

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

**TRUCKIN':
THE GENETIC ALGORITHM WAY**

JEFFREY EDELSTEIN

A THESIS

IN

THE DEPARTMENT

OF

COMPUTER SCIENCE

PRESENTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF COMPUTER SCIENCE AT
CONCORDIA UNIVERSITY
MONTREAL, QUEBEC, CANADA

APRIL 2001

© JEFFREY EDELSTEIN, 2001



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-59320-7

Canada

ABSTRACT

Truckin': The Genetic Algorithm Way

Jeffrey Edelstein

Over the past 25 years, a new form of optimization and search technique was refined using the same theories developed to explain human evolution, called the Genetic Algorithm. The algorithm is based on the premise that, when individuals exist in a competitive environment where resources are limited, only the fittest individuals will survive.

This thesis attempts to use genetic algorithms to solve a complex optimization problem, which involves a real world simulation where trucks compete to make a profit, by buying and selling commodities in a country filled with different types of dealers. The genetic algorithm is used to evolve the trucks over generations, by modifying the strategies that they use to control their behaviour, in an attempt to produce more profitable trucks.

The project was implemented using the C++ programming language. The purpose of the thesis and its implementation was to see if one could use object oriented techniques, such as inheritance and dynamic binding to achieve the genetic variation of the trucks.

The results of the project are moderately successful. This is likely due to the limited number of strategies available to the trucks. By increasing the number of strategies, the algorithm may produce trucks of superior genetic structure.

Acknowledgements

I would like to thank my supervisor Dr. Peter Grogono for his constant support, direction, and overwhelming patience, which enabled me to complete my thesis.

I would like to acknowledge the work of my colleague Debbie Papoulis who spent many long hours with me coding this project.

Finally, I would like to thank my loving wife and family for their support and understanding throughout my studies and I dedicate this thesis to them with love.

Table of Contents

LIST OF FIGURES.....	VIII
1. INTRODUCTION.....	1
2. BACKGROUND	3
2.1 DEFINITION	3
2.2 HISTORY	4
2.3 INTRODUCTION TO SIMPLE GENETIC ALGORITHMS	5
2.3.1 <i>Algorithm Components</i>	5
2.3.2 <i>Algorithm Component Definitions</i>	6
2.3.2.1 Encoding Mechanism.....	6
2.3.2.2 Fitness Function	7
2.3.2.3 Selection.....	8
2.3.2.4 Crossover	10
2.3.2.5 Mutation	11
2.3.2.6 Generation Cycle.....	12
2.3.2.7 Termination Condition	13
2.3.3 <i>Algorithm Pseudocode</i>	13
2.4 POTENTIAL PROBLEMS WITH GENETIC ALGORITHMS	14
2.5 APPLICATIONS OF GENETIC ALGORITHMS	15
3. OBJECT ORIENTED TRUCKIN'	17

3.1	INTRODUCTION	17
3.2	COMPONENTS.....	17
3.2.1	<i>The Country</i>	18
3.2.2	<i>Commodities</i>	18
3.2.3	<i>Trucks</i>	18
3.2.4	<i>Controllers</i>	19
3.2.5	<i>Dealers</i>	20
3.2.5.1	Producers.....	20
3.2.5.2	Retailers	21
3.2.5.3	Consumers.....	21
3.2.5.4	Gas Stations.....	22
3.2.6	<i>Managers</i>	22
3.3	THE SIMULATION	22
4.	DESIGN	24
4.1	COMPONENTS.....	25
4.1.1	<i>The Country</i>	25
4.1.2	<i>Trucks</i>	26
4.1.2.1	Initialize	27
4.1.2.2	Move	27
4.1.2.3	Go.....	28
4.1.2.4	Gas	28
4.1.2.5	Trade	30
4.1.2.6	Deal	31
4.1.2.7	Buy.....	33
4.1.2.8	Sell	34
4.1.3	<i>Controllers</i>	34
4.1.4	<i>Dealers</i>	35
4.1.5	<i>Managers</i>	35

5.	OOT: THE GENETIC ALGORITHM	36
5.1	ENCODING MECHANISM.....	36
5.2	FITNESS FUNCTION.....	37
5.3	SELECTION	38
5.4	CROSSOVER	38
5.5	MUTATION	38
5.6	GENERATION CYCLE	39
5.7	TERMINATION CONDITION.....	39
6.	IMPLEMENTATION	40
7.	RESULTS	45
8.	CONCLUSIONS	51
9.	REFERENCES.....	53
10.	BIBLIOGRAPHY	54
11.	APPENDIX A: CODE SAMPLES.....	55
11.1	POINTERS	55
11.2	ARRAYS	56
11.3	INHERITANCE AND VIRTUAL FUNCTIONS.....	59
11.4	OPERATOR OVERLOADING	60
12.	APPENDIX B: RAW RESULTS	61

List of Figures

FIGURE 1: ENCODING MECHANISM EXAMPLE	7
FIGURE 2: FITNESS FUNCTION EXAMPLE	8
FIGURE 3: FITNESS & SELECTION EXAMPLE.....	9
FIGURE 4: POPULATION AFTER SELECTION EXAMPLE	9
FIGURE 5: CROSSOVER EXAMPLE.....	11
FIGURE 6: MUTATION EXAMPLE.....	12
FIGURE 7: GENERATION CYCLE EXAMPLE	13
FIGURE 8: LOCAL MAXIMUMS.....	15
FIGURE 9: OOT PROJECT DESIGN	25
FIGURE 10: SUMMARY OF GENES.....	37
FIGURE 11: OOT PROJECT DESIGN (IMPLEMENTATION PHASE 1)	41
FIGURE 12: OOT PROJECT DESIGN (IMPLEMENTATION PHASE 2)	42
FIGURE 13: OOT PROJECT DESIGN (IMPLEMENTATION PHASE 3)	43
FIGURE 14: SIMULATION RESULTS	48

1. Introduction

“Genetic algorithms have emerged as practical, robust optimization and search methods.”¹

Over the past 25 years, a new form of optimization and search technique was refined using the same theories developed to explain human evolution. The algorithm is based on the premise that, when individuals exist in a competitive environment where resources are limited, only the fittest individuals will survive and reproduce.

My research of genetic algorithms involves a real world simulation where trucks compete to earn money, by buying and selling products in a simulated market. Trucks are controlled by their genetic makeup. The simulation is run several times, where each run is considered a new generation. Between each generation, a genetic algorithm is used to find the “fittest” trucks and the algorithm then uses crossover to create new trucks (the term “crossover” and other technical terms are explained in section two). The purpose is to see whether the algorithm can be used to solve the optimization problem of finding the most profitable truck from a finite, although large, search space. The hypothesis is that, future generations will have superior genetic structure, which will allow them to be more profitable.

The rest of this thesis is structured in the following manner: **Section 2** provides a background on the principles of human evolution and how it was first used as a model to solve complex problems in computer science. The section also includes a brief description of the Simple Genetic Algorithm. **Section 3** contains a detailed description of the Object Oriented Truckin' (OOT) project and its components, which is used to demonstrate the theories of this thesis. **Section 4** describes the design of the project using the object diagram of the OOT project as its basis. **Section 5** includes the analogy between the OOT project and the different steps of the Simple Genetic Algorithm, as well as a portion of its implementation, while **Section 6** provides additional implementation strategies for the project. **Section 7** discusses the results of the simulation, and finally **Section 8** presents the conclusions of this thesis.

2. Background

2.1 Definition

Charles Darwin presented one of the many theories on the evolution of mankind in his 1859 book entitled “The Origin of Species”. In his book, Darwin stated that all living species have evolved over thousands of centuries. In the case of the human species, man has evolved from the prehistoric Australopithecus through the Homo erectus to the Homo sapiens. We now know that this evolution is achieved by means of genetic evolution, which Darwin called “natural selection”.²

For centuries we have known that when organisms reproduce, the new generation resembles its parents. The transferal of information from a parent to its offspring is called *inheritance*. More recently, scientists have discovered that genetic information is transferred by means of chromosomes within the cell nucleus. These chromosomes contain thousands of genes, which carry the basic units of information, which will affect the characteristics of the organism.³

Thomas Malthus, an English clergyman and devotee of the natural sciences, was a great inspiration to Darwin. Malthus stated that all species are capable of reproducing

offspring faster than the food supply can increase. Darwin used this statement to explain that since more individuals are born than can possibly survive, there is a fierce struggle for existence. Those with favourable size, strength, running ability, or other characteristics necessary for survival will possess an advantage over others. According to Darwin's theory, as a population ages, fitter individuals are more likely to reproduce than less fit individuals. Over time, the average fitness level of the population increases.⁴

Based on an analogy with the theories of genetics and natural selection, genetic algorithms can be used to solve complex optimization problems. The analogy is made between the genetic structure of a living species and the solution search space to a complex problem. All living species have a chromosome. Breaking down the chromosome into smaller pieces, called *genes*, reveals the individual traits of that species. As in species, every possible solution to the complex problem can be mapped to a unique binary string viewed as a chromosome. Breaking down these strings into bits reveals the individual traits of the possible solution, similar to a gene.

