innovate achieve lead

BITS Pilani

Machine Learning (IS ZC464) Session 12:

Genetic Algorithms

Evolutionary Algorithms

- Nature inspired
- Used for optimization
- Use a heuristic
- Known as meta heuristic algorithms
- Major techniques – Genetic Algorithm, Particle Swarm Optimization
- Other recent techniques – Firefly algorithm, cuckoo search, bat search etc.

Applications

- Game problems
- Function approximation
- Weight learning
- Unsupervised clustering
- Feature selection etc.

Genetic Algorithm

- Population based algorithm – search begins with initial population or collection of hypotheses.
- Basic unit : chromosome (Name nature inspired) – a bit string representing hypothesis
- Current population of chromosomes produces better offsprings
- The solution of the given problem refers to the maximum fitness value corresponding to a chromosome evolved after a number of generations.

How to use GA in problem solving

- Problem is represented as a search problem.
- The parameters of the search space represent the dimensions of the hypothetical search space.
- An instance of all the parameters is represented as a chromosome.
- Example :

101100011

010111001

1.2 2.3 3.4 5.4 2.9 1.0

Chromosome

- The representation of chromosome is problem specific.
- It represents the parameter instances or problem solution if optimal.
- Corresponding to each chromosome is an associated fitness value.
- The chromosomes are said to be fit if they represent solution close to the optimal solution.
- Chromosomes evolve over the generations.
- Each generation consists of obtaining new chromosome from two fit parent chromosomes by applying crossover operation

When the training data is large, then we can search for patterns

- Let us view the bit strings as points in a n-dimensional space. A bit string is said to have an associates value called fitness value. Examples

  Fitness value of bit string 100 11 01 = 1

  Fitness value of bit string 010 11 10 = 0

- We may be interested in a decision problem to know a combination of attribute values and instead of making a DT, we explore the combinations with appropriate fitness values.

Finding good combinations using Genetic Algorithms

- The bit strings can be represented as a whole in the form

  1001101

  0101111

  0010111

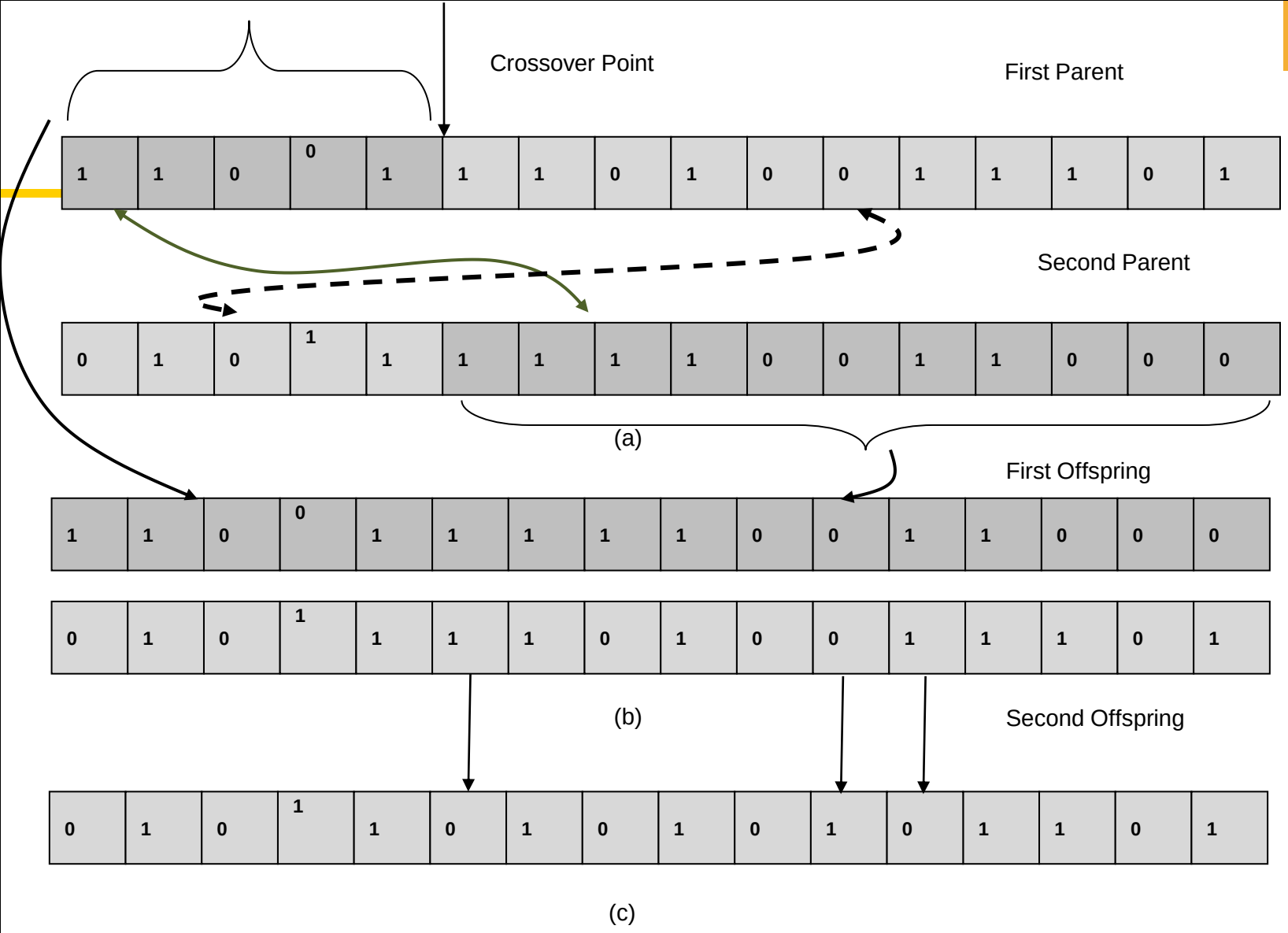
- If the problem is viewed as an optimization problem, then we search for a string which is the best combination.
- In the above example, we just have two values defining the goodness of a combination.

