Genetic Algorithm: The Happiness Pursuit (Knapsack Problem)

Tapadyuti Maiti | Simi Golbert Foss

**Problem:**

We have chosen the specific problem of finding the exact trail of activities list to achieve the maximum happiness factor through the day through Genetic Algorithm. Though this problem is essentially maps into Knapsack Problem, but we had faced few major roadblocks to convert it into solid code.

**Inspiration:**

There has been lot of work going on to lead a beautiful and happy, calm, poised life as a Human being. Our initial thought was to solve this problem through genetic algorithm as it follows how our gene gets mutated naturally to cope up with the current nature. We gave a try to solve this problem by finding out the most optimal list of activities along with its reward points associated with Happiness. Albeit, by doing these list of activity can’t guarantee to become Happy, but it’s a try is a worth.

We have found out list of 78 Activities and ranked them accordingly with its reward points associated with the approximate time taken to finish it. We accept the fact that these reward point can vary person to person.

**Implementation Detail:**

Flow of Data: Sequentially Downwards which wraps all the downwards entity.

**Generation:** It contains whole population.

**Population:** Consists of Many Colonies

**Colonies:** Constraint of containing a no more than specific number of Persons in it.

**Persons:** It contains a Chromosome

**Chromosome:** Defining Character of Person through List of Activity

**Set of Genes:** Represent Specific Activity Encoded by specific A-G-T-C Sequence which may or may not be same with another Person, as happens Naturally.

Following Steps were implemented sequentially to solve the problem through genetic algorithm:

1. Load Data from Excel Sheet and initialise Gene Symbol Table
2. Initilialise the Weighted Activity Graph
3. Make Generation 0 of 4 initial members
4. Do Evolution till 100.
5. Plot the result in JChart.

**Genetic Code:**

In our model each Person has a chromosome, which contains sequence of Gene. Each Person’s Activities total summation is equal to 16 hours. Each Activity from the excel list is encoded in terms of “A”,“G”,”C”,”T” sequence(length of 4) for representation purpose. Essentially each activity denotes a Gene in the chromosome. Here all the activities are represented as the nodes and edge weight is defined as the reward point of each activity. All the activities are inter- connected in the weighted edge graph with each other.

All the incoming edges to a node have the same weight as that of its node(activity)’s reward point.

**Gene Expression:**

It is the sequence of Each Activities encode in terms of “AGCT” whose length can vary person to person.

**Person 1:**

AAGG-ATGT-ATCT-ACTA-ATCT-AACT-AAAA-AATA-AAAG-CAAG-CAAT-AAAA-CACT-AGAC-ATTG-AATT-CATA-ATCA-ACGT-CACT-AACT-CATA-ACTA-ATAG-AAAG-AAGC-AGTT-AGCC-AAAG-ATGT-CAAA-CACG-AACG-ACTA-AATA-AAAG

Which is essentially coded in a format of list of activity:

Eating healthy food| health management| Eating Ice Cream |Smoking | Cycling |Cycling |Maintaining cleanliness| Non-Alcoholic Party |Watching Comedy shows | Interacting with friends |Non-Alcoholic Party | Donate Something. |Gratitude feelings |yoga |Visit to Salon |Getting News Update |Reading Magazine| health management Planning daily Activities |Singing |Maintaining cleanliness |Going out for a coffee |Gratitude feelings| Doing adventure sports|Singing|Calculating Budget | Read A Good Book. |Eating healthy food| Exposure to sunlight| Swimming |Decorating Home |Giving| Donate Something. |Smiling | Baking Cookies |Gratitude feelings| Getting News Update | Reading Magazine| Visiting Friend |Going out for a coffee |

Total Reward Point 116.89 || Total Time :960.0

**Fitness Function:**

It defined as the:

Fitness of a Person = \_\_\_\_\_\_\_\_\_\_\_\_ Sum of total reward points of a Person\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Maximum Reward point that can be achieved (960)

In our case, all the activity have reward points associated with range of 1-10 points and time taken to complete each activity in span of 5 – 60 minutes. So, in span of 960 minutes, one achieves maximum of 960 reward point, which is highly impossible case in experiment.

