



Inspire...Educate...Transform.

Simulations, Evolutionary searches

Dr. K. V. Dakshinamurthy

President, International School of Engineering



Random and guided random searches

WORKING WITH RANDOM SAMPLES

Measuring central tendencies

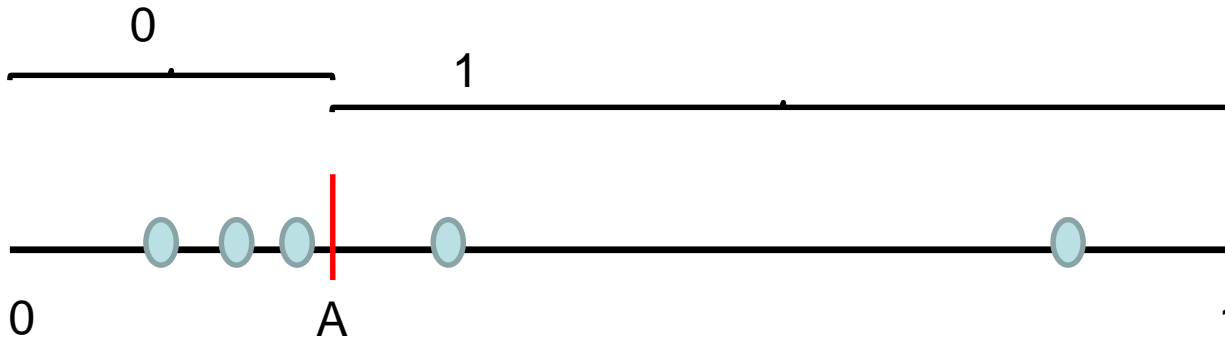
- Computing mean median and mode is a lot of calculus for a continuous distribution
 - Σ must be replaced by \int sign.
 - $Mean = \int xp(x)dx$
 - $Median = \int_0^M p(x)dx = \int_M^k p(x)dx = \frac{1}{2}$
 - $Mode$ is where $\frac{d(p(x))}{dx} = 0$

Estimating from samples

- Let us have IID (identically, independently distributed samples) of a distribution
- Then, mean is

$$\frac{1}{M} \sum_{m=1}^M x[m]$$

Creating samples from a discrete distribution



$$P(0) = A \text{ and } P(1) = 1 - A$$

Randomly generate a number between 0 and 1; If it is less than A, sample is 0 else it is 1

Sample =(0,1,0,0,1)

Discretizing a continuous variable

- Equal bin, equal frequency
- With class variable, sort and decide where to cut (cut wherever the class variable changes its values)
- Many such algorithms exist

Age	25	32	34	35	35	37	37	38
Loan	0	1	1	0	0	0	1	1



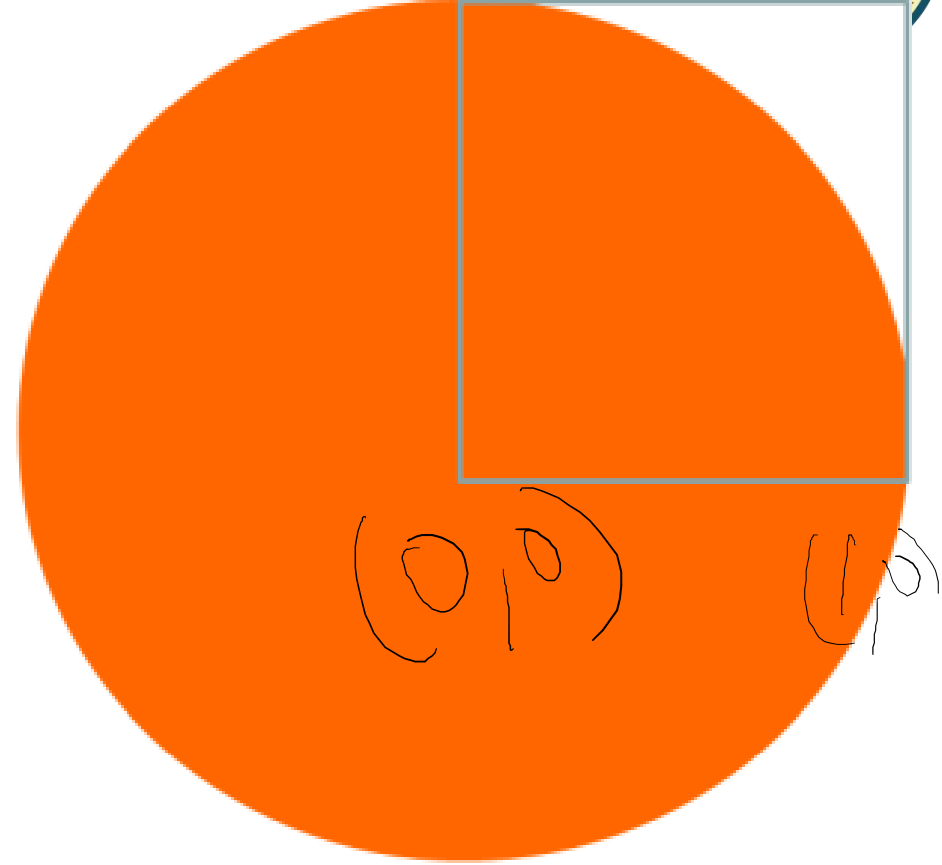
Computing central tendencies

APPLICATIONS OF SAMPLING: STRUCTURED DATA

Monte Carlo Simulations

Value of Pi

$$x^2 + y^2 \leq 1$$

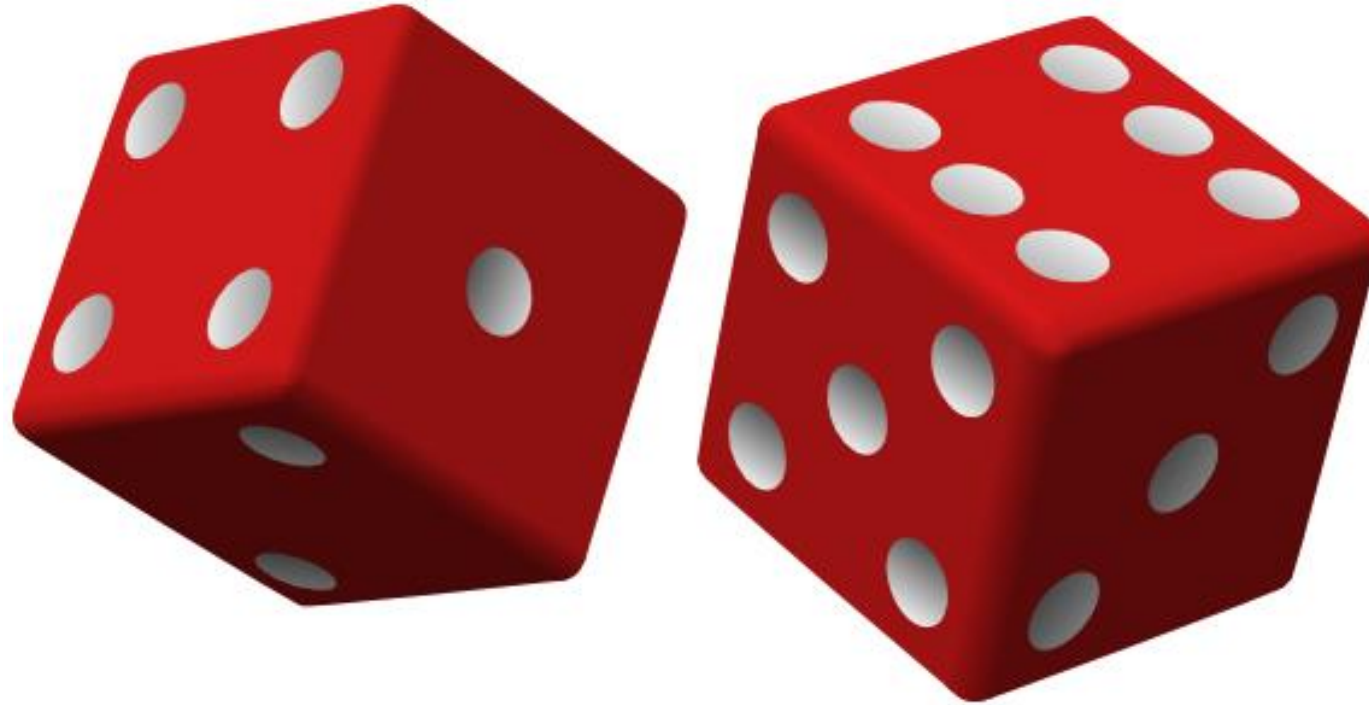




When to stop?

- Define a small tolerance δ delta
- If the difference between n^{th} iteration and $n+1^{\text{th}}$ iteration is less than delta, stop.
Else, continue

Probability



Capital markets



Outcome	Green	Red	White
1	0.8	0.06	0.9
2	0.9	0.2	1
3	1.05	1	1
4	1.1	3	1
5	1.2	3	1
6	1.4	3	1.1

Montecarlo and Linear optimization

A better sensitivity analysis



- Generate a set of samples from distribution
 - Optimize
 - Iterate and look at the distribution of the result
 - Stop at convergence



APPLICATIONS OF SAMPLING

**UNSTRUCTURED DATA: GENERATING
AND UNDERSTANDING BEHAVIOR OF
TEXT, IMAGES, AUDIO**

Transition probabilities

- Process
 - Learn the distribution from a large corpus
 - Collate counted probabilities as a matrix (columns, current state and rows future state)
 - Text, images, music etc.