Fitness Values

- Game : How close are we to the goal
- Pattern Recognition: How large is the accuracy of recognition
- Path finding: How less costly is the path.
- And so on…

Genetic Algorithm : Operators

- Two operators known as crossover and mutation are applied to the chromosomes to obtain new generation chromosome.
- **Crossover Operator:** applied on two parent chromosomes to obtain a new offspring.
- **Mutation:** applied on one chromosome to exploit the neighborhood.

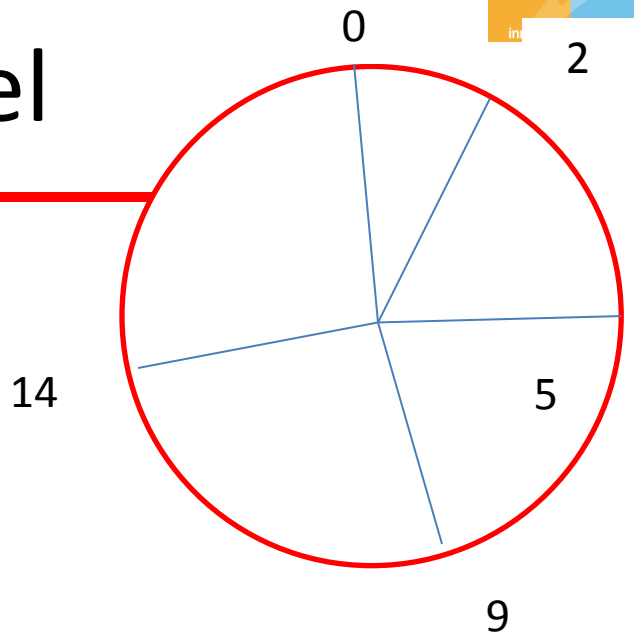


Crossover and Mutation (a) Previous Generation Parent Chromosomes selected for Crossover (b) Two new off springs generated by crossover using one crossover point. (c) Mutation applied at three places in second offspring

Fittest Parents: Exploration

- **Roulette wheel** is one of the commonly used methods to select the chromosomes to be treated as **fittest parents from the pool of fit parent.**
- The fitness values of all eligible chromosomes (with fitness greater than a **threshold**) are added and a random number decides the fittest parent because of its high favorable chances due to its individual large fitness value

Roulette Wheel



chromosome	fitness	cumulative	probability	Cumulative prob.
1	2	2	2/20 = 0.1	0.1
2	3	5	3/20 =0.15	0.25
3	4	9	4/20 = 0.2	0.45
4	5	14	5/20 = 0.25	0.7
5	6	20	6/20 = 0.3	1.0

Generate a random number in [0,1] and select a fit parent according to its probability