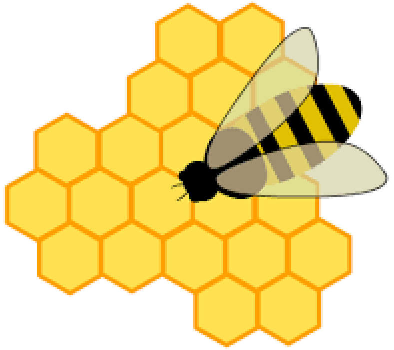
Thursday Learning Hour

Nature Inspired Optimizers

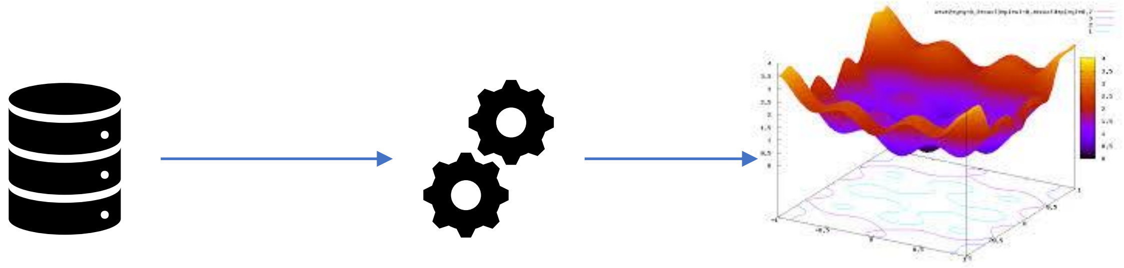


Prabakaran Chandran

19 Nov 2020

What is Optimization?:

- The selection of a best element (regarding some criterion) from some set of available alternatives
- In the simplest case, an optimization problem consists of maximizing or minimizing a real function by systematically choosing input values from within an allowed set and computing the value of the function.



Data

Statistical / ML model

Function which explains the scenario

This Function is our Objective Function

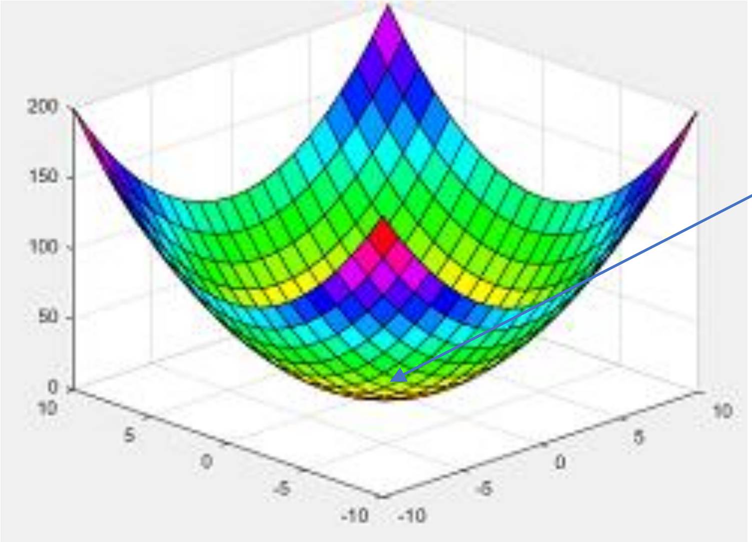
Our Objective : To find the input value which give us the minimum / maximum point of the function

Process of Reaching maxima or minima of the function is called optimization.
Sometimes , the process might be subjected to constraints.

Major Types of Optimization: Convex Optimization Vs Non-Convex Optimization:

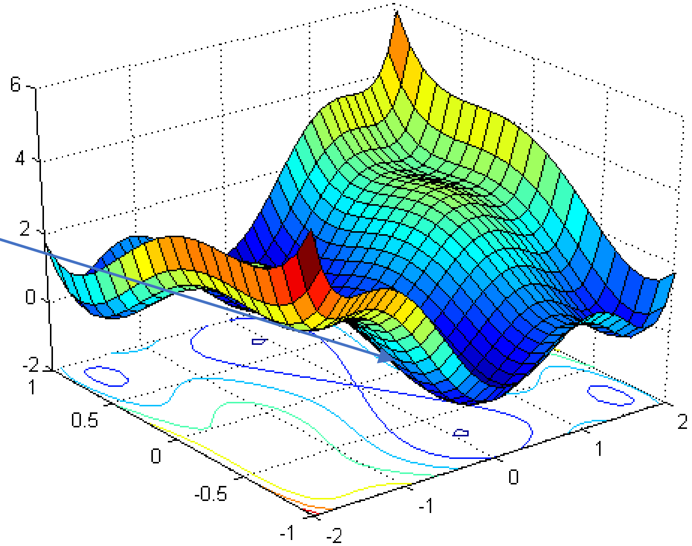
- A *convex optimisation problem* is a problem where all of the constraints are convex functions, and the objective is a convex function if minimising, or a concave function if maximising. A convex function can be described as a smooth surface with a single global minimum.