Learning distributions



	a	b	c	d	e	f	g	h	i
A	0.08	0.09							
B									

Larger corporuses yield better descriptions

Character generation

- Random
 - No use
- First order (character-Character transitions)
 - t I amy, vin. id wht omanly heay atuss n macon
aresethe hired boutwhe t, tl, ad torurest t plur I wit
hengamind tarer-plarody thishand.
- Second order (Previous 2 characters, di-gram)
 - Ther I the heingoind of-pleat, blur it dwere wing
waske hat trooss. Yout lar on wassing, an sit."
"Yould," "I that vide was nots ther.

Applications of Markov chains



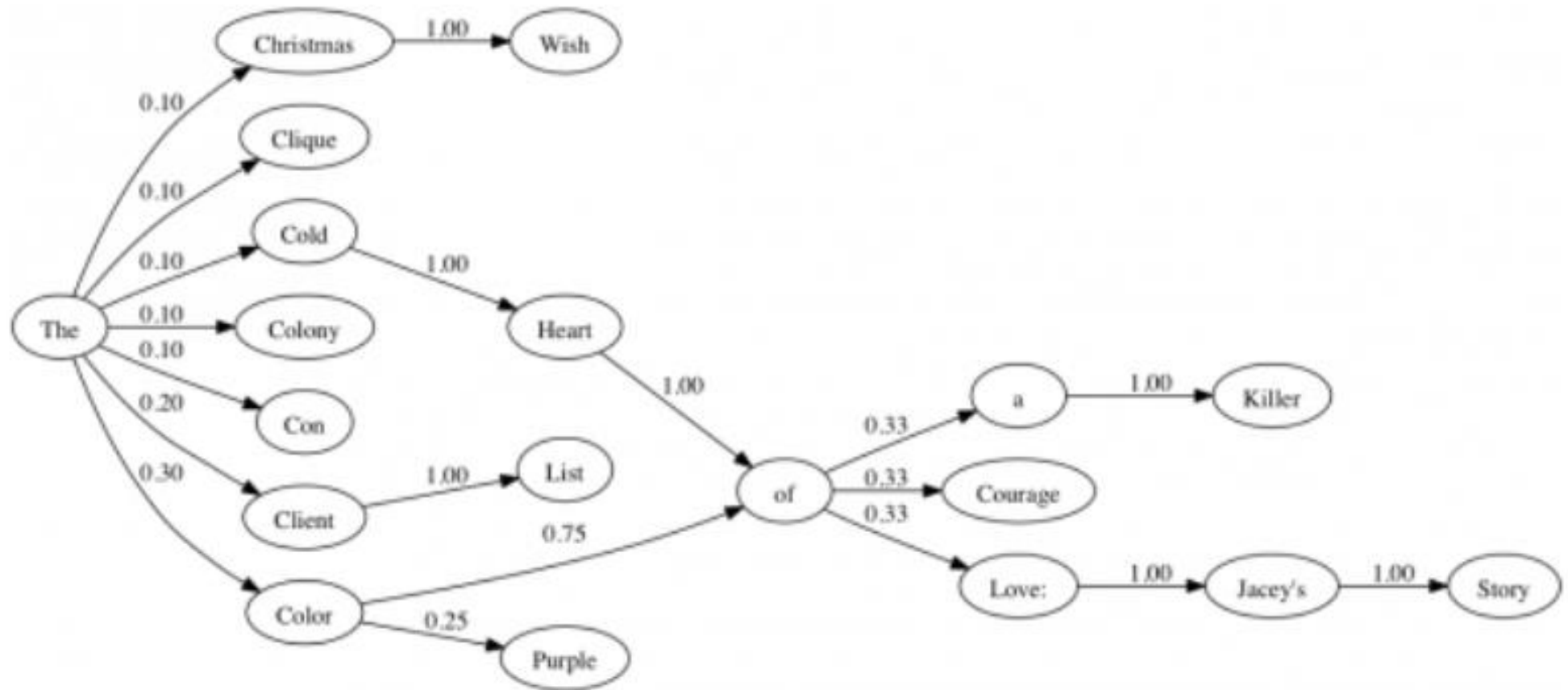
- 3rd order
 - I has them the saw the secorrow. And wintails on my my ent, thinks, fore voyager lanated the been elsed helder was of him a very free bottlemarkable,
- 4 gram
 - His heard." "Exactly he very glad trouble, and by Hopkins! That it on of the who difficentralia. He rushed likely?" "Blood night that.

Text generation

- First order word markov chains
 - Construct a matrix with transition probabilities of all word pairs and the probability of succeeding
 - Start random and select the words obeying the distribution

Let us generate new movie titles

- <http://www.soliantconsulting.com/blog/2013/02/title-generator-using-markov-chains>
 - The Christmas Wish, The Client, The Client List, The Clique, The Cold Heart of a Killer, The Colony, The Color of Courage, The Color of Love: Jacey's Story, The Color Purple, The Con
 - Constructing a list of unique words that appear in these titles: "The", "Christmas", "Wish", "Client", "List", "Clique", "Cold", "Heart", "of", "a", "Killer", "Colony", "Color", "Courage", "Love:", "Jacey's", "Story", "Purple", "Con".



Probability diagram

Process



- We can start with the word “The” and for each new word desired, we’ll choose from the distribution which next node to move to accordingly.
- Some possible generated strings include new titles “The Color of a Killer” or “The Cold Heart of Love: Jacey’s Story”.
- With more data, a more complicated graph can be constructed, and more interesting and varied strings can be generated as a result. You should use as much data in your corpus as you can find!
- Small tweaks like setting a minimum length, choosing to start with start words and end with end words might help



AN INTERESTING PROPERTY OF TRANSITION PROBABILITY MATRICES



Understanding Markov Chain

- Let us say, my class has the following properties
 - 90% of the interested students will remain interested; 10% doze off
 - 50% of the dozing students get interested; remaining stay as it is



What is the steady state?

- Let us start with 100 interested and 0 boring
- Work out what happens after 5 periods

Using matrices to solve the problem



- Transition matrix
 - Each columns: one state transition
- Transition matrix * current distribution =
Next distribution

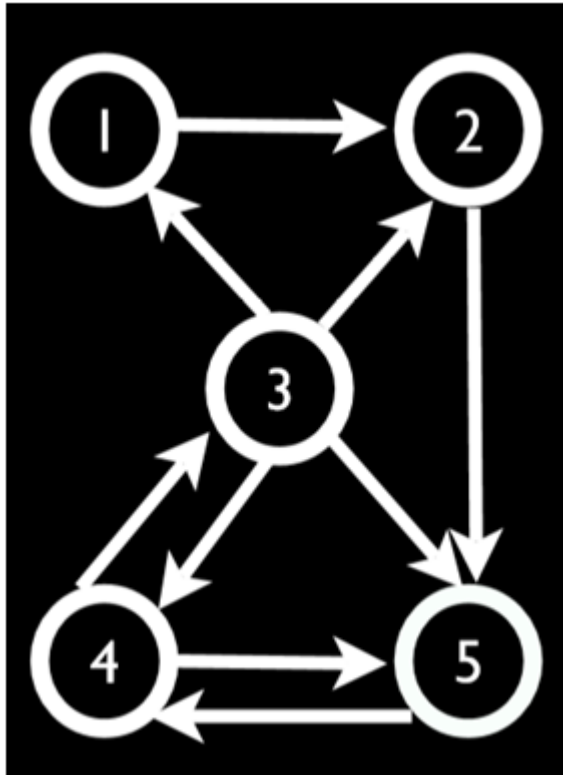


This is a Markov chain

- Finite states
- Fixed transition probabilities
- From one state, I should be able to go to any state
- Column is a stochastic matrices

Other applications

- Google Page rank



$$\begin{bmatrix}
 0 & 0 & 1/4 & 0 & 0 \\
 1 & 0 & 1/4 & 0 & 0 \\
 0 & 0 & 0 & 1/2 & 0 \\
 0 & 0 & 1/4 & 0 & 1 \\
 0 & 1 & 1/4 & 1/2 & 0
 \end{bmatrix}$$

$$\begin{bmatrix} 1/20 \\ 5/20 \\ 1/10 \\ 5/20 \\ 7/20 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1/4 & 0 & 0 \\ 1 & 0 & 1/4 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1/4 & 0 & 1 \\ 0 & 1 & 1/4 & 1/2 & 0 \end{bmatrix} \begin{bmatrix} 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \end{bmatrix}$$

$$r^{(t+1)} = H r^{(t)}$$

	Iteration 0	Iteration 1	Iteration 2	Page Rank
P_1	1/5	1/20	1/40	5
P_2	1/5	5/20	3/40	4
P_3	1/5	1/10	5/40	3
P_4	1/5	5/20	15/40	2
P_5	1/5	7/20	16/40	1



