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

Representing Hypotheses

- Consider an attribute 'outlook' which can take on any of the three values – *Sunny, overcast, rain*
- Another attribute say 'wind' with two values – *strong or weak*
- Suppose the rule is

  If Wind = strong then playTennis = yes

- *Having no constraint in this rule on outlook, we can represent the above rules as*

  111 10 10

Respectively for outlook values, wind values and playtennis values

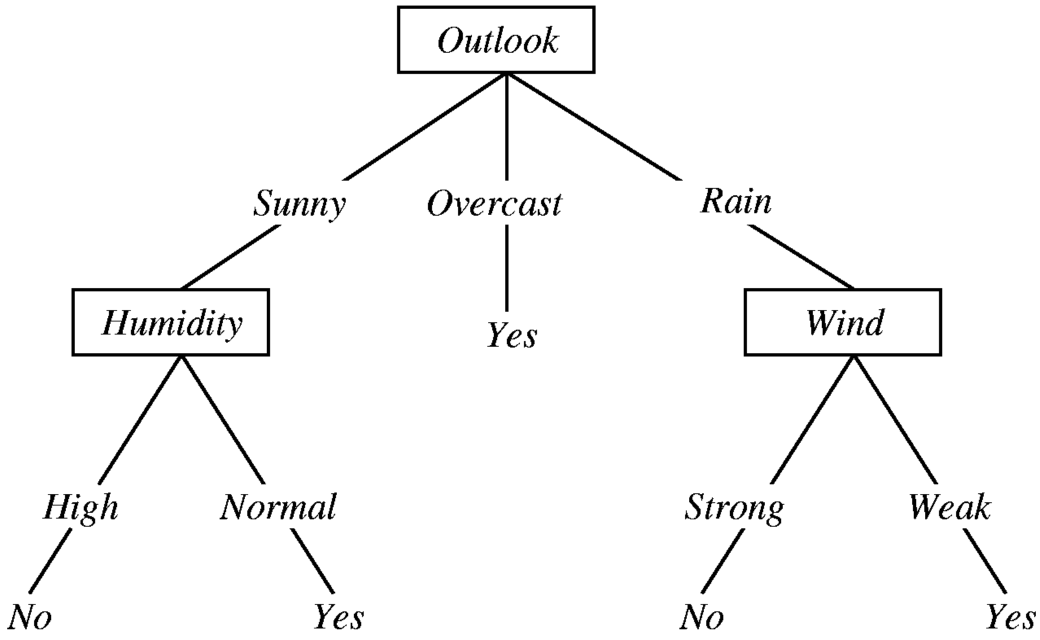
Representing rules as bit strings

- If (outlook = overcast ∨ Rain)
- if (Wind = Weak)
- Humidity = normal
- The bit string representation is

  011 01 01

- In a decision tree we can mark the leaf nodes with decision yes or no. A path from the root node to the leaf tells the combination of attribute values.

Decision Tree



There are three decisions with yes and two with no.

