Master's Degree in Informatics Engineering
Dissertation
Final Report

Evolutionary Creativity

Paulo Jorge Reis Pereira
paulojp@student.dei.uc.pt

Advisor:
Fernando Penousal Machado

Date: July 1, 2016

FCTUC DEPARTAMENTO
DE ENGENHARIA INFORMÁTICA
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

UNIVERSITY OF COIMBRA

FACULTY OF SCIENCES AND TECHNOLOGY

DEPARTMENT OF INFORMATICS ENGINEER

MASTER'S DEGREE IN INFORMATICS ENGINEERING

Evolutionary Creativity

Author:
Paulo PEREIRA

Supervisor:
Prof Dr. Fernando Penousal
MACHADO

Jury:
Prof Dr. Jorge GRANJAL

Prof Dr. Amilcar CARDOSO

July 1, 2016

"Why did humans lose their body hair? Why did they start walking on their hind legs? Why did they develop big brains? I think that the answer to all three questions is sexual selection."

Richard Dawkins

UNIVERSITY OF COIMBRA

Abstract

Faculty of Sciences and Technology
Department of Informatics Engineer

Master's Degree in Informatics Engineering

Evolutionary Creativity

by Paulo Jorge Reis Pereira

When Darwin published his theories about evolution, he presented two key mechanisms responsible for the evolution (natural selection and sexual selection). While the natural selection has widely accepted by the scientific community, the sexual selection has highly criticised and so it was forgotten over time. It was only nearly a century later, that the sexual selection began to be acknowledged, mainly due the researches of Fisher and Zahavi. In the last decades, other matter that has intrigued the scientific community is the reasons that lead an individual to join a herd and how to describe the movements observed in these herds. Some authors, among which Reynolds, suggest that the flocking behaviour emerges through simple motion rules.

Afterwards, hybrid algorithms combining the ideas from evolution and motion of flocks were presented. In such algorithms, the motion rules evolve over the time. Inspired in these algorithms, in this dissertation is proposed a new algorithm, which inserts in these hybrid models the ideas from sexual selection, in particular mate choice. In the end, the emergence of a flocking behaviour is expected and the effects of using mate choice instead of the traditional approaches are analysed.

Keywords. flocking behaviour, evolutionary algorithms, sexual selection, mate choice, boid, artificial life, co-evolutionary, prey, predator

Acknowledgements

I would like to acknowledge my thesis advisor, professor doctor Fernando Penousal Machado, for his support, patience, and guidance. He helped me to improve my knowledge on the area and to improve the developed work. Whenever I ran into a trouble spot or had doubts, he gave me suggestions which allowed me to steer in the right the direction.

I would also like to acknowledge my colleague António Leitão for his guidance, support and patience, especially with my writing. I am gratefully indebted to his valuable comments and suggestions, which allowed me to improve the developed work. I also like to thank his help with his knowledge on the area and his revision of documents.

I would like to give a special thanks to my close friends for their ideas, suggestions and company. We were not only able to support each other by deliberating over our problems and findings, but also happily by talking about other things than just our dissertations.

Last but not least, I am truly grateful to my parents and sister for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Contents

List of Figures

List of Tables

List of Algorithms

Abbreviations

EC	**E**volutionary **C**omputation
EA	**E**volutionary **A**lgorithm
GA	**G**enetic **A**lgorithm
GP	**G**enetic **P**rogramming
ES	**E**volution **S**trategy
EP	**E**volutionary **P**rogramming
CI	**C**ollective **I**ntelligence
Boid	**B**ird-**oid** Object
PSO	**P**article **S**warm **O**ptimisation
SGA	**S**imple **G**enetic **A**lgorithm
GUI	**G**raphical **U**ser **I**nterface
XP	**E**xtreme **P**rogramming

Chapter 1

Introduction

When *Charles Darwin* gathered his notes and published *"On the origin of species"* [1], he presented a revolutionary scientific theory stating that the species evolve over time, in order to be better suited to the environment which they inhabit. In his book, he presents two main mechanisms responsible for the evolution of species: the natural selection and the sexual selection. However, while the natural selection was widely accepted as a fact by the scientific community, the sexual selection (in particular the female mate choice) was widely criticised and so, over the time, this mechanism was forgotten. It was only nearly a century later, that the sexual selection came back on the spotlight through the presentation of a few researches, in which excel the works of *Fisher* [2, 3] and *Zahavi* [4]. In the last decades, sexual selection originated several discussions around the scientific community, contributing for an increase of evidences that support this theory.

Nowadays, the sexual selection is widely accepted by the community, but it has not established itself on evolutionary computation. Some researches, among which *Miller and Todd* [5], concluded that there are advantages of using sexual selection instead of the traditional algorithms. The evolutionary computation is a research field within artificial intelligence which is influenced by the ideas of *Darwin* [1, 6] about the evolution of species through natural selection. This area was mainly influenced by the natural selection, once when the evolutionary computation was stabilised, the sexual selection was still not accepted as a core mechanism responsible for the evolution of species. The algorithms inspired by the natural selection are called evolutionary algorithms and although their differences all of them have in common the same principles. In each generation, this kind of algorithm will try to improve the fitness of a given population, via processes of selection and variation.

Another matter that has intrigued the scientific community, more specifically the ethologists, is the cooperative behaviour of several individuals that can be seen in flocks. They

have question themselves about the reasons that can lead several individuals to group together in a herd and how to describe mathematically the different movements seen in herds. A behaviour study presented by *Barnard and Thompson* [7] suggests greatest benefits for an individual to join a flock instead of surviving by itself. By the other hand, several authors, in which excel *Reynolds* [8], suggested that the behaviours seen in flocks may emerge through the application of simple movement rules in each individual that belongs to the flock.

Recently, a new kind of algorithm was proposed, combining the ideas of evolution theory and motion of flocks [9, 10]. In this new category of hybrid algorithm, the movement rules of the individuals evolve over the time, wherein the motion rules are evolving to adapt to the surrounding environment and to the neighbourhood of each individual. Therefore, it should be interesting to combine the ideas of the two aforementioned algorithms (female mate choice and motion of flocks) in one algorithm. One idea that is not yet found in the literature. In this new type of algorithm, the mating parents aren't only selected based on the fitness but also based in the preferences of each individual.

1.1   Scope

Darwin [1, 6] introduced two methods for sexual selection: male competition and female mate choice. However in his researches, he has mainly focused on the method of female mate choice, following this inclination also the few presented researches combining evolutionary computation with sexual selection were mainly focused in the second method. *Leitão and Machado* [11] proposed a new design for female mate choice inspired in what occurs in the nature, in which: the individuals pick their mates according their preferences, the mates preferences are inherited (similar with the physical genotypes and they evolve over time) and the mate selection introduces its own pressure but is subject to the selection pressure.

The motion of flocks can be described through the simple movement rules proposed by *Reynolds* [8], which are a weighted sum of three accelerations (separation, cohesion and alignment) within the neighbourhood of each individual. *Spector et al.* [9] inspired by the ideas proposed by *Reynolds* [8], they developed two models constituted by individuals of several species. Wherein, each specie struggles with each other in order to survive. The movement rules are evolving over time and since that the components of the movement rules proposed by *Reynolds* are part of the available genotype, it is expected the emergence of a motion of flocks among the population of each specie.

1.2   Research Objectives

The main objective of this dissertation is to develop a model, combining the ideas of motion of flocks with the ideas of a variant of evolutionary algorithms that uses sexual selection instead of the traditional selection methods. Therefore, there are two intentions of developing such algorithm:

- Observe the existence of cooperative movements between individuals of the same specie;

- Analyse the effects on the results of using sexual selection (female mate choice) in comparison with the traditional selection methods.

Papers suggesting the visualisation of motion of flocks in hybrid algorithms and researches suggesting benefits of using methods of sexual selection in evolutionary algorithms can be found in the literature. Therefore since this two intentions are observed in each one of these algorithms, it's expected that in the proposed algorithm, which is a combinations of the previous algorithms, it is possible to observe the same evidences.

1.3   Outline

The remainder of this document is structured as follows: the main ideas behind this research are presented in chapter 2, where the evolutionary algorithms are presented in section 2.1, highlighting the recent approaches using sexual selection instead of the traditional approaches, in subsection 2.1.4. Afterwards, algorithms able to simulate motion of flocks are presented in section 2.2.1. Subsequently, hybrid models, models which combine the ideas of two fields of artificial intelligence: evolutionary computation and motion of flocks, are referred in section 2.2.3.

The developed simulator is described in chapter 3. In section 3.1, the framework used to develop the simulator is presented and then the components of the simulator are described in the section 3.2. Afterwards, in section 3.3, the evaluation system developed to evaluate each simulation is presented. The tests performed are analysed on chapter 4, where the results of five experiments are studied in the sections 4.1, 4.2, 4.3, 4.4.a and 4.4.b, respectively. The last section of this chapter, section 4.5, presents a global analysis of the five experiments. Finally, the conclusions of the obtained results are presented in the last chapter, chapter 5. This chapter also presents the proposed future work.

Chapter 2

State of the Art

The present chapter contains the state of the art of the dissertation. This chapter is divided in two sections: the fundamental concepts of evolutionary computation, a computational field inspired by the evolution of species, are presented in the first section and, in the second section, the main ideas behind collective intelligence are presented, in particular, the algorithms able to simulate motion of flocks, such as birds and fishes.

The first section begins by introduce the fundamental ideas behind the concepts of evolutionary computation, as well how they arose and who presented them. Afterwards, the algorithms inspired by the evolution of species, the evolutionary algorithms, are introduced. Subsequently, the basic algorithm of an EA is presented, followed by an explanation of the algorithm. Each of the main components of a classic EA are described in the following subsections, presenting the most common algorithms of each component. In the next subsection, the four classical algorithms, which belong to the group of evolutionary algorithms, are presented. Afterwards, the main concepts of sexual selection are introduced in a new subsection, giving special attention to the female mate choice over mate competition. Subsequently, an overview regarding those concepts is presented, followed by the introduction of variants of evolutionary algorithms that use mate choice selection instead of the traditional methods. Finally, in the end of that section, a recent proposed design for this variant of evolutionary algorithms is presented.

The main concepts related to collective intelligence are introduced in the following section, giving special consideration to algorithms related with motion of flocks. In the first subsection, the main concepts are presented in an overview, containing the suggested greatest advantages for an individual join a herd instead of survive by itself. Afterwards, some algorithms inspired on motion of flocks, such as the Boid algorithm and Particle Swarm Optimisation, are presented. Subsequently, some models which include more than one specie in the environment are shown below, pointing the benefits that the

competition between species may bring to a simulation. Finally, a new kind of algorithm which combines the ideas of evolutionary computation with the ideas of motion of flocks, is presented and discussed, in the last subsection.

2.1   Evolutionary Computation

Evolutionary Computation (EC) [12–15] is a research area within artificial intelligence which is heavily inspired by a nature process, the evolution of species by natural selection. This natural process was first introduced by *Charles Darwin* in *"On the origin of species"* [1]. This area results from the efforts of bring together several researches that had in common the simulation of evolution, such like, genetic algorithms (GA) [16], evolutionary strategies (ES) [17, 18], evolutionary programming (EP) [19] and genetic programming (GP) [20].

The evolutionary algorithms (EAs) [12–14, 21, 22] are algorithms that, despite their differences, they share a common inspiration. The evolution is an iterative process that given a population of individuals, this process will try to improve their individuals, via processes of selection and variation, to the surrounding environment. In other words, the individuals more suited have higher probabilities to survive to the next generations and they also have higher probabilities to reproduce and thus pass their genetic material to future generations.

In an EA, each individual represents a candidate solution to a given problem, in which its fitness is classified by its outcome to the problem. It's intended that along generations the current best candidate solution owns a better fitness when compared with the solutions of the former best candidates. Once the EAs belong to the group of stochastic algorithms, there are no guarantee of finding the optimal solution of the problem. Even find a satisfactory solution can take a large number of generations. Since this kind of algorithms interact over a group of candidate solutions, the EA belong to the family of trial and error problem solvers [14, 22], which via the processes of selection and variation it is considered a global optimiser.

2.1.1   Evolutionary Algorithms

Evolutionary algorithms (EAs) are a group of algorithms which **simulate evolution**. EAs are inspired in biological evolution and took with particular emphasis methods, such as recombination, mutation and selection. These algorithms are used to **search candidate solutions** of complex problems in a given search space [23]. The EAs

are divided into several different algorithms [16–20], but there are a common model kept among themselves, figure 2.1. In an EA, the only existing interactions between individuals are through the comparison of the fitness (this during the selection methods), and during the reproductions (while crossing the genotype of individuals to create new offsprings). Except those cases, typically, there aren't interactions between individuals.
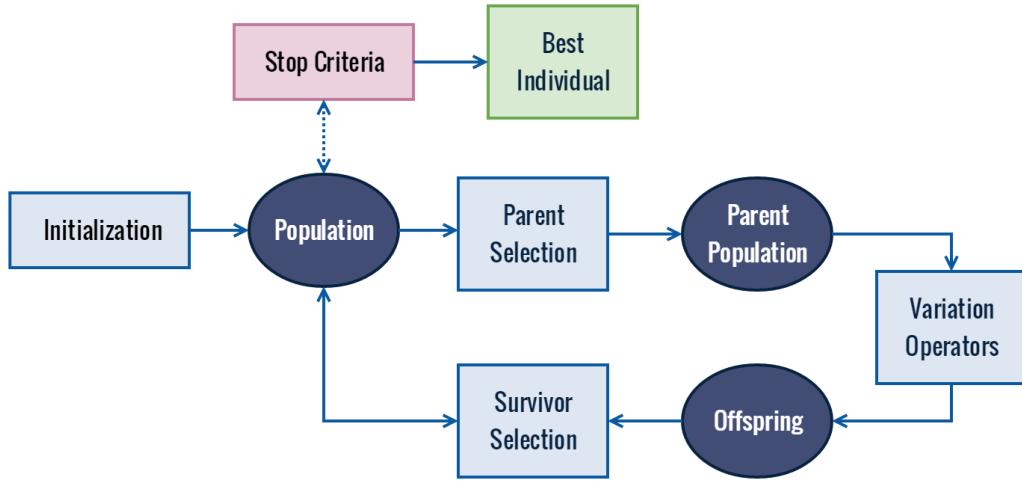
FIGURE 2.1: Common sketch of an EA

The several EAs have in common the same idea. Across generations, the population will have better suited individuals improving the fitness of the population, hence the fitness of the best individual cannot get worse when compared with the previous.

Given a complex problem, it is necessary to analyse the issue at hand, in order to find a function that describes the quality of each solution. The values returned by that fitness function will be used to evaluate and to compare the various candidate solutions. Normally, it is intended to maximise the **fitness function** and so the higher the better [12–14]. An EA begins to generate several random individuals, during the **initialisation** phase, and then, they are evaluated according the fitness function. There are various methods that can be used to generate the initial population, each one with their own characteristics affecting the diversity of the initial candidate solutions. However, they all shared the necessity to know the **search space** of the problem, this to generate the candidate solutions within that space. The search space [23] is the space of all feasible solutions, which means that an EA is only looking for candidate solutions within this space. Normally, the search space is known but when it's not, there are methods that can be used to find suitable solutions for the search space, such as hill climbing [24], tabu search [25] and simulated annealing [26]. Those solutions are not necessarily the optimal solutions but they can be used to define limits for the search space.

Based on the fitness of the candidate solutions, some of the most suited candidates are **picked** out to generate new **offsprings**. The generation of offsprings is performed in two steps: first the **recombination** (reproduction) and then the **mutation** [14]. Finally, by comparing the candidate solutions of the last generation with the offsprings, the most suited solutions are selected to **survive** to the next generation. This routine is repeated until a **stop criteria** is manifested. Finally, the candidate solution with the highest fitness is returned by the algorithm. This routine is displayed in the following algorithm, algorithm 1.

Algorithm 1: Basic algorithm of an EA
___
Data: Problem, Fitness function, Stop Criteria

Result: Best individual

Population ← RandomPopulation(Problem);

Population ← EvaluatePopulation(Population, FitnessFunction);

while *does not meet the Stop Criteria* **do**

    Parents ← SelectParents(Population);

    Offspring ← VariationOperators(Parents);

    Offspring ← EvaluatePopulation(Offspring, FitnessFunction);

    Population ← SelectSurvivors(Population, Offspring);

end

return *bestIndividual(Population)*

The improvement of the population's fitness occurs due the interrelationship between two forces [14]: the diversity and the selection pressure.

Diversity - It measures the number of different solutions. An increment in the diversity facilitates the entry of novelty in the population, which helps to further explore the search space. Due to this force, the EAs are less likely to get trapped in local optimums. This force is the main responsible for the **exploration** [27]. The exploration consists in probing the search space, with the hope of finding promising solutions which later will have to be refined. At the beginning, the diversity is promoted by the generation of the initial population and later, in each generation, by the variation operators.

Selection pressure - The selection of candidate solutions, during the reproduction and survival phases, increases the pressure between individuals, which struggle with each other in order to have a greater chance of reproduction and survival. This force is responsible by the improvement of the average fitness

of the population, through the **exploitation** [27]. The exploitation consists in probing a promising limited area (the neighbourhood of the most suited solutions) with the hope of improving the current solutions.

2.1.2   Components of EA

In this subsection, the main common components and operators used in practically every EA are presented. Each variant of the classic EAs has specific peculiarities for the components and operators that are discussed in this subsection.

2.1.2.1   Representation

One of the first steps during the development of an EA is to define how to represent the individuals, that is, link the problem context to the search space of candidate solutions. This often involves simplifying or abstracting some aspects of the real world [12–14, 23].
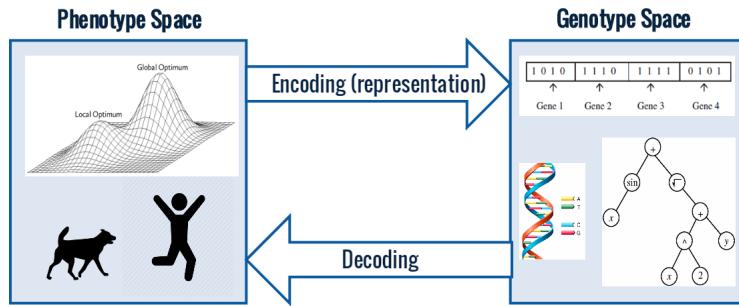
FIGURE 2.2: Spaces and mappings of candidate solutions

The possible solutions within the original problem context are called **phenotypes** and their encoding, that is, the individuals within the EA are referred as **genotypes**. The process of mapping an individual from the phenotype space to the genotype space (search space) is called **representation**, also known as encoding. Its inverse mapping is known as **decoding**. The word "representation" has also another meaning in EC, it may be used to refer the data structure used to store the genotypes.

A candidate solution can be represented by one of two forms:

- **Phenotype** - the original problem context, the object. E.g. human.

- **Genotype** - the representation of the object, the chromosome. E.g. DNA.

2.1.2.2   Population

The population [14] is a **multiset**[1] that is used to **store** candidate solutions, usually under the form of **genotypes**. These candidate solutions are **static** and therefore during the course of an EA they cannot change or adapt. It's the **population** that **changes** through the **entry** of new candidates and the **departure** of older **candidates**.

A population is defined by the representation of individuals and the size of the multiset. Normally, the **size** is **constant**, and so it does not change during the execution. However, in some variants of EA (adaptive EAs), the size of the population can change [28]. In these EAs, the size may vary according the diversity or fitness found within the population, or simply over the time.

During the evolutionary search, the population as a whole only **interacts** with the **selection operators**, in contrast with the variation operators that interact only with some individuals. In some EA oriented to dynamic environments, there are several sub-populations in which they also interact with each other, with the goal of improving their individuals. The Self-Organising Scouts (SOS) [29] and the Shift Balance Theory [30] are some examples of this variant of EA.

In other variants, the diversity within the population is measure and if a new individual is very similar to a candidate that belongs the population, then only one of them will belong to the population. The solution with the higher fitness is chosen, while the other is discarded [31].

2.1.2.3   Fitness Function

The fitness function [13–15] **measures** the **quality** of a **candidate solution** to a given problem, it represents what the solutions should adapt themselves.

The quality function **represents** the **problem** to be solved. Normally, this function is **composed** by a **decoding function** to recreate the phenotype, followed by a **quality measure** in the phenotype space.

The value returned by the fitness function will be used to evaluate and to compare the various candidate solutions to a given problem. That value will be **used** during the **selection operators** and at the end of the evolutionary search to retrieve the most suited individual among the population.

---

[1]A multiset is a set in which an object can appear more than once.

2.1.2.4 Initialisation

Usually, the initialisation is kept as simple as possible, where the first population is seeded by **randomly generated** candidate solutions. Depending on the problem, it may be possible to use some **heuristics** [14, 32] to generate the initial population, resulting in a higher average fitness. However, these heuristics are computationally heavy and, therefore, they should only be used when necessary.

The simplest method used to randomly generated solutions is the **random uniform sampling** [14, 32], where the candidates are generated randomly within the search space. This method does not distributed the individuals equally in the search space, which leads to a lesser diversity in the population. There are other methods that distribute better the individuals, such as **simple sequential inhibition** [32] and **latin hypercube** [32].

2.1.2.5 Parent Selection

The parent selection [12, 14], also known as mate selection, is a procedure that **picks individuals** based on their fitness, this to allow them **to be parents** of new individuals. This procedure can be **deterministic** or **stochastic**. In EC, the parent selection is usually stochastic, given two individual, the individual with the worst fitness can eventually be chosen but with a lower probability, in order to the search does not get too greedy. In the determinist methods, given two individuals, the individual with the higher fitness will always be the one chosen. Some of the most common algorithms used for mate selection are described below.

