Evolutionary Search

Genetic Algorithm II

GENETIC ALGORITHM

5. Selection Algorithm

It is used to select parents from the current population to be mated with one another to create children

The selection is primarily based on the fitness. The better the fitness of a chromosome, the greater its chance of being selected to be a parent

The rate at which a selection algorithm selects individuals with above average fitness is *selective pressure*

If there is not enough *selective pressure*, the population will fail to converge upon a solution. If there is too much, the population may not have enough diversity and converge prematurely

GENETIC ALGORITHM

5. Selection Algorithm

Basically four types of selection algorithms:

- **random selection**
- **proportionate selection**
- **linear rank selection**
- **tournament selection**

GENETIC ALGORITHM

Selection algorithm

Random Selection

Individuals are selected randomly with no reference to fitness at all

All the individuals, good or bad, have an equal chance of being selected

GENETIC ALGORITHM

Selection Algorithm

Proportionate selection

Chromosomes are selected based on their fitness relative to the fitness of all other chromosomes

For this all the fitness are added to form a sum S and each chromosome is assigned a relative fitness (which is its fitness divided by the total fitness S)

A process similar to spinning a roulette wheel is adopted to choose a parent; the better a chromosome's relative fitness, the higher its chances of selection

Roulette Wheel Selection

- This can be simulated by following algorithm:
  - **[Sum]** Calculate sum **S** of all chromosome fitnesses in population
  - **[Select]** Generate random number **r** from interval (0,S)
  - **[Loop]** Go through the population and sum fitnesses from 0 - sum s
    - When the sum s is greater than r, stop and return the chromosome where you are

GENETIC ALGORITHM

Selection Algorithm

The selection of only the most fittest chromosomes may result in the loss of a correct gene value which may be present in a less fit member (and then the only chance of getting it back is by mutation)

One way to overcome this risk is to assign probability of selection to each chromosome based on its fitness

In this way even the less fit members have some chance of surviving into the next generation

GENETIC ALGORITHM

Selection Algorithm

The probability of selection of a chromosome "i" may be calculated as

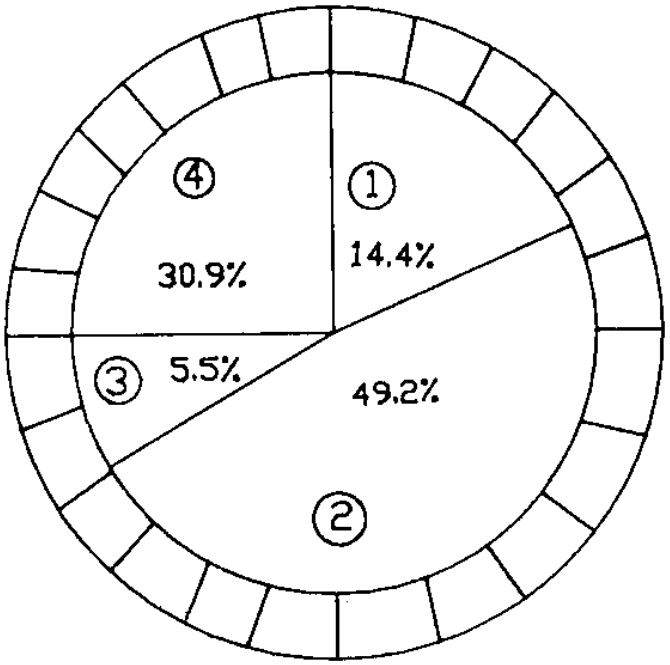
$$p_i = \text{fitness}_i / \sum_j \text{fitness}_j$$

Example

Chromosome	Fitness	Selection Probability
1	7	7/14
2	4	4/14
3	2	2/14
4	1	1/14

GENETIC ALGORITHM

Selection Algorithm



GENETIC ALGORITHM

5. Selection Algorithm

Advantage

Selective pressure varies with the distribution of fitness within a population. If there is a lot of fitness difference between the more fit and less fit chromosomes, then the selective pressure will be higher

Disadvantage

As the population converges upon a solution, the selective pressure decreases, which may hinder the GA to find better solutions

GENETIC ALGORITHM

5. Selection Algorithm

Linear rank selection

Even after selecting chromosomes by their probabilities there is a risk that a chromosome may have a vital gene but its fitness may be zero and hence its selection probability will be zero

The linear rank method assigns a small non-zero probability to even those chromosome which have a zero fitness