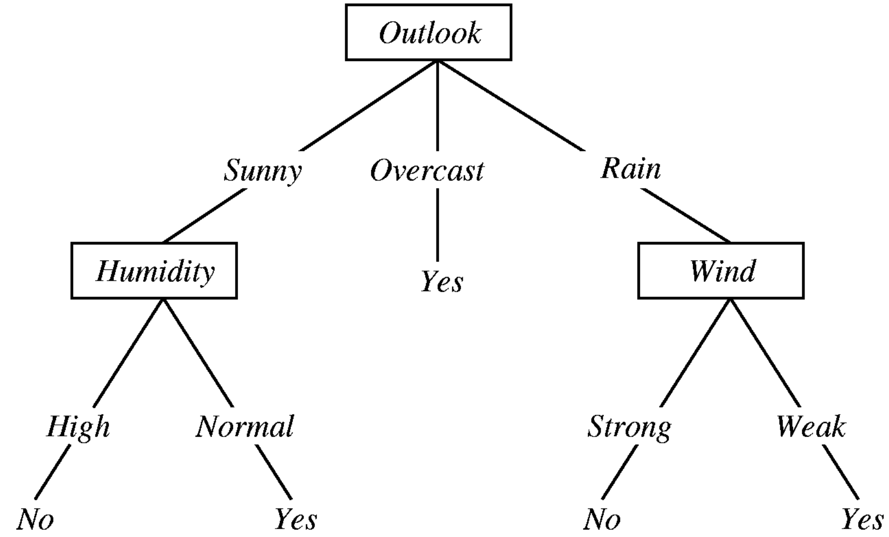
Representing DT as a search problem

- Find the combination of the attribute values for which the decision key is ‘yes’ (or no) in the given training data.
- All the hypothesis representing ‘yes’ are represented as bit strings as follows