Roulette Wheel [14] - The probability of choosing each candidate solution is calculated according the cumulative probability distribution. Afterwards, a random real within the interval of probabilities, $[0, 1]$ is picked. That number matches a certain individual, which will be chosen to be a parent. This method is repeated until a certain desired number of mates (n) are chosen.

Stochastic Universal Sampling [14] - This algorithm is similar with the previous, in the way that the probability of each candidate is calculated according the cumulative probability distribution. However it's only need to run once to get n parents. This because, a random real is picked within the interval $[0, 1/n]$ and to this number is cumulatively added the value $1/n$ until there are n cumulative probabilities. Each of these matches a certain individual.

Tournament Selection [14] - This algorithm starts to pick randomly a certain number (k) of individuals from the population. Then in the deterministic version

the individual with the highest fitness is chosen, in the stochastic versions the others solutions can also be chosen but with a lower probability. This method is repeated until a certain desired number of mates (n) are chosen.

2.1.2.6   Variation Operators

The variation operators [12–14] consist in operations that are used to **create new individuals** from the individuals selected during the mate selection. These operations are divided in two types depending in the number of parents:

- **Mutation** - generates a new individual from only one parent.

- **Recombination** - generates a new individual from more than one parent.

Usually, the recombination is applied first to generate offsprings and then the mutation is applied over those new individuals. The variation operators [12–14] are usually applied probabilistically according a **mutation rate** and a **recombination rate**.

The generation of offsprings can produce **infeasible solutions**. Solutions which infringe any existing restriction or that don't belong to the search space. Some EAs have implemented a repair algorithm [13, 14], that in the presence of infeasible solutions, the repair algorithm corrects them into feasible solutions.

Recombination

The recombination is an operator that **blends** the **chromosomes** of two or more individuals in order to generate new individuals. Consequently, the new individuals have features in common with each of their parents.

Normally, the recombination occurs between two parents [14], as it can be observed in nature. However in an EA, the recombination may also occur between more than two parents, depending on the issue at hand.

Mutation

Given a genotype of one individual, this operator **modifies randomly** its **genes**. So the genotype of the new individual results of a chain of random choices [14]. Normally, it is intended that the new individual does not differ a lot from the original individual.

2.1.2.7 Survivor Selection

Usually, the size of the population remains constant during an evolutionary search. As the **size** of the population is **limited**, after a generation it is necessary to **choose** which **individuals** (initial population + offsprings) will **survive** to the next generation.

At this stage, it is intended that the most suited individuals survive but it's also intended that some of the individuals with a worse fitness survive, in order to inject novelty in the population and thus increase its diversity. Some of the most common algorithms used for the survivor selection [12] are listed below.

Generational [14] - The older population is discarded and only the offsprings survive to the next generation.

Elitism [14] - Considering n the size of the population and $k < n$. The k most suited candidate solutions of the older population are picked out to survive to the next generation. Those are grouped with the $n - k$ most suited individuals of the offsprings.

Steady State [14] - This algorithm starts by grouping the older population with the offsprings in a multiset. Then the n individuals with the higher fitness are picked to survive to the next generation.

2.1.2.8 Stop Criteria

An EA is a computationally heavy algorithm, which can take a **long time** to return a **reasonable outcome**. As referred before, EAs are **stochastic** algorithms which means that there is **no guarantee** of reaching the **ideal solution**. Reason why it is necessary to have a stop criteria [14] to halt the evolutionary search.

The most common stop criterias are: a maximum limit number of generations; until the best fitness, average fitness and/or diversity exceed a certain value; and until the best fitness, mean fitness and/or diversity remains under a determined threshold during a certain number of generations.

2.1.3 EA variants

The EAs are a type of algorithms heavily inspired in evolution that have a common model among them. However each of their variants have their particularities, altering some components of the traditional EA.

In one kind of those variants during the variation operators, besides the offsprings are also added new random individuals (**immigrants**) [14]. This mechanism provokes an increase in the diversity. Other algorithms deal with a **dynamic environment** [14], in other words one environment wherein the problem may change over the time. In the case of a dynamic environment, the algorithm must have mechanisms to increase the diversity when a change is detected, favouring the exploration [27] under the exploitation [27]. In other algorithms, the **mutation rate** and/or **recombination rate** may **change** during an evolutionary search, depending on the value of diversity, best fitness and/or average fitness.

There are **various** kind of **algorithms** that belong to the group of EA being inconceivable mention them all. For that reason only some of them are referred in this subsection, giving relief to the **classical EA algorithms**. It is worth mentioning, that each of those classical algorithms gave rise to countless new algorithms. In which some components may be different from the original algorithms. However in the following subsections, only the classic algorithms are mentioned.

2.1.3.1   Genetic Algorithms

Genetic algorithms (GA) [12, 14, 16] were first proposed by *John Holland* [16] in 1975. In the first years, this approach was mistakenly considered as an optimisation method. However, the thesis of *De John* [33] helped to define what today is considered the traditional genetic algorithm. The features which distinguish the GA proposed by *Holland*, SGA, of the other EAs are described in the table below.

Representation	Bit-strings
Recombination	1-Point Crossover
Mutation	Bit flip (low probability of mutation)
Parent Selection	Fitness proportional (Roulette Wheel)
Survival Selection	Generational

TABLE 2.1: Main features of the Simple GA (SGA)

2.1.3.2   Evolution Strategies

In the early 1960s, evolutionary strategies (ES) [12, 14, 16] were introduced by *Rechenberg* and *Schwefel* [17, 18]. This type of EA is specialised in self-adaptation, by having a method to regulate the standard deviation, σ, of the Gaussian distribution. The values returned by the Gaussian distribution are then used to mutate individuals [14]. ESs

have more one singularity, during the survival selection each individual is not compared with the real fitness but with a fitness raking. Also, the survivor methods are different from the mentioned above. The most common methods are: the (μ, λ) which discards the older population and only the stronger offsprings survive and the $(\mu + \lambda)$ in which the offsprings are compared with their parents and only the best individuals survive.

Representation	Real-valued vectors
Recombination	Discrete (one of the parents) or intermediary (average of the parents)
Mutation	Gaussian perturbation (adjustment of σ by 1/5 success rule)
Parent Selection	Uniform random
Survival Selection	Deterministic elitist replacement by (μ, λ) or $(\mu + \lambda)$

TABLE 2.2: Main features of an ES

2.1.3.3 Evolutionary Programming

The Evolutionary Programming (EP) [12, 14] was introduced by *Fogel et al.* [19] in 1960. In first versions of this paradigm, the genotype was represented through finite state machines, but nowadays the traditional EP uses real-values vectors as representation [14]. Similarly with the ES, the EP is also specialised in self-adaptation and both have the same representation. However the variation operators and the selection methods are different. In an EP, all individuals are chosen to generate offsprings, which are generated only by mutation. Therefore each individual creates a new individual via mutation. The survivor selection begins by grouping the offspring with the parents, in a multiset, and then the survivors are chosen through tournaments.

Representation	Real-valued vectors
Recombination	None
Mutation	Gaussian perturbation (adjustment of σ)
Parent Selection	Deterministic (each parent create one offspring)
Survival Selection	Stochastic Round-Robin tournaments $(\mu + \lambda)$

TABLE 2.3: Main features of an EP

2.1.3.4 Genetic Programming

Genetic Programming (GP) was first introduced by *Barricelli and Nils Aall* [34] in 1954, but the main pioneer of GP was *John R. Koza* [20]. Unlike most of the other EAs, which are typically used to optimise problems, GPs are usually used to breed computer programs, so they can be classified as an algorithm of machine learning. While most of

EAs are used to find an optimum point of a problem, GPs are used to search models for an optimum fit [12, 14]. To accomplish such task each program of a GP is represented through a parse tree, in which each component of GP iterates over that representation.

Representation	Tree structures
Recombination	Exchange of sub-trees
Mutation	Random changes in trees
Parent Selection	Fitness proportional
Survival Selection	Generation replacement

TABLE 2.4: Main features of a GP

2.1.4   Sexual Selection

Since young, *Charles Darwin* showed great appreciation for natural history which possibly led him years later to embark on a five year survey voyage on board of the HMS Beagle. *Darwin* had a unique opportunity to closely observe principles of botany, geology and zoology. Over the voyage, he collected and observed a variety of specimens, including birds and fossils. Afterwards on board of Beagle, he described what he saw in letters and speculated about the similar features seen in different specimens, what lead him to speculate about the origin of life. After the voyage on Beagle, he gathered his notes and published *"On the origin of species"* [1], where he introduced a revolutionary scientific theory, the **evolutionary theory**. That theory holds that the species evolve over time, in order to be better suited to the environment that they inhabit. In his book, he also presents two main mechanisms responsible for the evolution of species, those mechanisms are able to influence the survival rate of an individual between and within species. The two introduced mechanisms were the natural selection and the sexual selection, afterwards *Darwin* developed the concept of sexual selection in *"The descent of man and selection in relation to sex"* [6].

The **natural selection** is a process where the most suited individuals, to the environment they inhabit, tend to survive more time and thus produce more offsprings. Since that the most suited individuals have higher probabilities to survive and to reproduce, it provokes that there is a greater probability that their traits are passed across generations and therefore the offsprings will be better adapted to the environment. This mechanism increases the selection pressure, which provokes a greater competition within individuals from the same species, since that the most suited individuals have higher probabilities to survive and to reproduce. However, it also promotes competition between species,

once the species with most suited individuals have superior probabilities to survive and to dominate the other species.

The natural selection explains many of the traits that can be seen in the nature, but there are some traits and behaviours that this mechanism **fails to explain**. For example, the large and garish tails of some birds, such like the plumage of male peacocks, don't seem to increase their odds of survival. Also, the courtship behaviours of some species and the parental investment by the males to their offsprings that can be seen in some species don't seem to provide any gain, since that the males could use that time to reproduce with other females. By the contrary, these traits and behaviours are costly to the individuals who have to bear them. An individual with such traits has to spend more energy to maintain them and these traits may also affect their mobility comparing with individuals without any of those traits.

According the natural selection, individuals with such traits should be less likely to survive and to reproduce. Since they manifest useless traits and therefore those traits should be less likely to be observed in future generations. However that is not observed in nature, wherein several ornamentations and behaviours are visible in different populations of species, which may suggest the existence of another evolution process responsible of this fact. *Darwin* explains this phenomenon by the **sexual selection**, in others words, traits that may not provide any gain in terms to make the offsprings more adapted to the environment can be passed between generations. However it increases the selection pressure within a population, by the individuals that manifest certain traits and behaviours are more attractive and so they have a higher probability to reproduce. So the magnitude of the manifested traits and behaviours increases the mortality (since they are costly) but also increases the reproduction rates.

Darwin proposed two main processes for sexual selection: male competition and mate choice. The **male competition** promotes evolution of traits of the strongest males. Therefore, the males compete with each other in order to have a higher chance to reproduce with the females, for example this behaviour is seen in prides wherein the lions struggle to be the alpha which it's the only lion that can reproduce with any lioness of that pride. The **female mate choice** promotes the evolution of traits that the females consider attractive and so a male with such traits have a higher rate of reproduction. This process is observed in ostentations in which the peacocks with the more attractive plumage and courtship dance have greater probabilities to reproduce. *Darwin* in his researchers about sexual selection had mainly focused on the female mate choice.

2.1.4.1   Overview

When *Darwin* presented his ideas in the *"On the origin of species"* [1], it caused a shock in the society. Over the time, his ideas were widely accepted by the scientific community as the most plausible explanation for the origin of life. However, while the **natural selection** was largely **accepted** by the scientific community; the **sexual selection**, in particular the female mate choice, was widely **criticised** and so it was forgotten over the time. In part, this was due the treatment of *Darwin* over sexual selection that sometimes it was confusing and it lacked a general framework in which, he could relate the various concepts (sex-linked inheritance, sex ratio at conception, differential mortality, parent care and form of the breeding systems) [35]. Some of the authors that criticised the female mate choice, such as *Alfred Russel Wallace* [36], defended that the male competition is a process that exists within the natural selection and therefore the natural selection and the sexual selection are part of only one evolutionary mechanism. It was only **nearly a century later**, that the sexual selection **came back** on the **spotlight** through the presentation of a few researches, in which excel *Fisher* [2, 3] and *Zahavi* [4].

It was only after the works of *Ronald Fisher* [2, 3] that the sexual selection fell again in the public domain. He presented new important concepts, among which the *sexy son hypothesis* (females tend to choose the males with higher probabilities of reproduction rate, also the care giving and other direct benefits are appreciated), *Fisher's principle* (the sex ratio of most species through sex selection is about 1:1) and the *Fisherian runaway* (male ornamentations are the result of a positive feedback ("runaway") by the preferences of females for those exaggerated ornamentations).

Years later, *Amotz Zahavi* [4] consolidated the ideas presented by *Fisher*. By introducing, the *handicap principle* which holds that the traits and behaviours that don't seem to contribute for the survival of a specie can be seen as indicators of fitness. Once those traits and behaviours are costly, they can be seen as handicaps. So those handicaps couldn't be maintained by weak individuals and therefore only the most suited individuals to the environment, with a higher fitness, are able to maintain such traits. One year later, *Richard Dawkins* [37] also defended this idea.

Over time, the researches mentioned above and, as well as, others researches have gained ground in the community and so the sexual selection began to gain relevance. In the last decades, these ideas originated several discussions between biologists, psychologists and anthropologists; contributing for an increase of evidences supporting sexual selection.

Nowadays, these theories are widely **accepted** by the community, but they have not established itself on evolutionary computation.

Evolutionary computation is a research field within artificial intelligence which is heavily influenced by the ideas of *Darwin* [1, 6] about the evolution of species through natural selection. The sexual selection wasn't able to influence the development of EAs, once when the evolutionary computation was stabilised the sexual selection was still not accepted as a core mechanism responsible for the evolution of species. In the last decades, some researches tried to bring the concepts of **sexual selection** to **EC**, instead of the traditional parent selection methods, they used methods based on sexual selection, in particularly based on mate choice. The results of those experiments [38, 39] suggest that they **outperformed** the **classic strategies**. Algorithms based on mate choice can bring [11]: an increase in the accuracy (when mapping phenotypes into fitness), an increase in the reproduction variance (since individuals, with no survival relevance, may be able to reproduce), by the stochastic selection be able to escape from local optimums and it promotes sympatric specification, diversity and parallel evolutionary searches. Similar with the researches of *Charles Darwin* related with sexual selection, also in the evolutionary computation, the algorithms based on sexual selection were mostly influenced by the female mate choice in relation to male competition.

2.1.4.2 Mate Choice in EC

In a traditional EA, each parent is independently **chosen** based only on the **fitness value**, this method is shown in the image 2.3. Wherein, it is considering that the reproduction occurs between two organisms and so it's necessary to choose two individuals to generate a new individual.
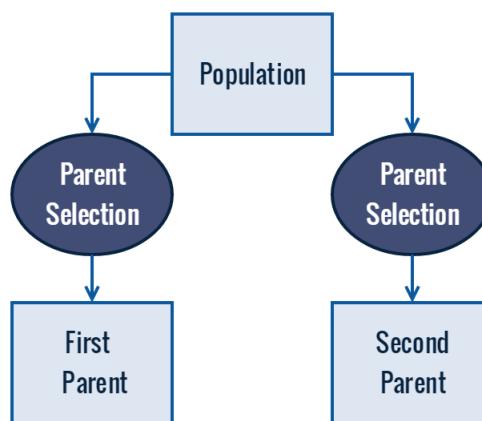

FIGURE 2.3: Classic Parent Selection

The evolutionary algorithms inspired by sexual selection through mate choice are very different from each other, but they seem to follow a same model [11], that model can be seen in the image 2.4. Most of those algorithms have in common that the **physical traits** are **inherited** from the parents and that the **females pick** their mates **according** their **levels of attractiveness**. The differences verified between those algorithms are mainly differences: in the attractiveness function, in the traits that are passed between generations (physical and/or mental) and in the reproduction of individuals (asexual or sexual).
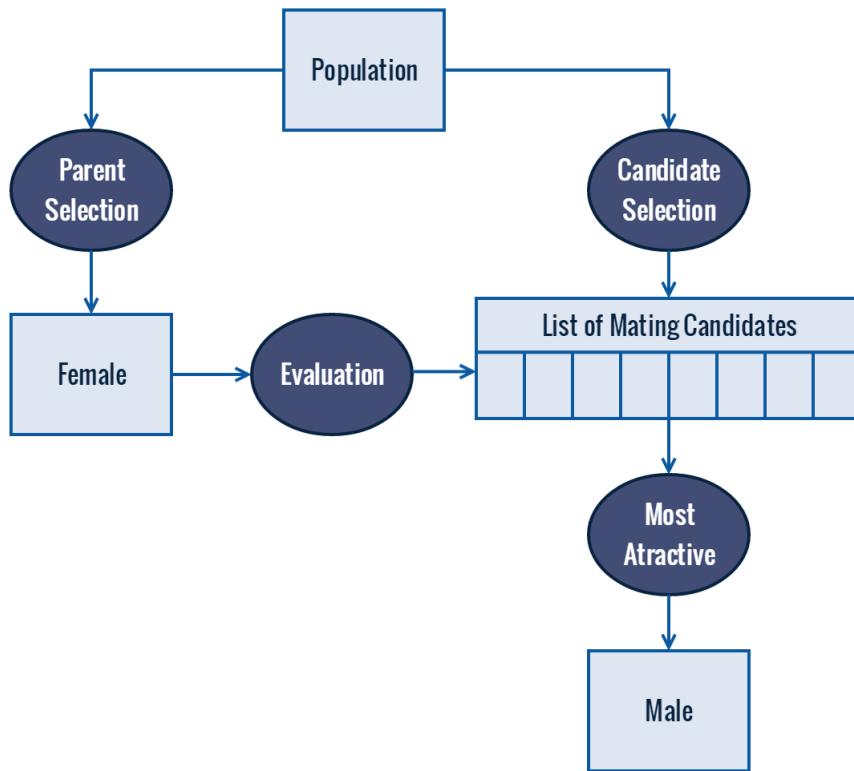
FIGURE 2.4: Parent Selection through Mate Choice

In the traditional EAs, picture 2.3, the mechanism for mate selection is used to pick the parents among the population, this only based on the fitness values. In the EAs influenced by mate choice, picture 2.4, the parent selection is a more complex process wherein they are selected based on the fitness but also in the mating preferences.

The algorithms inspired on mate choice began to use a similar method with the classic selection methods to **select** the first parent, the **female**. Then a predetermined number of **candidates are selected** from the remaining population, the method used to pick those candidates may be similar with the method used to pick the first parent. Some authors [40] presented algorithms that used exactly the same method to select the first

parent and then to select the candidates. Finally, the **candidates** are **evaluated** according to the features that the female considered attractive and the candidate with the **higher level of attractiveness** is **chosen** to be the second parent. This routine is described in the following algorithm, algorithm 2.

Worth mentioning, that the methods used to select the candidates are usually stochastic this to avoid the return of the same individuals when this method is called.

Algorithm 2: Mate Choice algorithm

Data: Population, Number of candidates (n)

Result: Parent_1, Parent_2

Parent_1 ← SelectParent(Population);

Parent_2 ← SelectCandidate(Population);

Most_Attractive ← evaluateCandidate(Parent_2, Parent_1);

for $i \leftarrow 2, n$ **do**

    Candidate ← SelectCandidate(Population);

    Attractiveness ← evaluateCandidate(Candidate, Parent_1);

    if *Attractiveness > Most_Attractive* **then**

        Parent_2 ← Candidate;

        Most_Attractive ← Attractiveness;

    end

end

return *parent_1, parent_2*

2.1.4.3 Mate Choice variants

A female selects an individual among several candidates according its level of attractiveness. Then the variation operators are applied to the couple, in order to generate offsprings. The level of attractiveness, also known as **mating preference**, is a set of traits which an individual considers attractive. Therefore, a female will pick an individual that manifests such particular traits during the mate selection. Those traits can be classified as **static**, the set of traits considered attractive does not change during its lifetime, or **dynamic**, such set of traits can change during the lifetime of that individual.

Several authors used fixed parameters and fixed mating selection functions (preferences) as **static preferences**, such as *hamming distances* [41, 42], *Euclidean distances* [42, 43] and using the *fitness value as distance* [41, 42]. Another two of those selection functions are the *assortative mating selection* [44] and the *disassortative mating selection* [44]. These two functions are opposites, while in the assortative mating selection the individuals with similar genotypes and/or phenotypes are more often paired than others.

In the disassortative mating selection, the individuals with dissimilar genotypes and/or phenotypes are more often paired than others. The difference between those functions are the traits that each individual considers attractive.

Static preferences can achieve competing results and desired behaviours. However, dynamic preferences can be extremely valuable. The dynamic function can be divided in three distinct groups [11]: **deterministic**, **adaptive** and **self-adaptive**.

Deterministic approaches are the least common of the dynamic preferences. *Ratford et al.* [45] proposed a deterministic preference, in which dissimilar individuals are favoured in the first generations, but afterwards, they are increasingly penalised over the generations.