Guided random searches

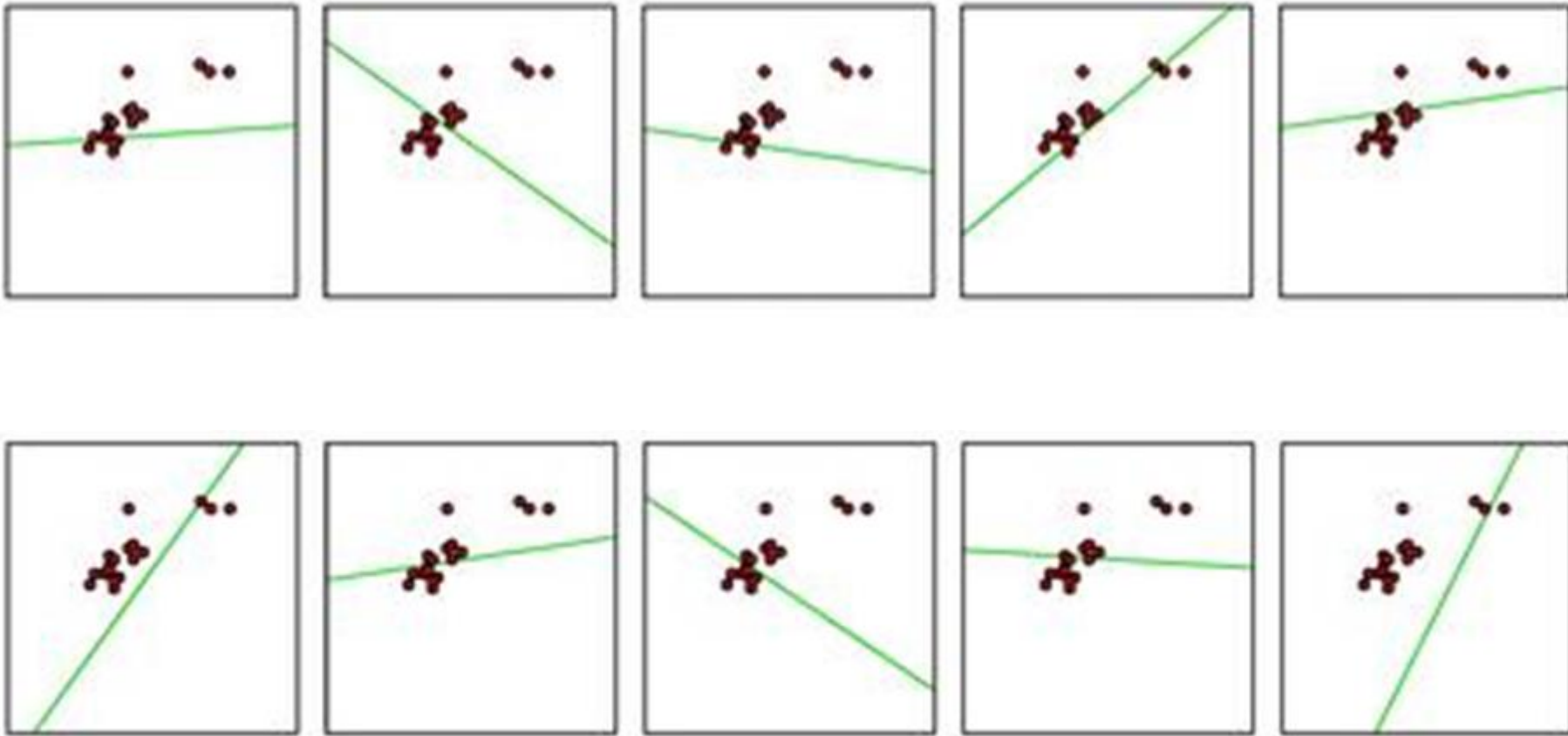
CAN SAMPLING BE USED TO SEARCH FOR MAXIMA AND MINIMA?



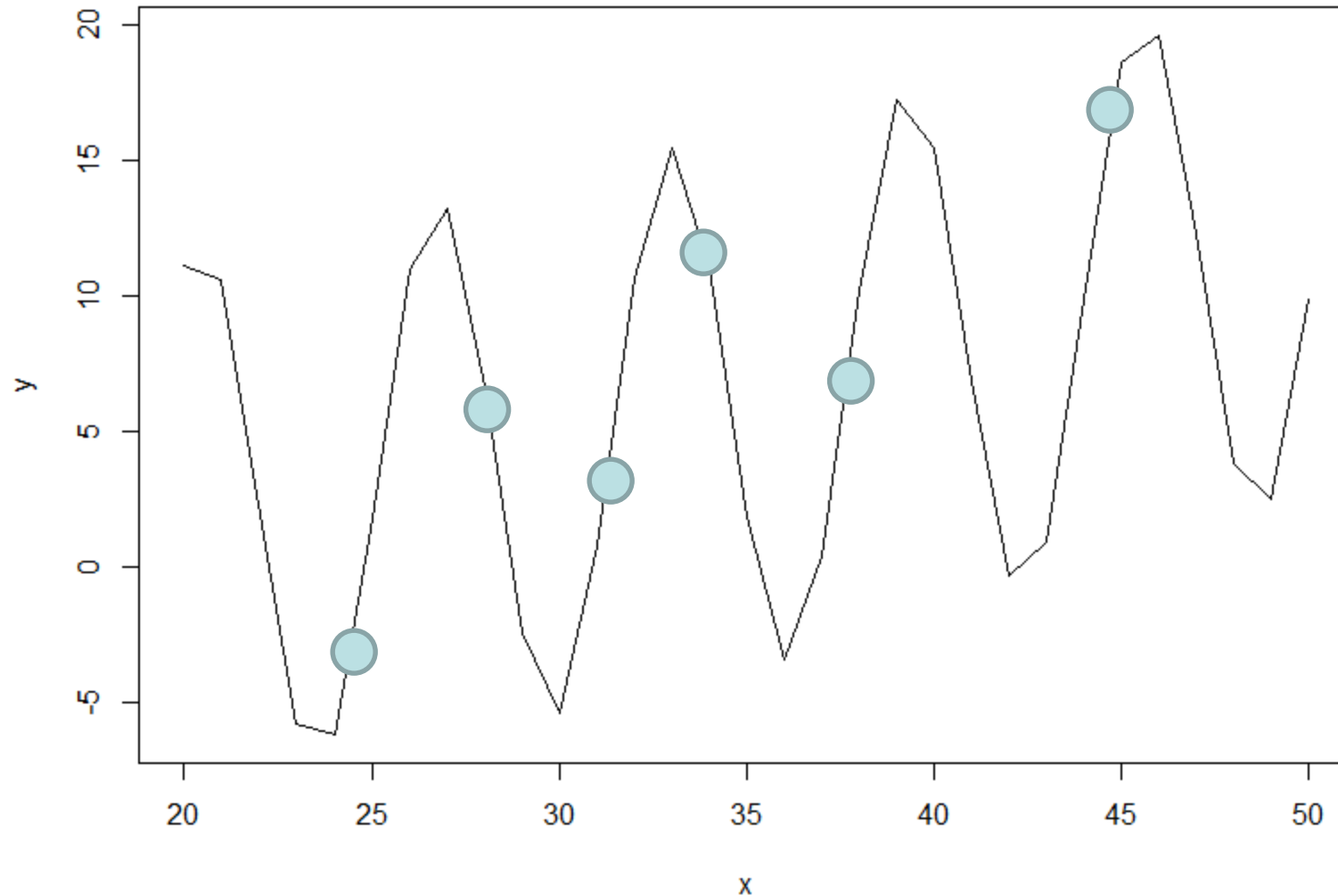
What Sampling and MonteCarlo are not

- They are random searches
- But, not optimization methods
- Can you do linear regression with Montecarlo?

Linear regression using Monte Carlo



A complex function and we need to find minima





A possible approach

- Do a million simulations and pick the most optimal solution
- Not guaranteed that we get anywhere close to the solution
- What to do?
 - Have memory



Genetic algorithms

GUIDED RANDOM SEARCHES

What does survival of the fittest mean



Summary of biology



- Not all members survive until reproduction. Only a few learn to earn lunch. Others are lunch!
- The strong ones then need to attract and mate other members of the species that survived.
- The offspring generation hence is normally better than the parents generation
- Over generations, species develop

Can a similar theory be applied to optimization

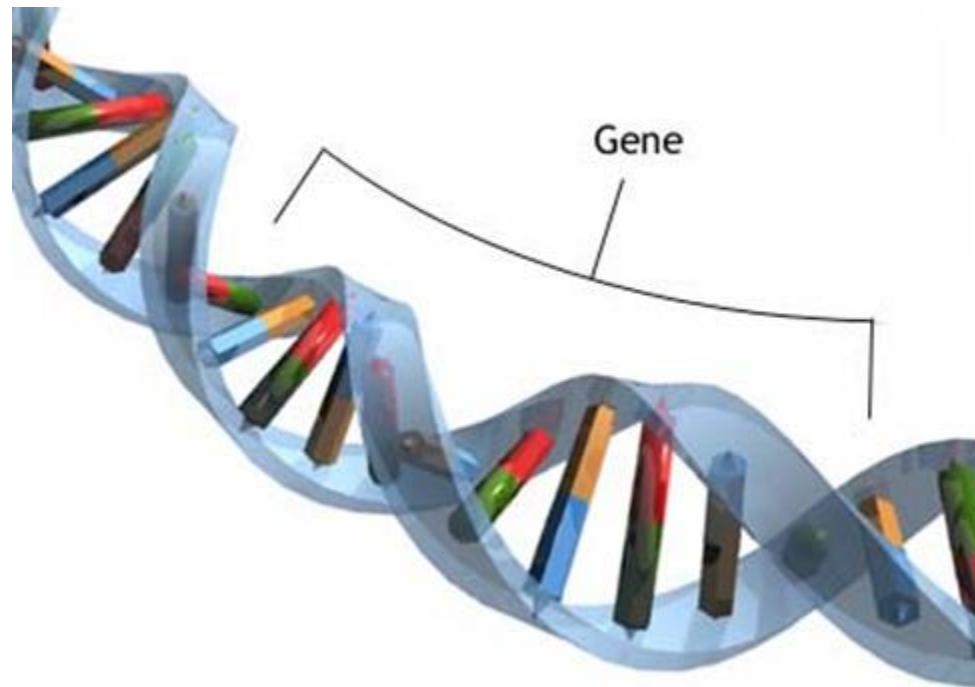


- Start with random solutions
- Kill the weak ones (memory!)
- Reproduce from the good ones
- Will the newer set be more optimal than its parents?