    100 11 01

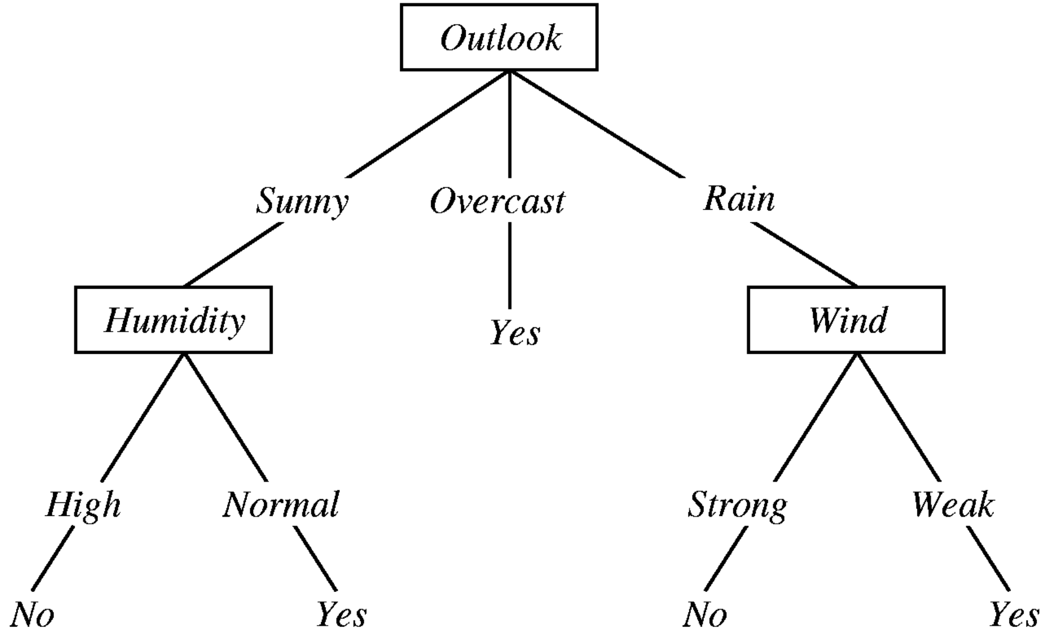
    010 11 11

    001 01 11



Class Assignment

- Represent the bit strings for ‘no’ decisions



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

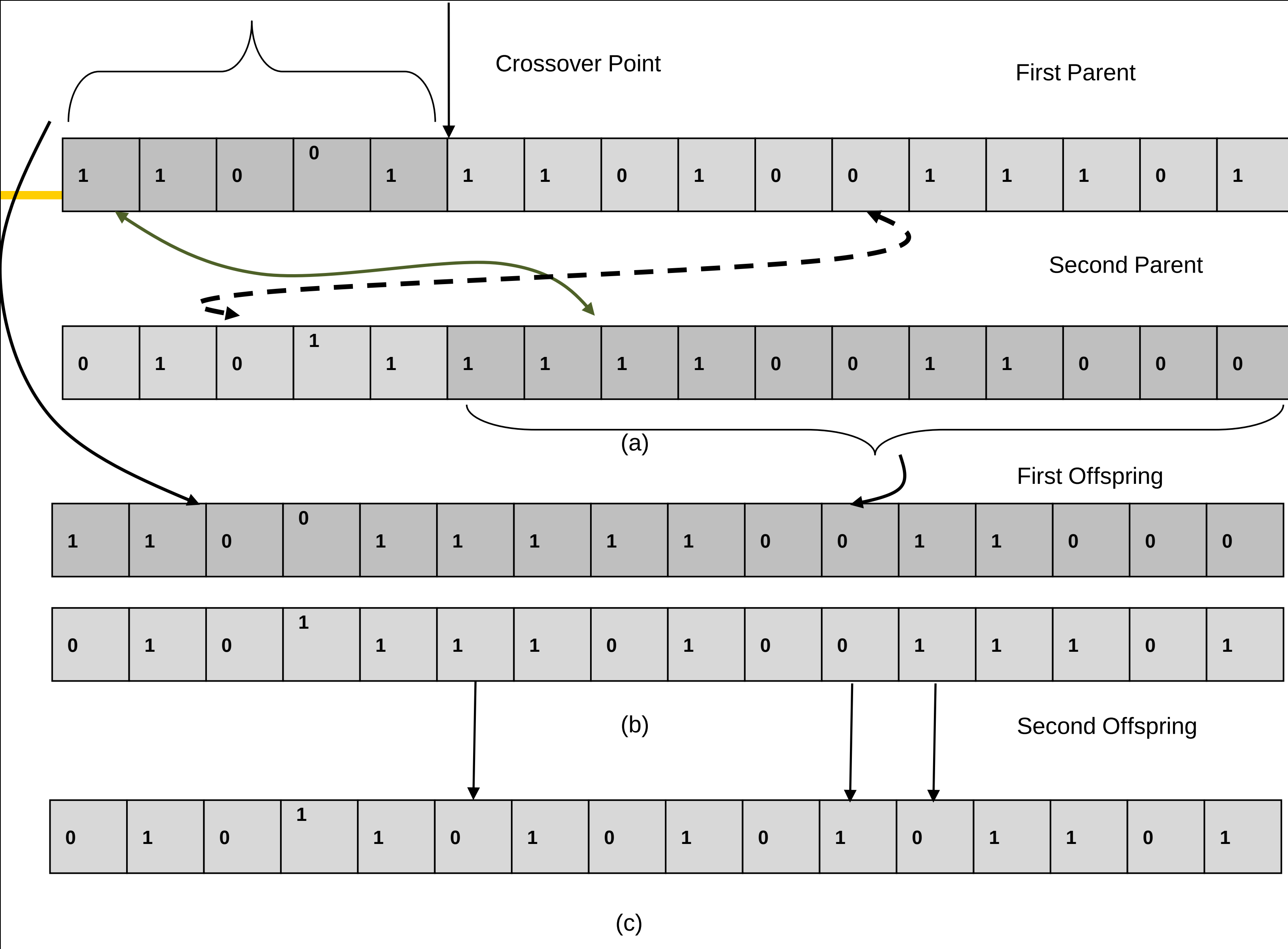
- If the problem is viewed as an optimization problem, then we search for a string which is the best combination.
- In the above example, we just have two values defining the goodness of a combination.

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.