Adaptive approaches entrust on the information obtained from past generations to control the parameters and/or the preferences. *Fry et al.* [46] introduced a strategy in which the probability of an individual, for a mate choice operator changes according the relative success of generating better offsprings, this in the previous generations. *Sánchez-Velazco and Bullinaria* [47] presented a model which combines three metrics: age, fitness and likelihood to produce offsprings. While the age and the fitness are deterministic metrics, the likelihood to produce offsprings corresponds to the feedback of other individuals.

The two previous groups don't correspond to what is observed in nature, wherein similar to what happens with the physical traits also the mating preferences change over time, during the evolutionary process. The concept of **self-adaptive** is the group most similar to what is seen in nature, wherein the parameters and/or the preference are encoded in individuals as it happens with the encoding of a solution. In that way, the parameters and/or preferences may change with the evolutionary process, adapting to the surrounding environment. *Fry et al.* [46] presented a second model wherein its probability of selecting a mate partner is also encoded in each individual. This probability is inherited by offsprings and through the comparison of fitness, it can be adapted. *Galan et al.* [43] introduced an algorithm that sorts the mating candidates by a Euclidean distance and by the fitness values. Then it chooses the candidate placed in a encoded position, that position is encoded in the first parent together with the genotype, which can be seen as an extra gene.

Looking at nature, the mate choice [11] is performed by the female individuals. Due to the costly **reproductive investments** of females (gestation and parental care), they pick males that can provide the better genes in order to create most suited offsprings. By other side depending on the species, some males want to reproduce with the largest

possible number of females, in order to increase the number of offsprings which share their genes. In other species, behaviours like care of youth by males are considered attractive, and so, also the males have reproductive investments. Normally, those males only reproduce with one female by mating season, once they have to invest their time to provide care of offsprings.

In the EAs influenced by mate choice, there are various approaches to establish the **gender roles** of each individual. The most common approaches are: the genre is determined randomly at the beginning of each generation [48], the genre is attributed alternatively during the reproduction cycle [49], the gender role is decided during the initialisation and then it is static [48, 49] and each individual participate as female during the female selection and then the remaining individuals participate as males in the candidate selection [40].

Leitão and Machado [11] suggested a **new design** for **female mate choice** inspired in what occurs in nature. The model proposed by them follows three nature inspired rules:

- Individuals choose their mates according their own mating preferences;

- Mating preference (mental traits) are inherited and they evolve from the mating preferences of their parents, as occurs with the physical traits;

- Mate selection introduces its own selection pressure but it is itself subject to the pressure of selection (once a male is selected from the previous selected candidates).

In this new proposed design, each individual is composed by two chromosomes, one for the physical traits and other for the mental traits. They suggested **two** different **models** to **represent** the **mating preferences** (mental traits): the first, by representing the chromosomes of the ideal mating (the individuals are mapped into the phenotype space and then, they are compared with the ideal mating, the most similar candidate is chosen to be the other parent), and the other, by representing an evaluation function (it's necessary to design a perceptive system to evaluate the candidates according to the number of displayed traits).

Similar to the work of *Holdener* [40], in their approach each individual can take the role of female or male. The female is chosen among the whole population by a traditional parent selection method and the male candidates are chosen randomly among the remaining individuals. This approach allows to select individuals with a low probability of survival, which can bring novel genes to the population.

2.2   Collective Intelligence

Another matter that has intrigued the scientific community, in particular the ethologists, is the formation of high structure groups of animal seen in nature. Those groups result from the grouping of several individuals of a same specie in a flock and from the movement of each individual emerges a global behaviour in the flock. The scientific community [50], have question themselves about the reasons that can lead several individuals to group together in a herd and how to describe mathematically the different movements seen in herds.

A behaviour study presented by *Barnard and Thompson* [7] suggests **greatest benefits** for an individual to **join a flock** instead of surviving by itself. The main proposed advantages, for an individual join a herd instead of survive by itself, follow:

- Avoiding predation ("safety in numbers");

- Deterrence (groups are able to defend themselves from predators);

- Easier to find resources;

- Easier to find a sexual partner.

In some flocks of animals, such as birds and fishes, seem to exist a strong coordination between all the individuals that belong to the flock, which may suggests that each individual is aware of all the remaining individuals. Some authors [8, 51] proposed that the **global behaviour** seen in flocks may **emerge** from **simple local rules** that are applied at each individual. From these simple local rules, a global behaviour emerges in the flock. *The whole is greater than the sum of its parts.*

Over the time, some algorithms influenced by the motion of flocks of some species were presented, although the behaviours of such algorithms are simpler when compared to the behaviours observed in nature. Some examples of such algorithms are: Particle Swarm Optimisation (PSO) [52], Ant Colony Optimisation (ACO) [53] and Bees algorithm [54].

2.2.1   Motion of flocks

In 1980s, different researches were presented suggesting that coordination of flocks may be achieved by an application of mathematics nonlinear dynamics [55] and that the behaviours seen in flocks may emerge from individuals following simple rules of movement [51]. Simultaneously but independently, *Reynolds* [8] presented a computer simulation

representing the flight of a flock of birds. In that simulation, the behaviour of each individual (Boid[2]) was determined by the **weighted sum** of **three simple movement rules**, based on geometrical calculations.

1. **Collision Avoidance** - avoids collisions with nearby Boids;

2. **Flock Centering** - attempts to stay close to nearby Boids;

3. **Velocity Matching** - attempts to match velocity with nearby Boids.

The following picture was withdrawn from *"Bio-Inspired Artificial Intelligence"* [56] and it describes the simple movement rules proposed by *Reynolds* [8], figure 2.5.
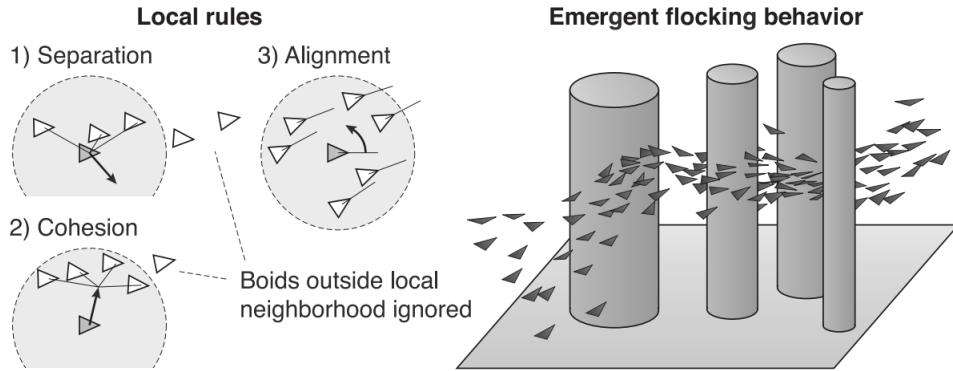

FIGURE 2.5: Simple movement rules proposed by Reynolds

The algorithm presented by *Reynolds* [8], algorithm 3, allows to simulate the motion of flocks. It starts by initialise the population. Afterwards, it determines the neighbourhood of each Boid, which are later used to calculate the three simple movement rules (separation, cohesion and alignment). Then those accelerations are added to update the velocity. Finally, the position of each Boid is updated using the new velocity.

---

[2]*Reynolds* refers to each individual as Boid

Algorithm 3: Boids algorithm

Data: Environment, Stop Criteria

Result: Flocking Simulation

Population ← initialisePopulation(Environment);

while *does not meet the Stop Criteria* **do**

> **for** *i ← 1, Population* **do**
>
>> *Neighbourhood ← getNeighbourhood(Population[i], Population);*
>>
>> *v1 ← separation(Population[i], Neighbourhood);*
>>
>> *v2 ← cohesion(Population[i], Neighbourhood);*
>>
>> *v3 ← alignment(Population[i], Neighbourhood);*
>>
>> *Population[i].velocity ← Population[i].velocity + v1 + v2 + v3;*
>>
>> *Population[i].position ← Population[i].position + Population[i].velocity;*
>
> **end**

end

Couzin [57], in his researches, explored different weights for those three simple movement rules (separation, cohesion and alignment). He discovered that **depending** on the **weights** given to each rule, **different behaviours emerged** in the population, figure 2.6 [58]. The main observed behaviours that emerged, depending on the weights of each of those simple movement rules, are: **swarm**, **torus** and **flock**.
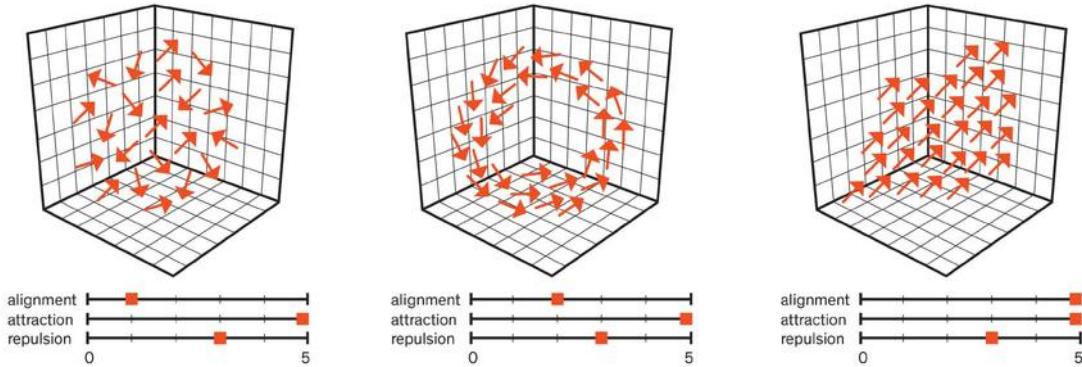
FIGURE 2.6: Different behaviours observed by *Couzin*

The motion of animal herds, in particular of birds and fishes, have originated other algorithms which use different approaches. *IL Bajec et al.* [50] introduced an algorithm based on *fuzzy logic*, which allows to use uncertain knowledge and it does not requires a transition from linguistic descriptions to mathematical formula. In 1985, two years before the work of *Reynolds* [8], a short animation called "Eurythmy" [59] was presented showing motion of a flock of birds. However, the algorithm used to create that simulation

was not specifically intended for flock modelling, since it utilised a "force field animation system" to create all the behaviours seen in that simulation.

Kennedy and Eberhart [52] inspired in the algorithms representing motion of flocks, they presented a new algorithm to **optimise nonlinear functions** by iteratively try to improve candidate solutions. The particle swarm optimisation (PSO) consists in a population of candidate solutions which collectively move in the search space, searching for the optimum point. Each individual [56] is characterised by its position, velocity and fitness.

The algorithm of the particle swarm optimisation is shown below, algorithm 4. It starts by randomly initialise the position and velocity of each particle. Then each particle is compared with the problem in order to get its evaluation and the solution with the highest fitness is recorded. Afterwards, until a certain restriction is found true, the velocity of each particle is updated, based on its own velocity and on an acceleration vector towards the best particle. Then the position of each particle is updated with the new velocity. In the next step, each particle is re-evaluated and if appropriate the best particle is updated.

Algorithm 4: PSO basic algorithm

Data: Problem, Fitness function, Stop Criteria
Result: Best individual
Population ← RandomPopulation(Problem);
Population ← EvaluatePopulation(Population, FitnessFunction);
BestParticle ← findBestParticle(Population);
while *does not meet the Stop Criteria* **do**
    for *i ← 1, Population* **do**
        *Population[i].velocity ← Population[i].velocity + randomPositiveNumber * (Population[i].position - BestParticle.position)*;
        Population[i].position ← Population[i].position + Population[i].velocity;
        UpdateEvaluation(Population[i], FitnessFunction);
        BestParticle ← compareEvaluation(Population[i], BestParticle);
    end
end
return *BestParticle*

The basic algorithm of PSO has been the root for the creation of new variants. The main differences found in those variations are: different neighbourhoods (geological vs social and/or global vs local), consider the best particle of the neighbourhood instead of the global best particle, implement also a random component, introduce a maximum velocity, give a weight to each component when the velocity is updated and also consider the previous best position of each particle.

2.2.2 Predator-Prey Competition

In the nature, a **balance** tends to **exist between** the numbers of **preys** and **predators** within the same environment. They evolve together over time, this to help them to survive in the environment through the development of features. While the **preys develop traits** and behaviours to help them **to escape** from predators, the **predators develop traits** and behaviours to be more effective at **catching** their preys. Some preys group together in order to improve their odds of survival. Some of the benefits for an individual to join a flock instead of surviving by itself were mentioned earlier (behaviour study of *Barnard and Thompson* [7]). Similar to the behaviours observed in some preys, also some predators use similar strategies to help them catching their preys. A fine example of this occurrence is the pack of wolves, in which through the coordination of several wolves, they are able to hunt animals bigger than them, which would be impossible if they were alone.

Reynolds [8] suggested for future work that would be interesting to include other components in his model, such as the inclusion of other types of Boid, e.g. preys and predators, and rules for a Boid seeks food resources. Years later, *Lee Spector et al.* [9] included some of those behaviours in their swarm simulation. Their simulation allows the existence of **several species**, wherein each specie **competes** with the **others** in order to have **access** to **food resources**. The food resources are scattered in the environment and are finite, the sum of food resources remains constant by the appearance of new food resources. The observed species can be categorised as preys, since they don't attack the others and they only gain energy from the food resources. Each Boid belongs to one of the existing species and its movement is decided by a set of four basic rules (separation, cohesion, alignment and a vector towards the closest food resource). Therefore, a specie to survive has to track its food resources and should try to control the food resources in order to deny their access to Boids from other species.

Adam Birt and Samuel Shaw [10] presented a similar simulation but with preys and predators. They have studied the impacts that the addition of predators might bring to an environment of Boids. Their results suggest that if the number of predators is suitable to the population of preys, it seemed to stabilise the population maintaining the gene value high.

2.2.3 Evolutionary Behaviour

Various authors tried to **combine** the ideas of the algorithms of **motion of flocks** with the **evolutionary algorithms**, creating new hybrid models. One of the biggest differences between the algorithms representing the motion of flocks and evolutionary algorithms happen at the level of interactions between individuals. In EAs, the individuals don't interact with others except in selection methods and variation operators. However, in the flocking algorithms there are a large interaction between particles, once the movement of each particle depends on the movement of other particles.

Several authors [60–63] proposed hybrid models which combine the ideas from **EAs** and **PSO**. They have in common the fact that both algorithms are used to optimise non-linear functions. The new hybrid algorithms had interesting results suggesting greater advantages when compared with the classic algorithms, which have a faster convergence and outperform the classic algorithms.

Other researches [9, 10, 64] proposed hybrid models which combine **EAs** with **Boid algorithm**. The approaches chosen by most of them were similar, such as *Reynolds* [8] suggested the movement of each individual is described through a **weighted sum** of **simple rules**. However the weights of each rule are not equal for all individuals, where the **weights evolve** through a **genetic algorithm**. *Adam Birt and Samuel Shaw* [10] introduced a similar model but with preys and predators, in their simulation only the preys evolve over generations, the predators are limited to chase the closest preys. *Lee Spector et al.* [9] presented two new models for swarm simulation. Their simulations have several species, wherein each specie competes with the others in order to have access to food resources. Each specie has to evolve aiming to track the food resources and also attempting to deny the access of food resources to Boids from different species. The first model is similar to the earlier described. The other model is similar with a combination of a **GP** with the **Boid algorithm**, it is used to evolve programs. In other words, it is used to **evolve the movement rules** of each Boid. However, instead of a GP they use Push [65], which is a programming language (a based-stack GP) that represents the evolving programs through stacks of finite size.

Chapter 3

Experimental Setup

In this dissertation, the author proposes himself to develop a new model for **swarm simulation**, which in addition of the **combination** of concepts of **evolutionary algorithms** and of algorithms inspired in **motion of flocks** (Boid algorithm), it also combines the concepts of **sexual selection**, more specifically the concepts of female mate choice. The proposed model consists in a group of multi-agents, where the agents can belong to one of **two species** (preys or predators). The individuals of both species have similar features, such as a set of motion rules, position, velocity, maximum velocity, acceleration, maximum acceleration, vision angle, maximum steering angle, energy and a function describing the energy consumption.

As it happens in nature, in the proposed model, **both** species **evolve** over time in order to become better **adapted** to the **environment** and to help them to improve their odds of **survival against** the adaptations of the **other specie**. The **preys** have the intentions to develop traits that help them to catch the food resources and to escape from predators; while the **predators** aim in traits which help them to be more effective at catching preys. As the observation of the nature suggests in the long term, the existence of a **balance** between the number of preys and predators is expected, which means that both species are adapted against the other specie and also adapted to the surrounding environment.

In the long run, it is excepted the **emergence** of a **coordinate behaviour** among all agents of each specie, once the **simple motion rules** proposed by *Reynolds* [8] are part of the **available genotypes** to form the set of motion rules for each particle. Subsequently, the observation of nature and also the literature [7] suggest the existence of biggest benefits for an individual to join a flock instead of surviving by itself, such as by working together is easier to find and catch their food resources and to escape from predators.

Algorithms inspired in both **EC** and **motion of flocks** are **not** a **novelty**. *Lee Spector et al.* [9] presented two models inspired in those ideas, wherein the motion rules of different species are evolving over generations. Those species are competing with each other to have access to food resources. In one of those models, the weights of a static motion rule, a weighted sum, are evolving. While in the other model, the own motion rules of each Boid is evolving. Also, *Adam Birt and Samuel Shaw* [10] presented an algorithm, in which the preys are evolving through a Genetic Algorithm. The GA is used to evolve the weights of a static motion rule. They have also added another specie of predators, although the predators have a static motion rule that doesn't evolve. The authors found that the increment in pressure by the introduction of predators brings benefits for the swarm model, such as stabilisation of the average fitness, but this only if the number of predators is adequate for the number of preys.

However, the **literature lacks** of an algorithm which **combines** the concepts of **EAs**, algorithms inspired in **motion of flocks** and **sexual selection**. Notwithstanding, some researches inspired in sexual selection, in particular female mate choice, suggest greatest benefits in EAS when compared with the results of traditional selection methods, such as *Miller and Todd* [5] and *Penousal Machado et al.* [38, 39].

The author of this dissertation proposed to develop a **swarm model inspired** in **evolutionary algorithms** and algorithms inspired in **motion of flocks**. The parent selection, during each generation, is performed with a selection method inspired in **female mate choice** instead of the traditional parent selection methods. The effects of using sexual selection are later analysed, in order to study the effects of using such algorithm instead of the traditional methods.

3.1   Framework

The **framework chosen** to develop the simulator of this dissertation was **BREVE**[1] [66]. There are other frameworks that implement decentralised system simulations, such as Swarm [67] and StarLogo [68]. However, they are not well suited for simulations of realistic artificial life. Swarm [67] is not a single integrated application, it is a collection of libraries written in Objective-C and Java. So, it is not easy to develop new simulations. Also, Swarm does not provide frameworks for 3D simulations and visualisation. The StarLogo [68] is considered easy to use. However, the simulations are based on a 2D landscape of 'patches', and therefore, this package is more indicated for discrete models. The simulations of complex 3D models and physical models are a tough task.

---

[1]http://www.spiderland.org/s/

BREVE is a 3D **simulation** environment designed to be used in decentralised systems and artificial life. It simulates continuous time and space and the simulations can be written in one of **two** available **programming languages**: in "**steve**" (a simple interpreted object-oriented language created especially for BREVE, which borrows many features from popular languages such as C, SmallTalk and Objective-C) or in **Python**. Additionally, BREVE allows the developers to use external libraries or languages, through plugins that can be accessed in steve.

This framework allows the development and simulation of realistic creatures, having its own **OpenGL** display **engine**, **methods** for detection and resolution of **collisions**, an experimental support for articulated body **physical simulation** and mechanisms able to **simulate** the breeding and **evolution** of artificial life, such as genetic algorithm, genetic programming and Push [65]. Push is a programming language used to simulate genetic programming (a stack-based GP), where the genes are represented in a stack of finite size instead of the traditional tree structure.

BREVE allows the **customisation** of each **agent**, such as its shape, properties and constants. It also offers the **customisation** of **simulations**, where the users of the framework can change its velocity, control the visualisation camera, the light settings, the graphic rendering and user interface. The users can take snapshots of simulations and record them in a video clip. Additionally, BREVE has other features to **help** the **development** of new simulations, such as tools for logging and debugging.

One of the **biggest advantages** of BREVE when compared with its competitors is the number, variety and quality of the **examples** accompanying the **framework**, which exhibit the power of the simulation engine and language. BREVE provides **several demos**, such as game of life, Braitenberg vehicles, motion of flocks, evolution of walking behaviours and physical examples, such as fountain, demolition, and springs. From those examples **stand out two algorithms** inspired in the **motion of flocks**, "*SwarmEvolve-1.0*" and "*SwarmEvolve-2.0*". Those two examples were created by *Lee Spector et al.* [9], wherein they both are inspired in evolutionary algorithms, and thus the motion of particles evolve over time. In "*SwarmEvolve-1.0*", the weights of a fixed motion rule (weighted sum) will adapt to the surrounding environment, while in "*SwarmEvolve-2.0*", the own motion rule of each particle is evolving over generations.

3.2   Design

In this section, the main components and features of the proposed simulator are detailed. The new model combines the main ideas of evolutionary computation, sexual selection and motion of flocks. The environment is populate by agents that can belong to one of two species (prey or predator). The **simulator** was **developed** in **BREVE** simulation environment and was written in **Python** programming language.