Gene



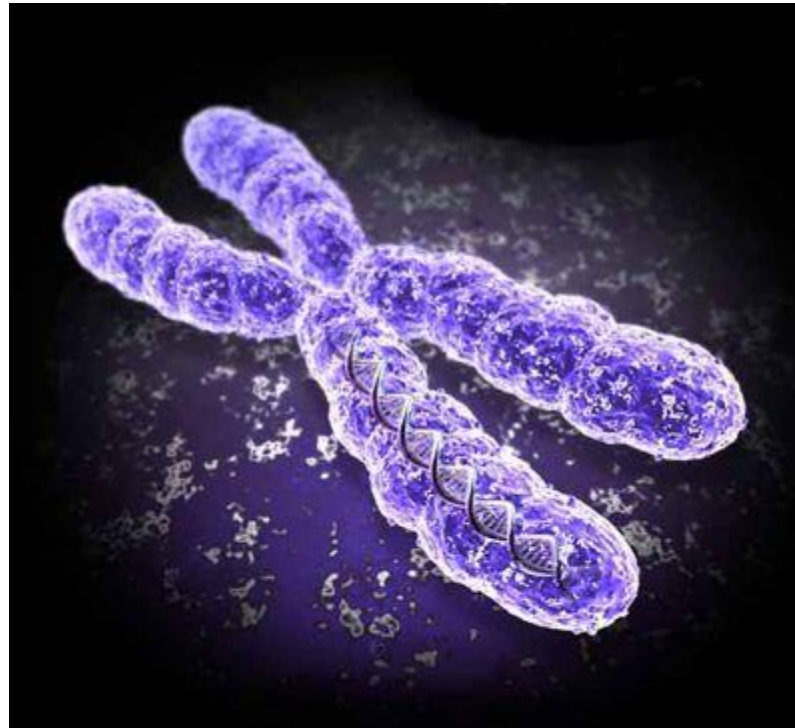
Gene: The reproducible building block of chromosome



Beg, Borrow or Steel



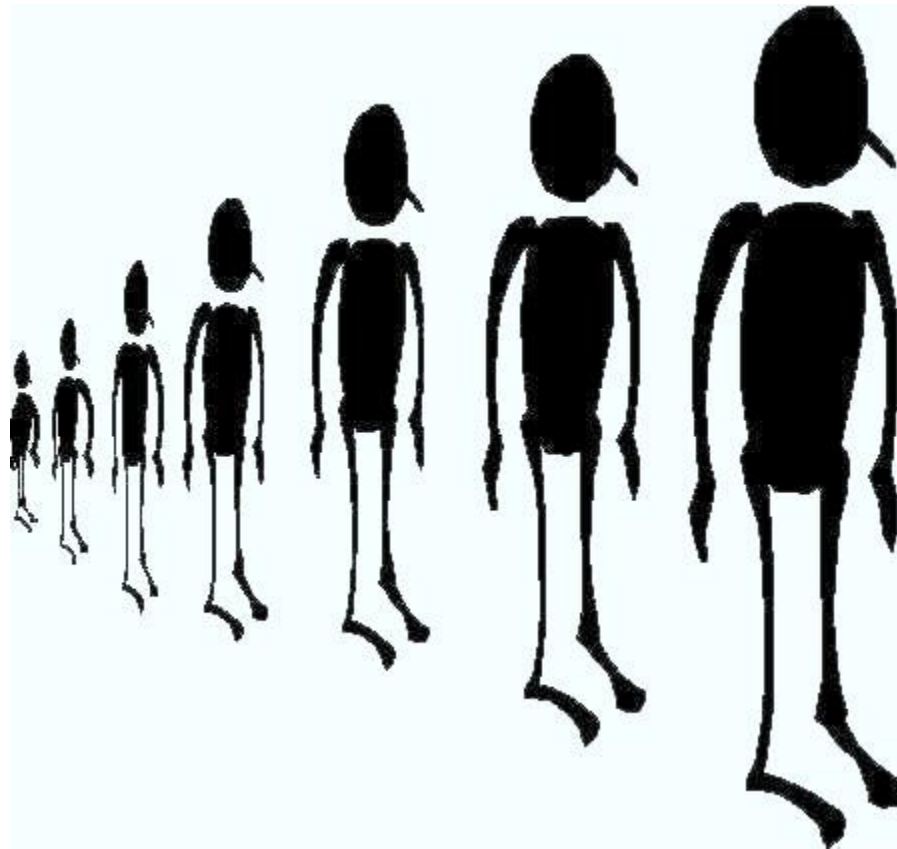
- Chromosome: A possible solution



Population



- A set of chromosomes



Generation



- A population derived from the fittest chromosomes of the previous population



Reproduction and Mutation

- Create a new solution from the old fitter solutions



```
{  
    initialize population;  
    evaluate population;  
    while TerminationCriteriaNotSatisfied  
    {  
        select parents for reproduction;  
        perform recombination and mutation;  
        evaluate population;  
    }  
}
```

Some unique differentiators

- GAs work with a coding of the parameters and not the parameters themselves
- GAs search from a population of points and not a single point
- GAs work with the objective function and not the derivatives
- GAs transition based on random rules and not on deterministic rules



OPERATIONAL ASPECTS



Innovations: Creating a coding scheme

- Knap Sack
- Polynomial fit
- TSP
- Portfolio allocation
- Define objective function and solution

Population



Chromosomes could be:

- Bit strings
- Real numbers
- Permutations of element
- Lists of rules
- Program elements
- ... any data structure ...

(0101 ... 1100)

(43.2 -33.1 ... 0.0 89.2)

(E11 E3 E7 ... E1 E15)

(R1 R2 R3 ... R22 R23)

(genetic programming)

How do we decide mating pool

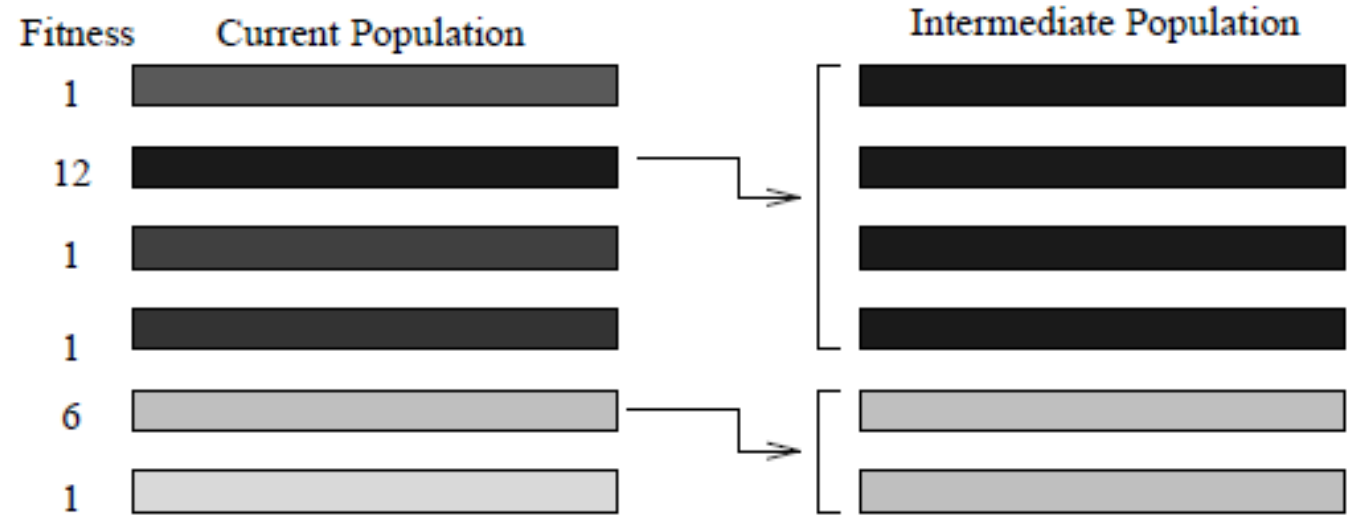
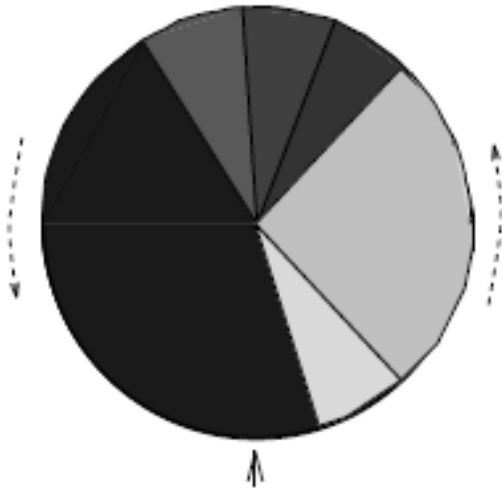


Figure 2. Selection application



Mutation or mating?