2.2 History

John Holland first proposed the notion of Genetic Algorithms (GAs) in 1975. Holland's goals were to explain the adaptive process of natural systems and to design computing systems embodying their important mechanisms. The study of GA's has grown tremendously over the past 26 years due to the work by Holland, Goldberg, DeJong, Grefenstette, Davis, Muhlenbein and others.⁵

2.3 Introduction to Simple Genetic Algorithms

The Simple Genetic Algorithm, created by Holland, uses binary strings to represent the traits of a population of programs. Each string, encoded using 0's and 1's, denotes a member of the population. Each member of the population is evaluated using a set of criteria or function, usually called a fitness function. Using the values returned by the fitness function, the algorithm can determine which programs are better. Applying genetic operators, such as selection, crossover and mutation, creates the next generation of strings in the population. Until a termination condition is met, these operators are continuously exercised.⁶

2.3.1 Algorithm Components

GA's use a simple iteration to perform the task of solving complex problems⁷:

- 1) Random creation of binary strings to form the initial population of genes
- 2) Evaluation of the string resulting in its fitness
- 3) Selection of the best strings for possible reproduction
- 4) Manipulation of the strings to create the next population

These four steps can be further broken down into seven basic steps to fully explain how the algorithm works⁸:

- 1) Encoding Mechanism
- 2) Fitness Function

- 3) Selection
- 4) Crossover
- 5) Mutation
- 6) Generation Cycle
- 7) Termination Condition

These seven steps will be described in more detail in the following section. To help explain these steps, a simple optimization problem will be used for illustrative purposes.

2.3.2 Algorithm Component Definitions

2.3.2.1 Encoding Mechanism

The encoding mechanism is used to map potential solutions of the optimization problem to binary strings. The encoding is chosen in such a way that every distinct potential solution maps to a distinct string. Given a string, the corresponding potential solution can be reconstructed.

The example will focus on integers between the values 0 and 255. To encode each of the 256 possible integers into binary strings, these strings will need to be eight bits in length. The length is important so that each integer can be mapped, using its binary value, to a unique string. An initial population is generated by randomly selecting binary strings of length eight. The table below shows the values of a population of size four.

<i>Label</i>	<i>Mapping</i>	<i>Potential Solution</i>
A	00001000	8
B	10010001	145
C	11100100	228
D	00010101	21

Figure 1: Encoding Mechanism Example

2.3.2.2 Fitness Function

The fitness function indicates the effectiveness of the potential solution being evaluated to solve the optimization problem. Each encoded solution to the problem is evaluated to see how well it solves the problem. Based on its results in comparison to other solutions, the solution's strengths and weaknesses can be determined. These strengths and weaknesses are referred to as the solution's fitness. Once each string's fitness has been calculated, the selection mechanism determines which solutions will remain in the system and which solutions will be omitted.

We assume for our example that the fitness function is defined as the number of 1's found in the encoded string of each member of the population. If the objective is to have the most number of 1's as possible, then a string whose fitness is 3 is less likely to make it into the next generation than a string whose fitness is evaluated as 6. Using the table in the previous section, here are the fitness values evaluated for each member of the population.

<i>Label</i>	<i>Mapping</i>	<i>Fitness</i>
A	00001000	1
B	10010001	3
C	11100100	4
D	00010101	3

Figure 2: Fitness Function Example

2.3.2.3 Selection

The basic purpose of the selection mechanism is to model nature's selection process. The weak perish while the strong survive. This phenomenon has been coined "survival of the fittest". Strings with higher fitness function values will persevere, while others will perish. Survival is dictated by having a higher number of offspring in the ensuing generation.

In Simple Genetic Algorithms, the proportionate selection scheme is used. The number of offspring of a particular string is calculated with respect to the average fitness value of the generation to which the string belongs. A string with a higher than average fitness value will produce more than one offspring in the next generation, while a string with a lower than average fitness value will produce less than one offspring in the next generation.⁹

Continuing our example, the average fitness of the initial population is $\frac{1}{4}(1+3+4+3)=2.75$. Using a proportionate selection scheme, each string is allocated a number of offspring equal to its fitness divided by the average fitness of the population (f/f' , where f is the fitness of an individual member of the population and f' is the average fitness of all the members of the population). The table below shows these proportions. However, it is difficult to produce a fraction of an offspring. Due to this difficulty, the algorithm rounds these values off to whole numbers. Notice how the strings with greater fitness have a greater chance for survival.

<i>Label</i>	<i>Fitness</i>	<i>Calculated Offspring (f/f')</i>	<i>Actual Offspring in next generation</i>
A	1	0.36	0
B	3	1.09	1
C	4	1.45	2
D	3	1.09	1

Figure 3: Fitness & Selection Example

The following table indicates the population of strings after selection is complete.

<i>Label</i>	<i>Mapping</i>	<i>Fitness</i>
A	10010001	3
B	11100100	4
C	11100100	4
D	00010101	3

Figure 4: Population After Selection Example

2.3.2.4 Crossover

Once the selection process is completed and the algorithm knows which strings will spawn offspring in the next generation, it is time to generate these offspring. The algorithm must randomly select pairs of strings from the population to be subjected to crossover. This simulates human reproduction. Several types of crossover schemes have been studied: single-point crossover, two-point crossover, and others. Of these, the simplest is single-point crossover. Assuming a string length of l , a random crossover point between 1 and $l-1$ is determined. The portions of the paired strings following the crossover point are exchanged, consequently, creating two new strings.

There are cases, however, where no crossover is performed on a pair of strings. In this case, the original pair of strings is kept in the next generation. For each pair of strings a crossover rate, p_c , is used to determine whether or not crossover is performed.

The theory behind crossover is that mating strings with above average fitness values will increase the overall fitness level of the population.

For our example, assume a crossover rate, $p_c=1.0$, using single point crossover with a crossover point between the third and fourth bits of the strings. As a result, all attempts to perform crossover on pairs of strings in the population are completed. Having a crossover rate of 0.5, would indicate that half of all attempts would fail and the offspring generated from the strings would be the strings themselves, unchanged. The table below shows the resulting offspring when strings are randomly paired up as follows: A' and C';

B' and D'. Notice how the average fitness of the population after crossover has increased from 2.75 to 3.5.

<i>Label</i>	<i>Parent 1</i>	<i>Parent 2</i>	<i>New String</i>	<i>New Fitness</i>
A'	A' (10010001)	C' (11100100)	10000100	2
B''	A' (10010001)	C' (11100100)	11110001	5
C''	B' (11100100)	D' (00010101)	11110101	6
D'	B' (11100100)	D' (00010101)	00000100	1

Figure 5: Crossover Example

2.3.2.5 Mutation

After crossover, strings are subjected to mutation. Bits in a string affected by mutation have their value inverted. For instance a string '1111', becomes '1011' after mutation is performed on its second bit. Mutation is applied on a bit by bit basis based on a mutation rate p_m , typically used with a low value.

Mutation is used to introduce new strings into the algorithm that may not have otherwise been attainable through crossover. The initial, randomly generated, population therefore does not limit the algorithm.

Using a mutation rate, $p_m=0.065$, the following table shows the resulting population after mutation is performed. Out of the 32 bits available for mutation only two are mutated.

The two bits mutated are the second bit in the string labeled A'' and the fourth bit in the

string labeled C”. Notice how mutation can affect the fitness of a string positively or negatively.

<i>Label</i>	<i>String Before Mutation</i>	<i>Fitness Before Mutation</i>	<i>String After Mutation</i>	<i>Fitness After Mutation</i>
A	10000100	2	11000100	3
B	11110001	5	11110001	5
C	11110101	6	11100101	5
D	00000100	1	00000100	1

Figure 6: Mutation Example

2.3.2.6 Generation Cycle

The generation cycle defines the population from one generation to the next. Starting with the initial population, performing the genetic operators results in the next generation in the cycle.

Using the example to better understand what a Generation Cycle is, the cycle is composed of every generation of strings from the initial population when the algorithm commences to the final population when the algorithm completes including every intermediate population. An intermediate population is the population of strings within a generation, after selection, crossover and mutation are performed. Each of these stages shows how the population has evolved and is therefore part of the Generation Cycle.

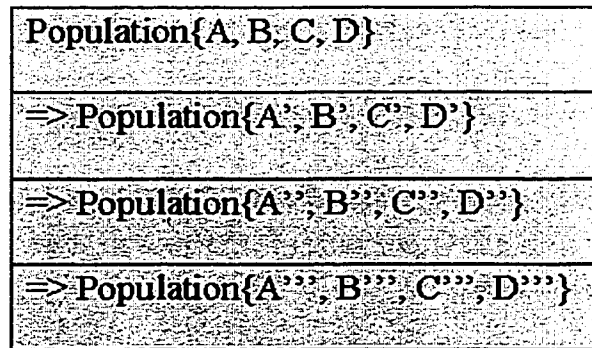


Figure 7: Generation Cycle Example

2.3.2.7 Termination Condition

To complete the algorithm, a termination condition must be set and met. Different types of conditions exist and are in use in different genetic algorithms. These types include: reaching a predetermined number of generations, attaining a predefined fitness level with one or more strings in the population; and lastly, attaining a predefined degree of homogeneity (a large number of strings have identical bits in most positions).¹⁰

2.3.3 Algorithm Pseudocode

The basic algorithm for the Simple Genetic Algorithm has the following form:

```
Simple Genetic Algorithm()
{
    initialize population;
    evaluate population;
    while termination condition not met
    {
        select solutions for next population;
        perform crossover on solutions for next population;
        perform mutation on solutions for next population;
        evaluate next population;
    }
}
```

2.4 Potential Problems with Genetic Algorithms

The genetic algorithm is very simple, yet it performs well on many different types of problems. This has prompted the wide spread use of the algorithm. However, there exist several potential problems with the algorithm that need to be addressed.

The first potential problem of the algorithm is based on its flexibility and ability to be modified easily. There are many ways to modify the basic algorithm, and many parameters that can be 'tweaked'. Depending on the parameter settings used and the modifications made to the basic algorithms, the algorithm can produce varying results. An example of a modification to the algorithm that can vary the results is the scheme used for selection and crossover. Often the crossover operator and selection methods turn out to be too effective and they end up driving the genetic algorithm to create a population of individuals that are almost exactly the same. When the population consists of similar individuals, the likelihood of finding new solutions typically decreases. Another example of a parameter that can affect the results of the algorithm is the size of the population. The size of the population should be large enough and random enough to ensure good coverage of the search space.