Each **specie** has **different objectives** and they will evolve to better accomplish those objectives. While the preys have to catch food resources, that are scattered in the environment, and move away from predators. The predators have to approach preys in order to steal their energy. Each agent has a limited energy and when its energy is equal or less than 0 the agent dies and thus it disappears from the environment. Therefore, in order **to survive** the agents **have to evolve** to the surrounding environment in order to accomplish their objectives.
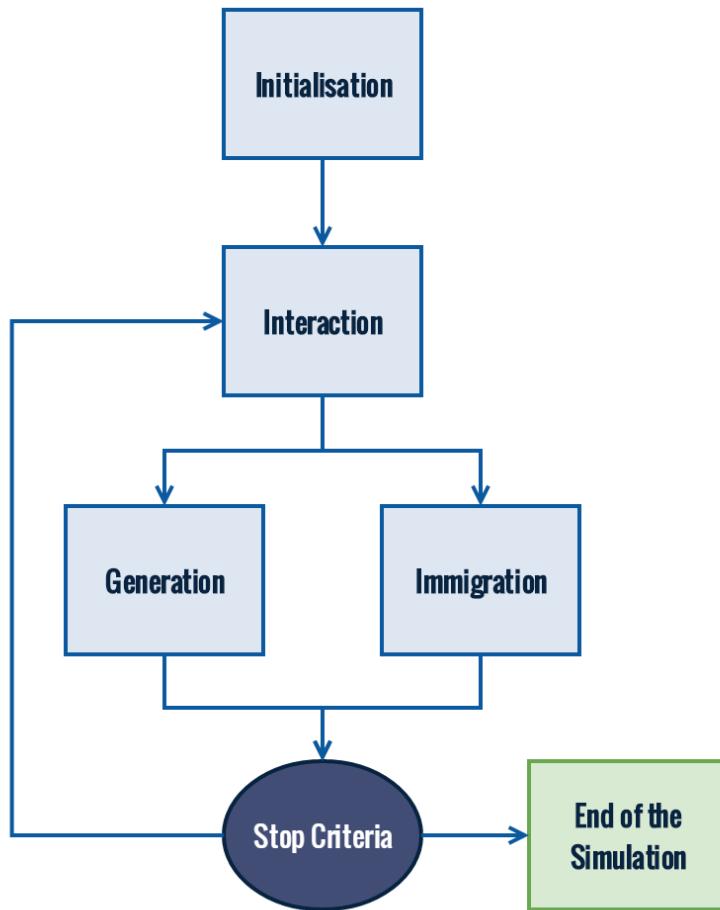
FIGURE 3.1: Draft of the proposed algorithm

The algorithm starts by **initialise** the various agents. It begins to spread the food resources around the environment and then the environment is populated with preys and predators. The size of each agent depends of its energy. Afterwards, the first cycle of this algorithm begins. In the first step, **interaction**, the position, velocity and energy of each agent is updated, according their neighbourhoods and their motion rules. The energy of each Boid is decremented in each interaction depending on its age and its level of attractiveness. Also, the energy of each agent is checked, if it is equal or lesser than 0 the agent dies and then its instance is added to a list of dead agents. There are two lists of dead agents, one for each specie. Afterwards, depending in the number of the interaction, it can be executed one of two steps: generation or immigration. The **generational** step is a periodic step that occurs between a certain numbers of interactions. In this step, new offsprings are generated, after chosen their parents. If the model does not enter in the generation step then it enters in the **immigration**. The size of each population is checked, if it is lower than a determined threshold, new immigrants are generated and scattered around the environment. One part of those immigrants consist in new random agents and, the other part, in previously dead agents that are brought back to the environment. This cycle is repeated until a certain **stop criteria** is manifested, occurrence that stops the execution of the algorithm. The proposed algorithm is represented in the previous image, figure 3.1.

It is worth mentioning the regarded meanings by the author for the words **interaction** and **generation**. Those words mean different steps, in which a **interaction** is any cycle in the simulation (any step in which agent can move). The **generation** is considered a periodically step which can occur after a interaction, depending in the count of interactions. In a generation, new offsprings are created.

The **main components** of the algorithm are detailed in the **following subsections**. The developed model has other features, such as **save** and **load simulations**. Before run a simulation, it can be decided to store a simulation or load a previously stored simulation. If it was decided to save a simulation, all the agents of that simulation will be stored in checkpoints, periodical events between a determined number of interactions, as well as at the end of the simulation. Also, the developed simulator allows to **record** a **simulation** in a **movie clip**. A simulation may take a long time until the simulation stops, reason why this feature is the most appropriate way to observe the execution of the simulation.

3.2.1   Representation

In order to represent a **Boid** (phenotype) in the model, it is necessary to have a representation for the genotype. Each Boid is **characterised** by **several parameters**, such as position, velocity, acceleration, vision angle and energy. However, since it is desired that the motion rules of each species evolve over time. The **genotype** representation has to focus in the **motion rules** of each Boid.

Besides the physical genotype, each prey has also other genotype to represent the **features** of their **tails** (size and brightness), but the tails are not represented graphically. This genotype, the **sexual genotype**, is similar with the physical genotype, in the way that both are **inherited** from their parents and they **evolve** over generations. The values returned by this structure (size and brightness) act as **handicaps**. The size affects the mobility of the preys and the brightness increases the consumption of energy. Those values will also constrain the parent selection, but only when the selection method inspired on mate choice is being used.

3.2.1.1   Physical Genotype

In the proposed model, the representation for the motion rules can be chosen between **three available representations**: **GA** (genetic algorithm), **GP** (genetic programming) or **Push** [65]. The framework has implemented methods for evolutionary algorithms (GA and GP), but the author had to develop their own algorithms of GA and GP, once the implemented methods are not suitable for the proposed model. In those methods, it is not possible for a element to interact with other elements and so, it is impossible to calculate the motion of each element since it depends of the neighbouring agents. Due their similarity, first the GA is presented, followed by the description of GP and Push.

The genotype of the **GA** consists in **real-valued vectors**. Those real values **represent** the **weights** of a **static motion rule**, wherein they are evolving in order to improve the motion of the individuals of each species. Similarly with the works of *Adam Birt and Samuel Shaw* [10] and *Lee Spector et al.* [9].

$$acceleration = \boldsymbol{w1} \cdot cohesion + \boldsymbol{w2} \cdot alignment + \boldsymbol{w3} \cdot separation +$$

$$+ \boldsymbol{w4} \cdot target + \boldsymbol{w5} \cdot randVector + \boldsymbol{w6} \cdot flee$$

The static motion rule is shown above, wherein the last expression is only applied on the preys. There are six components which consist in six vectors. **Cohesion** is an attempt to stay close to the neighbour Boids of the same species. A Boid attempts to match the velocity of the neighbour Boids of the same species, through the **alignment**. The **separation** is a vector aiming to avoid the collisions with the neighbour Boids of the same species. The **target** vector represents the direction of the closest source of energy within its neighbourhood. If the Boid is a prey then it points the closest food resource and if it is a predator, it is points the closest prey. To introduce novelty, a **random vector** is added. Finally, the **flee** rule represents the escape movement, its vector is the opposite direction to the predators that exist in the neighbourhood.

The representations based on **GP** and **Push** are similar. Both of them are **used** to **evolve** the **own motion rule** of each individual based on the motion rules of its parents. Those motion rules result of a combination of vectors of a given set of vectors with another set of operators. The structure of the genotype of a **GP** is **well defined**. The genotype is stored in a **binary tree** with a defined depth, where each **leaf** has to be a **vector** of the given set of possible vectors. The remaining **nodes** can be any of the available **operators**. Contrariwise, the data structure of **Push** is not well defined. The genotype is saved in a **stack of finite size**, where it exists a stack for each data type (vector, integer, float and string). It also allows sub-trees within the stack, but there are not a structure and so, the **operators** and **vectors** are **placed** in **random** positions. The result vector of each motion rule is stored on the top of the vector stack.

The set of operators and vectors used by GP and Push are the same. The chosen operators and vectors have been influenced by the rules proposed by *Reynolds* [8] and by the rules used by *Lee Spector et al.* [9], which beyond the rules proposed by *Reynolds*, they also used other new rules. The sets of the available operators and vectors are displayed in following table, table 3.1.

Vectors	Operators
separation	+
aligment	-
cohesion	*
target	/
flee	
randVector	
currentVelocity	
currentEnergy	

TABLE 3.1: Sets of available vectors and operators for the physical genotype

Most of those vectors were explained before. Therefore, since there are two new vectors, the currentVelocity and currentEnergy, they need to be decribed. As the name regards, the **currentVelocity** is a vector with the current velocity of the Boid and the **currentEnergy** is a vector in which all the positions have the current energy of the Boid.

Respectively the operations in vectors, while the **addition** and **subtraction** of vectors of the same size (3×1) is a **permitted** operation and it results in a vector of the same size. The **multiplication** and **division** of vectors of the **same size** are **not allowed**. Due to that motive, the operators of multiplication and division were respectively **replaced** by the **point-wise product** and **point-wise division**. The result of those operations is a vector of the same size of the input vectors.

3.2.1.2   Sexual Genotype

Usually, the traits and behaviours seen in the nature, which are associated to the **sexual selection** (female mate choice), are **observed** in **preys**. Normally, the **predators** deal with the most challenged conditions, which don't allow them to waste energy in **features** that don't help them to **protect** themselves or to **catch** their **preys**. On the other hand, the **preys** usually have most favourable conditions than the predators, which allow them to have a **larger number of individuals**. Due this increment on the selection pressure, the preys developed other methods to select their mates, the sexual selection. Those individuals to stand out of the others developed **eccentric** and **costly features** with the goal of **increasing** their **odds of reproduction**.

Due this fact, the author of this dissertation decided that the **mate selection** of the **preys** will be performed with **female mate choice**, while the **traditional methods** for mate selection will be used to select the parents of the **predators**. The **female mate choice** performed in the preys is conducted **having regard** to some features of the **tail** of each prey. The chosen characteristics of the tail are the **size** of its tail and its **brightness**. The **size** of the **tail** will **affect** the **mobility** of the prey, the greater the tail the lower the maximum speed. To maintain its **tail brighter** a prey need to **spend** more **energy** than the necessary and so, the brighter the tail the bigger the energy spend.

The chosen female mate choice was inspired in the new design proposed by *Leitão and Machado* [11], wherein the sexual genotype is **self-adapted** (it is inherited and evolves over generations) and the individuals **pick** their **mates** according their own **mating**

preferences. To accomplish such task, it was chosen that the sexual genotype is **represented** through a **tree-based GP**, which during the variation operators the operators of crossover and mutation will operate on this representation. This structure **returns** the **size** and the **brightness** of the **tail** of each prey in each interaction, also during the mate choice, those values will be used to select the mates of the preys. The genotype is stored in a **binary tree** with a defined depth, where each **leaf** has to be a **float** of the given set of possible floats. The remaining **nodes** can be any of the available **operators**. The sets of the available operators and floats are displayed in following table, table 3.2.

Floats	Operators
tailSize	+
tailBrightness	-
randFloat	*
currentEnergy	/
distanceToMatingSeason	
numPreysNeighbourhood	
numPredatorsNeighbourhood	

TABLE 3.2: Sets of available floats and operators for the sexual genotype

As the name regards, the **tailSize** returns the size of the tail and the **tailBrightness** returns the brightness of the tail. A random float is introduced through the **randFloat**. The **currentEnergy** returns the current energy of the prey. The preys are able to know the number of interactions remaining until the mating season through the **distanceToMatingSeason**. The **numPreysNeighbourhood** and **numPredatorsNeighbourhood** return the number of preys and predators, respectively, which are found in the neighbourhood of each prey.

3.2.2   Population

The population consists in a **list** of **all alive agents** from the **same species**. Those agents are scattered around the environment. The **environment** chosen for this model consists in an 2D environment with limited size, which has **wrap boundaries**. In other words, if a Boid overcomes the barriers it will appear in the opposite side of the environment. However, a Boid will only perceive the opposite side when it gets there. There are **three** kind of **agents** (**preys**, **predators** and **food resources**), which are scattered around the environment. The size of each agent will depend on its energy.

Each population **does not have any restriction** around its **size** and so, each population **grows** and **decreases according** the advances of the **other specie** and **its own features**, which should adapt to the environment.

Each Boid perceives the others according a **local neighbourhood**. Each Boid can only **perceives physical features** and so, they are unable to perceive the internal features of its neighbours, such as their motion rules. The **size** of each neighbourhood is similar between species, but its size varies according the specie. Due the advantageous traits to follow trails seen in the predators, the "target zone" of the **predators** is **slightly bigger**. The actual proportion of the neighbourhood of a prey towards the environment is shown in the following picture, figure 3.2.
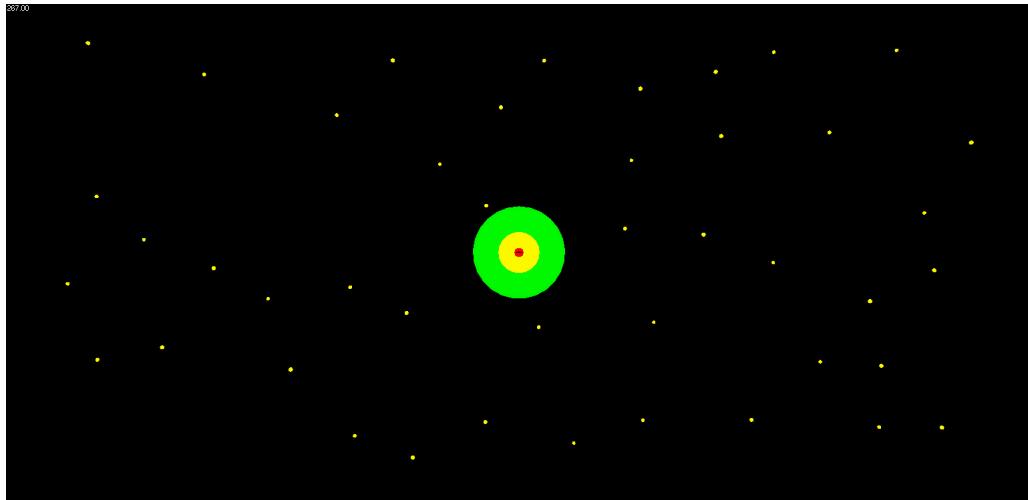
FIGURE 3.2: Proportion size of the neighbourhood of a prey towards the environment

The previous picture shows a simplistically vision of the neighbourhood, where only the size of each neighbourhood is regarded. In reality, the proposed **neighbourhood** is a **combination** of **three neighbourhoods** which **depend** on the **direction** and **vision angle** of each Boid, as well as, the **distance** to the Boid, figure 3.3.
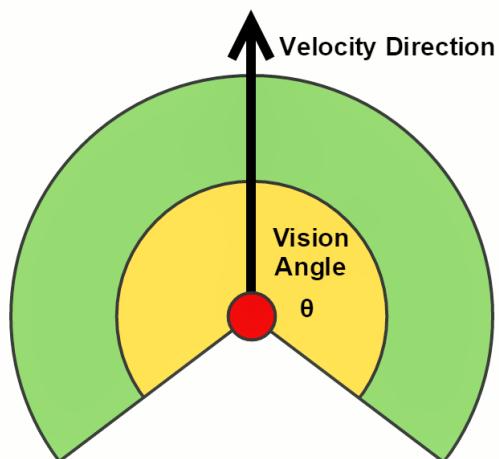
FIGURE 3.3: Neighbourhoods of each Boid

The size of each of the three **neighbourhoods** will **affect** the number of Boids which will be considered by each **motion rule**.

- **Separation Zone** - only the Boids of the same specie within the **red** neighbourhood will be considered for the **separation** rule;

- **Social Zone** - the Boids of the same specie within the **red and yellow** neighbourhoods will be considered for the **alignment** and **cohesion** rules;

- **Target Zone** - the agents within the **red, yellow and green** neighbourhoods will be considered for the **target** and **flee** rules.

3.2.3   Fitness Function

Among the various parameters that a Boid have, the **energy** is the parameter that **better represents** the **quality** of **each Boid**. However, between generations the energy can vary a lot depending of the surrounding environment and so, a Boid with a **higher energy** at the **mating season** (generational step), it's not necessary a Boid that had a **great performance** throughout all the time. Reason why the chosen fitness function for a Boid is the average of the energy between generational steps.

$$fitnessFunction = average(Boid.energy)$$

However, the **energy** is not a trivial measure of quality, once the energy is **not only subject** to the **performance** of **its own individual** but also on other factors. The chosen fitness function is **subject** to **its own survivor** which **depends** on **three factors**. The survivor of each Boid will depend on the **advances** of the **Boids** of **other specie**, since if the other Boids are adapted to the movements the opposing species, so the Boids of that specie are less likely to survive. Other factor, is the **energy** that each Boid **loses** in **each interaction**. Which increases, depending on the increase of its age or on the increase of its level of attractiveness. The last factor is the **adaptation** of the **population** of the **same specie** than the Boid, since the motion of each Boid may depend on the motion of other Boids of the same specie, which may negatively influence a Boid with a good genotype.

3.2.4   Initialisation

The initialisation is the first step of a running simulation, in this step the initial population is generated. **By default** the initial population is **created randomly**, but it can be decided to **load** a previously **stored simulation** and therefore, the simulation is able to retrieve the stored agents and to progress the simulation from the moment that was saved.

As referred before, the chosen environment consists in a limited area with wrap boundaries, area that will be randomly populated by preys, predators and food resources. The initial sizes of the populations of preys and predators are a set of given numbers. The **positions** of **preys** and **predators** are randomly generated using the latin hypercube method, which allows to distribute reasonably a predetermined number of agents in a finite known space. The **velocity** is a random real vector that is generated between the two given numbers. In the beginning, the **acceleration** starts null. The generation of the **physical genotype** of each Boid depends on the chosen representation. However, they have in common the fact that they are initialised randomly, also the **sexual genotype** is initialised randomly.

In the **GA**, each real-valued vector is populated by random real numbers between two predetermined values. In **GP**, each node can be populated with an operator or a vector. If the node was populated with an operator then he will have two child nodes and the process is repeated for those nodes. In the **Push** representation, the genotype is created randomly between the following choices: operators, vectors and creating sub-trees. The only restriction is the length size of the genotype.

The **energy** of each **food resource** is a random real number that is generated between two predetermined values, which don't differ a lot from each other. The sum of the energies of all food resources is predetermined and constant and so, the total number of food resources depends of their cumulative energy. New resources are created until the predetermined total energy is reached. Once the total **number** of **food resources** is **not known** at the beginning of a simulation, the latin hypercube method cannot be used to generate the initial positions of the food resources, to use that method is necessary to know the size of the population. In order to distribute reasonably the food resources in the environment, the **simple sequential inhibition** was chosen to generate the initial positions, this method ensures that there are a minimum distance between each resource.

3.2.5 Interaction

The first step of each interaction is the step with its same name, "interaction". This step starts by **checking** the **energy** of each agent, followed by the **update** of each **agent** (position, velocity, acceleration, energy, size of the tail and its brightness).

The energy of each **prey** and **predator** is checked, if it is **below or equal** than **zero** then the Boid **dies** and its instance is added to a list of dead instances of the same specie than it. During the immigration, these instances can return to the environment. By default but it can be changed, there is an **animation** when a Boid dies, where its shape will **gradually disappear**. The energy of each **food resource** is also checked, if its energy is below or equal than zero it will disappear from the environment. However, once it was decided that **total energy** of all food resources is **constant**. When the total energy decreases new food resources are added to compensate for the lost energy. Those new food resources are created using the method referred before, in the initialisation sub-section. However, they don't simple appear with all of their energy. Instead, there is an **animation** were the new food resources **will grow** until its real size. The real size of each food resource depends of its energy.

The update of each **Boid** (prey and predator) starts by the **update** of its **acceleration**, followed by the **update** of its **maximum velocity**. A new acceleration is calculated through the motion rules presented in its physical genotype and in its neighbourhood. Through this vector and the former **velocity**, the new velocity for this interaction is found. The acceleration and the velocity are restricted by maximum and minimum values. Finally, the previous **position** is added with the new velocity to find the new position for the Boid.

$$energy = Boid.energy + collectedEnergy - 0.01 \times (1 + logsig(\frac{Boid.age - 100}{12}) + 2 \times logsig(Boid.tailBrigh - 5))$$

The **energy** is **updated** in each interaction according the **previous energy**. At this energy is added the **energy collect of its targets**, if the Boid is close to its targets. Then a **static amount** of energy is decremented. **Depending** on the **values** of its **age** and **brightness of its tail** more energy is decremented. The functions used to retrieve those values are **sigmoid functions**, which depend on its age and brightness of its tail. Those sigmoid functions were carefully chosen, taking into account the range of values that the the inputs may vary. The values of the size and brightness of the tails may vary from 0 to 10 and the sigmoid functions are limited to return values between 0 and

1, where they return a value near 0, if the passed parameters are equal or below than 0, and with the increase of the parameters also increase the returned value.

$$maximum\_velocity = maxVel * (1 - logsig(Boid.tailSize - 5))$$

As referred before, the **size** and the **brightness** of the **tail** of each **prey** are returned in each interaction by the sexual genotype. Their values will affect the mobility and the consumption of energy by each interaction. The **size** of the **tail** of each prey will **affect** its **maximum velocity**, which can vary between zero and a given threshold. The value of the size is obtained through a **sigmoid function**.