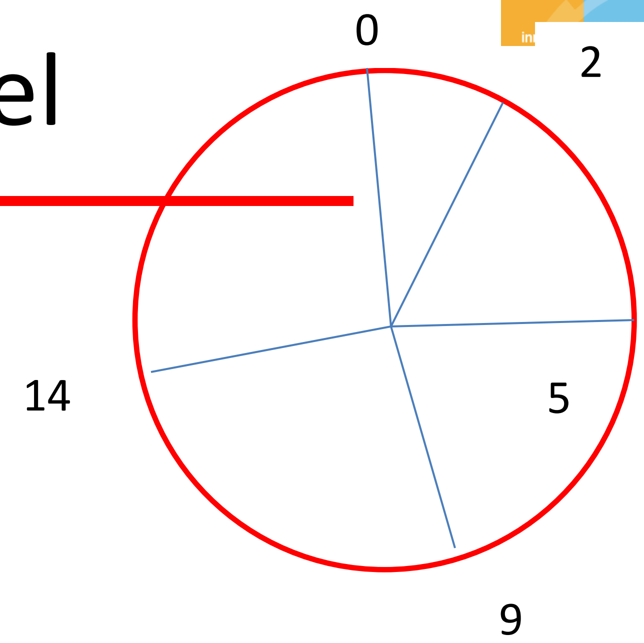


Crossover and Mutation (a) Previous Generation Parent Chromosomes selected for Crossover (b) Two new off springs generated by crossover using one crossover point (c) Mutation applied at three places in second offspring

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

Roulette wheel

```
sum = 0.0
for all members of population
        sum += fitness of this individual
end for
sum of probabilities = 0.0
for all members of population
        probability = sum of probabilities + (fitness / sum)
        sum of probabilities += probability
end for
loop until new population is full
        do this twice
                number = Random between 0 and 1
                for all members of population
                        if number > probability but less than next probability
                                then this member gets selected
                end for
        end
        create offspring
end loop
```

How many fit chromosomes should be selected

- If the number of chromosomes with fitness value greater than a specified threshold in generation t is p and if the crossover probability is p_c, then $p_c \times p$ chromosomes are selected as parents
- These participate pair wise in crossover to produce pair of offsprings of the new generation.

Mutation:Exploitation

- If the mutation probability is p_m then $p_m \times p$ chromosomes undergo the mutation process.

Crossover operation

- One point crossover
- Two point crossover
- Crossover mask
- Uniform crossover – combines bits sampled from the two parents

Feature Selection

- Given a set of features, select few most informative and discriminative features
- **Class assignment:**
  - Represent as a search problem
  - Represent as a chromosome
  - Define fitness function

23	29	10	2
5	30	12	25
52	16	9	2
15	11	12	13

(a)

c1	c2	c3	c4
c5	c6	c7	c8
c9	c10	c11	c12
c13	c14	c15	c16

(b)

c1	**c2**	**c3**	**c4**	**c5**	**c6**	**c7**	**c8**	**c9**	**c10**	**c11**	**c12**	**c13**	**c14**	**c15**	**c16**

(c)

23	**29**	**10**	**2**	**5**	**30**	**12**	**25**	**52**	**16**	**9**	**2**	**15**	**11**	**12**	**13**

(d)

1	**1**	**0**	**0**	**1**	**1**	**1**	**0**	**1**	**0**	**0**	**1**	**1**	**1**	**0**	**1**

(e)

0	**1**	**0**	**1**	**1**	**1**	**1**	**1**	**1**	**0**	**0**	**1**	**1**	**0**	**0**	**0**

(f)

Genetic Encoding of Face Image (a) 4×4 window of 16 features as the initial pool of features (b) coding of features (c) codes as a vector (d) Feature vector (e)and (f) Two different Chromosomes with 1's representing inclusion of the corresponding features

Comparison of the Evolutionary Algorithms

Characteristic	GA	PSO	FA
Population	Chromosomes	Particles	Fireflies
Search Heuristic	Survival of Fittest	Swarm Behavior	Attractiveness
Exploration Process	Crossover	Global Best	Attractiveness
Exploitation Process	Mutation	Local Best	Randomization
Modality	Single	Single	Multiple
Convergence	Slowest	Slow	Fast
Overall merit in finding optimal solution	Weak: May get trapped to Local Optima	Moderate: Only single optimal solution can be obtained	Strong: Reaches Global, Multimodal Optimal Solutions and is very fast