- In most of the machine learning problems we come across loss surfaces which are non-convex in nature. Hence, they will have multiple local minimum. Which is called as Non-Convex optimization



Single Optimum point

Multiple Optima includes Global and Local optima



Convex Function

Non-Convex Function

Constrained and Unconstrained Optimization

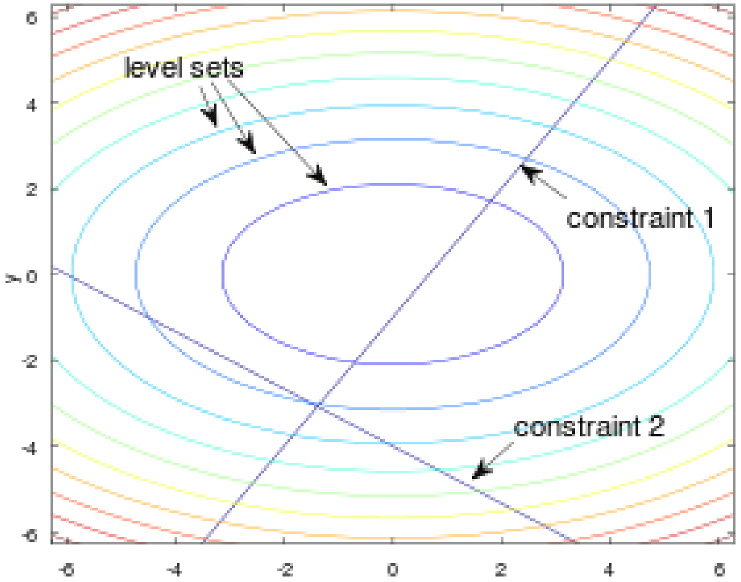
Constrained Optimization includes some conditions/limitations on the solutions

For Example: A company wants to maximize the sales , whereas they have a condition on price limit.

=> Sales = 1000-10price + 45.6units+26.5 marketing-0.25Competetor

1. Objective : max Sales
2. subjects to price > 10$ and marketing < 5000$

If we don't have any of the constraints --→ Un Constrained Optimization



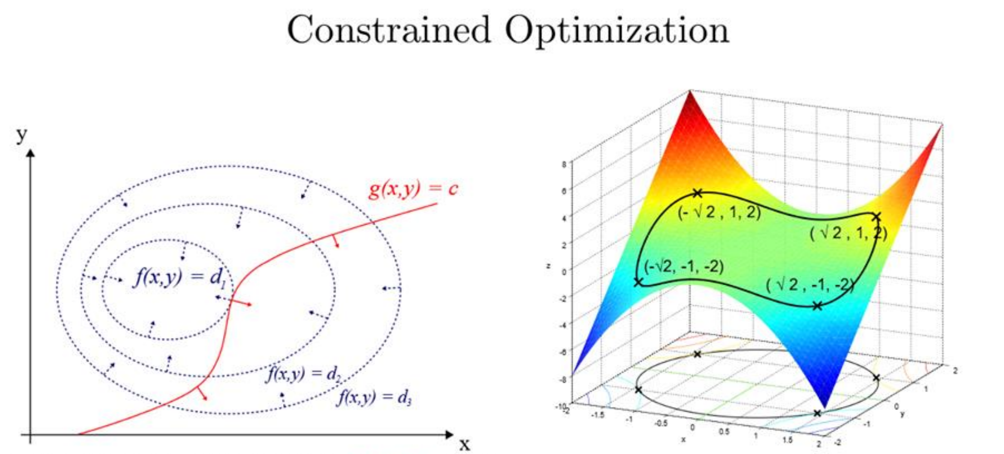


Figure 1: The red curve shows the constraint $g(x, y) = c$. The blue curves are contours of $f(x, y)$. The point where the red constraint tangentially touches a blue contour is the maximum of $f(x, y)$ along the constraint, since $d_1 > d_2$.

Figure 2. Illustration of the constrained optimization problem

Nature-Inspired Optimizers:

Evolutionary Algorithms

Population Behavior/ Individual Behavior

Nature-inspired algorithms are a set of novel problem-solving methodologies and approaches derived from natural processes