3.2.6   Generation

Depending in the count of interactions, the generational step (mating season) can be executed. This step is a **periodic step** which **occurs between** a **certain** number of **interactions**. In this step, offsprings are generated based on the selected parents.
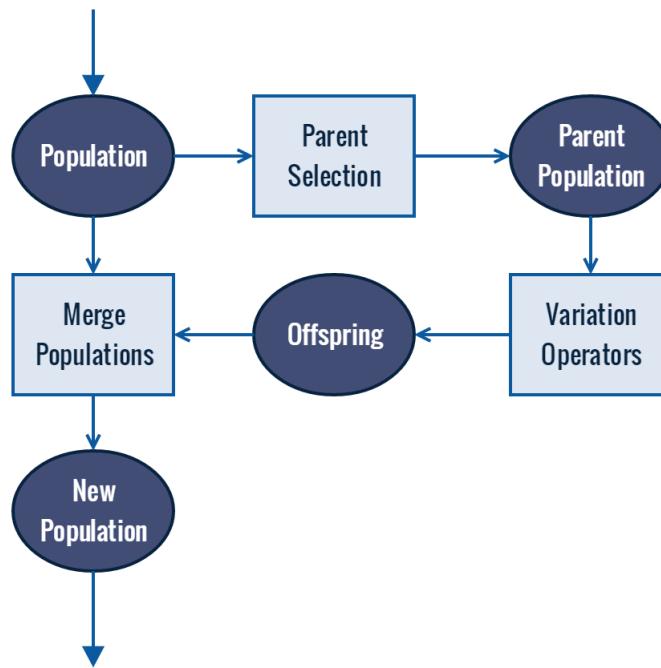
FIGURE 3.4: Generational step

3.2.6.1   Parent Selection

As observed in nature, the **preys** may **choose** its mates not only by its **fitness** but also by **other features** that it may have (mate choice). By the other side, the **predators** tend to **struggle** with each other in order to reproduce (male competition). In a EC, the

simulation of the male competition is really hard to mimic, reason why the **methods** decided to select the mates of the Boids **vary depending** on the **specie**. To the **predators**, the chosen mate selection is a **traditional method** for mate selection, similar to the method represented in picture 2.3, in which both parents are chosen through a **tournament** selection method (five elements of the same specie are randomly picked and the element with the higher fitness is chosen to be a parent).

The mate selection method used to select the parent of **preys** is similar to the method represented in the picture 2.4. The chosen mate selection method is a **female mate choice**, which was inspired in the new design proposed by *Leitão and Machado* [11]. This mate selection method introduces its own selection pressure, by the selection of the most attractive candidate, but it is itself subject to the selection pressure, by the selection of candidates.

The chosen mate selection method starts by **choosing** the first parent (**female**) through a **tournament** selection method (five elements are randomly picked between all the preys and the element with the higher fitness is chosen to be a parent). Afterwards, **five candidates** are **chosen randomly** between the remaining preys. The **tail's traits** (size and brightness) of the **candidates** are **evaluated with** the traits of the **female**, and the candidate with the **slightest difference** is chosen to be the other parent (**male**). As referred before, the traits of the tail act as handicaps, since they constrain the preys. The size affects the mobility (maximum velocity) and the brightness constrains the spent energy by interaction.

3.2.6.2   Variation Operators

The variation operators are used on the parent population to generate new individuals. First, the recombination method is applied to generate offsprings and then the mutation operator is used on them. The used methods for **recombination** and for **mutation** will **depend** on the **representation** chosen by the user, this for the **physical genotype**. Since the author decided to represent the **sexual genotype** only by a tree-based GP representation, there are **only one method** for the **recombination** and other for the **mutation** that will operate on this data structure.

Recombination

If it was decided to use a **genetic algorithm** (real-value vector) to represent the physical genotype then the recombination method that will be used is the **one point crossover**. In this method, the vectors representing the physical genotype of the parents are likewise

divided in two vectors, the genotype of the descendants results of the grouping of the two different parts of the two parents.

The recombination method used in a **tree-based** GP representation (sexual genotype and physical genotype) is the **exchange** of **sub-trees**. One random sub-tree of a parent is exchanged by a random sub-tree of the other parent, in order to construct the genotype of the offspring. The documentation of Push lacks the description of the recombination method used by the framework, but the observed results suggest that this **stack-based** GP representation (**Push**) uses a **similar method** to the method explained above.

Mutation

If it was decided to use a **genetic algorithm** (real-value vector) to represent the physical genotype then the mutation method that will be used is the **uniform mutation**. In this method, according to a given probability, each real value in the vector might be added with a uniform random value.

The recombination method used in a **tree-based** GP representation (sexual genotype and physical genotype) is the **exchange** of a **sub-tree** with a **random sub-tree**. One random sub-tree of a individual might be exchanged with a new random sub-tree, according to a given probability. The documentation of Push lacks the description of the mutation method used by the framework, but the observed results suggest that this **stack-based** GP representation (**Push**) uses a **similar method** to the method explained above.

3.2.6.3 Grouping of populations

The last generational step corresponds to the merging of the current population with the offsprings. In a **EA**, this step **matches** the **survival selection**, where the most suited individuals between the current population and the offsprings are more likely to survive to the next generation.

However, the author decided that in the proposed model the **entry** and **depart** of **Boids** only happen, **respectively**, with the steps of **generation** and **immigration** and when the **energy** of a Boid is **equal** or **lesser** than **zero**, which means that the size of the population is not constant over time. Therefore, in this last step, the **offsprings** are simply **grouped** with the **current population**.

3.2.7   Immigration

This step is **executed if** the **generational** step is **not performed**. In other words, this is a step that is practically always performed, except when the modulus of the count of iterations is equal to a given number. In that situation, the generational step will be executed instead of this step.

In this step, the **size** of **each population** (preys and predators) is **checked**, if it is **lower** than a **given threshold** then **immigrants** are **added** to that population, in order to **maintain** the **balance between species**. Without this step, the population of one specie could vanish. If the preys disappear, then the predators will die since they could not steal the energy of the preys. The disappearance of the predators would cause an exponential growth in the number of preys, once there are no predators to control their numbers.

The new **immigrants** are **generated** through **two** different **mechanisms** and afterwards they are **scattered** around the **environment** via latin hypercube method. One part of those immigrants consist in **new random Boids** in order to **introduce novelty** and the other part, in **dead Boids** that are brought back to life. This last mechanism will bring the last dead Boids to the simulation, this mechanism aims to **reuse Boids** which **might** have a **good performance** but when they were alive their **surrounding environment** was too **hostile**.

3.2.8   Stop Criteria

The proposed model can take a long time to return a reasonable outcome and, even afterwards, it can run continuously. Reasons why it is necessary to have a stop criteria [14] to halt the evolutionary search.

The chosen **stop criteria** is a **threshold** for the **number of interactions**. That threshold was refined after the performance of some simulations, in order to improve that limit.

3.3   Evaluation Mechanism

One of the **biggest challenges** of a model like the one proposed is the process of **evaluation** and **comparison** of results from several simulations. This **due** the **co-evolution** of **two** different **species**. This model is constituted by agents of two species which evolve at the same time. Therefore, it's necessary to find reliable values to be compared

with the returned values of other simulations, in order to compare the efficiency of each simulation.

One drastic **change** at the values of the evaluation metrics **may not mean** a drastic **change in** the quality of **that specie**. Instead, it can have a different meaning, such as a **change** in the **other specie** (either in quality or in their numbers) or that the **surrounding environment** is too **hostile** (for some reason the individuals of that specie cannot find and/or get close to any energy source in their neighbourhood).

In order to get reliable results, it is **necessary** to **change** some **mechanisms** of the model described above, once that **model** is **not suitable** for this objective. In that model, it's not possible to **analyse** the evolution of **only one specie** and the **evolution** of each **specie** is **constrained by** the **other** specie, which can negatively influence the results. Therefore, the evaluation process of simulations has to take in account two important aspects:

- The genotype of each population

- The size of each population

In order to be **coherent** with the results and **resolve** the **problem** of **co-evolution**, it's necessary to **evaluate** only **one specie** by simulation. Therefore, the **genotype** of the **other specie** have to remain **static** during the simulation. Also, the **size** of each population will **affect** the **evolution** of the species. Reason why it is necessary to implement a mechanism to control the size of each specie.

3.3.1   Design

Accordingly, the proposed model is constituted by two different systems. The "**free world**" is the system that is described in the section above. This system does not have any restriction, either in the size or in the genotypes of the populations. The other system is the **evaluation system** which is used to evaluate simulations. This last system has restrictions in the size of each population and in the genotypes of species. In order to have the same initial conditions, it is preferable that the evaluated simulations have always the same random seed to generate numbers.

The **evaluation system** is divided in **two steps**. Initially, it is executed a simulation wherein **both species evolve** over time and **without** any **constrain**. After a certain

number of generations, that simulation finishes and the **genotypes** of the **most suited individuals** of both species are **stored** in a data-set. Afterwards, **two simulations** are executed in order to **evaluate** the **species separately** and **one at a time**. The individuals belonging to the **evaluated specie** will evolve over time. In order to **control** their **size**, it was decided an **initial size** for the population. This size is also the maximum size of the population and at each **mating season**, new **individuals** are added until the **size** of the population is **equal** to this **initial size**.

If the size is bigger than **60%** of the initial size then the **elements** with a **worst** fitness are **removed** until the size is equal to 60%. Afterwards, new **offsprings** are **added** to the population, until the size of the population is equal to its initial size. Therefore, the size of the survival population will consist in at least **40%** of **offsprings**. The described mechanism used to control the size is very **similar** with the **elitism**.

In those two simulations, the genotypes of **individuals** belonging to the **specie** that is **not** being **evaluated** remain static over the simulations. The **genotype** of those individuals is **loaded** from the **data-set** stored in the **initial step**. To control their numbers there are a determined initial size for that population and **when** an individual **starves**, is out of energy, he dies and then he **reappears randomly** in the environment.

3.3.2   Metrics

Each **simulation** will be **evaluated** through **four metrics**. **Additionally**, the **main actions** of each agent (birth, death, reproduction and immigration) are **recorded** in a **log file**. Through the analyses of that file, it's possible to find the relationships between agents and, thus, determine the number of offsprings of each agent. Also, if the **preys** are the analysed specie, the values of the **traits** of their **tails** (size and brightness) are **stored** in order to allow the analysis of its evolution.

- Average fitness

- Fitness of the best individual

- Diversity of the physical and mental (attractiveness) genotype

- Number of deaths by specie

In the **average fitness** metric, the average of the energy of all the individuals that are being evaluated is stored. Afterwards, it is performed the average of those values in order to have the average of the energy between mating seasons. At the end of each

mating season, the fitness of the most suited individual is stored, in order to record the values for the fitness of the **best individual**. The **diversity metrics** provide a global information about how the genotypes differ in the population and about the stabilisation of the population, both in terms of making the individuals more suited to the environment and in terms of making the individuals more attractive for the females. Finally, the **number of deaths** by specie provides information about the evolution of the number of individuals between mating seasons. During the course of the simulation, it is intended that the number of deaths decreases and stabilises, due the adaptation of the specie.

Chapter 4

Experimental Results and Discussion

The present chapter contains the results of this dissertation. Those results were obtained from different simulations, after several runs. First, the results are shown and afterwards, they are statistically compared with other results. Finally, a discussion of the results is presented, in which is discussed if the results are statistically equal within a given confidence interval and, if not the case, how they differ from each other.

The author has decided to conduct **five** different **experiments**. The **first** experiment aims to study possible **effects** that the **representation** used for the physical genotypes may have in a simulation. Therefore, three simulations were tested each one with 30 runs. Those simulations have the same initial conditions, but they differ in the representation used for the motion rules, which can be genetic algorithm (GA), tree-based genetic programming (GP) or stack-based genetic programming (Push).

The **following two experiments** aim to observe possible **effects** of using **sexual selection**. In the **second** experiment, the results of two simulations are compared in order to **analyse** possible **effects** of the using **sexual selection** instead of using traditional methods for the parent selection. The two simulations have the exact same initial conditions and representation for the motion rules. The difference between them is the method used for parent selection, one uses a **traditional method** and the other uses a method inspired on **female mate choice**. Afterwards, the **third** experiment is intended to **observe** possible **effects** of **sexual selection** on **different** initial **conditions**. Therefore, this experiment is very similar with the previous except on the initial conditions which are different and **harsher** for the agents.

Considering the **results** of the **previous experiments**, several experiments were conducted in order to further **explore other solutions** which may have different conclusions. Several experiments were carried out and since it is not possible to report all of them. The author decided to only report and analyse two of those experiments in this dissertation.

As it was explained in the above section, subsection 3.3, in order to test a simulation, first it is needed to train a population of agents (preys and predators) with the imposed constraints and then each population is tested only one at a time, this because of the problem of co-evolution. In order to get a sufficient number of replicates each **test** is **repeated 30 times**.

Afterwards, the results are compared with each others. The **first hypothesis** to be **tested** is whether there are **significant differences** between the results, this within a **confidence interval** of **95%**. If the first hypothesis does not manifest, then a **second hypothesis** is tested to discover what **results** have statistically the **higher values**, this also within a confidence interval of 95%.

Normally, the results of **EC don't follow** a **normal** (Gaussian) **distribution** for the underlying data. Reason why it was used **nonparametric** statistics to compare the results. Due to the characteristics of data and experiments, the author used the **Mann-Whitney test**, since he only compared two simulations at a time. In the first hypothesis, it was used a two-tailed test and, for the second hypothesis it was used a one-tailed test.

4.1   Experiment I

As referred above, this first experiment aims to **observe possible disparities** between the results of several **simulations**, which have the same initial conditions but they have **different representations** for the motion rules. The motion rules can be represented through genetic algorithm (**GA**), tree-based genetic programming (**GP**) or stack-based genetic programming (**Push**). The method used for the parent selection was a traditional selection method. Only **six metrics** were considered for the comparison and analysis of the results (best fitness, average fitness, number of deaths, diversity of the sexual genotype, tail's brightness and tail's size). The data from the **diversity** of the **physical genotype** (motion rules) were not regarded, once each **simulation** uses a **different representation** and so, the results of each representation may not have a direct relationship with the results of other representations, which makes their comparison **meaningless**.

Firstly, the initial conditions used in the simulations are described, followed by the discussion of the results of the preys and then of the predators. The graphics shown in the discussion represent the evolution of the values of the regarded metrics between generations (mating seasons), in which the vertical bars don't represent the exact values but an estimation for the 95% confidence interval.

The following list shows the **values** used in this experiment for the **initial conditions**.

- **Number of runs:** 30
- **Number of generations by run:** 100
- **Interval of interactions between generations:** 50
- **Mutation probability:** 10%
- **Elitism proportion:** 60%
- **Size of tournament:** 5
- **Number of candidates:** 5
- **Maximum depth of a tree:** 5
- **Initial number of preys in each generation:** 80
- **Initial number of predators in each generation:** 20
- **Size of the environment:** [400, 200]
- **Maximum amount of food resources:** 350
- **Minimum distance between food resources:** 25
- **Size of the separation zone:** 2
- **Size of the social zone:** 10
- **Size of the prey's target zone:** 25
- **Size of the predator's target zone:** 30
- **Maximum acceleration:** 2
- **Maximum prey's velocity:** 0.5
- **Vision angle of preys:** 150°
- **Maximum steering angle of preys:** 30°
- **Maximum predator's velocity:** 0.73
- **Vision angle of predators:** 130°
- **Maximum steering angle of predators:** 20°

Those **values** result from previous values that the author found appropriate, which were later **tuned** after the observation of several simulations. It worth mentions the **differences between preys** and **predators**, as referred before, the developed simulator is strongly inspired by nature and normally in the nature the **predators** have a **greater**

freedom in their **movements** than the preys. Reason why in this simulator they have a higher maximum velocity and a greater neighbourhood. However, in order to better focus their preys, the predators have their eyes more close to each other which reduce them angle of view.

In the nature, a **balance** between the number of **preys** and **predators** tends to **exist**, in order to **maintain both** species **alive**. In this simulation, the **size** of each **population** was **chosen** after the observation of the evolution of their number in a "**free world**" simulation, over several generations. It was chosen the sizes that the populations seem to converge.

Preys

In the following graphics, it is possible to observe the outcomes of the evolution of preys through six metrics: best fitness, average fitness, number of deaths, diversity of the sexual genotype, tail's brightness and tail's size.
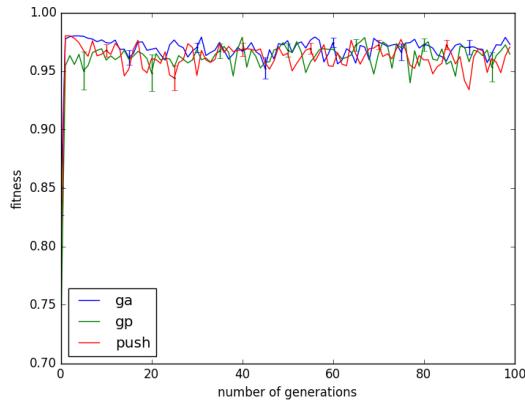
FIGURE 4.1: Expt 1.1 - Best fitness of the preys
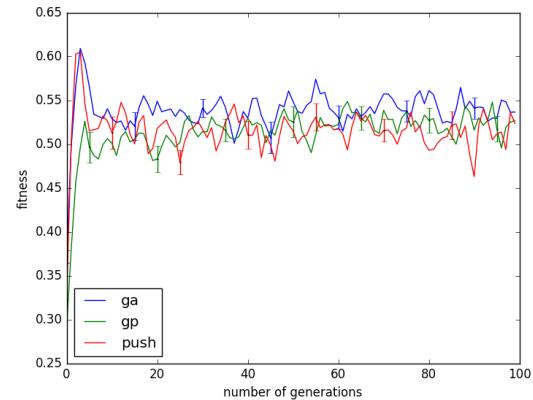
FIGURE 4.2: Expt 1.1 - Average fitness of the preys
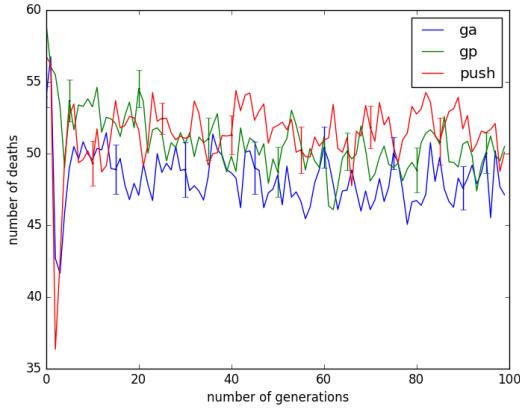
FIGURE 4.3: Expt 1.1 - Number of deaths of the preys
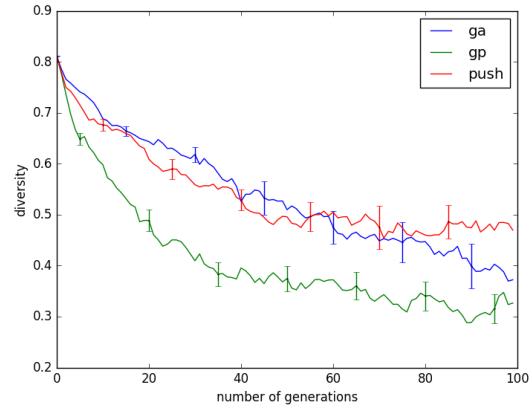
FIGURE 4.4: Expt 1.1 - Diversity of the sexual genotype
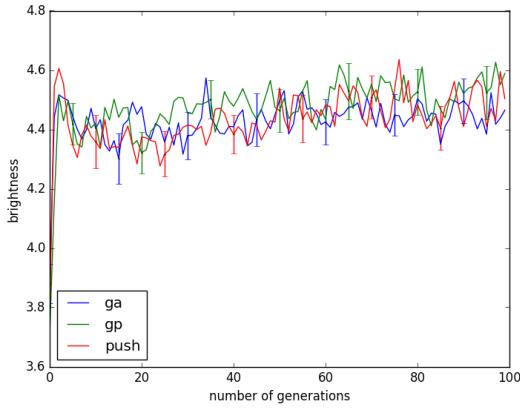
FIGURE 4.5: Expt 1.1 - Tail's Brightness of the preys
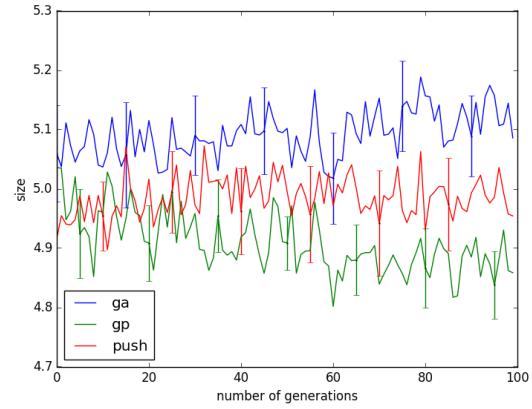
FIGURE 4.6: Expt 1.1 - Tail's Size of the preys

It is possible to observe a **rapid rise** on the fitness in the graphics of the **average** (figure 4.1) and **best fitness** (figure 4.2). Also, during that time is observed a **large decline** on the **number of deaths** (figure 4.3) which may suggest that until that moment the preys are evolving their motion in order to get access to the food, until they converge to have effective motion rules. After that moment, the competition among prey increases causing a lower average fitness and a increase in the number of deaths.

In the **diversity** of the **sexual genotype** (figure 4.4), it is possible to observe that the results from **GP** are **statistically smaller** that results from the other representations. It may be caused by the partition of the data. Other interesting facts are the values of the **brightness** (figure 4.5) and **size** (figure 4.6) of the tail. Since they both act as handicaps, it was expected that their values decrease over the generations which didn't happen. Both results may vary between 0 and 10 and since the values of the **size** seem to oscillate between the value 5, they seem to be random and so the size seem to not

affect the evolution of the preys. Also, the values of the **brightness** seem to start smaller (around 3.8) and then, they increase (around 4.5), which may suggest that in the beginning the brightness influences the preys, but after they have effective motion rules, they seem to ignore that handicap since they can bear it.