**Evolution Process:**

In each Generation the population doubles up. As we are starting from 4 individual we are restricting the cull process until it reaches a certain number of required of population (in our case -500).

Each Colony can contain only limited number of person i.e. 100.

Breeding Process happens in each Colony (Intra – Colony Breeding) and at the end of Breeding Process, if the total population crosses 500, we select only the fittest one with cull ratio of 0.5.

In the end, after sorting the whole population based on there fitness score in Minimum Priority Queue, we randomize the whole population and put it into the colonies back again. We have used Fisher-Yates randomization process in our case.

**Mutation and Cross Over Process:**

Crossover process is done .1% in each generation’s breeding process. It been done by cutting the Parents Gene in two Parts and replacing the second half from another Parent’s gene.

We have constructed the Chromosome precisely to make half of its gene sequences’s summation of activity time equals exactly 960/2 i.e 480 minutes(16 hours = 960 minutes).

Parent1:

|  |  |
| --- | --- |
| AAGG-ATGT-ATCT-ACTA-ATCT- AAAA-AATA-AAAG | AAGC-AGTT-AGCC-AAAG-ATGT-CAAA-CACG |

Parent 2:

|  |  |
| --- | --- |
| CATC-ACAT-ATAG-CAAC-ATCT-ACGT-AAAA-CACG | ATCC-AAAA-ACTA-AGAC-AATA-AATA-CATA-AATC |

Child 1

|  |  |
| --- | --- |
| AAGG-ATGT-ATCT-ACTA-ATCT- AAAA-AATA-AAAG | ATCC-AAAA-ACTA-AGAC-AATA-AATA-CATA-AATC |

Child2:

|  |  |
| --- | --- |
| CATC-ACAT-ATAG-CAAC-ATCT-ACGT-AAAA-CACG | AAGC-AGTT-AGCC-AAAG-ATGT-CAAA-CACG |

While half of the remaining 90% of breeding is mutated in the following manner:

1.Selecting random 10 positions of Array List of Genes and then mutating it with some other activity of same time span from the Gene Pool table (where all the Genes are present for reference)

Parent:

|  |
| --- |
| CATC-**ACAT**-ATAG-CAAC-**ATCT**-ACGT-AAAA-CACG-AGGA-**AAAG**-AAAA-ATCC-AAAA-**ACTA**-AGAC-AATA |

Child:

|  |
| --- |
| CATC-**AGTC**-ATAG-CAAC-**AAAY**-ACGT-AAAA-CACG-AGGA-**ATTG**-AAAA-ATCC-AAAA-**GCTA**-AGAC-AATA |

For illustration I have mutated only 4 out all the genes of a person.

In my experiment I have range of varying number of genes 40-50 , that amounts ¼ to be point mutation.

2.Other Half of the breeding population are simply copied from the previous generation.

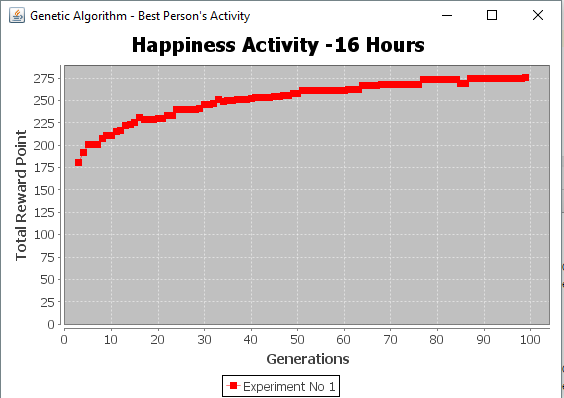
**Findings:**

We have ran our GA to check how our results are coming up , below are given data of 5 experiment associated with its reward point.

|  |  |  |
| --- | --- | --- |
| Experiment | Fitness Score | Total Reward Point |
| 1 | 0.287083333 | 275.59 |
| 2 | 0.306302083 | 294.04 |
| 3 | 0.3471875 | 333.3 |
| 4 | 0.373489583 | 358.549 |
| 5 | 0.400885417 | 384.85 |