A second potential problem of genetic algorithm, are local maximums. Genetic algorithms perform better than most other search methods when the search space has many local maximums, but it is not the perfect method. The potential problem of the algorithm is that the optimal solution indicated by the algorithm could be a local maximum that is not the optimal solution. A search space is said to have many local

maximums when the problem has many solutions that are considerably better than nearby solutions. When these solutions are evaluated, they can appear to be optimal solutions, when in reality there are still other solutions that are actually better. Figure 8 shows a set of possible solutions to a problem. Note that amongst the 20 different solutions along the x-axis there are three points considered local maximums, namely 4, 11 and 16. They are called local maximums because they are the optimal solutions based on the other solutions in their proximity. It is possible that the genetic algorithm never finds the solution with the highest peak in the graph. Rather the algorithm may end up concentrating its efforts around a lower peak that is a local maximum.

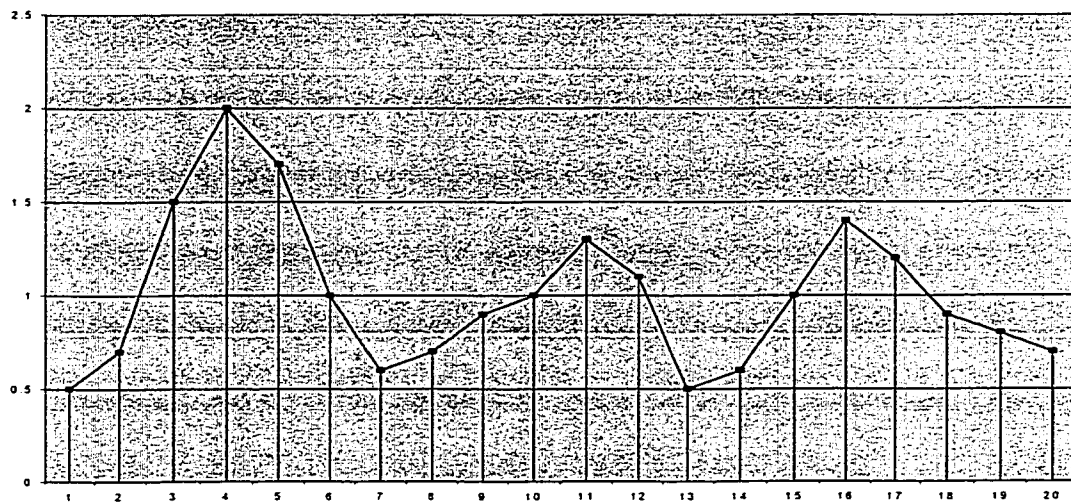


Figure 8: Local Maximums

2.5 Applications of Genetic Algorithms

Genetic algorithms have been applied across many diverse fields to help solve practical application problems. The most prominent examples of these fields include engineering, natural sciences, economics, and business. Often these application problems cannot be

solved by other methods commonly used, either because these methods cannot be applied easily or the results turn out to be unsatisfactory.¹¹

Examples of practical application problems where genetic algorithms have been successfully used include:

- Engineering design application problems, such as aircraft design
- Scheduling application problems, such as job scheduling and exam scheduling
- Routing application problems, such as vehicle routing and telephone call routing
- Packing application problems, such as packing pallets on a truck

3. Object Oriented Truckin'

3.1 Introduction

Object Oriented Truckin' (OOT) is a framework for developing programs that adapt to their environment. It is loosely based on a game invented at the Xerox Palo Alto Research Center by Mark Stefik and others during the eighties. OOT models a country ("OOTLand") in which commodities are distributed by trucks. Trucks negotiate with dealers to buy and sell commodities.¹²

3.2 Components

The OOT framework simulates trucks driving around a country distributing commodities. Distribution is accomplished either by buying commodities from Producers and selling them to Retailers, or by buying commodities from Retailers and selling them to Consumers. To add complexity to the simulation, trucks must be aware of their gas consumption and buy gas from gas stations whenever necessary.

The OOT framework is made up of the Country, Commodities (Gas, Crates, Items), Trucks, and Dealers (Producers, Retailers, Consumers, Gas Stations). In the next few subsections each of these components will be described further.

3.2.1 The Country

The country, modeled as a 10 by 10 matrix of highways, is made up of avenues running North-South and streets running East-West. Intersections are either rented to producers, retailers, consumers or gas stations, or they are left as empty lots. Trucks traverse the country moving from intersection to intersection trying to distribute commodities and make a profit. All streets and avenues may be traversed in both directions, however, since the country is a square, and not a torus, trucks must stop once they have reached the edge or corner of the country and change directions. A torus is cylindrical in shape. Had the country been designed using the torus, trucks would be able to loop around from the Eastern most point on the country to the Western most point on the country in one move. Using the square design produces a more realistic country for trucks to travel.

3.2.2 Commodities

Three types of commodities exist in the current system - Gas, Crates and Items. A Crate as defined in the current simulation is equal to 20 Items sold together as a single commodity. Crates can only be unpacked in the simulation by Retailers to yield Items, which can then be resold individually. Trucks may buy Gas from Gas Stations, Crates from Producers and Items from Retailers. Trucks may also sell Crates to Retailers, Items to Consumers, but never sell Gas.

3.2.3 Trucks

Trucks must travel around the country trading commodities between dealers in search of profitable deals. A profitable deal is obtained when a truck is able to buy a commodity from a dealer at a certain price and sell it to another dealer at a higher price. Depending

on how smart the truck is, it can also take into account the cost of the gas required to travel from dealer to dealer when evaluating a profitable deal. Trucks have limited carrying capacities and must therefore constantly sell to Retailers and Consumers so that they can further purchase commodities to increase profits.

At the beginning of the simulation, each truck is allotted a fixed amount of money and a full tank of gas. As trucks move around the country trying to increase their capital by trading, gas, money, and time are expended. Once a truck runs out of gas it becomes immobile and ceases to trade, thus completing the simulation for the truck. At the end of the simulation, the truck with the most capital is the winner.

3.2.4 Controllers

Although not mentioned above, the Controller is an integral component of the simulation. The function of the Controller is to provide services to the trucks as they search for rewarding deals. Each truck is assigned its own Controller at the start of the simulation.

The Controller can be considered the eyes and ears of the truck, providing information to the truck on timing (time elapsed since the start of simulation, time remaining in the current slot, and the simulation time remaining), positioning (where in the country is the truck currently located), capital (how much money does the truck currently have), stock (how much stock is currently owned by the truck), and dealers (dealer information at each intersection previously visited or making a phone call to intersection not previously visited). On the other hand, the Controller also facilitates certain restrictions on the

trucks in the simulation. Due to the Controller, trucks may not purchase quantities of gas, crates, and items if the cost of the quantities exceeds the amount of capital owned by the truck. In addition to this, Controllers are responsible for limiting the quantity of stock a truck can carry at one time.

Controllers are not absolutely necessary. They were introduced as a software mechanism to limit a truck's behaviour and to prevent cheating when Trucking was viewed as a competition between programmers. The objective of the competition was to develop several different trucks to run in the simulation. Programmers were asked to develop these trucks using different algorithms to manipulate the trucks. The trucks would then run in the simulation to see which programmer built the most profitable truck.

3.2.5 Dealers

The simulation is made up of four different types of dealers - Producers, Retailers, Consumers, and Gas Stations. Each of these will be explained in the following subsections.

3.2.5.1 Producers

Producers are used to introduce commodities into the simulation. Increasing or decreasing their production rate varies the amount of commodities in the simulation at a given time. Modifying the production rate will affect the profitability of the trucks in the simulation in a similar fashion. For this simulation the rate will be fixed. Producers produce Crates at a rate of 100 Crates per simulation time unit and have no restrictions on

stock capacity. Trucks buy Crates from Producers at a fixed price of \$20 per Crate.

Producers do not buy any commodity from trucks.

3.2.5.2 Retailers

Retailers are used as a middleman to buy Crates from trucks, and split the Crates into Items to be sold individually to other trucks for a profit. Retailers make a profit by keeping their buying price for Crates lower than their selling price for individual Items when multiplied by the number of Items per Crate. In the current simulation, \$24.00 was the buying price per Crate, and \$1.30 was the selling price per Item (\$26.00 for an unpacked Crate containing 20 Items).

As with Trucks, Retailers are limited in the amount of space available to them for storage. It is for this reason that Retailers must check their current available space before buying Crates from a truck. Retailers reduce their stock by splitting up the Crates into Items and selling them back to trucks. In selling Items to trucks, Retailers are only limited to the amount of stock they own.

3.2.5.3 Consumers

Consumers perform an opposite task to that of the Producers of the simulation, draining the simulation of its commodities. Consumers purchase Items from trucks at a fixed rate of \$1.50 per Item. Consumers have no limits due to capacity, however, they are restricted to a maximum size purchase of 100000 items, each time they buy.

3.2.5.4 Gas Stations

To add complexity to the simulation and give it a more realistic feel, trucks were required to run on gas. As trucks explore the country and make deals they expend gas. The simulation, therefore, required a new kind of dealer, the Gas Station. The new dealer has a single function in the simulation, to sell gas to the trucks so they can continue to move about the country. Gas Stations have an unlimited Gas commodity and have no restrictions on amounts they can sell to trucks. The price of gas varies between Gas Stations but each Gas Station has a fixed price. Prices range between \$0.80 and \$1.05 per litre of gas.

3.2.6 Managers

As in the case of the controller, another essential component in the simulation, not yet mentioned, is the Manager. The function of the Manager is to provide services to the dealers as they buy and sell commodities, in addition to enforcing certain restrictions on the dealers. Each dealer is assigned its own Manager at the start of the simulation.