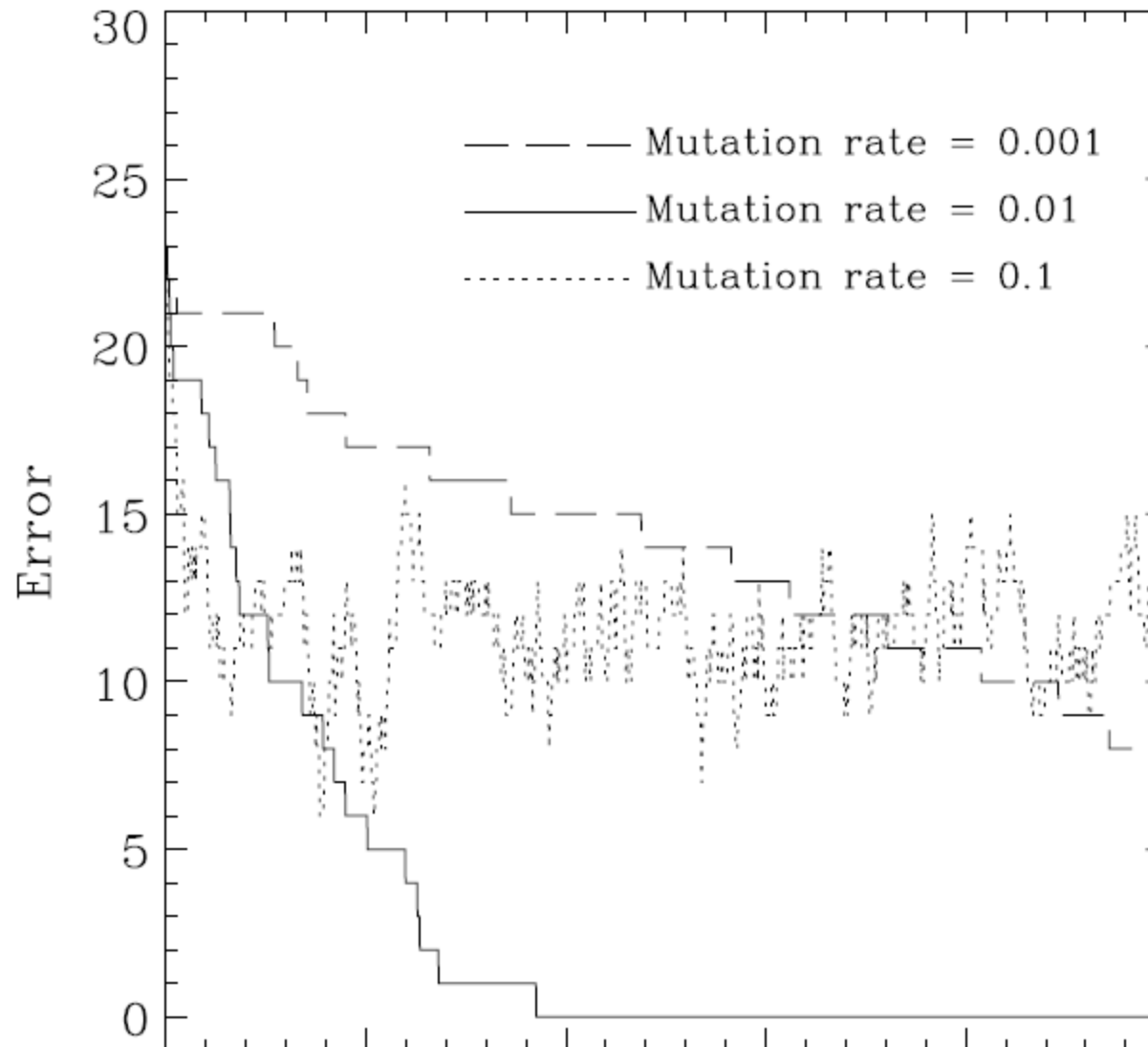
- A low probability of mutation
- Come down with generation
- ✓ Crossover : Requires 2 Parents.
- ✓ Mutation : Only 1 parent.



Non uniform mutation

- Initially mutate heavily
- As the number of generations increase mutate less

What about mutation?



How do we determine reproduction

- Most common: Cross over
 - npoint crossover: n crossover points are randomly selected and the segments of the parents, defined by them, are exchanged for generating the offspring.
 - uniform crossover: the values of each gene in the offspring are determined by the uniform random choice of the values of this gene in the parents.

For real coding

Arithmetical crossover (Michalewicz, 1992)

Two offspring, $H_k = (h_1^k, \dots, h_i^k, \dots, h_n^k)$ $k = 1, 2$, are generated, where $h_i^1 = \lambda c_i^1 + (1 - \lambda)c_i^2$ and $h_i^2 = \lambda c_i^2 + (1 - \lambda)c_i^1$. λ is a constant (uniform arithmetical crossover) or varies with regard to the number of generations made (non-uniform arithmetical crossover).

- Flat crossover: An offspring, H is generated, where h_i is a randomly (uniformly) chosen value of the interval $[c1_i; c2_i]$

Cross over



- Knap Sack
- Polynomial fit
- TSP
- Portfolio allocation



Inversion or Mutation

- Flip
- Random generation
- An unexpected operation



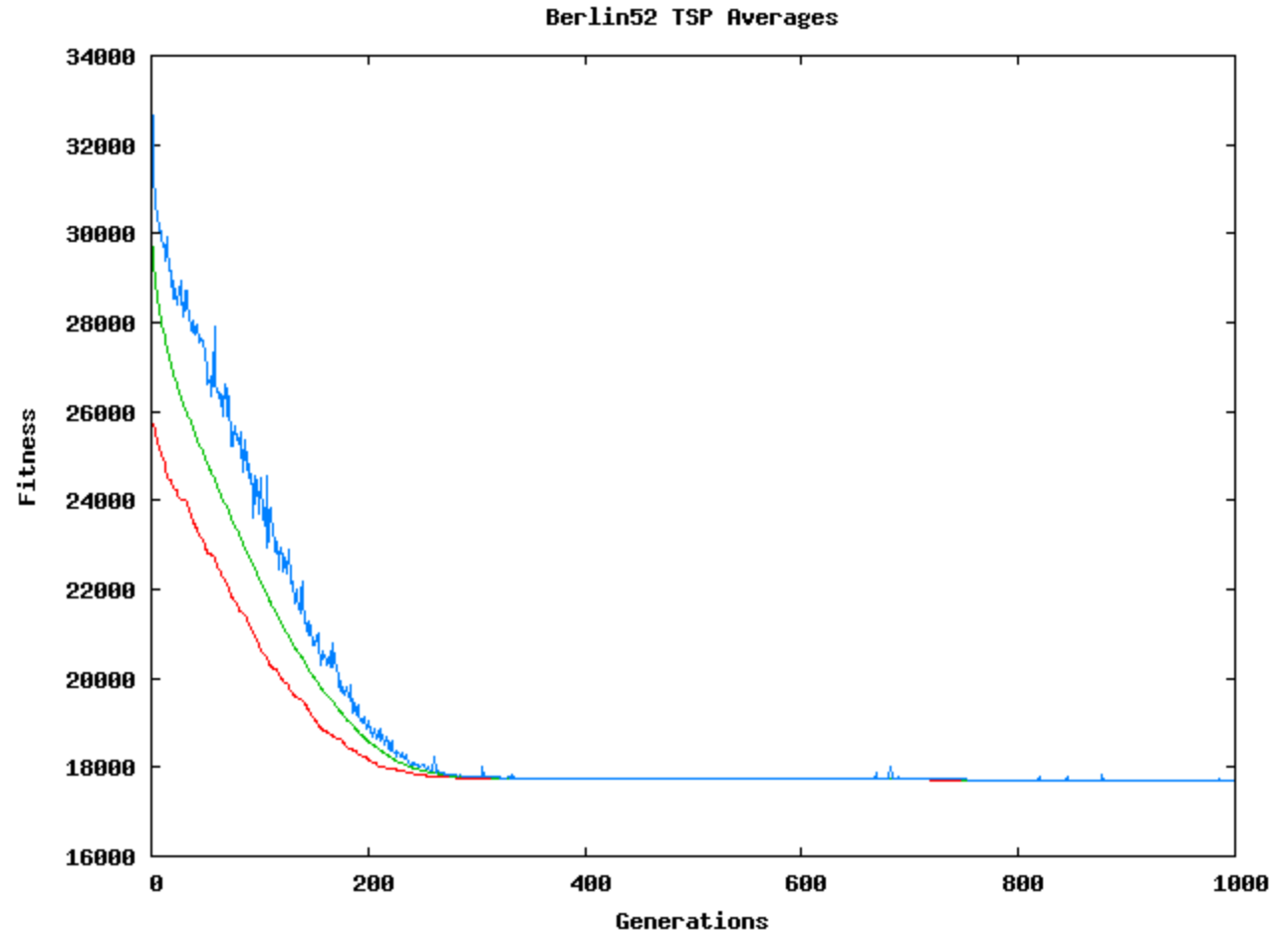
Mutation scheme

- Knap Sack
- Polynomial fit
- TSP
- Portfolio allocation

Convergence of a GA



When to stop?