Except for the values of the sexual diversity, within a **95% confidence interval**, it cannot be stated that exist significant differences between the results shown above.

Predators

In the following graphics, it is possible to observe the outcomes of the evolution of predators through three metrics: best fitness, average fitness and number of deaths.
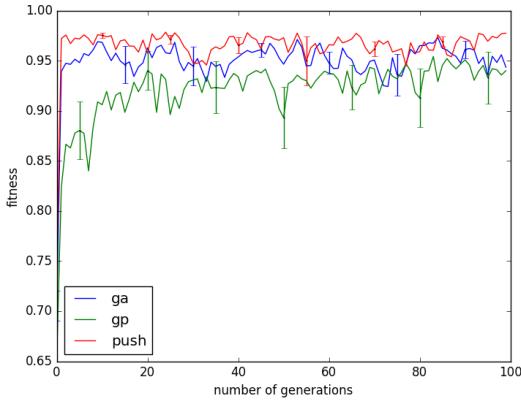
FIGURE 4.7: Expt 1.2 - Best fitness of the predators
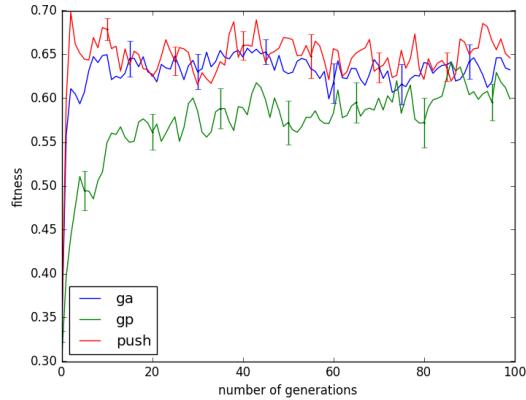
FIGURE 4.8: Expt 1.2 - Average fitness of the predators
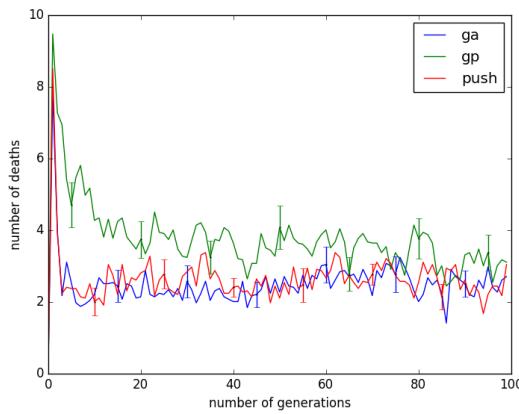
FIGURE 4.9: Expt 1.2 - Number of deaths of the predators

The results from GP stand out from the remaining results, its results from **average** (figure 4.8) and **best fitness** (figure 4.7) **start statistically lower** than the results from the other representations. However they increase until they reach the same values.

The same applies in values of the **number of deaths** (figure 4.9), wherein its values **start statistically higher** than the other values, but they also decrease until they reach the same values.

Those results were expected since the motion rules of GP are being generated, unlike the GA which has already a motion rule. However the results of the GP should also be observed on the results from Push since they are similar, which didn't happen likely because of the motion rules of Push. The results of the metrics didn't indicate, but the motion rules of Push are simpler than the other motion rules, wherein the Boids seem to only pursue their energy sources. Not existing a motion of flocks.

Analysis of the results

The comparison between the outcomes of different representations has revealed that, in **general**, there are **no** significant **differences between** the different **representations**. This observation was expected since both simulations start with random genotypes which will evolve over generations, in similar environments.

However, it exists a **difference** between the results of **GP** and the other representations. In the **preys**, the values of the **sexual diversity** seem to be **smaller**, but it may happen because of the partition of the data. While, the simulation that use GA and Push to represent the physical genotype, also have another data structure (tree-based GP) to represent the sexual genotype. The same does not happen with the simulations that use GP to represent the motion rules, wherein the sexual genotype is stored in the same data structure that is used to store the physical genotype and so, the sexual and physical genotype are directly related.

The use of **different data structures** can lead to **different partitions** of the data during the crossover operators. For example, in a simulation that uses GA or Push, a offspring can inherit 30% of the physical genotype of its mother and 80% of the sexual genotype of its father. However, in a simulation that uses GP, those percentages will be equal for both genotypes. The effects of those differences can be observed on the evolution of the sexual diversity over generations.

Also, in the **predators**, it is observed a **difference** in the **first generations** between the values of the GP and the other representations. Those results are within the expected, since the motion rules of the predators are evolving over the time, requiring a greater number of generations to stabilise than the results from GA. Those results should also be observed on Push. However, the motion rules of Push unlike the others, usually evolve to just pursue its target which provokes different results and motions.

4.2   Experiment II

The second experiment has the purpose to discuss possible **effects** that **may happen** in simulations **when** the **traditional parent selection** method (tournament) is **replaced** with a method inspired on **female mate choice**. In order to test this scenario, the results of simulations using traditional selection method and a method inspired on mate choice were collected after each simulation is repeated 30 times. Afterwards, the results of simulations having the same representation for the physical genotype are analysed.

All the simulations were executed with the **same initial conditions**, which are the same as those described in the first experiment. It should be noted, that the simulations that are compared only refer to simulations wherein the **preys** are being evaluated, once it was decided that only the preys can use a method inspired on mate choice for the parent selection.

Firstly, the results obtained through simulations using GA to represent the motion rules are analysed, followed by the analysis of simulation using GP for the representation and by the analysis of simulation using Push for the representation. Afterwards, a discussion covering general aspects of the results is presented.

GA

The following graphics were used to compare the outcomes of seven metrics: best fitness, average fitness, number of deaths, diversity of the physical genotype, diversity of the sexual genotype, tail's brightness and tail's size. The results were obtained from two set of simulations, one using a traditional method for the parent selection and the other a method inspired on female mate choice, wherein both simulations use GA to represent the physical genotype.
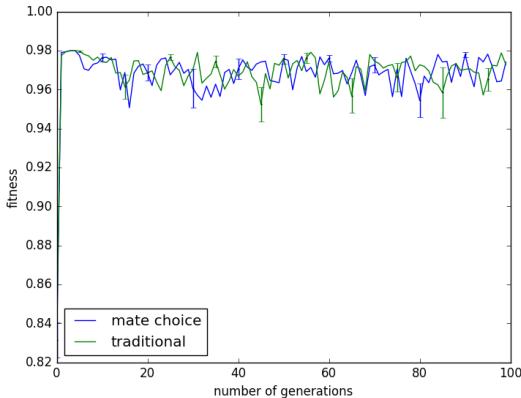

FIGURE 4.10: Expt 2.1 - Results of the best fitness
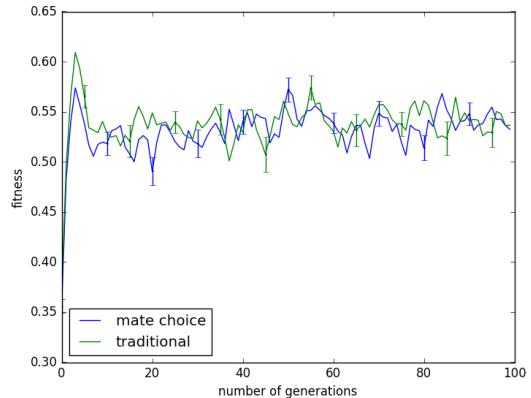
FIGURE 4.11: Expt 2.1 - Results of the average fitness
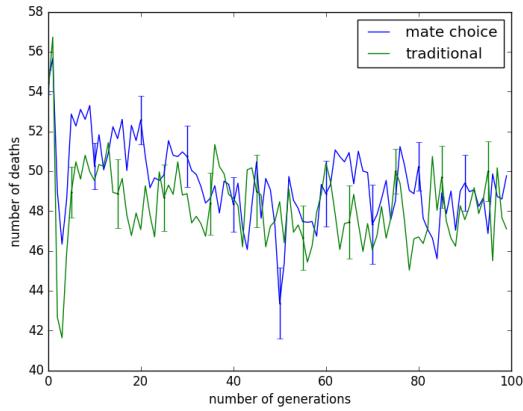
FIGURE 4.12: Expt 2.1 - Results of the number of deaths
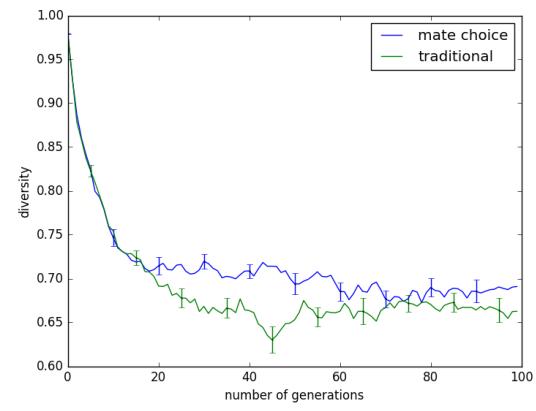
FIGURE 4.13: Expt 2.1 - Results of the physical diversity
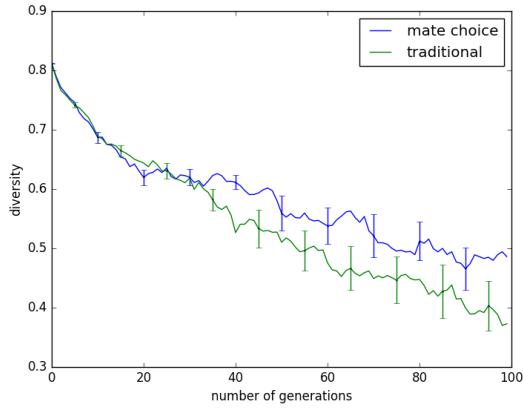
FIGURE 4.14: Expt 2.1 - Results of the sexual diversity
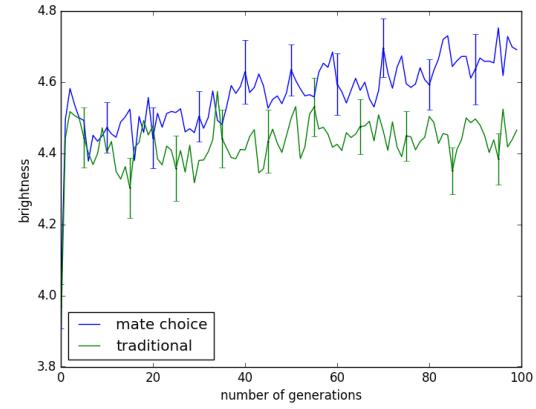

FIGURE 4.15: Expt 2.1 - Results of the tail's brightness
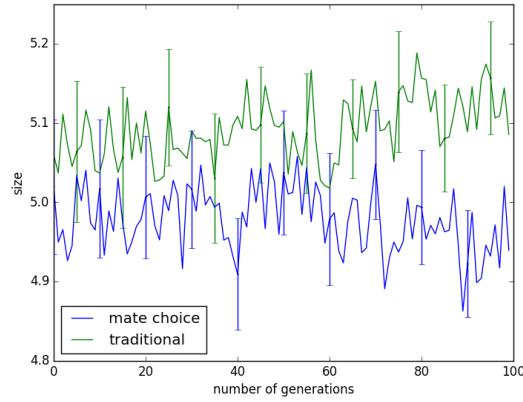
FIGURE 4.16: Expt 2.1 - Results of the tail's size

Among the results from **best fitness** (figure 4.10), **average fitness** (figure 4.11) and **number of deaths** (figure 4.12) do **not** seem to exist **significant differences** between

the simulations using traditional selection methods and the method inspired on mate choice, this within a **95% confidence interval**.

In the **physical diversity** (figure 4.13), the results of the simulation using mate choice from the generation 25 to the 50 are statistically higher than the results from the traditional method. Also, the results from the **sexual diversity** (figure 4.14) of mate choice seem to be higher than the other, but statistically there are no significant differences between them.

The same happens with the results of the **brightness** (figure 4.15), wherein the results from mate choice seem to be higher. However that hypothesis was not statistically manifested. The results from the **size** (figure 4.16) seem to not contribute for the evolution of the preys where is values appear to be random.

GP

The following graphics were used to compare the outcomes of seven metrics: best fitness, average fitness, number of deaths, diversity of the physical genotype, diversity of the sexual genotype, tail's brightness and tail's size. The results were obtained from two set of simulations, one using a traditional method for the parent selection and the other a method inspired on female mate choice, wherein both simulations use GP to represent the physical genotype.
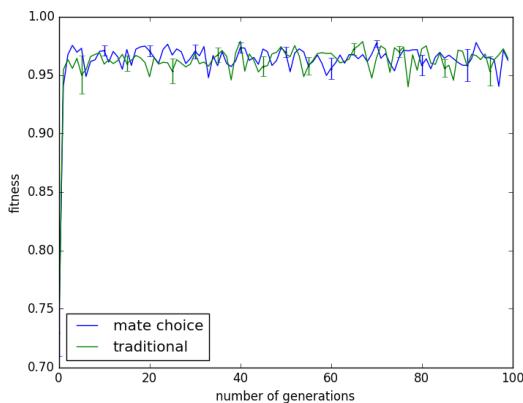
FIGURE 4.17: Expt 2.2 - Results of the best fitness
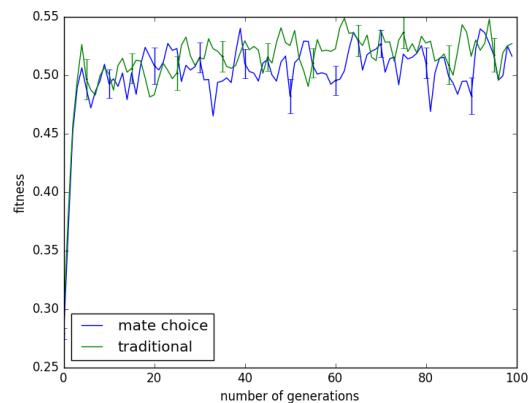
FIGURE 4.18: Expt 2.2 - Results of the average fitness
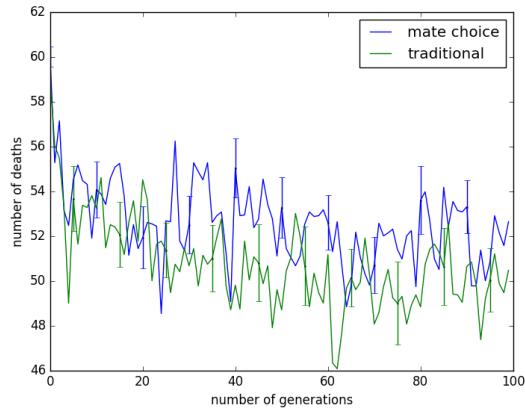
FIGURE 4.19: Expt 2.2 - Results of the number of deaths
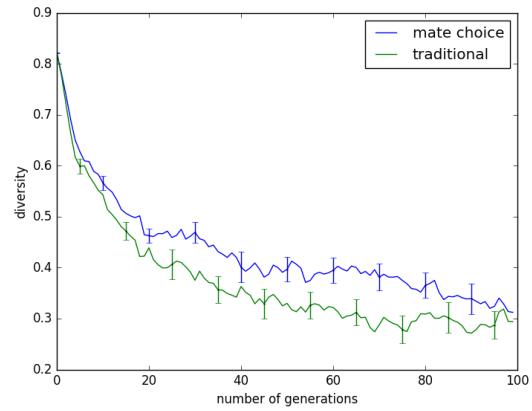
FIGURE 4.20: Expt 2.2 - Results of the physical diversity
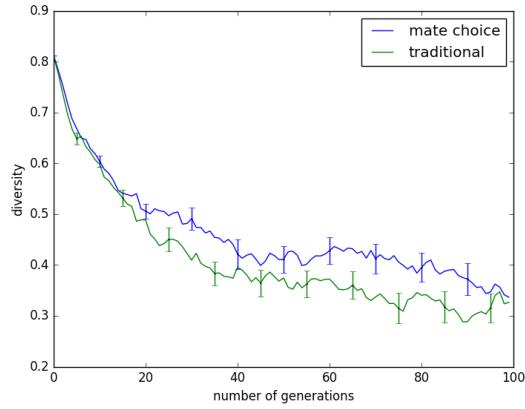
FIGURE 4.21: Expt 2.2 - Results of the sexual diversity
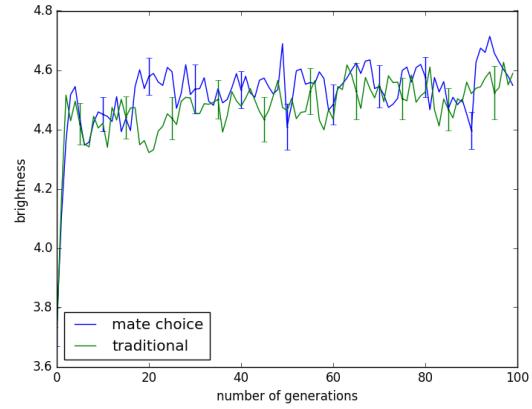
FIGURE 4.22: Expt 2.2 - Results of the tail's brightness
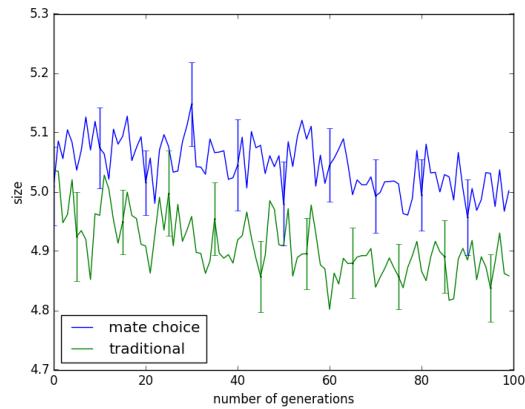
FIGURE 4.23: Expt 2.2 - Results of the tail's size

As it happened in the previous data, within a **95% confidence interval**, it seems that there are **no significant differences** between the results from **best fitness** (figure 4.17), **average fitness** (figure 4.18) and **number of deaths** (figure 4.19).

In the results of the **sexual** (figure 4.21) and **physical diversity** (figure 4.20), it is possible to observe a clear separation between the results of mate choice and the traditional method, wherein the results from mate choice seem to be higher. However, it cannot be stated that exist significant differences between those results.

The results from the **brightness** (figure 4.22) suggest that there are no differences between the selection methods. Once again the results from the **size** (figure 4.23) seem to be random.