As with controllers, Managers provide information to their dealers regarding their current stock, buying price and selling prices, as well as the ability to judge how much can be bought and sold at one time. All deals between a Truck and a Dealer are mediated by a Controller and Manager to prevent cheating in the competition.

3.3 The Simulation

Step one of the simulation is to define the country and create the dealers within it. To start, dealers are created as follows: 10 gas stations, 5 producers, 25 retailers, and 20

consumers. These dealers are then distributed across the country, each having a fixed position, which is maintained throughout each cycle of the simulation. A predetermined number of trucks are then created, each having a randomly generated genetic composition. The simulation then begins.

As the simulation progresses, trucks use their genetic compositions as blueprints to determine which strategies to use for their operation (how to buy and sell commodities, how to move around the country, and when to purchase gas, etc.).

At the end of the simulation cycle, trucks are evaluated based on their profits. Trucks with more money are said to be stronger, or more fit, than those with less money. Using this information, a new generation of trucks is created for another simulation cycle. The new generation of trucks, whose population remains constant each cycle, is defined as follows: 25% are new randomly generated trucks; 25% are trucks from the previous generation (those most profitable); and 50% are offspring of trucks from the previous generation.

This is repeated a fixed number of times in the hope that as the new generations are created they will become increasingly fit.

4. Design

As the name implies, the OOT project was built using an object oriented design approach.

Most of the components described in Section 3.2 have been designed as classes in the project.

Figure 9 analyses the components in a class diagram showing their subclass structure and how each of the classes interacts.

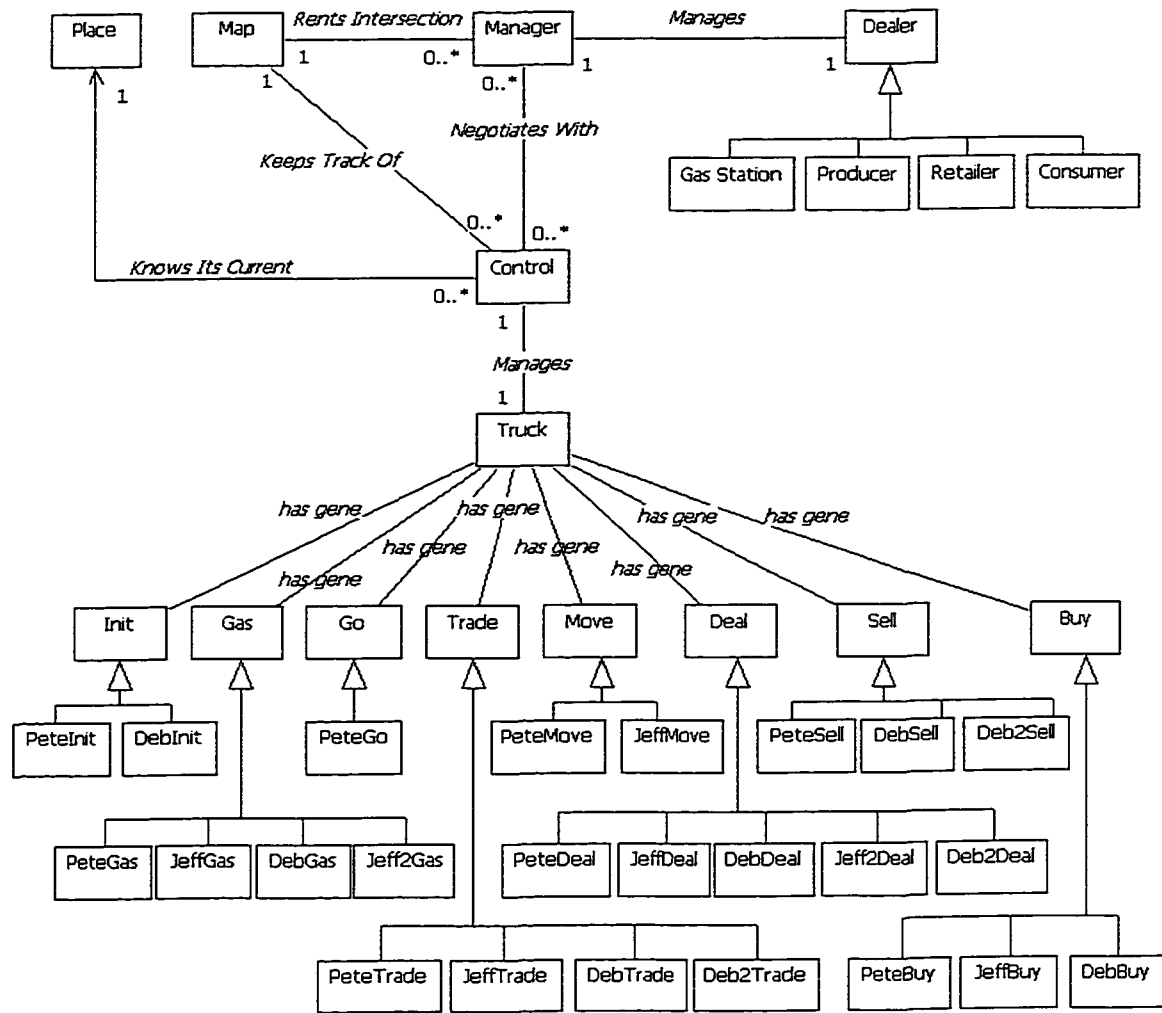


Figure 9: OOT Project Design

4.1 Components

4.1.1 The Country

The country “OOTLand” is designed as a class called Map. The Map class describes the topology of the simulation. It is used to create the country, assign managers at each intersection, and return information concerning the manager at a specific location to whoever requests it. Another feature of the map class, although not currently used in the

simulation, is the ability to draw itself to an output stream presenting a low level graphical view of the map with all its dealers.

To complement the Map class, a Place class is included in the design. A Place refers to a coordinate (street and avenue) on the map. The Place class is used to input and output coordinates of the map in the form “(Street, Avenue)”, and to calculate the distance between two given places.

4.1.2 Trucks

The optimization problem of the OOT project is summarized with the following question: “What strategies should a truck use to make the most profit during a run of the simulation?”. The truck must then be able to change its strategies easily after each generation, enabling it to search for a better combination of strategies. To do this, the truck class has been designed as a framework for the truck, storing pointers to its different strategies. These strategies can be viewed as best practices or areas of expertise. For each task that a truck must perform, an area has been defined. The current simulation has trucks with eight areas of expertise: Initialize, Gas, Trade, Buy, Sell, Move, Go and Deal. Each area of expertise has between one and five strategies, which can be used by trucks to perform the specific task of that area.

The design includes a separate class for each of the areas of expertise mentioned above, as well as, separate subclasses for each of the strategies within each area. This allows

instances of trucks to instantiate, and store as pointers, the classes it will use for its strategies while running the simulation.

In the next sections I will further explain the different areas of expertise and go into some detail on each subclass.

For a summary of these strategies please refer to section 5.1.

4.1.2.1 Initialize

The first task a truck must do when starting the simulation is to decide what to do first.

The Init class is used to define the initial thoughts, or plan of attack for the truck when the simulation begins. The current simulation has two strategies for this task, making the Init class act as the base class for two subclasses: PeteInit and DebInit.

The subclass PeteInit simply sets the initial direction of the truck, whereas the DebInit subclass sets the initial direction of the truck and tells the truck that it must travel to the top left corner of the map and scan the entire map as its initial action.

4.1.2.2 Move

The next task that is important to a truck is to know how to move around the country to collect information about the locations of dealers. This behaviour is triggered when the truck does not know of a good deal or a truck must build a knowledge base of the country and dealers prior to looking for deals. The Move class is a base class for two subclasses:

PeteMove and JeffMove. Each of these subclasses outlines a different strategy for trucks as they move from one intersection to another.

Both subclasses, PeteMove and JeffMove, perform the task of systematically moving the truck around the country while obtaining information about the dealers visited, in the exact same manner. The algorithm simply zigzags left to right and top to bottom or vice versa depending on the direction of the truck. The difference between the two is that JeffMove tries to sell surplus stock to dealers as it travels the country.

4.1.2.3 Go

Once a truck is able to move around the country one intersection at a time, it is important to be able to specify a specific intersection to go to. The Go class defines the strategy used by trucks to travel to a specific intersection of the country. In the current simulation, only one approach is used, which is why there is only one subclass, PeteGo, using the Go class as a base.

The algorithm used in PeteGo is quite simple. It moves the truck North or South until the truck is located on the requested street and then moves the truck East or West until the truck is located on the requested avenue.

4.1.2.4 Gas

As trucks move around the country, they consume gas. As a result, the task of knowing when and where to get gas is an important strategy for the trucks. A Gas class was created for this purpose, which is used as a base class by four subclasses: PeteGas,

JeffGas, DebGas, and Jeff2Gas. Once again, each of these subclasses provides a different strategy for trucks to know when and where to get gas.

From the four available strategies, the simplest is found using the subclass PeteGas.

Using this strategy, trucks determine that gas is required when their reserve falls below the amount needed to travel the longest distance across the country. This strategy will always buy the maximum amount of gas possible, based on the room left in the truck's tank and the amount of money available to spend.

The next strategy, found in the subclass JeffGas, decides that gas should be purchased when trucks find themselves located at a gas station or when their reserve falls below the amount needed to travel the longest distance across the country. When deciding the quantity of gas to purchase, the strategy bases this decision on the amount needed to fill the tank, the amount the truck can afford to purchase, and the amount of gas needed to continue driving before the simulation time runs out.

The subclass DebGas contains a strategy which decides that gas is required when the reserve falls below the amount of gas required to travel to the closest gas station with an additional 10 units to spare, however, gas is never required during the last 100 time units of the simulation. This strategy will always buy the maximum amount of gas possible, based on the room left in its tank and the amount of money available to spend.