EXAMPLES

A Simple Example

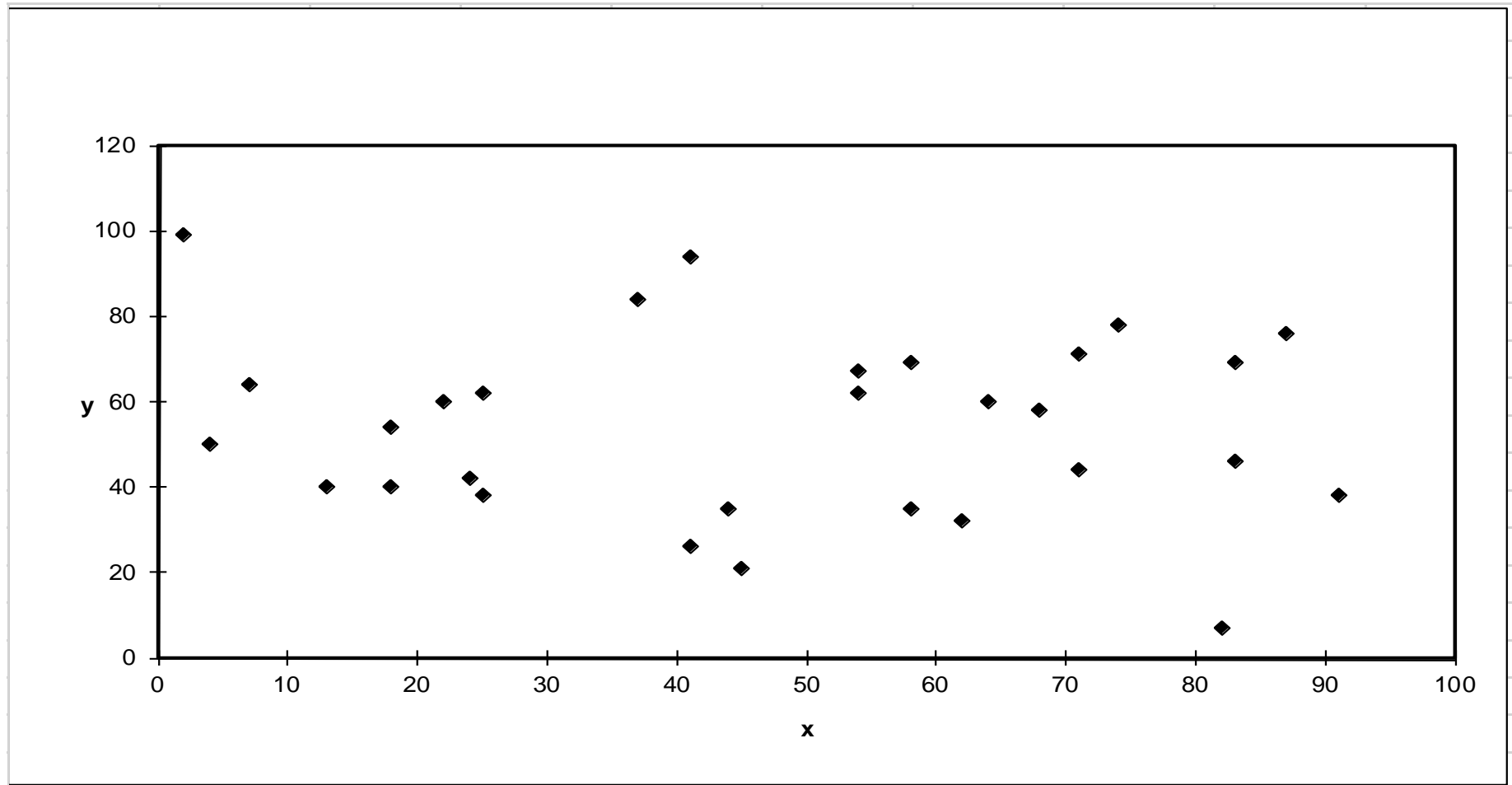


The Traveling Salesman Problem:

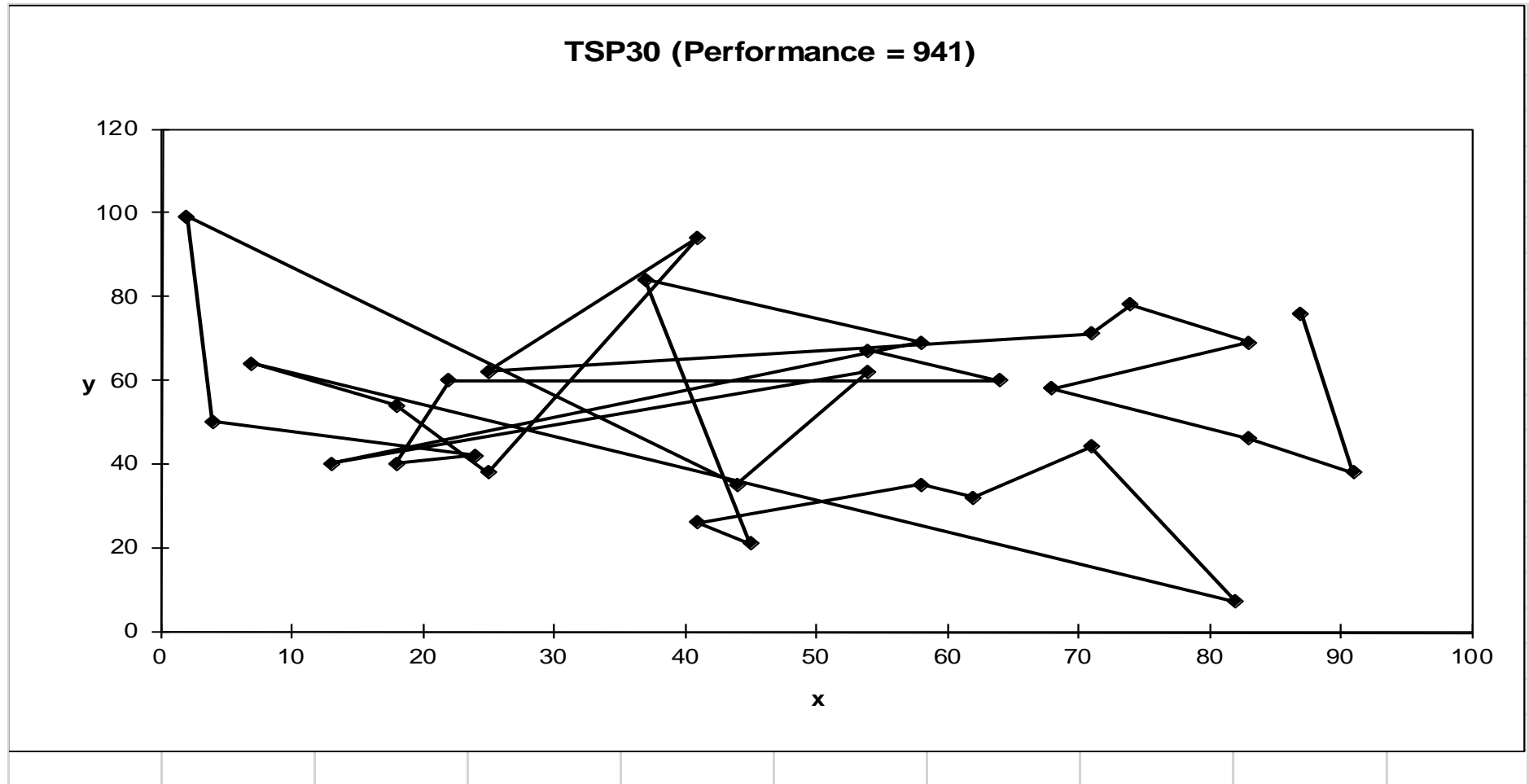
Find a tour of a given set of cities so that