Push

The following graphics were used to compare the outcomes of seven metrics: best fitness, average fitness, number of deaths, diversity of the physical genotype, diversity of the sexual genotype, tail's brightness and tail's size. The results were obtained from two set of simulations, one using a traditional method for the parent selection and the other a method inspired on female mate choice, wherein both simulations use Push to represent the physical genotype.
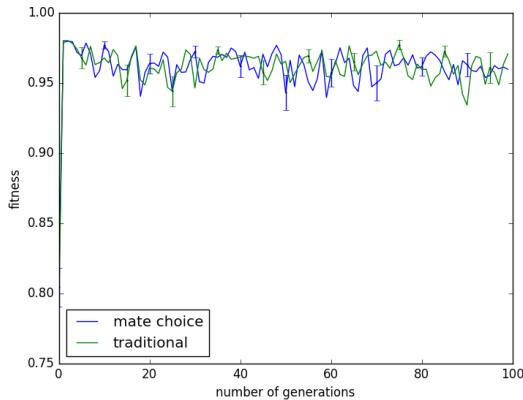

FIGURE 4.24: Expt 2.3 - Results of the best fitness
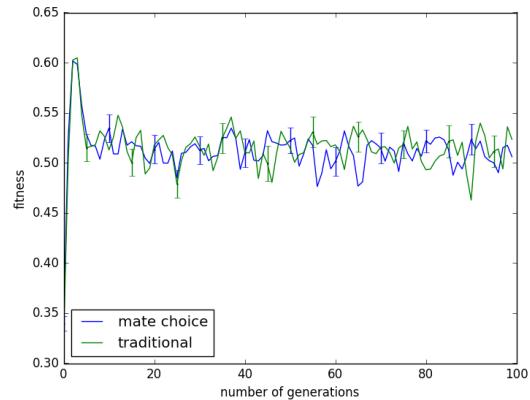

FIGURE 4.25: Expt 2.3 - Results of the average fitness
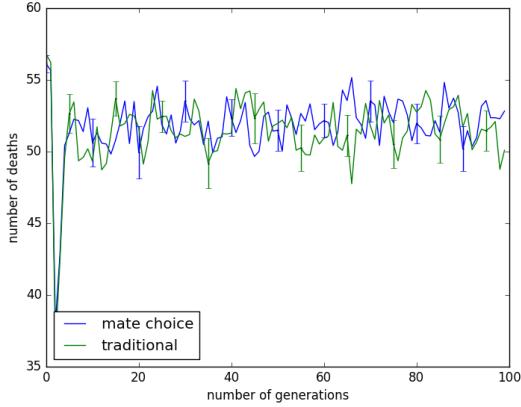
FIGURE 4.26: Expt 2.3 - Results of the number of deaths
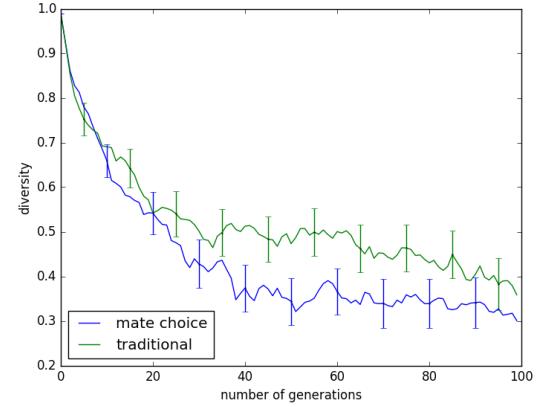
FIGURE 4.27: Expt 2.3 - Results of the physical diversity
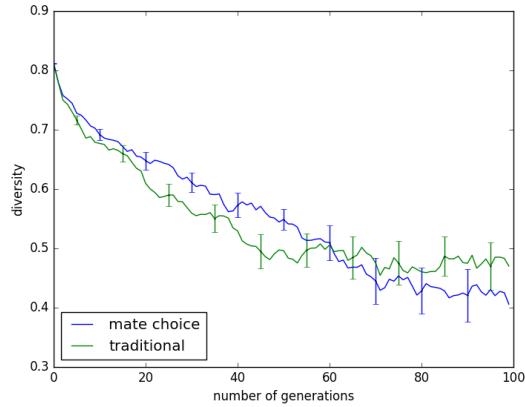
FIGURE 4.28: Expt 2.3 - Results of the sexual diversity
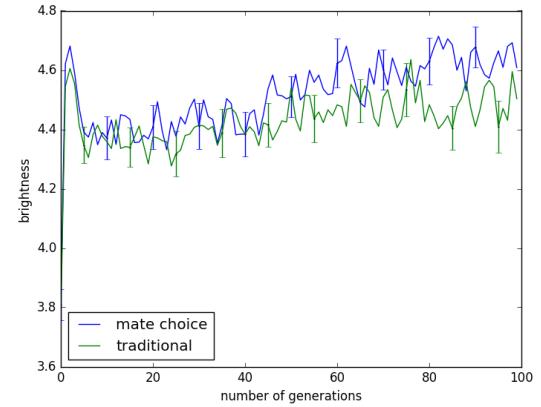
FIGURE 4.29: Expt 2.3 - Results of the tail's brightness
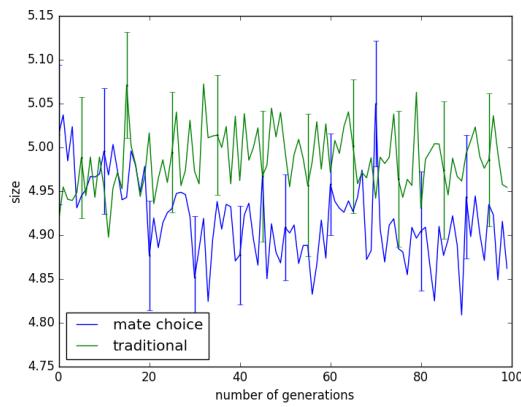
FIGURE 4.30: Expt 2.3 - Results of the tail's size

Within a **95% confidence interval**, it **doesn't** seem to be **significant differences** among the above graphics. **Except** for the data from **physical genotype** (figure 4.27), where contrary to what was observed in simulations with other representations, in this

experiment the physical diversity of the mate choice seems to be lower. That difference is not statistically valid and may be caused by the motion rules generated by Push which are simpler than the others. That fact may provoke unpredictable consequences.

Analysis of the results

In general, the statistical analysis suggests that there are **no significant differences between** using a **traditional method** for the parent selection and using a method inspired on **mate choice**, this for the results of best fitness, average fitness, number of deaths, tail's brightness and tail's size. Within a **95%** of **confidence interval**, it seems that there are no significant differences between those results.

Also, the results from **GA** and **GP** seem to suggest that the values of the **physical** and **sexual diversity** may be **higher** when using the method inspired on **mate choice**. That difference is not statistically verified, but there are a clear separation of the values, wherein the values from mate choice seem to be higher. **However**, the same didn't happen with the results from **Push**, wherein it seems that there are no differences between the values of **physical diversity**, also in the sexual diversity, the values of **mate choice** seem to be **lower** than the others.

However, since this observation is only observed in simulations using Push. It may be caused the own representation of Push. One possible explanation for the difference on the results can be its design for the genotype. During the execution of the code, the results of the different nodes and sub-trees are stored at the top of the same stack. At the end of the execution of the code, the final outcome is the value standing at the top of the stack. That value can be the result of a complex operation or simply the value in a leaf, which may provoke the simpler motions of the individuals when compared with the individuals from other representations.

4.3   Experiment III

In **similarity** of the **second experiment**, the third experiment has the purpose to **study** possible **effects** that the **replacement** of a **traditional** parent choice method by a method inspired on **mate choice** can cause on the results. The **difference** between this and the previous experiment **lies** on the **initial conditions**, in particular in the quantity of food resources and in the minimum distance between them. This difference aims to confirm if the conclusions found in the previous experiment remain true in different initial conditions.

Besides, the initial conditions are different than on the other experiments, they are also harsher in order to test the simulation in a more extreme environment. Therefore, the **maximum amount of food resources** decreases from 350 to **222** and the **minimum distance between food resources** increases from 25 to **35**.

Firstly, the results obtained through simulations using GA to represent the motion rules are analysed, followed by the analysis of simulation using GP for the representation and by the analysis of simulation using Push for the representation. Afterwards, a discussion covering general aspects of the results is presented, discussing in particular the evolution of the values obtained from the sexual genotype.

GA

The following graphics were used to compare the outcomes of seven metrics: best fitness, average fitness, number of deaths, diversity of the physical genotype, diversity of the sexual genotype, tail's brightness and tail's size. The results were obtained from two set of simulations, one using a traditional method for the parent selection and the other a method inspired on female mate choice, wherein both simulations use GA to represent the physical genotype.

FIGURE 4.31: Expt 3.1 - Results of the best fitness
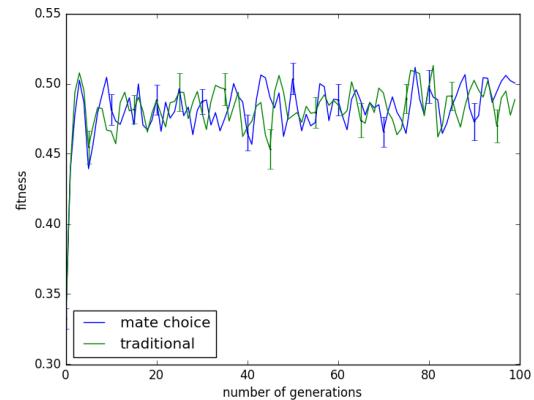
FIGURE 4.32: Expt 3.1 - Results of the average fitness
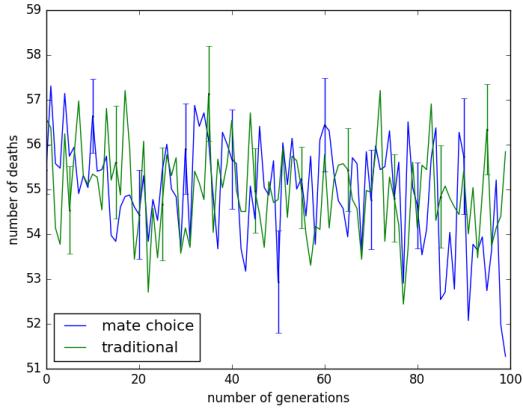
FIGURE 4.33: Expt 3.1 - Results of the
number of deaths
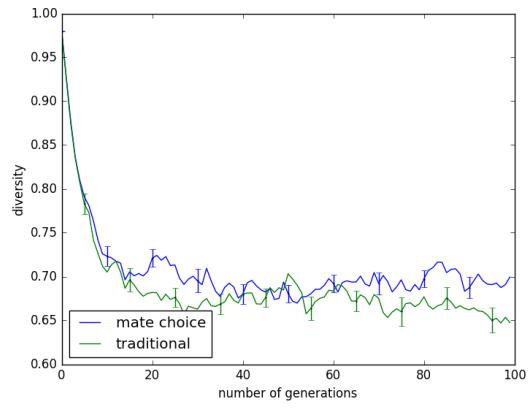
FIGURE 4.34: Expt 3.1 - Results of the
physical diversity
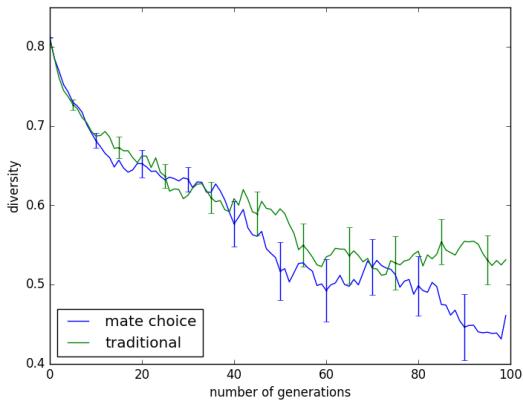
FIGURE 4.35: Expt 3.1 - Results of the
sexual diversity
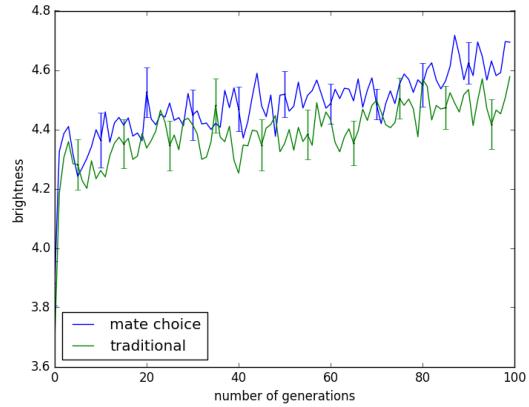
FIGURE 4.36: Expt 3.1 - Results of the
tail's brightness
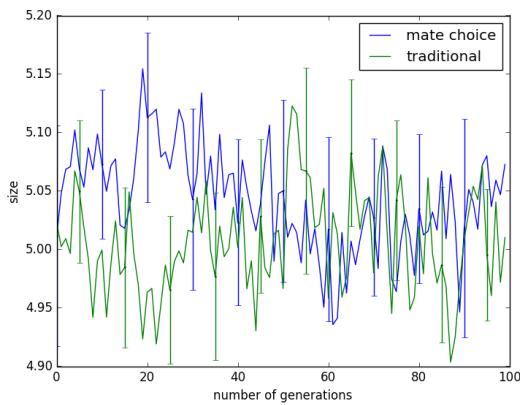
FIGURE 4.37: Expt 3.1 - Results of the tail's size

In these conditions, it seems that there are **no significant differences** between the results shown above, this within a **95% confidence interval**.

GP

The following graphics were used to compare the outcomes of seven metrics: best fitness, average fitness, number of deaths, diversity of the physical genotype, diversity of the sexual genotype, tail's brightness and tail's size. The results were obtained from two set of simulations, one using a traditional method for the parent selection and the other a method inspired on female mate choice, wherein both simulations use GP to represent the physical genotype.
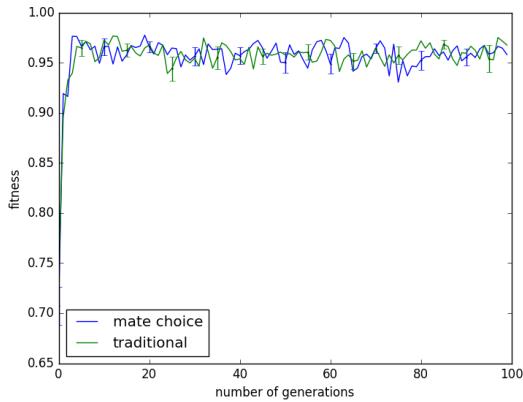
FIGURE 4.38: Expt 3.2 - Results of the best fitness
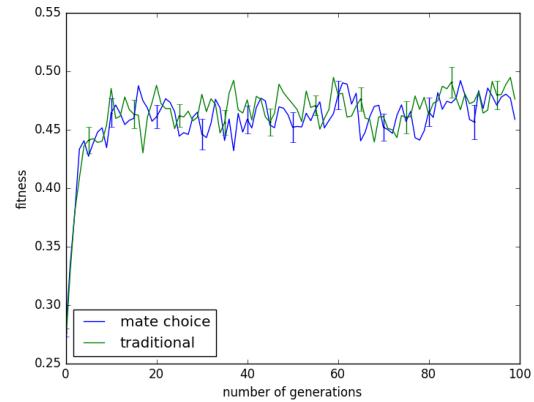
FIGURE 4.39: Expt 3.2 - Results of the average fitness
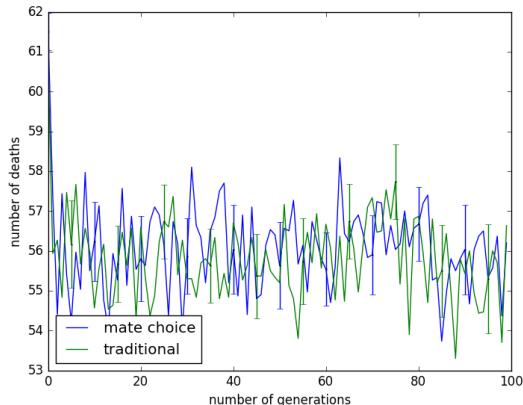
FIGURE 4.40: Expt 3.2 - Results of the number of deaths
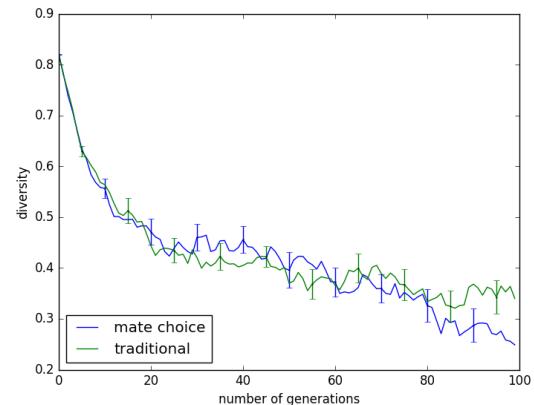
FIGURE 4.41: Expt 3.2 - Results of the physical diversity
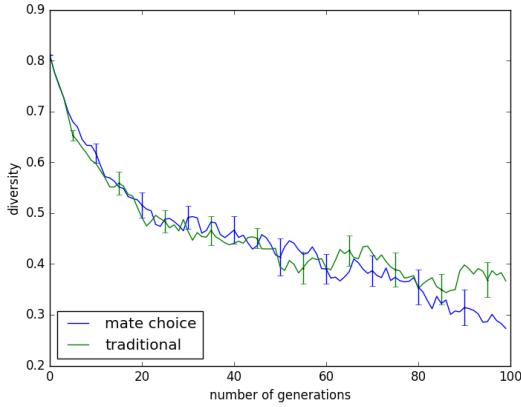
FIGURE 4.42: Expt 3.2 - Results of the
sexual diversity
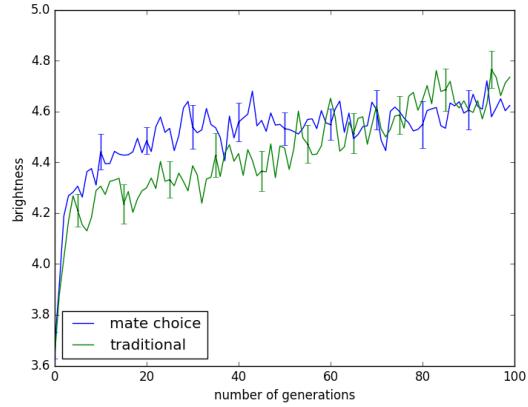
FIGURE 4.43: Expt 3.2 - Results of the
tail's brightness
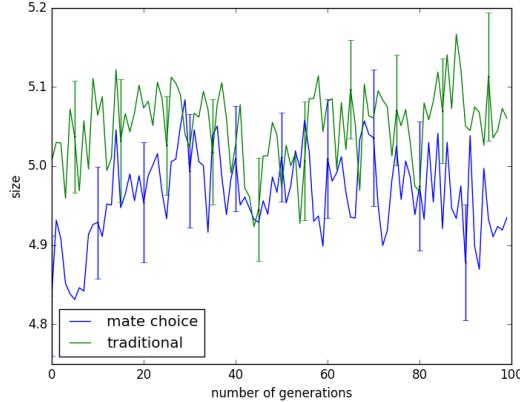
FIGURE 4.44: Expt 3.2 - Results of the tail's size

In these conditions, it seems that there are **no significant differences** between the
results shown above, this within a **95% confidence interval**.

Push

The following graphics were used to compare the outcomes of seven metrics: best fitness,
average fitness, number of deaths, diversity of the physical genotype, diversity of the
sexual genotype, tail's brightness and tail's size. The results were obtained from two set
of simulations, one using a traditional method for the parent selection and the other a
method inspired on female mate choice, wherein both simulations use Push to represent
the physical genotype.
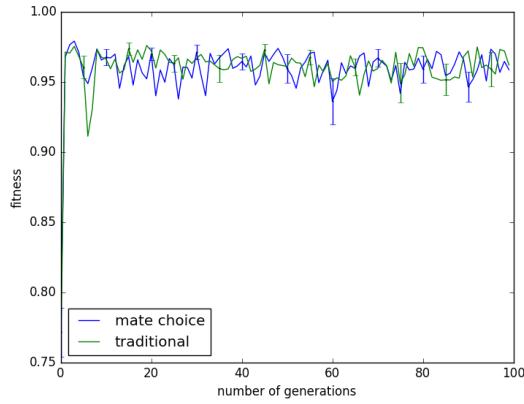
FIGURE 4.45: Expt 3.3 - Results of the best fitness
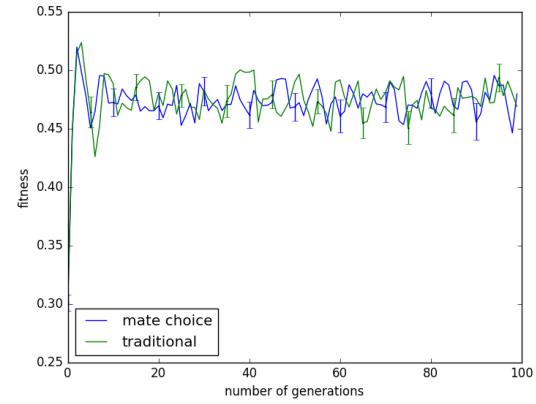
FIGURE 4.46: Expt 3.3 - Results of the average fitness
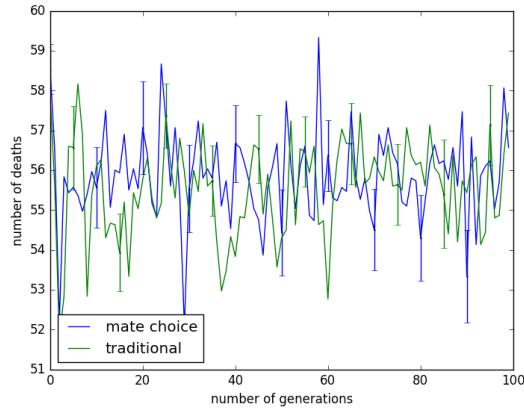
FIGURE 4.47: Expt 3.3 - Results of the number of deaths
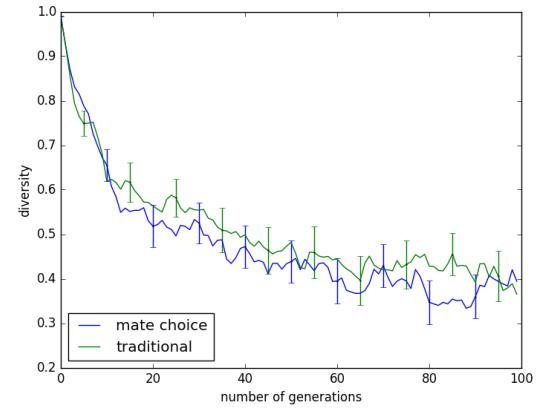
FIGURE 4.48: Expt 3.3 - Results of the physical diversity
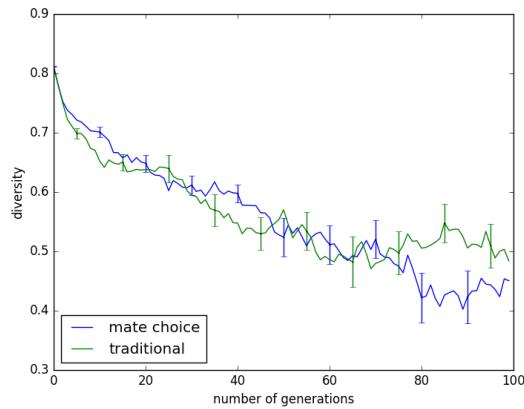
FIGURE 4.49: Expt 3.3 - Results of the sexual diversity
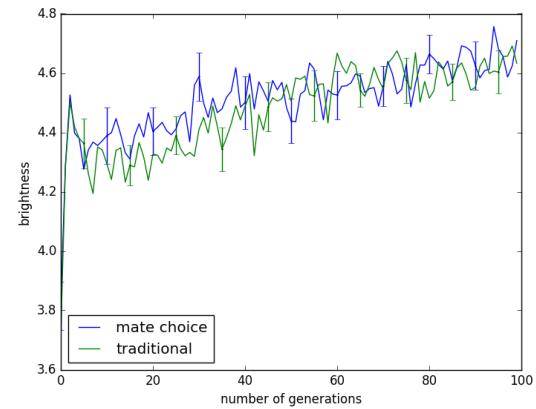
FIGURE 4.50: Expt 3.3 - Results of the tail's brightness
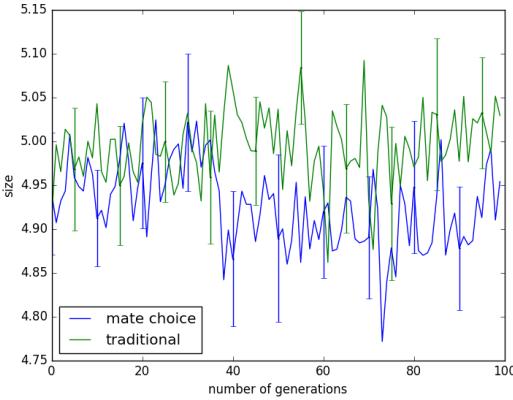
FIGURE 4.51: Expt 3.3 - Results of the tail's size

In these conditions, it seems that there are **no significant differences** between the results shown above, this within a **95% confidence interval**.