The last subclass, Jeff2Gas, contains a strategy which determines that gas is necessary when trucks are located at a gas station selling the cheapest gas seen to date, or when the reserve of gas will not be enough to permit the trucks to travel to a gas station if one more move is made in the opposite direction of the closest gas station. As in the case of JeffGas, the strategy bases the decision of the amount of gas to purchase on the amount needed to fill the tank, the amount the truck can afford to purchase, and the amount of gas needed to continue driving before the simulation time runs out.

4.1.2.5 Trade

Section 4.1.2.1 described the Init class which defined the initial thoughts, or plan of attack for the trucks when the simulation began. The Trade class was designed to provide the trucks with a way of knowing what they should be doing next. There are four strategies available to trucks in the current simulation: PeteTrade, JeffTrade, DebTrade, and Deb2Trade. Each of these subclasses of the Trade class defines how the truck resolves its next plan of action.

All four subclasses basically work the same way with slight differences. Each uses a case statement to decide what action should be taken next. The PeteTrade subclass strategy looks for deals when it is not explicitly in the middle of trying to complete one. If a deal is found, the strategy tells the truck to go buy the commodity from the dealer. Once the commodity is purchased, the strategy has the truck go sell it to the other dealer.

The JeffTrade subclass strategy differs from that of the PeteTrade subclass only in that it does not look for deals until it has found at least one gas station by traveling around the country.

The strategy used by DebTrade differs from that of the PeteTrade subclass in that it will move the truck to the next corner of the country from its current location before searching for a deal, and if a deal is found, it will only go buy the commodity from the dealer if more than 200 time units remain in the simulation.

The last strategy found in the subclass Deb2Trade differs from that of the PeteTrade subclass in that it will not look for a deal within the last 150 time units of the simulation.

4.1.2.6 Deal

This next class is important to the trucks if they wish to continue in the next generation of the simulation. The task defined in the Deal class is to find the most profitable deal at any given time. This class has five subclasses that use it as their base, the most of all the areas of expertise for the trucks. The subclasses in the design are: PeteDeal, JeffDeal, DebDeal, Jeff2Deal, and Deb2Deal. Each of these subclasses provides, with a different strategy, the most profitable deal available at the time the request is made.

The first strategy used in the PeteDeal subclass, simply uses the information obtained by the truck as it travels the country to find the most profitable deal. The most profitable

deal is defined as two dealers dealing in a common commodity, the first sells it for less money than the other will pay to buy it.

The strategy used by the subclass JeffDeal is also based on maximizing profit, however, there are some slight modifications. This strategy will phone the dealers to make sure that their prices have not been changed before deciding that the deal is profitable. The strategy also forces trucks to perform three deals buying and selling crates before it starts alternating the deals between crates and items.

The next strategy, DebDeal, works differently from the others in that it is not a strategy looking for a buyer and a seller. This strategy checks the current location for a dealer. If a dealer is found and there is a good deal or the truck is running low on capital (less than \$50) or there is more than 200 time units left in the simulation it will try and sell the appropriate commodity to the dealer. In this case, a good deal is defined as a dealer whose buying price is greater than the average price the truck paid for the commodity. If none of the above criteria is met, the truck will attempt to purchase the commodity from the dealer.

Similar to the JeffDeal subclass, the subclass, named Jeff2Deal, has a strategy which uses the phone to call dealers and verify the prices before declaring something a profitable deal. The difference between the two strategies is that this one calculates the time it will take to perform the deal (this includes buying and selling), as well as, the amount of gas consumed to complete the deal. If the simulation time will not permit the completion of

the deal or the reserve of gas in the truck will not permit the completion of the deal, the deal is not considered to be a good one.

The last strategy, in the `Deb2Deal` subclass, searches for the closest deal to the current truck location, which is profitable.

4.1.2.7 Buy

Once the truck has located a profitable deal, using the `Deal` class, the truck must navigate the country to the dealer selling the commodity in question and buy the commodity from that dealer. This is the task for which the `Buy` class and its three subclasses have been added. The three subclasses are `PeteBuy`, `JeffBuy`, and `DebBuy`. Each of these subclasses have a defined strategy for moving to the dealer and buying the commodity from the dealer.

All of the strategies in the above three subclasses are based on the same principal which is to continuously move closer to the seller until it is reached. Once the truck has located the dealer, the truck can request to purchase commodities from the dealer. The strategy of subclass `PeteBuy` will try to buy as much as possible, saving only enough capital to buy gas to be able to move from one extreme of the country to the other. The other two strategies (`JeffBuy` and `DebBuy`) try to purchase as much as possible, saving only enough money to fill a tank of gas. The difference between the strategies in `JeffBuy` and `DebBuy` is that `DebBuy` keeps track of the average price at which commodities were purchased.

This information is used later in the simulation to help sell off overstocked commodities to other dealers who will at least pay the truck what the commodity was bought for.

4.1.2.8 Sell

If the trucks do not know how to sell the commodities bought from other dealers to make a profit, the Buy class is of little use to the trucks. The Sell class is used to fulfill this need of the trucks. The current simulation contains three strategies to perform this task. The subclasses created for the three strategies are: PeteSell, DebSell, and Deb2Sell.

The strategy within the subclass PeteSell is very simple. The strategy travels to the dealer and attempts to sell all of its commodities to the dealer.

The DebSell subclass and Deb2Sell subclass, go one step further than the strategy in PeteSell. If the dealer does not purchase the truck's entire stock, these strategies will then search for the closest dealer to the truck's current location and go to sell the surplus to the dealer even if it means selling at a loss. The only difference between the two strategies is that the DebSell subclass will only attempt to sell its surplus three times, after which it will simply retain the surplus, whereas the Deb2Sell strategy will continue to try to unload its surplus.

4.1.3 Controllers

As mentioned in earlier sections, the controller was implemented as a software mechanism to limit a truck's behaviour and prevent cheating. The controller was designed as a class to be used as the interface that a truck must use for all interactions

during the simulation. A limited amount of the controller's functionality was made public to the trucks, thus preventing the trucks from performing illegal operations (i.e. buying more commodities than they can hold or have money for and getting information about a dealer at an unvisited intersection without phoning).

4.1.4 Dealers

All dealers needed the same basic functionality as a result, a class was created to be used as the base of all the different types of dealers in the simulation. This class was named the Dealer class. The basic functionality given to all dealers was the ability to buy and sell commodities. The four subclasses in the design using the Dealer class as a base are: Producer, Retailer, Consumer, and Gas_Stn.

4.1.5 Managers

The Manager class was added to the design for the same reason the Controller class was added for the trucks. The Manager class is used to support dealers at each intersection and is a software mechanism to limit a dealer's behaviour and prevent cheating. The manager helps the dealers buy and sell commodities, as well as inform them when to increase and decrease selling and buying prices (although this is not enabled in the current simulation).

5. OOT: The Genetic Algorithm

One of the best uses of Genetic Algorithms (GAs), as discussed in section two, is to search for an optimal solution, or at least a reasonably good solution, to a complex problem having a very large search space.

The OOT project is a complex optimization problem whose search space consists of sets of algorithms that form trucks. The object is to find an optimal set of trucks by evaluating their profits after running the simulation. The trucks differ in their strategies chosen from each area of expertise.

In the next sections, the different algorithm components designed into the OOT project will be discussed.

5.1 Encoding Mechanism

Each truck can be uniquely identified based on its selection of strategies from each of the areas of expertise. The trucks, therefore, have eight genes defining its genetic structure. Each gene has a numeric value between zero and the number of strategies for that gene less one. Figure 9 depicts the different genes, their influential traits on the trucks and the value range for each gene.

<i>Gene Number</i>	<i>Truck Trait</i>	<i>Gene Values</i>
0	Init	0:1
1	Gas	0:1:2:3
2	Trade	0:1:2:3
3	Buy	0:1:2
4	Sell	0:1:2
5	Go	0
6	Move	0:1
7	Deal	0:1:2:3:4

Figure 10: Summary of Genes

The total number of trucks in the search space can be calculated from the above table.

$$\text{Total Number of Trucks} = 2 * 4 * 4 * 3 * 3 * 1 * 2 * 5 = 2880$$

The total number of trucks in the search space of this example is not a typical size. When using genetic algorithms the size is usually larger. This is because for smaller search spaces it may be more efficient to use other search methods, which can analyze each potential solution instead of analyzing a sampling of them.

5.2 Fitness Function

Once the trucks are encoded the simulation can be run. In the case of OOT, the fitness function, or equation used to evaluate the trucks, is the simulation itself. The trucks are evaluated according to how profitable they are after finishing a run of the simulation.

5.3 Selection

After each run of the simulation is completed and the trucks are evaluated based on the fitness function, the current simulation selects the top 25% of trucks that were profitable. The other 75% of trucks are discarded and will not play again in the next run of the simulation.

5.4 Crossover

As the algorithm states, the next step is to mate the selected trucks and give birth to new trucks for the next generation. The OOT project performs a unique form of crossover. The program selects pairs of trucks from the preferred trucks of the previous generation and with each pair yields four new offspring for the simulation. The method used to create these offspring is by randomly selecting four genes from the first truck and the remaining genes are taken from the second truck. These genes are then combined to form the genetic structure of the new offspring.

5.5 Mutation

Mutation is sometimes used in GAs to allow new gene combinations to be introduced into the algorithm that may not be attainable from offspring of the original population. The OOT project does not directly support mutation, however, another method producing the same results is used. This method is simply the introduction of 25% new trucks whose genetic structures are defined by randomly selecting genes. In effect, it is mutation of each gene of these new trucks with a certainty factor of $p_m = 1.0$.