**Sequence of activity of Exp5**

Smelling fresh flowers|Cuddling|Exposure to sunlight|Listening to music|Giving|Gratitude feelings|Exposure to fresh air (nature)|Snuggle With Your Dog/Pet Care|Playing music|Photography|Cuddling|Cutting daily travel time|Donate Something. |Smiling|Gratitude feelings|Listening to music|Smelling fresh flowers|Gratitude feelings|yoga|Cuddling|yoga|Meditation|Smelling fresh flowers|Giving|Exposure to fresh air (nature)|Smiling|Empathazing with peers|Gratitude feelings|Helping people to cross Road |Meditation|health management|Watching Comedy shows|yoga|Gratitude feelings|Cuddling|Smelling fresh flowers|Playing music|Dancing|Gratitude feelings|Listening to music|Watching Comedy shows|Gratitude feelings|Exposure to sunlight|Physical Exercise|Meditation|Meditation|Cuddling|Smelling fresh flowers|Smiling|Smelling fresh flowers|Donate Something. |Exposure to sunlight|

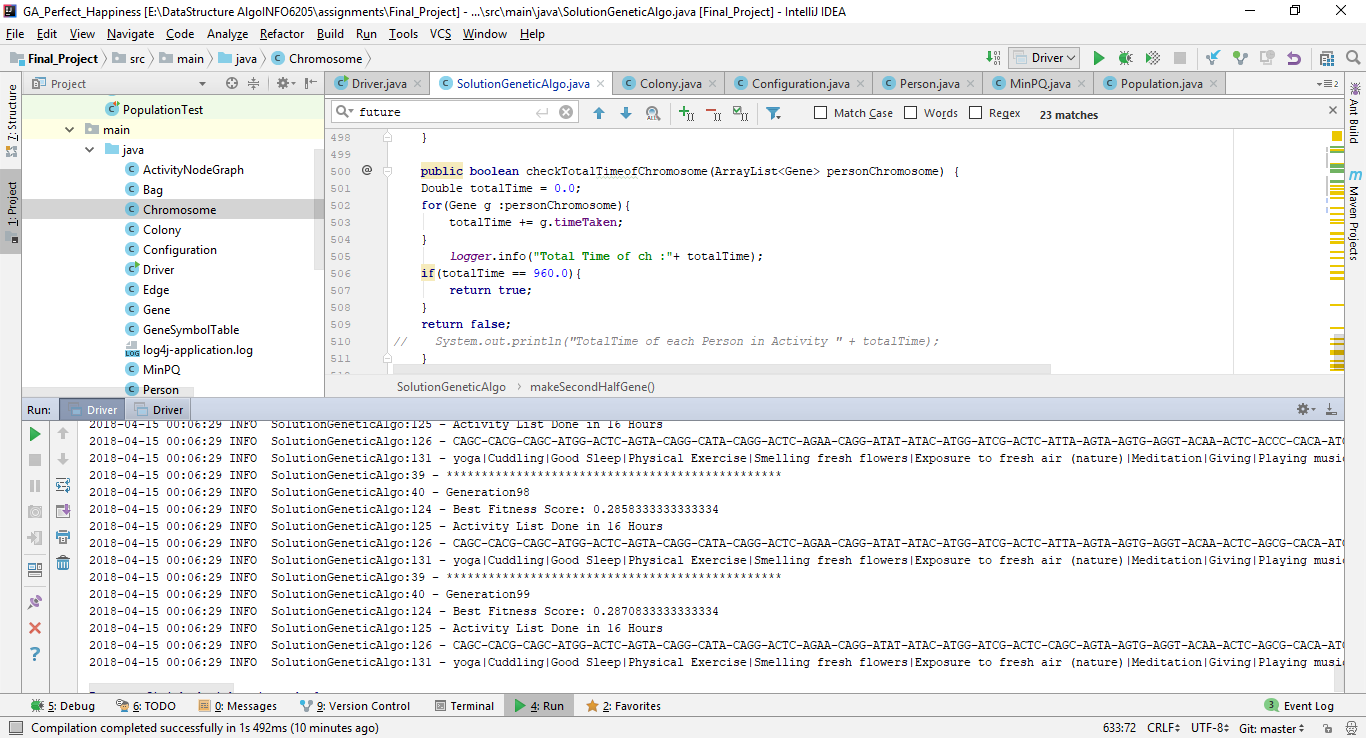


**Conclusion:**

As from the above graph we can see total reward points for happiness or our fitness function gets increasing as generations increases. We are culling with 0.5 ratio in each generation. After 100 generation we are getting approximately 300-400 reward points out which turns out to be an excellent solution as of activity listed out in our consideration. Thus, we can infer that Genetic Algorithm repeatedly modifies the population by mutation of individual solutions to get the best solution.