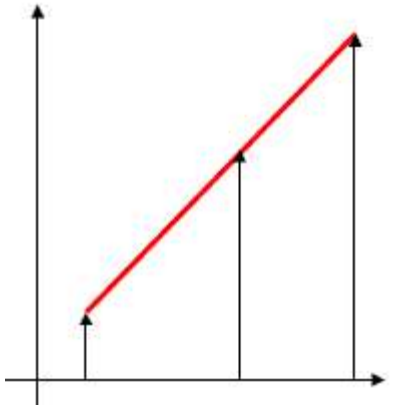
GENETIC ALGORITHM

5. Selection Algorithm

Linear rank selection

The population is sorted from best to worst according to the fitness

Each chromosome is then assigned a new fitness based on a *linear* ranking function



New Fitness = (P – r) + 1

where P = population size, r = fitness rank of the chromosome

If P = 11, then a chromosome of rank 1 will have a New Fitness of 10 + 1 = 11 & a chromosome of rank 6 will have 6

GENETIC ALGORITHM

5. Selection Algorithm

Linear rank selection

A user adjusted slope can also be incorporated

$$\text{New Fitness} = (P - r)\,(\max - \min)/(P - 1) + \min$$

where max and min are set by the user to determine the slope (max - min)/(P – 1) of the function

Let P = 11, max = 8, min = 3,
then a chromosome of rank 1 will have a New fitness of 10*5/10 + 3 = 8
& a chromosome of rank 6 will have 5*5/10 + 3 = 5.5

GENETIC ALGORITHM

5. Selection Algorithm

Linear rank selection

Once the new fitness is assigned, parents are selected by the same roulette wheel procedure used in proportionate selection

Advantage: selective pressure, once determined by the user, remains constant

Disadvantage:

- **Population must be sorted**
- **Chromosomes with the same fitness will not have the same probability of being selected**

GENETIC ALGORITHM

- Another way of linear rank selection is:

- The highest ranking chromosome has probability *p* of selection (e.g. 0.66)

- If the 1st chromosome fails to get selected, then it is taken out of competition and the second ranking chromosome gets the chance of selection with a probability *p* among the remaining chromosomes (hence its overall probability is 0.33 x 0.66 = 0.22)

- This process is continued for all the chromosome until one of them gets selected

GENETIC ALGORITHM

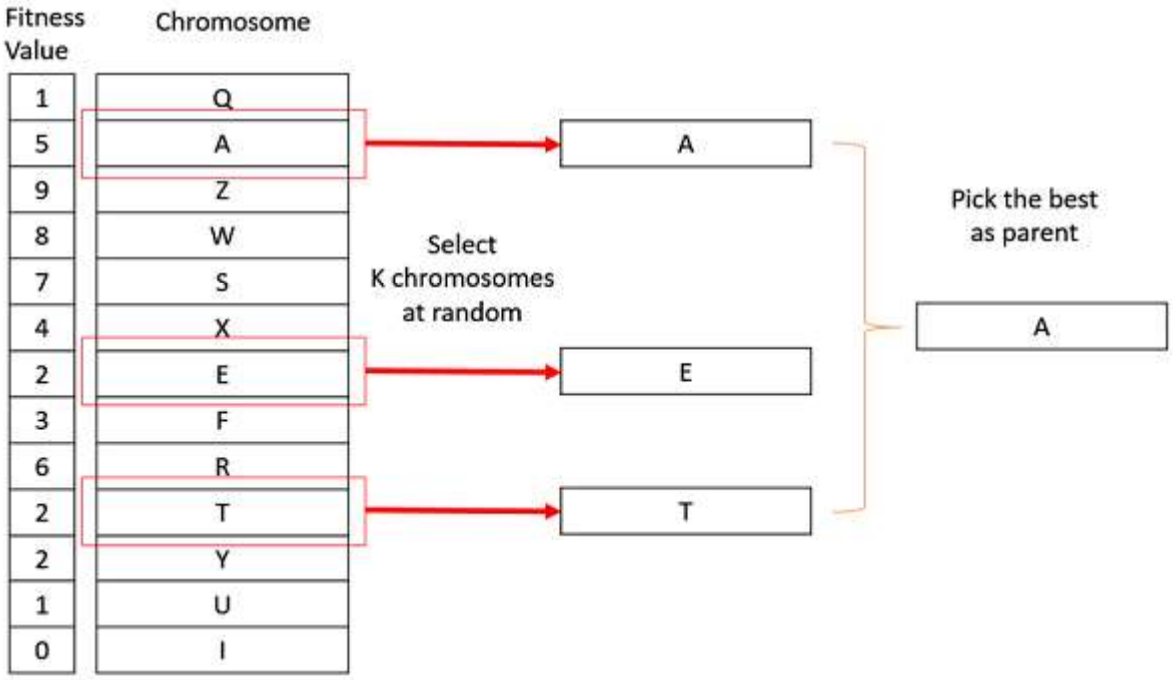
5. Selection Algorithm

Tournament selection

One parent is selected by comparing a subset *b* of the available chromosomes, and selecting the fittest; a second parent may be selected by repeating the process