Physics and chemical inspired

Immune Systems and Neural Networks

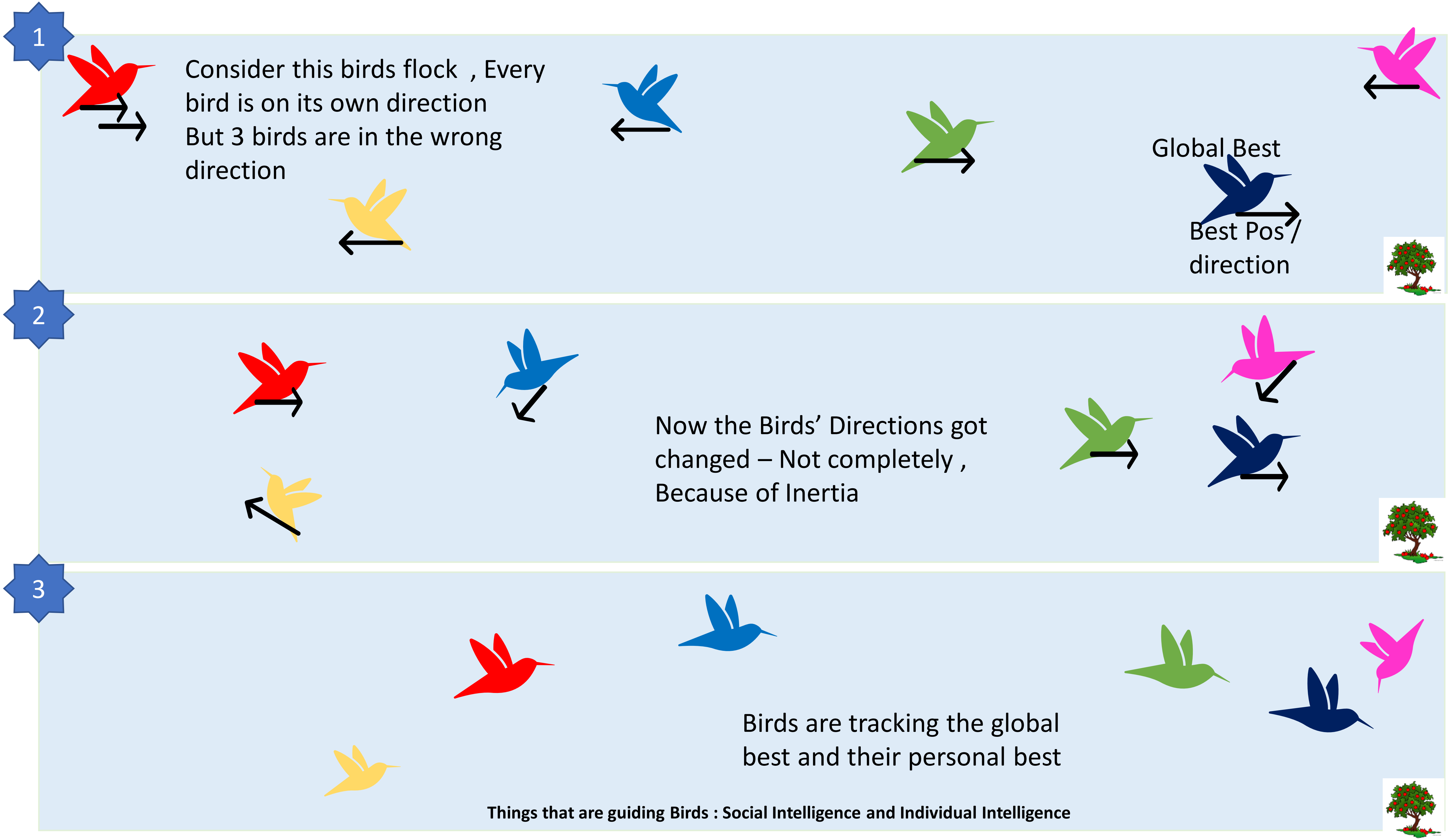
Algorithms inspired from Population behavior / Individual behavior:

Population Behavior :

- Algorithms which are inspired from the behavior of group of animals / birds / insects based on their interacting , searching , foraging , hunting nature.
- For Example :
  1. Particle Swarm Optimization
  2. Ant Colony Optimization
  3. Elephant Herding algorithm
- This is called as Swarm intelligence

Individual Behavior :

- Algorithms which are inspired from the behavior of individual animal / bird based on its searching , foraging , hunting nature.
- For Example :
  1. Cuckoo Search
  2. Crow Search
  3. Cat Search


1
Consider this birds flock , Every bird is on its own direction
But 3 birds are in the wrong direction
Global Best
Best Pos / direction
2
Now the Birds' Directions got changed – Not completely , Because of Inertia
3
Birds are tracking the global best and their personal best
Things that are guiding Birds : Social Intelligence and Individual Intelligence

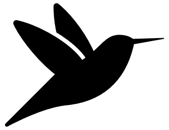
Particle Swarm Optimization: Mimic birds flocking

Let's Take an Example problem:

We have a Cost driver function

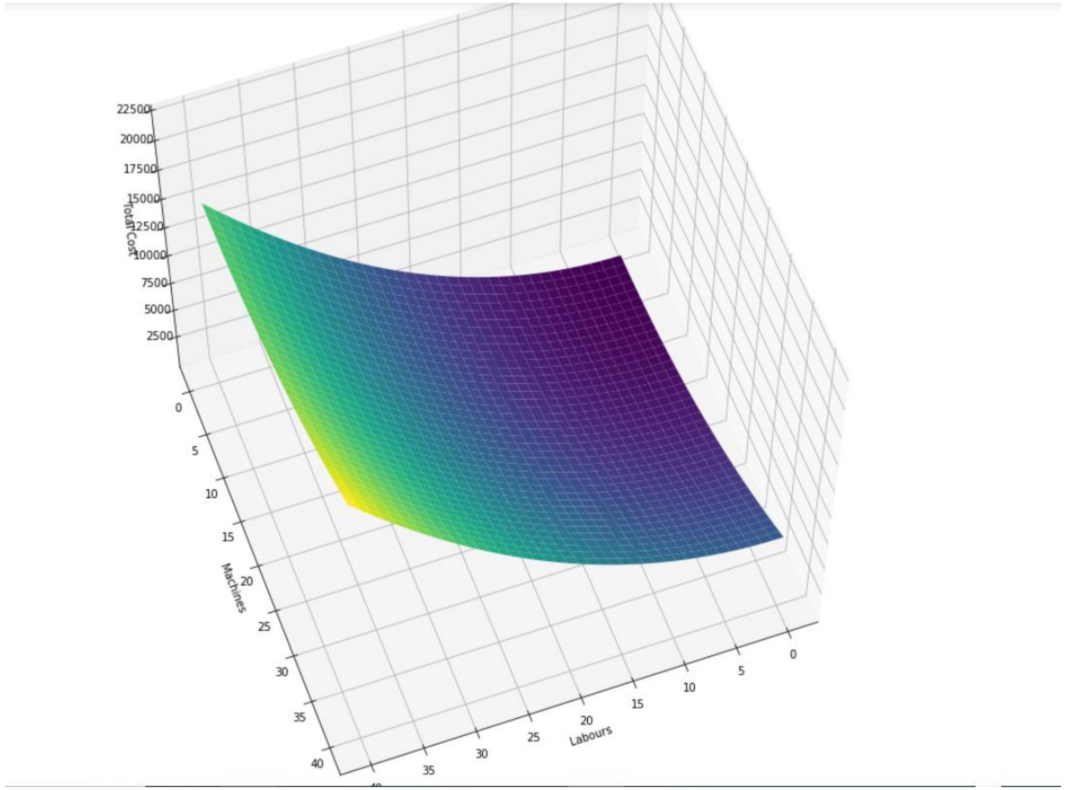
$$Min\ f(Rent\ , Labour) = 100 + 10\ \text{Labour}^2 + 4\ \text{Machine}^2$$

subject to Labour > =4 ;Machine >= 4



Particle 1
Position X1: (4 , 10)
Velocity V1: (4 , 4)
Fitness : 660
Personal Best : (4 ,10)

Particle 2
Position X2: (5, 5)
Velocity V2: (10 ,5)
Fitness : 450
Personal Best : (5,5)



So far, the Global Best : G best => (5,5) with 450 fitness function

If we let the birds behave individually , no one will reach the optima (which is a target)

The Role of Swarm intelligence (in the form of Global Best) helps individual particle to change their position and direction in order to move towards the food. The concept of PSO was published in 1995 by Kennedy et.al

Social behavior and individual behavior helps the whole flock to reach the correct target.

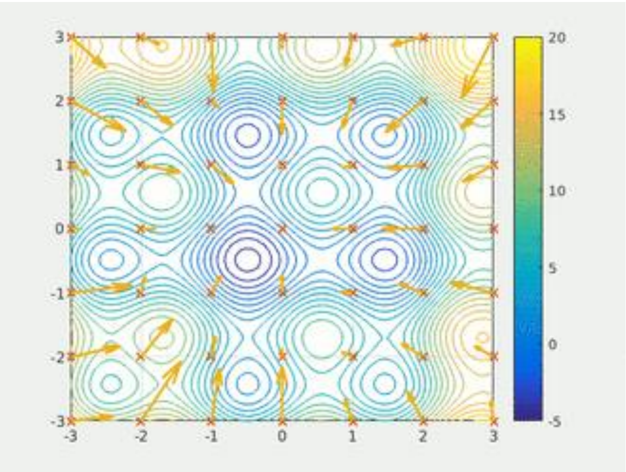
This can be achieved by :

Velocity Update : $v_{k+1}^{i} = wv_{k}^{i} + c_1 r_1 \left(Pbest\ -\ x_k^i\right) + c_2 r_2 (Gbest - x_k^i)$

Positional Update: $x_{k+1}^{i} = x_k^i + v_{k+1}^{i}$

$c_1 r_1 \left(Pbest\ -\ x_k^i\right)$ → Cognitive term (which comes from individual bird's experience)
$c_2 r_2 (Gbest - x_k^i)$ → Social term (which comes from group experience)

