : 0.8
#2 : 0
#3 : 1.5
#4 : 0.2
#5 : 6
#6 : 1.2
#7 : 0
#8 : 0
#9 : 0
#10 : 0.4
#11 : 6.1
#12 : 0
#13 : 0.8
#14 : 0.9
#15 : 0
#16 : 0.1
#17 : 0
#18 : 0.7
Score: 1.6 (12.3)
Mutation Rate: 5 %

Artificial Life

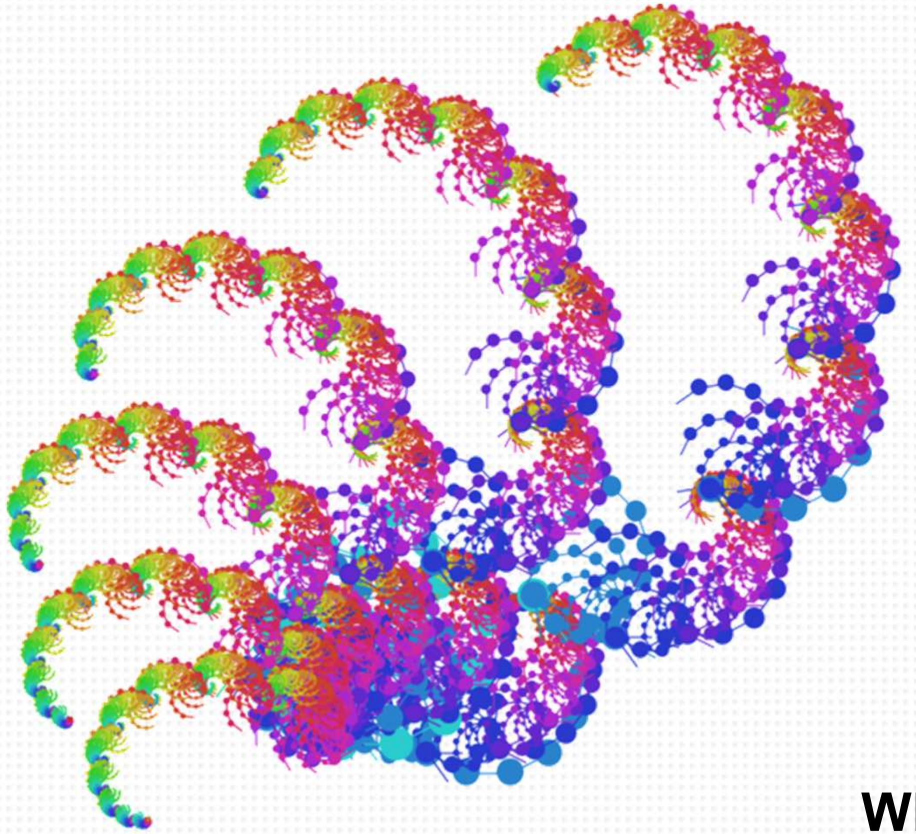






"Automatic design and manufacture of robotic lifeforms", *Nature* 406, 974-978.

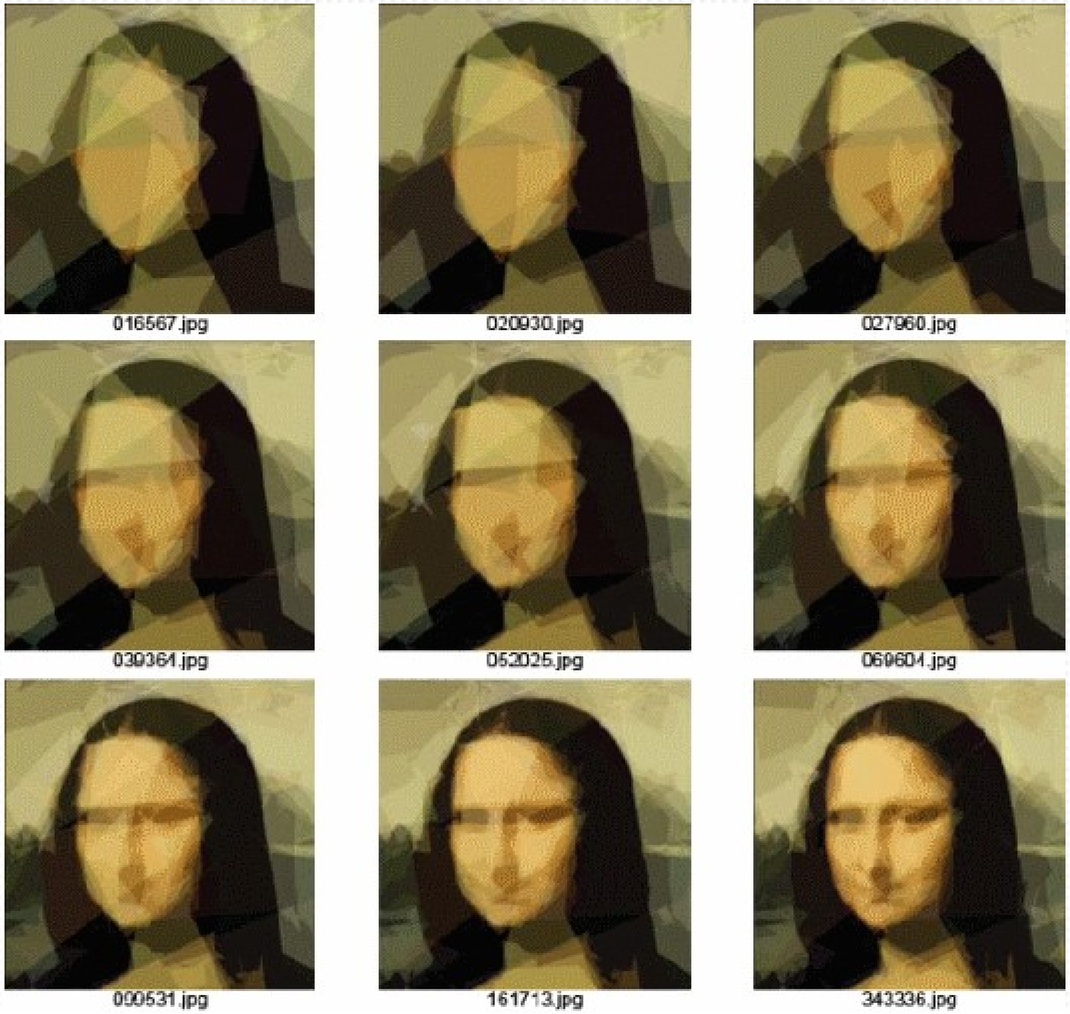
Evolutionary Arts



What is the major challenge?

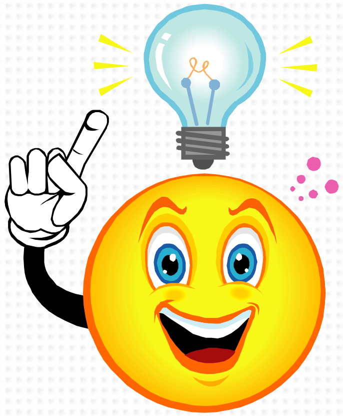
http://alteredqualia.com/visualization/evolve/

Evolving Mona Lisa



There is No Fate But What We Evolve …

Take Home Message



- Evolutionary Algorithms
  - A group of nature-inspired algorithms
  - General purpose techniques (Very Powerful!)
- Genetic Algorithms can do
  - Function Optimization
  - Feature Selection
  - Classification
  - Clustering
- Genetic Programming can do
  - Regression
  - Classification
  - Intelligent Design