The selection pressure increases as *b* increases

Value of *b* = 2 is most commonly used

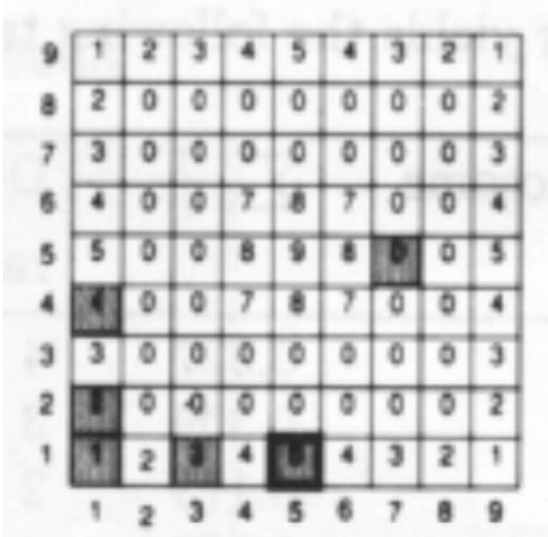

Fitness Value
Chromosome
1
Q
5
A
9
Z
8
W
7
S
4
X
2
E
3
F
6
R
2
T
2
Y
1
U
0
I
Select K chromosomes at random
A
E
T
Pick the best as parent
A

GENETIC ALGORITHM

Survival of the most diverse

If a chromosome is close to the best chromosome and has a higher score than a chromosome that is far away (diverse) but has a slightly lower score, then the far away chromosome will be wiped out

This results in uniformity, even in large populations



GENETIC ALGORITHM

Survival of the most diverse

Diversity of a chromosome from the already selected chromosomes can be calculated as

- **Take the distance of the chromosome from a selected chromosome**
- **Square the distance (to remove negative distance)**
- **Take the sum of all squared distances**
- **The larger this sum the more diverse is the chromosome**

$$\text{Diversity} = \sum_i \text{distances}_j^2$$

One way to incorporate diversity is to rank the chromosomes by the sum of quality rank and diversity rank

GENETIC ALGORITHM

Survival of the most diverse

Hence diversity (or being different) and can be incorporated into the selection procedure and a chromosome which is almost as good as another chromosome but is far away from the already selected chromosomes has a better chance of survival

GENETIC ALGORITHM

6. Generation Gap

It is a real number between zero and one that represents the fraction of the current population that gets replaced by the offspring

If it is 1.0, all of the current chromosomes are replaced. This implies that for a population P we would have to create P children

For a generation gap value between 0.0 and 1.0, it is necessary to create that many children, e.g. for a gap of 0.5 we would need to create 0.5*P children.

GENETIC ALGORITHM

6. Generation Gap

Another issue is how to determine which individuals in the current population die. The most common strategy is to replace the worst individuals of a population

We may have a generation gap which is not fixed. In this method, the fittest P chromosomes will be selected from the set of current population plus new children, and will form the new generation

GENETIC ALGORITHM

7. Amount of Elitism used

A real value between 0 and 1, which represents the fraction of the best individuals of a population that will not get selected to die

Example P = 20 and elitism = 0.1, then 2 best individuals of current population do not get replaced

It is used only if generation gap is not fixed.

If generation gap of 0.9 is used that means that 18 old chromosomes will definitely get replaced.

If elitism of 0.1 is used that means that any number of old chromosomes between 0 and 18 may get replaced

In genetic algorithm, a generation has 50 chromosomes, out of which 5 fittest ones are placed in the next generation without any competition from the offspring. The rest of the 45 chromosomes of the next generation are chosen from the pool of parents and offspring. Suppose 30 chromosomes of the offspring make it to the new generation. What is the generation gap and what is the elitism value. Duplicates are not allowed.

GENETIC ALGORITHM

8. Number of Duplicates Allowed

Chromosomes that represent the same candidate solution

If they are allowed then that chromosome has higher probability of producing an offspring, and may probably create many offspring

Eliminating duplicates increases the efficiency of the genetic search and reduces the danger of premature convergence

Eliminating duplicates means that if an offspring is created which is a duplicate of a chromosome of the current population, we terminate it immediately and create a new one. It increases the processing time in large populations