C1 → enabler of cognitive nature   r1 and r2 are random numbers
C2 → enabler of social nature   w → constant Inertia



Let's Initialize the parameters

- c1 = 1 , c2 =1 , r1 = [0.4, 0.4] r2 = [1,1], w=0.5
- $Min\ f(Rent\ , Labour) = 100 + 10\ \text{Labour}^2 + 4\ \text{Machine}^2$

Particle 1
Position X1: [4 , 10]
Velocity V1: [4 , 4]
Fitness : 660
Personal Best : [4 ,10]

Particle 2
Position X2: [5, 5]
Velocity V2: [10 ,5]
Fitness : 450
Personal Best : [5,5]

Iteration 1:

- Particle 1:
- New Velocity : [3,4]
- New Position : [7 , 14]
- New Fitness : 1246
- Pbest : [4, 10]
- Gbest: [5,5]

- Particle 2:
- New Velocity : [5,4]
- New Position : [10, 9]
- New Fitness : 1424
- Pbest : [5,5]
- Gbest: [5,5]

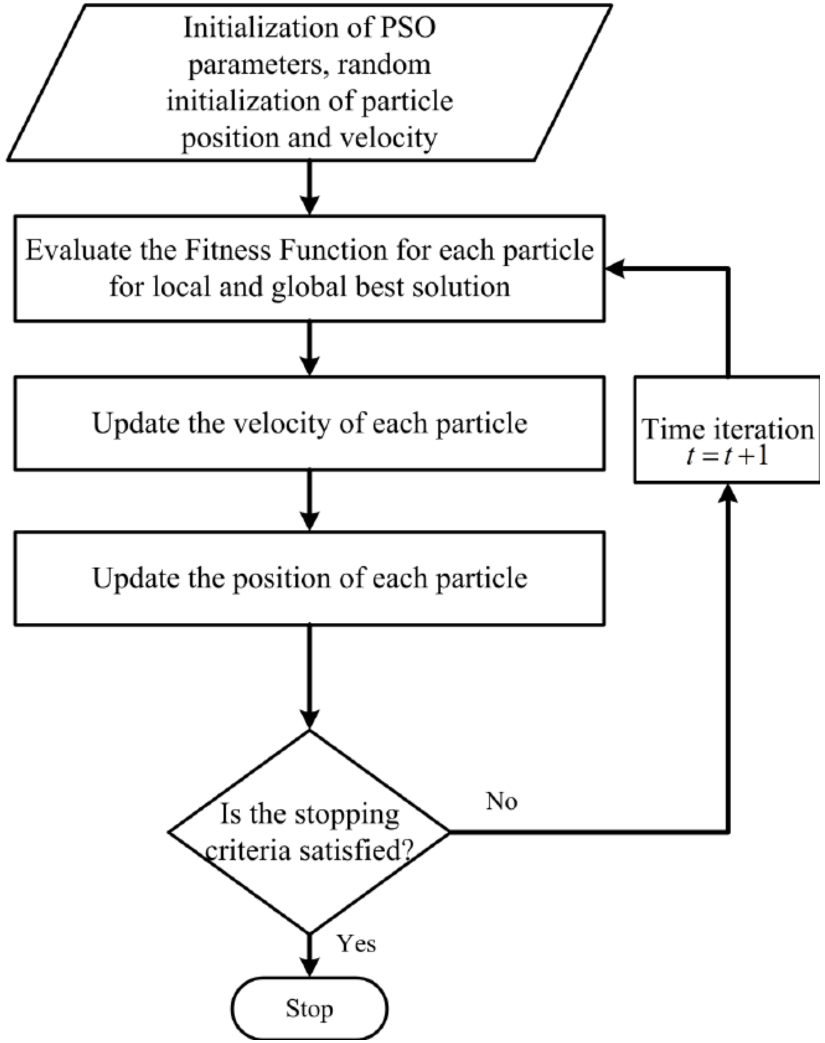
Iteration 2:

- Particle 1:
- New Velocity : [7.9,13.4]
- New Position : [14.9, 27.4]
- New Fitness : 5323.04
- Pbest : [4, 10]
- Gbest: [5,5]