- each city is visited only once
- the total distance traveled is minimized

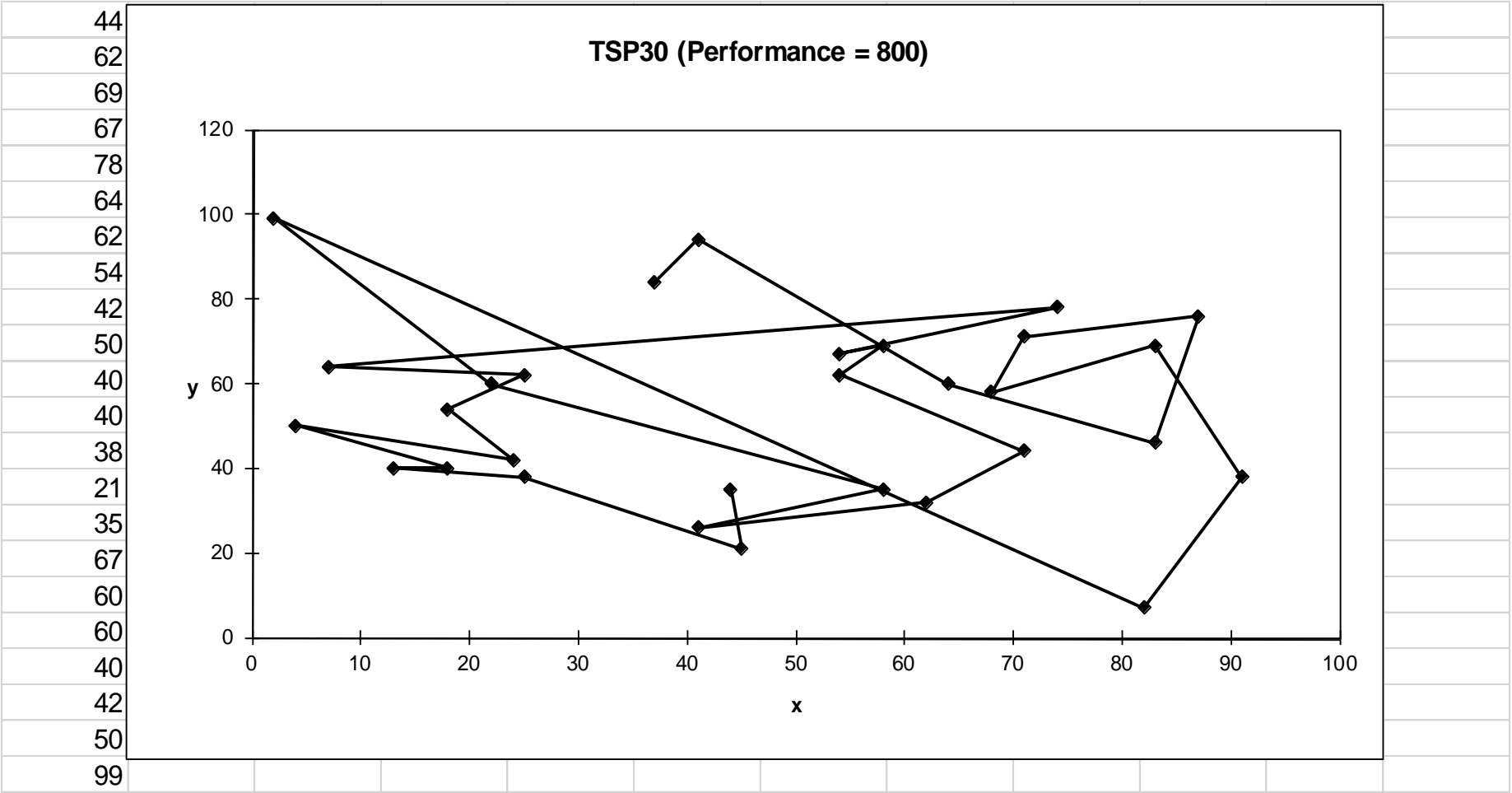
TSP Example: 30 Cities



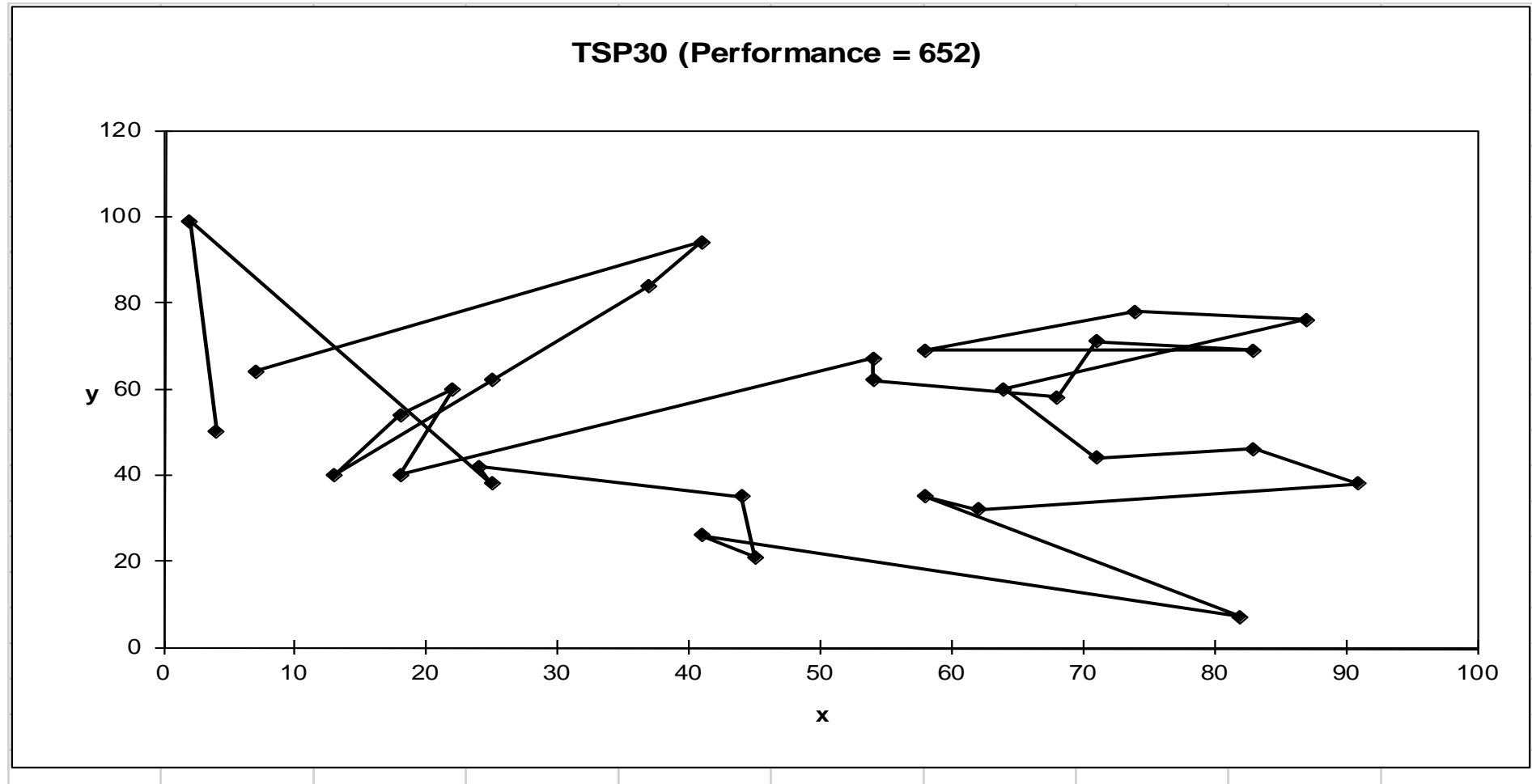
Solution _i (Distance = 941)



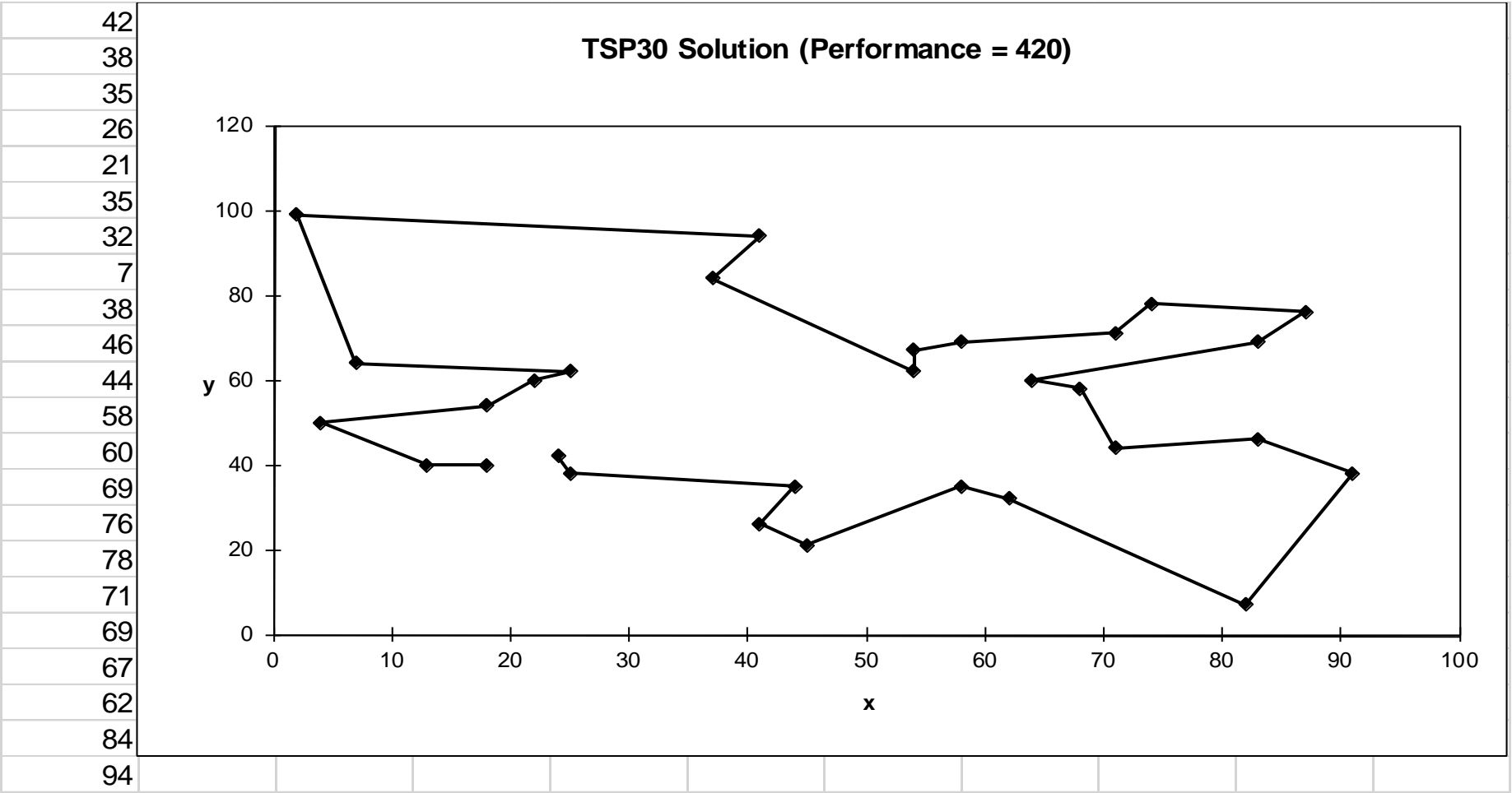
Solution _j (Distance = 800)



Solution k (Distance = 652)