5.6 Generation Cycle

Each run of the simulation is considered a generation. The simulation runs for a predetermined amount of time units, or until all trucks have run out of gas and are immobile.

The initial generation consists of trucks whose gene structures are randomly generated.

After this first generation, all subsequent generations consist of profitable trucks from the previous generation, offspring from these trucks created through crossover, and new trucks randomly generated.

The cycle continues until a termination condition is met.

5.7 Termination Condition

The termination condition of the OOT project is based on the number of generation cycles completed. The simulation is requested to cycle 100 times. After the 100th generation the program halts and the results can be evaluated.

6. Implementation

The OOT project was implemented using the C++ programming language. Due to the number of iterations of the simulation and length of time the simulation would run, speed was a top concern in the decision of what language to use for the implementation. It was therefore best to choose an efficient language. This section will describe the different implementation strategies used to complete the project. Two important questions answered in this section are, how were the different truck strategies created? and how were they made easily swappable with one another?

The implementation of the truck in the OOT project was completed in three phases. The original design of the project was used as a competition between similar trucks. The design consisted of a truck class containing several member functions. Each member function had a specific task and used a basic algorithm to fulfill its task. Figure 11 shows the object diagram of the original design of the project. The purpose of this phase was to create a working simulation. Could several trucks be created at one time and run in the simulation? Would they be able to profit during the simulation time?

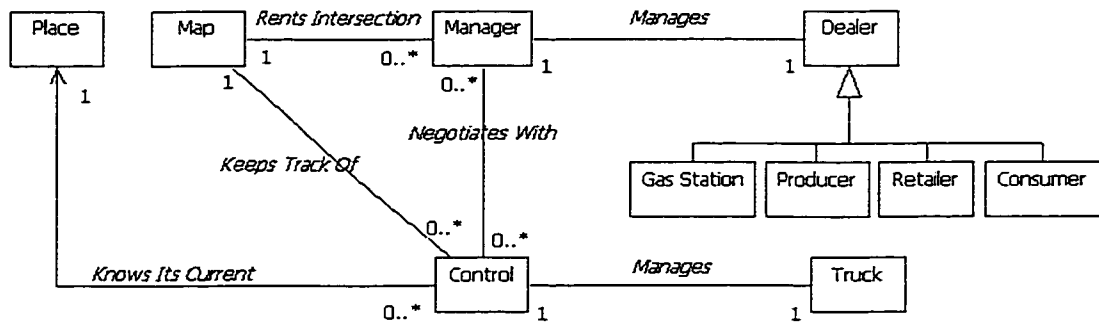


Figure 11: OOT Project Design (Implementation Phase 1)

The goal of the next phase of the project was to create several different trucks of varying intelligence to run in the simulation simultaneously. An important requirement for building these new trucks was that they all be derived from a common base class, thus ensuring compatibility of the trucks for the third and final phase. To fulfill this requirement, the original truck class was modified and each of the member functions was declared as a virtual function to be overridden in subclasses of this truck base class. The algorithms used in the original truck were then removed and a new class called PeteTruck was created made up of the original algorithms. A simple framework was in the making for a truck. A total of four more trucks were created. Each truck was built as a single class derived from the base truck class. For the most part, each new class implemented its member functions using different algorithms. Within the simulation, these classes, representing different trucks of varying intelligence, competed to see which would profit most, buying and selling commodities using its unique combination of strategies. The phase was complete when each truck was tested in the simulation to make sure it functioned properly. Figure 12 depicts the design at the end of this phase.

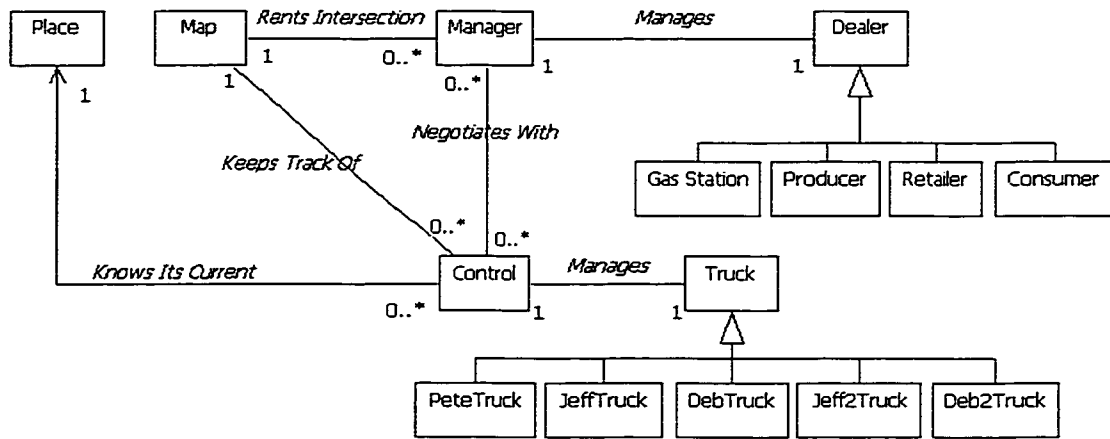


Figure 12: OOT Project Design (Implementation Phase 2)

The third and final phase of the implementation of the OOT project involved the introduction of the genetic algorithm. There were two requirements of this phase for the project implementation. The first requirement was the ability to mix and match the different algorithms of each of the trucks with one another to create hybrid trucks. The second requirement of the phase was to be able to create these truck hybrids at run time. To fulfill the first requirement, the design of the truck class was once again altered. The algorithms within the derived truck classes needed to be entities of their own and not groups of functions of a larger entity. The base class of the truck needed to be divided into several different base classes. The altered design made each of the virtual member functions of the truck class a base class of its own. Each one of these base classes was now responsible for a specific action or strategy of the truck, and therefore required a single member function. Having redesigned the base class of the truck, it was now necessary to perform the same redesign for each of the five derived truck classes. From this, a new truck class was defined. This new class was not a base class anymore,

however. This class would be used as the skeleton or framework for all trucks. The new truck class would manage pointers to each of the classes defined in the new design.

Figure 13 depicts the third and final design.

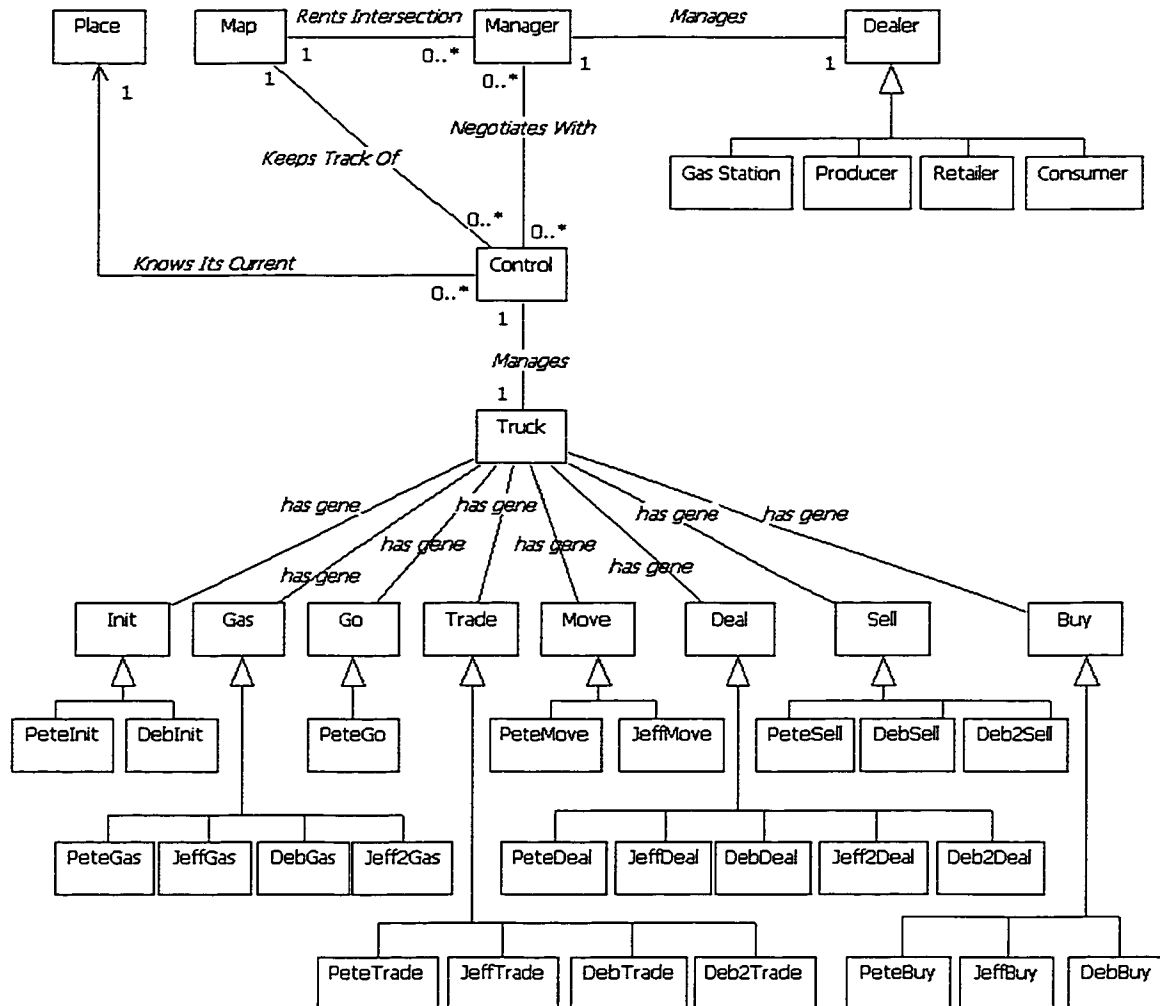


Figure 13: OOT Project Design (Implementation Phase 3)

For the implementation of the project, C++ provided the efficiency in a language that was required to produce fast code, as well as the key tools required to complete each phase.