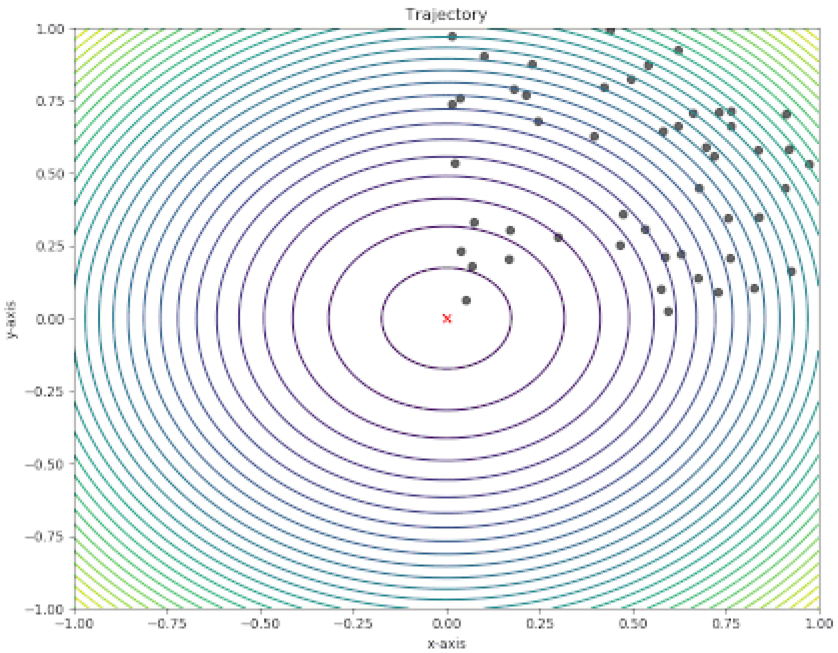
- Particle 2:
- New Velocity : [4,4]
- New Position : [14 , 13]
- New Fitness : 2736
- Pbest : [5, 5]
- Gbest: [5,5]

By Continuing the iterations by adjusting their position and velocity birds can be able to reach the lowest point

The flow chart given below explains the whole PSO algorithm:



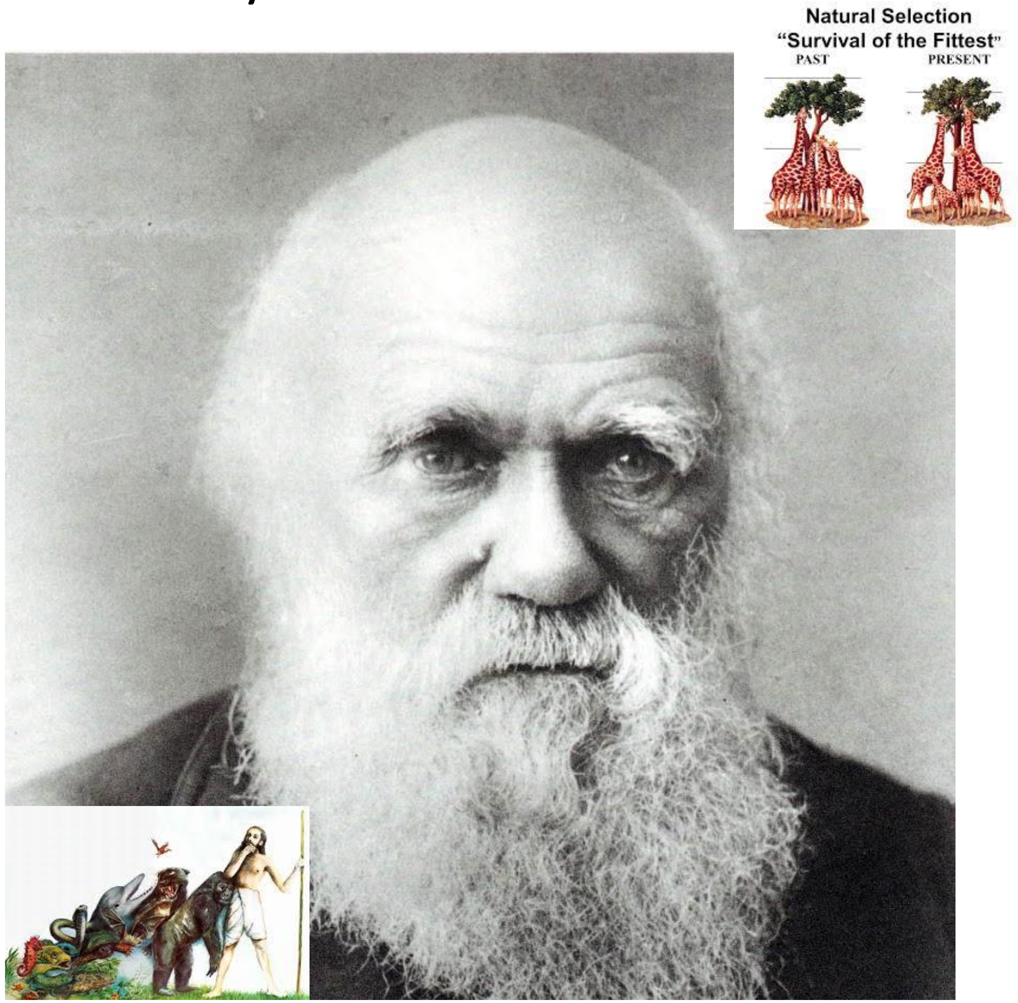
Applications of Particle Swarm optimization:
-- Hyper parameter tuning
-- Multimodal optimization
-- Global Search and Convex optimization



Evolutionary Algorithms : Inspired from Theory of Evolution:

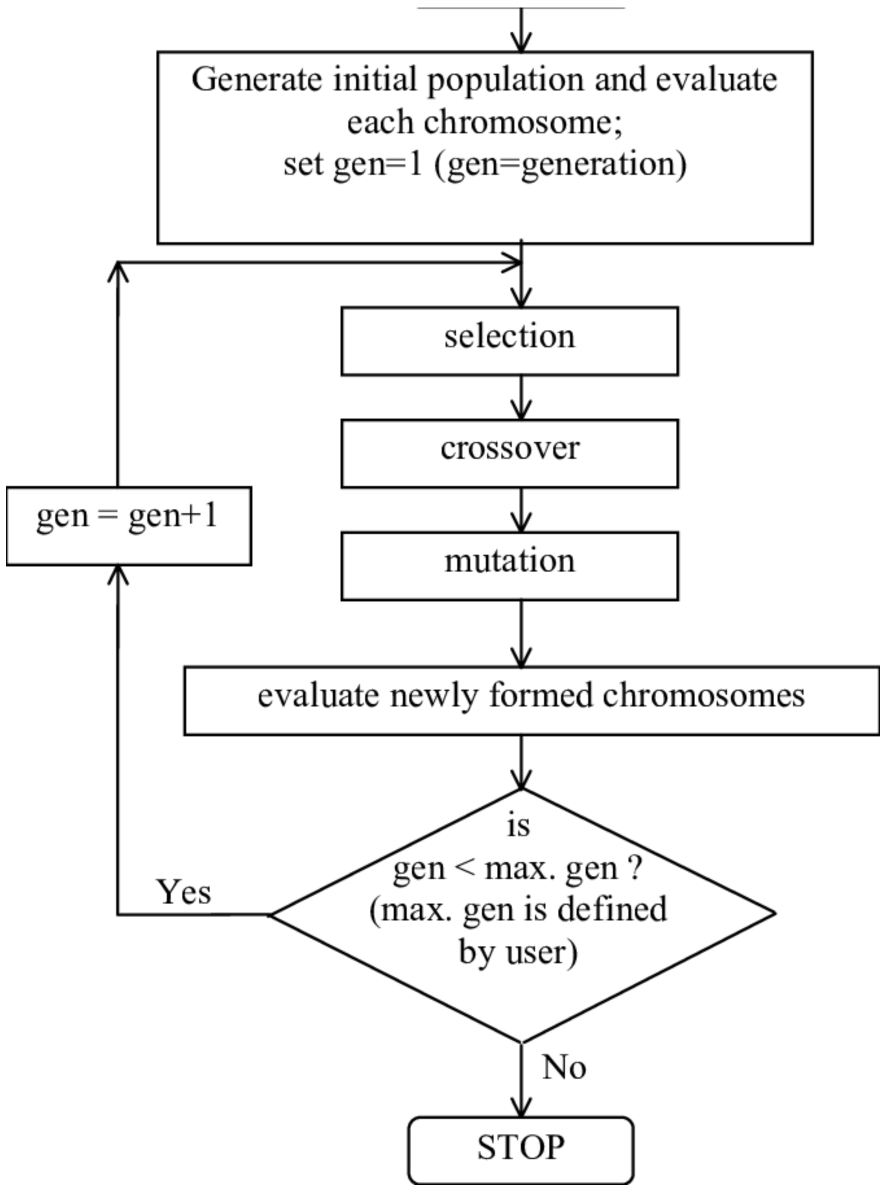
Genetic Algorithm:

- The genetic algorithm is a search-based optimization technique. It is frequently used to find the optimal or nearest optimal solution. It was introduced by John Holland in 1975.
- It is based on Darwin's Natural Selection Theory. Before explaining how the genetic algorithm works let me first explain Darwin's theory on natural selection.
- In his theory, he defined natural selection as the "principle by which each slight variation [of a trait], if useful, is preserved".
- The concept was simple but powerful: individuals best adapted to their environments are more likely to survive and reproduce
- "Survival of the Fittest"

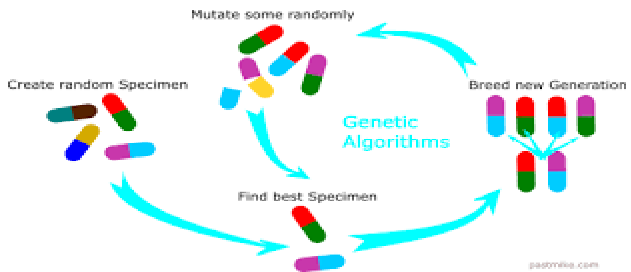


Nowadays Lot of Varieties of Genetic algorithms are available. Here we discuss the Very Fundamental one.