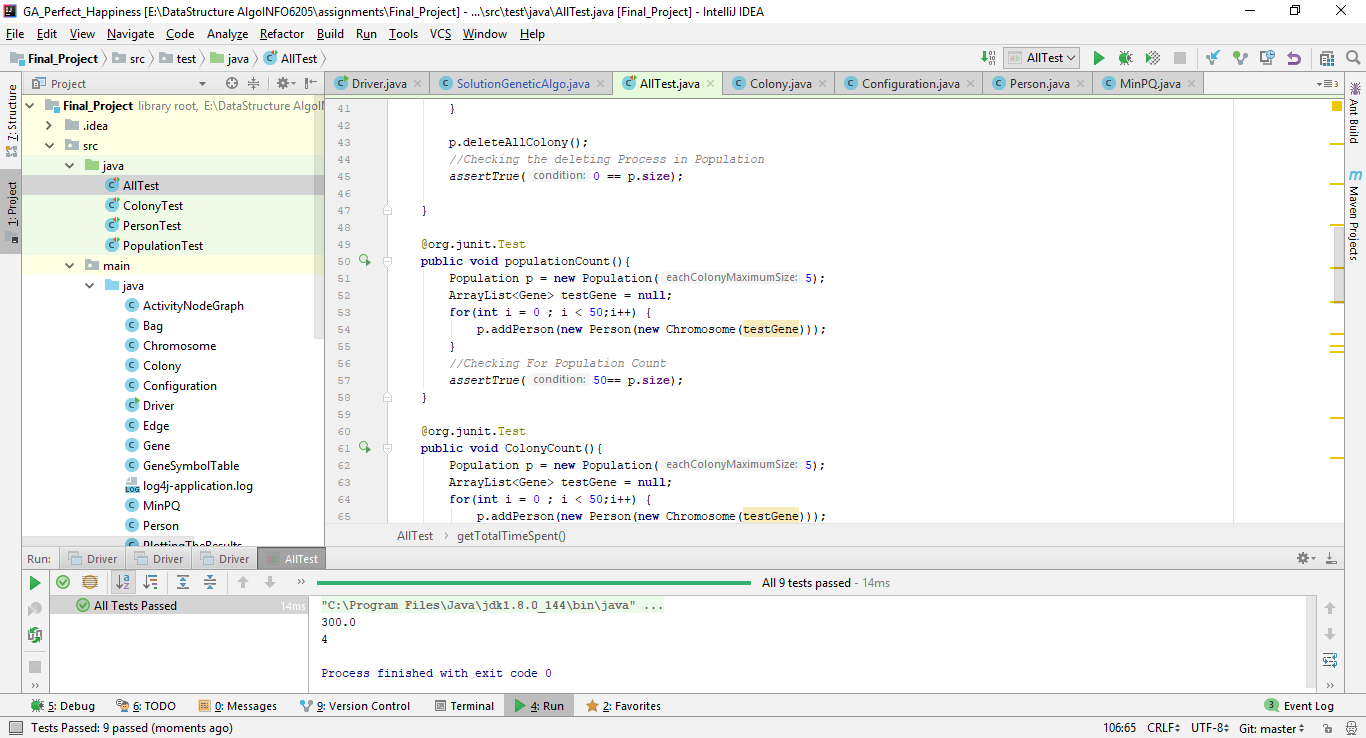
**Screenshot:**

**Running of Program:**



**Unit Tests:**

All the Test case are present in the ALL test file.



**Reference from Book:**

* Algorithms 4th Edition by Robert Sedgewick, Kevin Wayne

**Reference for Ranking the Activities:**

* happify.com
* https://www.hongkiat.com/blog/how-to-increase-happiness/
* <https://technologyadvice.com/blog/information-technology/activate-chemicals-gamify-happiness-nicole-lazzaro/>

--------END--------