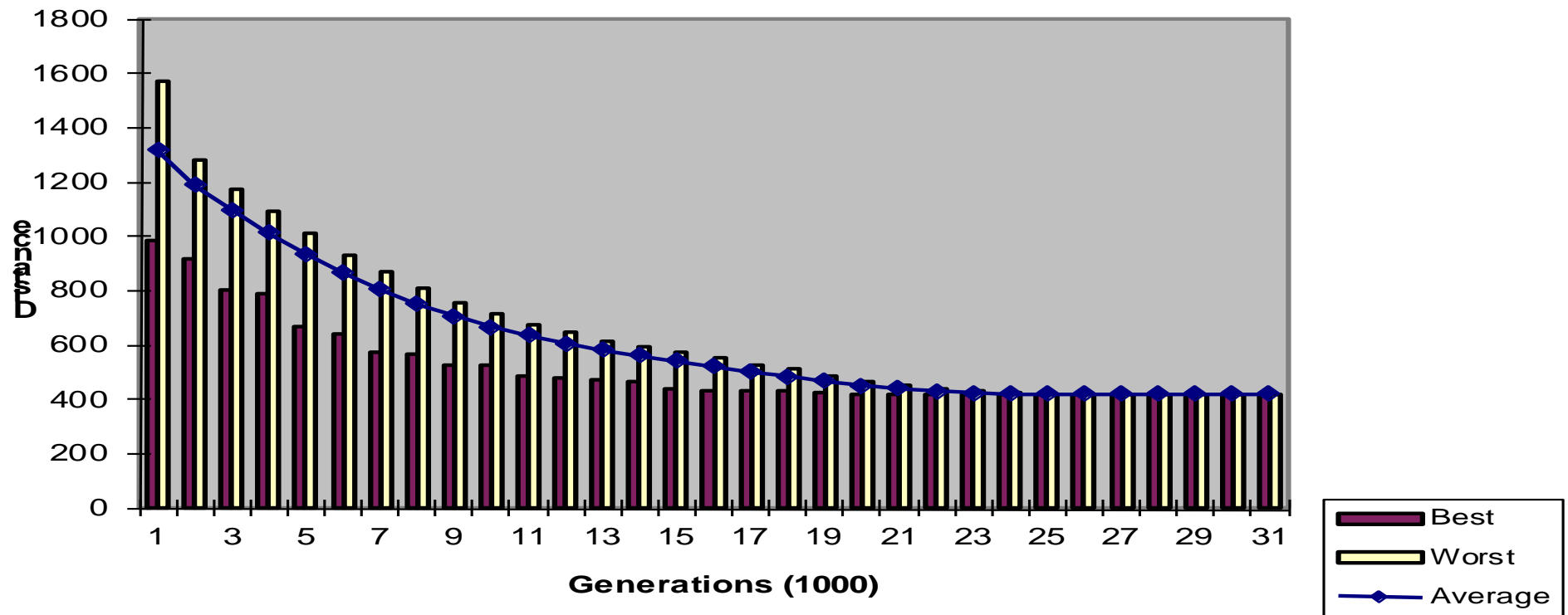
Best Solution (Distance = 420)



Overview of Performance



TSP30 - Overview of Performance





GA and Decision trees

- Create many trees
- Define mating (interchanging branches) and mutation (replacing a node with another)
- Define objective function (accuracy)
- Pick parents and evolve!!

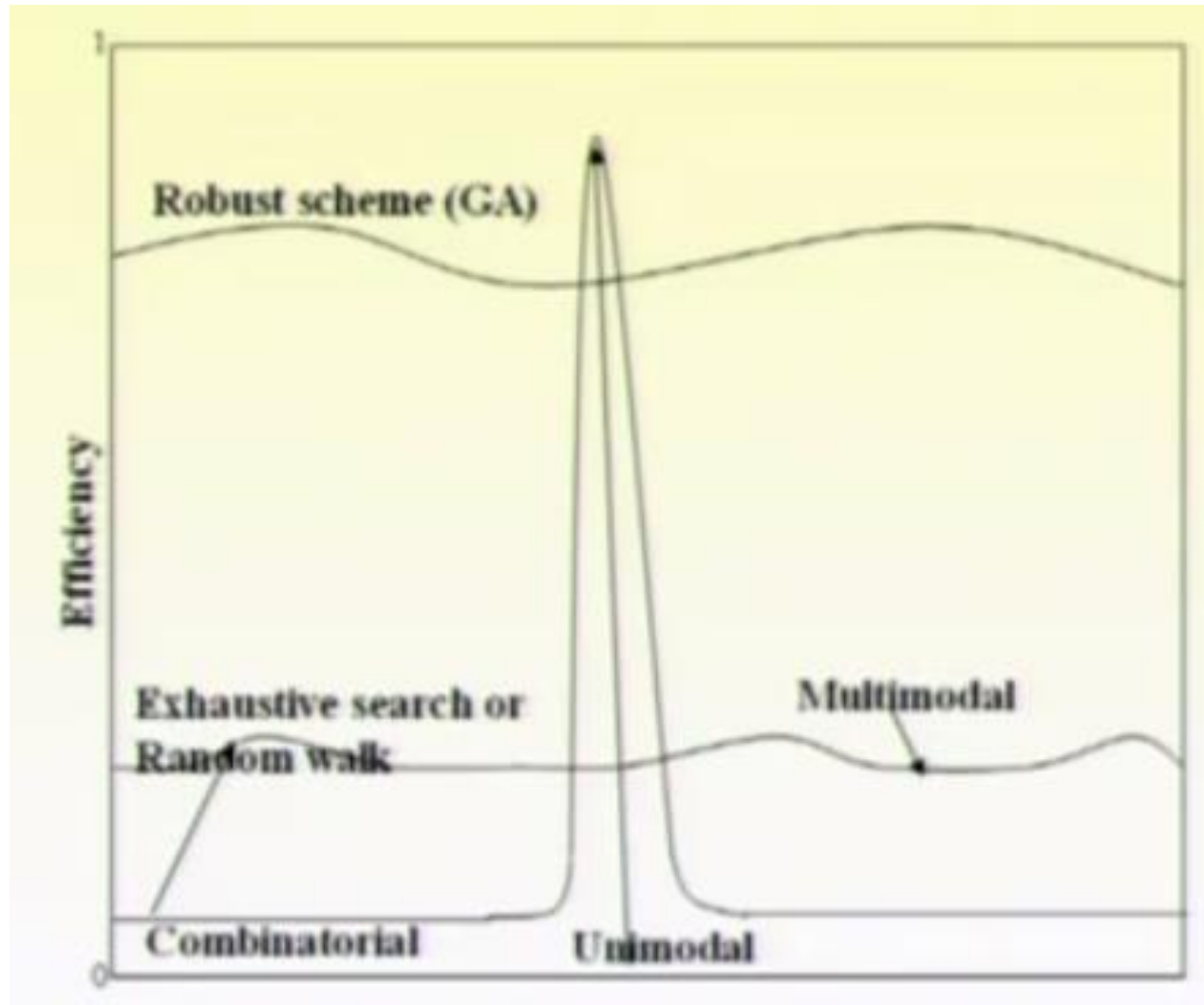
Comparison

Gradient descent and Traditional methods

- Require derivatives
- Gets stuck at local minima
- Quick where they work

Genetic algorithms

- Substitute in the function
- Can yield global minima
- Time consuming



Searching for optimal solution

- Enumerative search (search all possibilities): Works for very small data sets
- Random search: Works in cases where fluctuations are not much
- Calculus/Graphical driven search: Works well, fast but gets stuck at local minima
- Evolutionary search



HYDERABAD

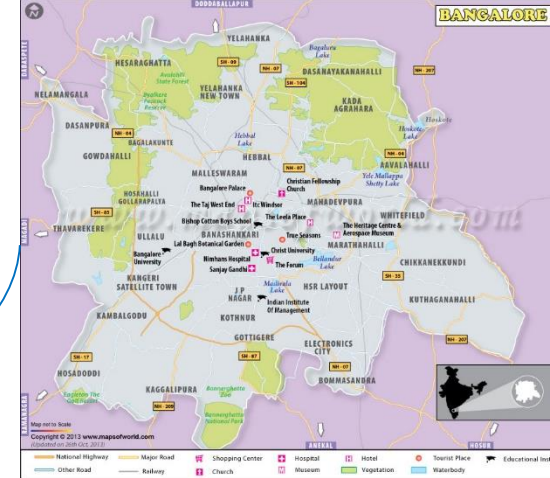
Office and Classrooms

Plot 63/A, Floors 1&2, Road # 13, Film Nagar,
Jubilee Hills, Hyderabad - 500 033
+91-9701685511 (Individuals)
+91-9618483483 (Corporates)

Social Media

Web: <http://www.insofe.edu.in>
Facebook: <https://www.facebook.com/insofe>
Twitter: <https://twitter.com/Insofeedu>
YouTube: <http://www.youtube.com/InsofeVideos>
SlideShare: <http://www.slideshare.net/INSOFE>
LinkedIn: <http://www.linkedin.com/company/international-school-of-engineering>

This presentation may contain references to findings of various reports available in the public domain. INSOFE makes no representation as to their accuracy or that the organization subscribes to those findings.



BENGALURU

Office

Incubex, #728, Grace Platina, 4th Floor, CMH Road,
Indira Nagar, 1st Stage, Bengaluru – 560038
+91-9502334561 (Individuals)
+91-9502799088 (Corporates)

Classroom

KnowledgeHut Solutions Pvt. Ltd., Reliable Plaza,
Jakkasandra Main Road, Teacher's Colony, 14th Main
Road, Sector – 5, HSR Layout, Bengaluru - 560102