These tools included pointers, which were used by the different classes making up a truck to talk to one another, arrays, which were heavily used to store the genes of the trucks and

also used to keep statistics of all the runs of the simulation, inheritance and virtual functions, both of which were used to help trucks easily evolve to use different strategies, and operator overloading, which was used in the Place class to redefine input/output and arithmetic operators. For code samples of these object-oriented techniques and tools, used in the implementation of the OOT project, please refer to Appendix A.

7. Results

There are several different factors which come into play in running the OOT simulation, which could affect the performance of the trucks from generation to generation, thereby affecting the results of using this approach to solving the optimization problem. The OOT project accepts these factors as input parameters to define how the simulation runs. Based on these parameters the results can vary. This section will discuss the different input parameters of the OOT simulation and why the parameter values affect the results of the simulation positively or negatively. Later in the section, the results of one simulation run will be discussed.

The first parameter of the simulation is the length of time allocated to each generation. It is evident that the more time a generation of trucks is given to run in the simulation, the better the trucks will perform. This is because there is more time to gain knowledge of the distribution of the dealers and gas stations across the map. This knowledge is important because otherwise, many of the trucks will simply trade between the same dealers within a limited section of the map. The length of time will also more clearly define the fitter trucks. If you were to compare the profits of a truck whose “init” algorithm says travel the entire map before beginning to look for deals, versus a truck whose “init” algorithm says look for a deal right away, the second truck is more likely to

be more profitable than the first in a short simulation. If the first truck had a much better algorithm for finding deals and buying gas compared to the second truck and the simulation time was longer, there is a good chance that the first truck could make up for the time it spent traveling across the entire map, especially with its expanded knowledge, potentially being more profitable than the second truck. The extended simulation time could therefore reverse the preconceptions of the second truck being more profitable than the first.

The second parameter of the simulation is the number of runs, or generations, within the simulation. This factor can greatly affect the end results of the simulation. Each generation added to the simulation increases the chances of finding a better solution to the problem. This is because each generation introduces more possible solutions from the search space to be evaluated.

Another important parameter is the number of trucks per generation. The map is a fixed size of 10 streets and 10 avenues. There are also a fixed number of dealers and gas stations distributed around the map. Although the number of crates produced by producers is unlimited and the number of items consumed by consumers is also unlimited, the retailers, who divide crates into items, have limited storage space. This restriction would not seem to cause a problem but in fact it does for most of the trucks. As a result of the restriction, too many trucks finding the same deal at the same time will purchase too much from a producer and will all travel to the same retailer to unload merchandise. Once they reach the retailer, they find that the retailer has already

purchased too much from other trucks and does not have any more storage space. This causes many trucks to be overstocked, making them less profitable. The simulation appears to run more effectively when fewer trucks are competing.

The last parameter of the simulation defines whether the trucks all start at the same point on the map at the beginning of each new generation, or if every truck starts at a randomly generated point on the map. This small factor of the simulation can also have a large affect on the simulation results. It has been found that by randomly placing trucks on the map, certain trucks have unfair advantages over others because they may have been placed in an area of the map heavily populated by dealers. The trucks seem to compete a lot better with their starting points being all the same.

To be able to analyze the results of a specific run of the simulation, key pieces of information are collected and stored in memory and then used to produce several different tables showing views of the progress of the algorithm. The following information is stored for each run of the simulation: the amount of money each truck has earned from each generation within the simulation, and the number of instances of a specific truck that is created within each generation. This data is then displayed two different ways. In the first, the data is sorted by genetic composition, while in the second, the data is sorted by the amount of money made throughout all of the generations.

Using the information provided by OOT, one could easily verify if the algorithm is working. In theory, if the algorithm is working, the trucks from one generation to the

next should be more profitable. Consequently, the amount of money made by the trucks from one generation to the next should increase. If the algorithm is working, you could also expect that trucks with similar genetic composition would be created, due to the effects of crossover.

To examine a run of the OOT simulation, the following parameters were used:

- Length of time allocated to each generation = 5000 units
- Number of generations within the simulation = 100
- Number of trucks per generation = 10
- All trucks start at the same point on the map at the same time each generation

Certain trends indicating the algorithm did work to a certain degree are evident when reviewing the data from the above simulation.

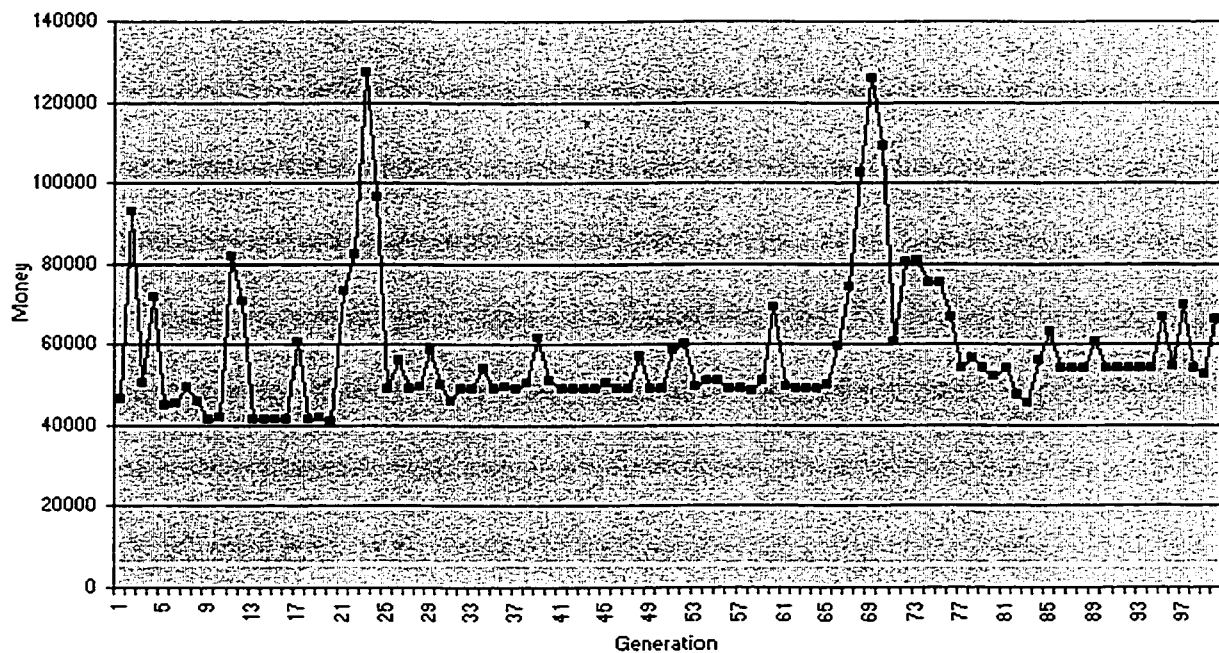


Figure 14: Simulation Results

The graph above represents the total profits earned by all trucks within the same generation across each of the 100 generations. The ideal graph of this data would have a constant upward sloping curve to indicate that the trucks within each generation were becoming increasingly profitable and “fit”. The data returned by the simulation shows a different type of graph. The graph indicates that at several points throughout the simulation, the curve did slope upwards and reach a peak, but for some reason the trucks that were very profitable, in one generation or more, suddenly became less profitable, thus accounting for the sudden drops in the graph.

Even though this was the case, subsequent generations of trucks are still more profitable than those trucks prior to the peak. This “peak/lull” trend will likely continue to occur in longer running simulations. In order to verify this, one would have to allow the simulation to run for more generations. A possible explanation for this phenomenon is that the population of trucks within a generation becomes smarter and the gap between the stronger and weaker trucks narrows. Once a population is composed of genetically similar trucks there is greater competition and no truck can perform exceptionally well.

Although the resulting graph is not ideal, the graph does not indicate that the algorithm failed to work. In fact, one can infer the exact opposite from this graph. Between generation 20 and 23, there are three generations that have totals that are constantly increasing. A fit truck was found and the next three generations each produced trucks whose profitability was greater than its preceding generation. This happened again

between generations 66 and 69 where the totals increased from 49076 through to 125653 within five generations.

Another interesting result of the simulation is that by looking at the genetic makeup of the trucks that made the most money during each of the generations, which are also the ones that were created most often, we see that there are similarities between them. Take for instance the top five money making trucks throughout the 100 generations of the simulation, whose gene structures are as follows: 13311011, 12011011, 11201014, 11001011, and 11001014. All five of these trucks are using the same algorithms to initialize themselves, to sell commodities to dealer, to go places, and to move around the map. Three out of the five trucks are using the same algorithm to find deals, while the remaining two also share another algorithm to find deals. This proves that as time passes and the generations change, more possible solutions from the search space are tested within our simulation and the ones that seem to do well have similar characteristics. This indicates once again that the algorithm seems to be working.

To view the raw results of the simulation run discussed above refer to appendix B.

8. Conclusions

This thesis attempted to use genetic algorithms to solve a complex optimization problem involving a real world simulation where trucks compete to make a profit, by buying and selling commodities in a country filled with different types of dealers. While the results are not ideal, the data does indicate that the algorithm helped in deriving potential solutions to the problem at hand.

The project was implemented using the C++ programming language. This language was selected because it is an efficient language producing fast running code. The performance of the simulation was a major concern, due to the number of iterations of the algorithm. For the simulation to run using the parameters discussed in section 7, the simulation took under 10 minutes to complete. Running the simulation with a larger population, longer simulation time, and larger number of generations could cause the program to run for several hours before completing. The use of C++ was also prompted because of the availability of several object-oriented techniques, such as inheritance and virtual functions. These techniques and the basic tools inherent in the language, such as pointers and arrays, made it the most appropriate language to use for this project.