Genetic Algorithm : Flow chart



- Population : Set of Solutions
- Chromosomes : Each solution is a Chromosome
- Gene : Granular level of Chromosome.
- Selection : Process of selecting best fits to produce the new generation (Parental Chromosomes)
- Crossover : process of producing new off springs – Children
- Genotype and Phenotype
- Mutation : Changes in Gene for maintaining the diversity from one generation to another

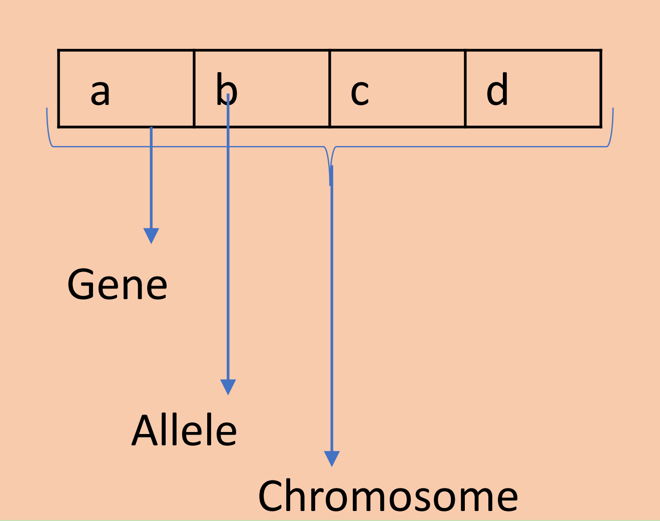


Example with Step-by-Step Explanation

- We have a Function : f(a:d) = f(x) = ((a + 2b + 3c + 4d) - 30) Our objective is to find the a:dvalues to minimize the f(a:d)
- Each chromosome can be interpreted as :

a	b	c	d

- Each variable is called as Gene (a : d) ;
- The value hold by gene is Allele
- We can perform the Genetic Algorithm with or without doing Encoding.



Step1: Initialize the population of chromosomes.

- Chromosome[1] = [a;b;c;d] = [12;05;23;08]
- Chromosome[2] = [a;b;c;d] = [02;21;18;03]
- Chromosome[3] = [a;b;c;d] = [10;04;13;14]
- Chromosome[4] = [a;b;c;d] = [20;01;10;06]
- Chromosome[5] = [a;b;c;d] = [01;04;13;19]
- Chromosome[6] = [a;b;c;d] = [20;05;17;01]

Selection of Parental chromosomes.

Step 2 : Evaluation using fitness function and Probability of selection calculation

Selection of Parental chromosome is depended on Probability which follows Roulette wheel method

In Realtime , Imagine you are performing media spend optimization / Project scheduling

f(a:d) = min f(x) = ((a + 2b + 3c + 4d) - 30)

Initial Population	F(x) Function	Fitness function (1/(1+F))	Probability count	Estimated Count	Actual Count
[12;05;23;8]	93	0.0106	0.125	0.7	1
[02;21;18;03]	80	0.0123	0.145	0.8	1
[10;04;13;14]	83	0.0119	0.1408	0.8	1
[20;01;10;06]	46	0.0213	0.2521	1.5	2
[01;04;13;19]	94	0.0105	0.1243	0.7	0
[20;05;17;01]	55	0.0179	0.2118	1.2	1

Total : 0.0845 Average : 0.0145

Selected Parental chromosomes

- Chromosome[1] = [a;b;c;d] = [12;05;23;08]
- Chromosome[2] = [a;b;c;d] = [02;21;18;03]
- Chromosome[3] = [a;b;c;d] = [10;04;13;14]
- Chromosome[4] = [a;b;c;d] = [20;01;10;06]
- Chromosome[5] = [a;b;c;d] = [20;01;10;06]
- Chromosome[6] = [a;b;c;d] = [20;05;17;01]

First Parental Chromosome pair

Second Parental Chromosome pair

Third Parental Chromosome pair

Cross over - Second Generation

- offspring[1] = [a;b;c;d] = [12;05;18;03] => fitness : 88
- offspring[2] = [a;b;c;d] = [02;21;23;08] => fitness: 145
- offspring[3] = [a;b;c;d] = [10;04;10;06] => fitness: 72
- offspring[4] = [a;b;c;d] = [20;01;13;14] => fitness: 98
- offspring[5] = [a;b;c;d] = [20;01;17;01] => fitness: 77
- offspring[6] = [a;b;c;d] = [20;05;10;06] => fitness: 84

Mutation:

- Mutation is a change happens inside a chromosome itself
- Mutation will not happen every time
- For Example : chromosome [12;05;18;03] can mutate into [12;07;18;09]

- Optimal solutions will not arrive in a single generation
- After 50 generations : Chromosome = [07; 05; 03; 01] has minimized f(x) into zero
- This might be very simple problem , but in real time problems like multi objective and multimodal problems will be more complex.

Applications of Genetic Algorithms :

1. Hyperparameter tuning
2. Global Optimization with constraints and multiple objectives

Questions

Thank you