Subsequent work on the project has shown that the strategies of certain genes may be too similar with one another, thereby causing the results of the simulation to be inconclusive. A possible improvement to the project, to improve the results, may be as simple as creating new diverse strategies for the trucks to use.

Another improvement to the project, to improve the results, is to increase the size of the country. As mentioned in section 7, the results of the algorithm were better when the number of trucks competing at one time was small. This was due to the size of the country and trucks all converging on the same deals. If the country was larger and there were more dealers spread out across it, the simulation would allow for more trucks to compete at one time. The increase in the population size would lower the risk that the solutions found by the algorithm would be local maximums because a larger percentage of the search space would be analyzed.

Along with the improvements discussed above, some ideas for future work on the project, to improve results, can include, using different forms of crossover and mutation between each generation cycle, and making the country a torus instead of a square.

9. References

1. M. Srinivas and L.M. Patnaik, "Genetic Algorithms: A Survey", *Computer*, Vol. 27, No. 6, June 1994, p. 17.
2. C.A. Villet, E.P. Solomon, C.E. Martin, et al., *Biology*, 2nd ed., Philadelphia: Saunders College Publishing, 1989, p. 20.
3. Villet, pp. 221-222.
4. Villet, pp. 417-421.
5. J.L. Ribeiro Filho, P.C. Treleaven and C. Alippi, "Genetic-Algorithm Programming Environments", *Computer*, Vol. 27, No. 6, June 1994, p. 28.
6. Srinivas, pp. 18-19.
7. Ribeiro Filho, p. 29.
8. Srinivas, pp. 19-20.
9. Srinivas, p. 19.
10. Srinivas, p. 20.
11. T. Back, U. Hammel and H.P. Schwefel, "Evolutionary Computation: Comments on the History and Current State", *IEEE Transactions on Evolutionary Computation*, Vol. 1, No. 1, April 1997, p. 10.
12. P. Grogono and G. Butler, "The Truckin' Project", August 1997

10. Bibliography

1. M. Srinivas and L.M.Patnaik, "Genetic Algorithms: A Survey", *Computer*, Vol. 27, No. 6, June 1994, pp. 17-26.
2. C.A. Villee, E.P. Solomon, C.E. Martin, et al., *Biology*, 2nd ed., Philadelphia: Saunders College Publishing, 1989.
3. J.L. Ribeiro Filho, P.C. Treleaven and C. Alippi, "Genetic-Algorithm Programming Environments", *Computer*, Vol. 27, No. 6, June 1994, pp. 28-43.
4. P. Grogono and G. Butler, "The Truckin' Project", August 1997
5. M. Mitchell, *An Introduction To Genetic Algorithms*, Cambridge, Massachusetts: The MIT Press, 1996.
6. P. J. Darwen and X. Yao, "Speciation as Automatic Categorical Modularization", *IEEE Transactions on the Evolutionary Computation*, Vol. 1, No. 2, July 1997, pp. 101-108.
7. T. Back, U. Hammel and H.P. Schwefel, "Evolutionary Computation: Comments on the History and Current State", *IEEE Transactions on Evolutionary Computation*, Vol. 1, No. 1, April 1997, pp. 3-17.

11. Appendix A: Code Samples

The OOT project makes use of several object-oriented techniques inherent in the C++ programming language, as well as some basic tools which originated in the C programming language. This appendix will show code examples to demonstrate the use of these techniques and tools in the implementation of the project.

11.1 Pointers

The following code sample is taken from the file *truck.h*. It contains the class definition of the truck class. As discussed in section 6, the truck class is used as the framework for creating trucks in the simulation. The class is responsible for maintaining pointers that reference the particular strategies that the instance of the truck uses to compete in the simulation.

```
class Truck {
public:
    Truck (int id_num, Control *controller, int genes[8]);
    void play ();

    Dir dir;          // current direction of the truck
    Dealer_info *info[NUM_AV][NUM_ST]; // Dealer data base
    // Plan data. The state tells what we are currently doing.
    // The plan consists of going to the 'buyer', buying something,
    // taking it to the 'seller', and selling it.
    Plan plan;
    Control *ctl;     // Pointer to truck's controller.
    int id;          // Truck's ID.
};
```

```

Init *init;        // Pointer to truck's init strategy
Gas *gas;         // Pointer to truck's gas strategy
Trade *trade;     // Pointer to truck's trade strategy
Buy *buy;        // Pointer to truck's buy strategy
Sell *sell;      // Pointer to truck's sell strategy
Go *go;         // Pointer to truck's go strategy
Move *move;     // Pointer to truck's move strategy
Deal *deal;     // Pointer to truck's deal strategy

void message (char *text);
    // Display truck's status and the given message.

void update_info ();
    // Update information about a dealer.

void Phone_Update (Place pointA);
    // Update information by phone about a dealer at pointA

long Gas_Cost (Place pointA, Place pointB);
    // returns the cost of the gas to go from here
    // to pointA, then to pointB, and finally to the
    // closest gas station to pointB.

long Closest_Gas_Price (Place pointA);
    // returns the price of gas at the closest
    // gas station to pointA.

Place Closest_Station (Place pointA);
    // returns the location of the closest
    // gas station to pointA.

int Cheapest_Station (long price);
    // given a price, return 1 if this is the
    // cheapest price for gas we have seen.
};

```

11.2 Arrays

This code sample demonstrates the use of arrays in the implementation of the OOT project. The sample is taken from the file *truck.c*. Arrays are used throughout the project to define the genetic structure of trucks. In this example, this information is then used by the truck class constructor to determine what strategies are linked to the truck being created through the truck class framework.

```

// Constructor: stores truck's ID, controller, and initial
// direction. Creates the truck as a collection of strategies
// according to the gene code passed.
Truck::Truck (int id_num, Control *controller, int genes[8])
{
    id = id_num;
    ctl = controller;

    // Create our own "map" of the country, with no information.
    for (int av = 0; av < NUM_AV; av++)
        for (int st = 0; st < NUM_ST; st++)
            info[av][st] = NULL;

    switch(genes[0])
    {
        case 0:
            init = new Peteinit(this);
            break;
        case 1:
            init = new Debinit(this);
            break;
    }

    switch(genes[1])
    {
        case 0:
            gas = new Petegas(this);
            break;
        case 1:
            gas = new Jeffgas(this);
            break;
        case 2:
            gas = new Debgas(this);
            break;
        case 3:
            gas = new Jeff2gas(this);
            break;
    }

    switch(genes[2])
    {
        case 0:
            trade = new Petetrade(this);
            break;
        case 1:
            trade = new Jefftrade(this);
            break;
        case 2:
            trade = new Debtrade(this);
            break;
    }
}

```

```
case 3:  
    trade = new Deb2trade(this);  
    break;  
}
```

```
switch(genes[3])  
{  
case 0:  
    buy = new Petebuy(this);  
    break;  
case 1:  
    buy = new Jeffbuy(this);  
    break;  
case 2:  
    buy = new Debbuy(this);  
    break;  
}
```

```
switch(genes[4])  
{  
case 0:  
    sell = new Petesell(this);  
    break;  
case 1:  
    sell = new Debsell(this);  
    break;  
case 2:  
    sell = new Deb2sell(this);  
    break;  
}
```

```
switch(genes[5])  
{  
case 0:  
    go = new Petego(this);  
    break;  
}
```

```
switch(genes[6])  
{  
case 0:  
    move = new Petemove(this);  
    break;  
case 1:  
    move = new Jeffmove(this);  
    break;  
}
```

```
switch(genes[7])  
{
```

```

case 0:
    deal = new Petedeal(this);
    break;
case 1:
    deal = new Jeffdeal(this);
    break;
case 2:
    deal = new Debdeal(this);
    break;
case 3:
    deal = new Jeff2deal(this);
    break;
case 4:
    deal = new Deb2deal(this);
    break;
}

init->init();
plan.price_bought = 0;
}

```

11.3 Inheritance and Virtual Functions

The use of inheritance in the OOT project was key to producing a framework for the trucks which would allow the trucks to easily and efficiently evolve over generations through the change of strategies. Inheritance was used to ensure that each strategy for a given task could be replaced by another strategy for the same task without having to modify the interface of the truck. This was accomplished by defining virtual functions in the base classes, which would need to be redefined when subclasses were created to implement new strategies. The code samples below are taken from the files *Go.h* and *PeteGo.h*. Each one defines its respective classes, Go and PeteGo. Class PeteGo is a subclass of Go and therefore, inherits the methods and properties from class Go. The Go class contains a single function declared as virtual within its definition. The function does not have an implementation at this level. It is up to the subclasses to implement the function.

```

class Go {
public:
    virtual int go_to (Place dest, int& done_it);
protected:
    Truck *truck;
};

class Petego : public Go {
public:
    Petego(Truck *tr) { truck = tr; }
private:
    int go_to (Place dest, int& done_it);
};

```

11.4 Operator Overloading

This last code sample contains code taken from the file *place.h*. This sample demonstrates the use of operator overloading in the OOT project implementation. There are three operators being overloaded, the input/output operators “<<” and “>>”, and the arithmetic operator “-”.

```

class Place {

    friend ostream& operator<< (ostream& os, Place pl);
    friend istream& operator>> (istream& is, Place& pl);

public:

    Place (int avenue = 0, int street = 0) {
        // Construct a new place.
        av = avenue;
        st = street;
    }

    int operator- (Place pl) {
        // Return Manhattan distance between here and place.
        // This is a symmetric difference.
        return abs(pl.av - av) + abs(pl.st - st);
    }

    int av, st;           // The location of this place.
};

```