Analysis of the results

The analysis of this experiment has revealed **some different evidences** than the results of the **previous experiment**. However the **conclusions remain** the **same**. For the proposed model, it seems that there are no significant differences between using mate choice or the traditional selection methods.

The **suggestions** found in the experiment II do not hold true when the conditions are harsher for the individuals. In the experiment II, it was suggested that could exist slight differences between using mate choice and the traditional methods, although they are not statistically relevant. In this experiment, those suggestions don't hold true, **none** of the **results stand out** from the others.

4.4   Experiment IV

As been seen, the **results** from the **above experiments** are **far from** the **expected** results, wherein there are no significant differences between using mate choice or the traditional selection methods. However, the most surprising fact is the values of characteristics of the tail (brightness and size). It is expected that their values decrease over generations, since they both act as handicaps, which didn't happen.

It is observed a **small contribution** of the **brightness** for the evolution of the preys. In the first generations, it starts smaller but it increases over the generations until it stabilises. Also, the size of the tail seems to randomly oscillate around the value 5, which

taking in account that both the brightness and the size may vary from 0 to 10. It reveals that the values of the **size** are **random**, it seems to not contribute for the evolution of the preys.

Those values suggest that the **preys** are in the beginning more **focused** in **developing** an **effective motion rule** in order to survive, wherein the brightness slightly influence their survival. So, the preys are more focused evolving its motion rules, over the brightness and the size of their tails. After they have an effective motion rule, they don't seem to need to evolve the features of the tail, since they are fit enough to bear those handicaps.

In order to **further explore** the **parameters** of the developed model, several experiments were conducted, varying various aspects of the simulations, such as the number of generations, the interval between mating seasons, the size of each population, the energy functions, the sexual genotype, the percentage of survivors and the quantity of food resources.

Among those experiments, **two** were chosen to be shown and analysed. The difference, in both **experiments** when compared with the second experiment, is the **energy function**, wherein the pressure under the brightness of the tail was increased in both experiments.

The energy function used in the above experiments follows below.

$$energy \mathrel{+}= collectedEnergy - 0.01 \times (1 + logsig(\frac{age - 100}{12}) + 2 \times logsig(tailBrigh - 5))$$

4.4.a   Experiment 4.1

As referred above, this experiment aims to **explore** other **energy function** which may causes different conclusions. As been seen in the previous experiments, the brightness of the tail seems to only have slightly effects in the beginning, after that its values increase until they stabilise. This may suggest that the individuals are **not sufficiently constrained** by the **brightness**. For that reason in this experiment, the energy function was replaced by the following.

$$energy \mathrel{+}= collectedEnergy - 0.01 \times (1 + logsig(\frac{age - 100}{12})) - Boid.energy \times (tailBrigh/10)$$

As it can been seen the values of the evaluated functions don't only depend on age of the individuals but also in brightness of their tails, which unlike the previous function

its **outcome** is **not limited**. In the previous equation the outcome of the function that has the brightness as a parameter could only vary between 0 and 0.02. In the evaluated function, that value can vary between 0 and 1 and it also **depends** on the value of the **current energy** of the individual. This with the purpose of the individual does not always lose the same amount of energy but a quantity directly related with its own energy. Another difference is the function used to determine the energy lost due to the brightness of the tail. In the previous experiment that function was a **sigmoid** function, but in this experiment that function was **replaced** by a **straight line**.

Firstly, the results obtained through simulations using GA to represent the motion rules are analysed, followed by the analysis of simulation using GP for the representation and by the analysis of simulation using Push for the representation. Afterwards, a discussion covering general aspects of the results is presented, discussing in particular the evolution of the values obtained from the sexual genotype.

GA

The following graphics were used to compare the outcomes of seven metrics: best fitness, average fitness, number of deaths, diversity of the physical genotype, diversity of the sexual genotype, tail's brightness and tail's size. The results were obtained from two set of simulations, one using a traditional method for the parent selection and the other a method inspired on female mate choice, wherein both simulations use GA to represent the physical genotype.

FIGURE 4.52: Expt 4.a.1 - Results of the best fitness
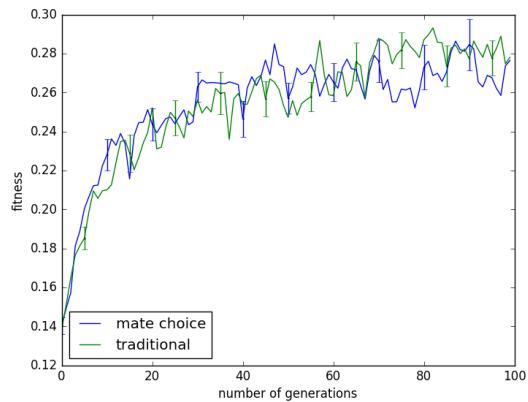
FIGURE 4.53: Expt 4.a.1 - Results of the average fitness
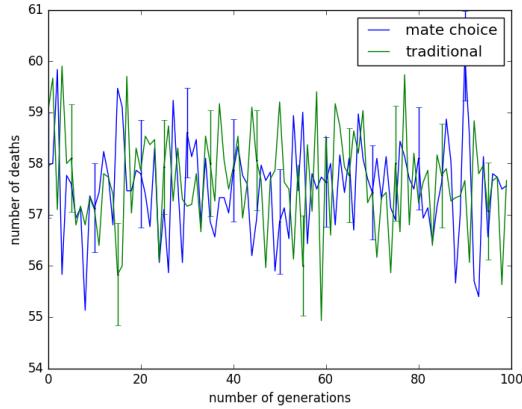
FIGURE 4.54: Expt 4.a.1 - Results of
the number of deaths
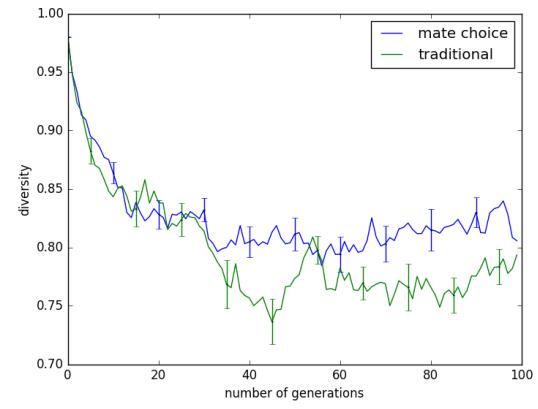
FIGURE 4.55: Expt 4.a.1 - Results of
the physical diversity
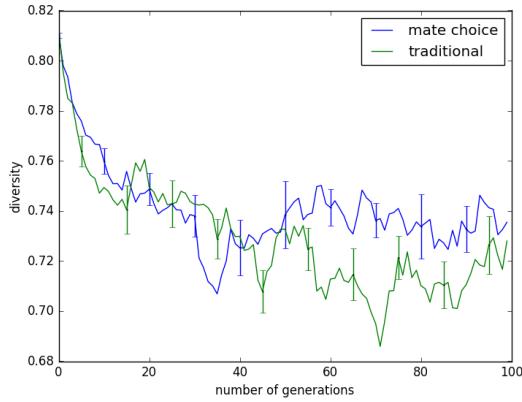
FIGURE 4.56: Expt 4.a.1 - Results of
the sexual diversity
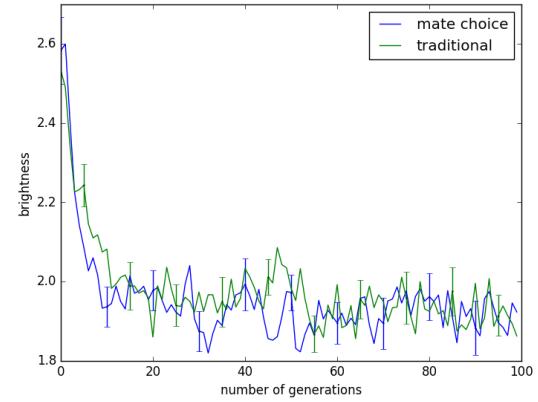
FIGURE 4.57: Expt 4.a.1 - Results of
the tail's brightness
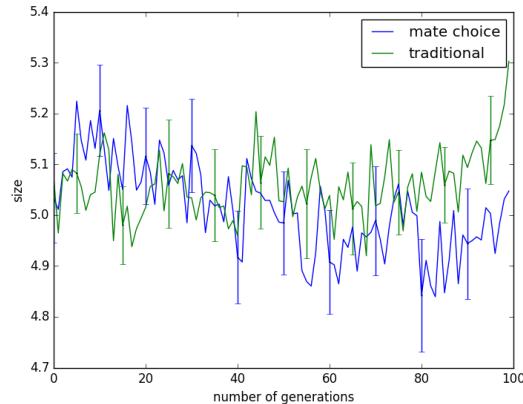
FIGURE 4.58: Expt 4.a.1 - Results of the tail's size

The results from the **best** (figure 4.52) and **average fitness** (figure 4.53) are smaller than the results from the second experiment, also they are taking more generations until they stabilise. By the other hand, the **number of deaths** (figure 4.54) seem to be higher

and they are more unstable. As it happened in the second experiment, the values of the **physical** (figure 4.55) and **sexual diversity** (figure 4.56) seem to be visually higher for the mate choice. However, that hypothesis did not manifest statistically. Also their values are vastly superior when compared with the values of the second experiment.

In this experiment, the **brightness** (figure 4.57) seems to influence the evolution of the individuals, where it goes down until it stabilises. Once again the **size** (figure 4.58) of the tail seems to be ignored during the evolution. Within a **95% confidence interval**, it seems that there are **no significant differences** between the results shown above.

GP

The following graphics were used to compare the outcomes of seven metrics: best fitness, average fitness, number of deaths, diversity of the physical genotype, diversity of the sexual genotype, tail's brightness and tail's size. The results were obtained from two set of simulations, one using a traditional method for the parent selection and the other a method inspired on female mate choice, wherein both simulations use GP to represent the physical genotype.
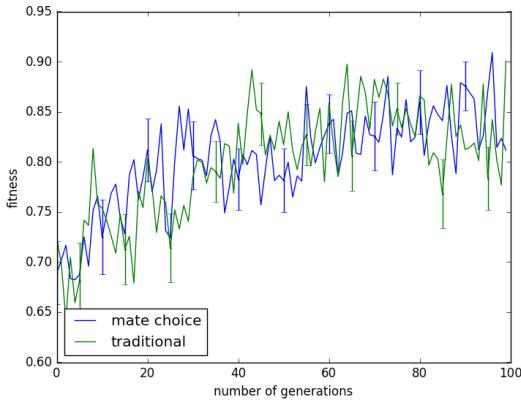

FIGURE 4.59: Expt 4.a.2 - Results of the best fitness
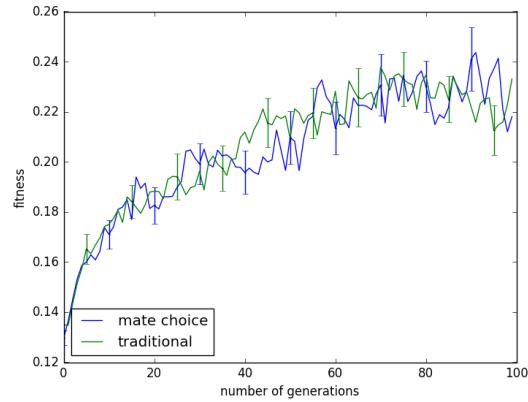
FIGURE 4.60: Expt 4.a.2 - Results of the average fitness
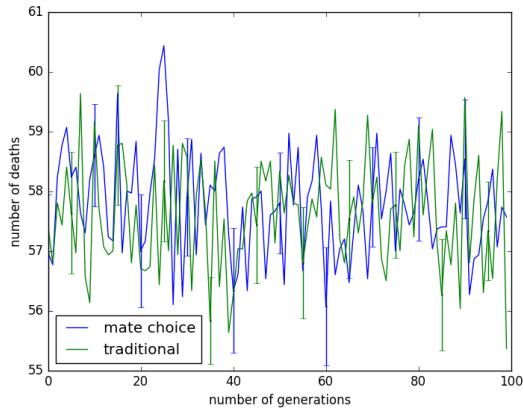
FIGURE 4.61: Expt 4.a.2 - Results of the number of deaths
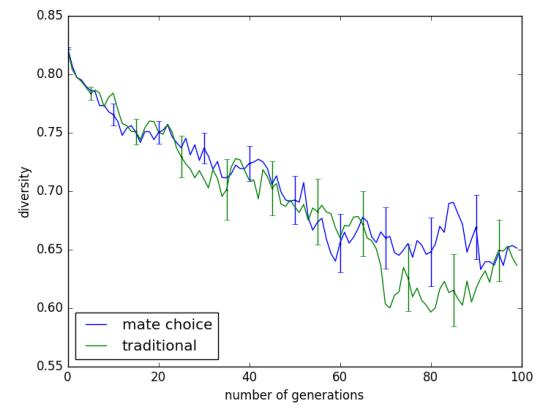

FIGURE 4.62: Expt 4.a.2 - Results of the physical diversity
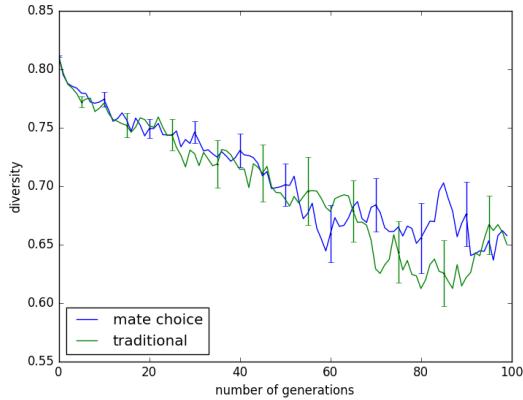

FIGURE 4.63: Expt 4.a.2 - Results of the sexual diversity
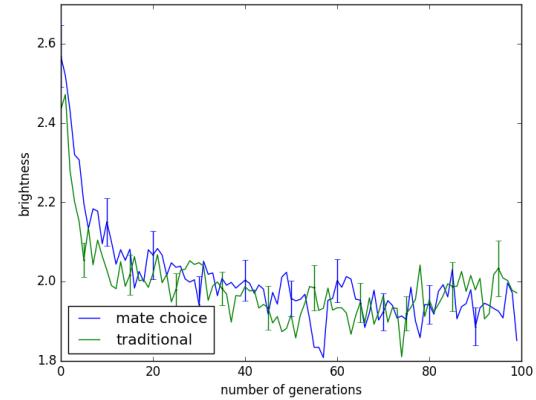

FIGURE 4.64: Expt 4.a.2 - Results of the tail's brightness
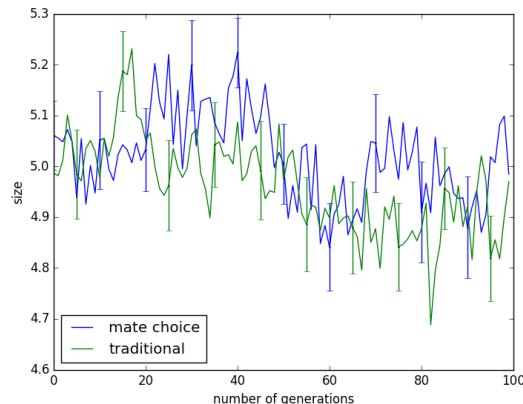

FIGURE 4.65: Expt 4.a.2 - Results of the tail's size

The **results** from this experiment are very **similar with** the results from the **previous** results (GA). **Except**, in the results from **physical** (figure 4.62) and **sexual diversity** (figure 4.63), in which there are no significant differences between the selection methods.

It seems that there are **no significant differences** between the results shown above, this within a **95% confidence interval**.

Push

The following graphics were used to compare the outcomes of seven metrics: best fitness, average fitness, number of deaths, diversity of the physical genotype, diversity of the sexual genotype, tail's brightness and tail's size. The results were obtained from two set of simulations, one using a traditional method for the parent selection and the other a method inspired on female mate choice, wherein both simulations use Push to represent the physical genotype.
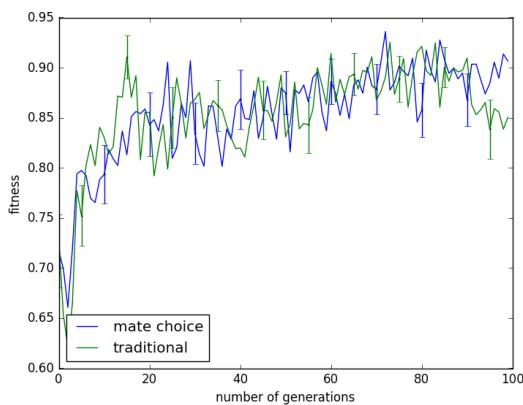

FIGURE 4.66: Expt 4.a.3 - Results of the best fitness
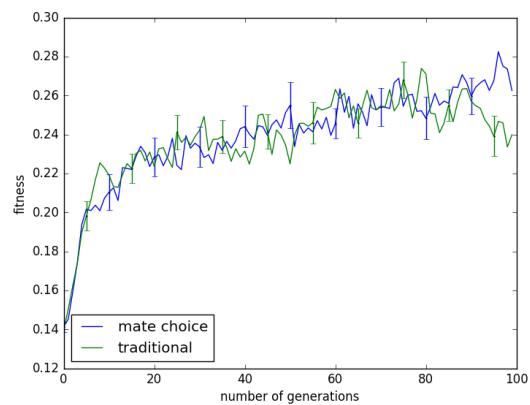
FIGURE 4.67: Expt 4.a.3 - Results of the average fitness
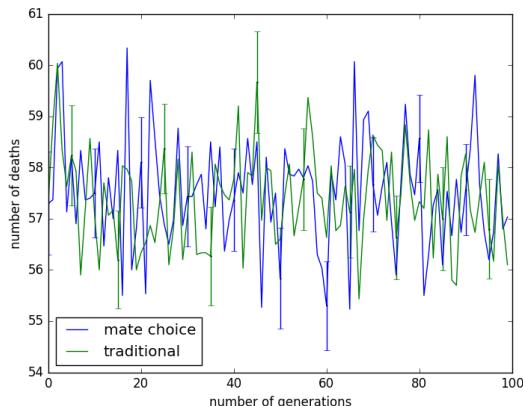
FIGURE 4.68: Expt 4.a.3 - Results of the number of deaths
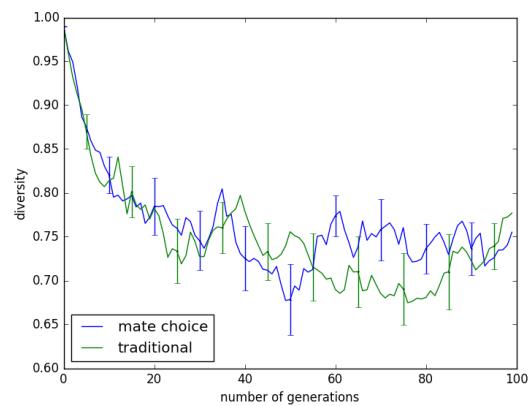
FIGURE 4.69: Expt 4.a.3 - Results of the physical diversity
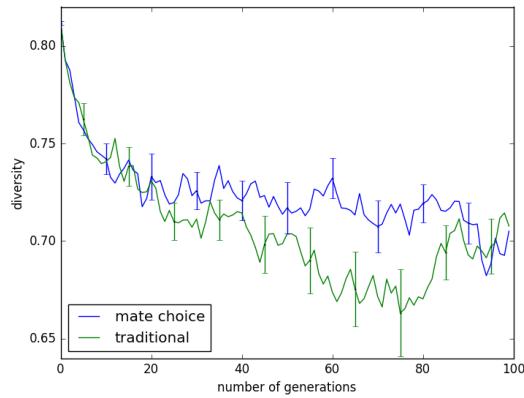
FIGURE 4.70: Expt 4.a.3 - Results of the sexual diversity
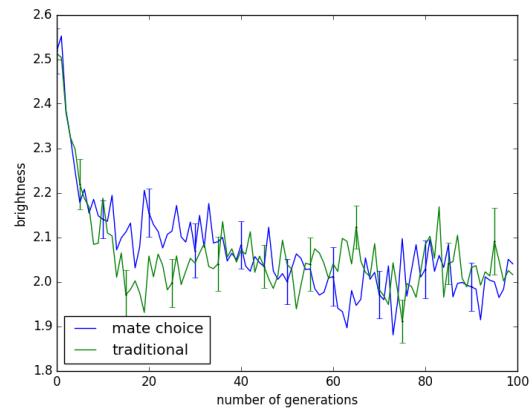
FIGURE 4.71: Expt 4.a.3 - Results of the tail's brightness
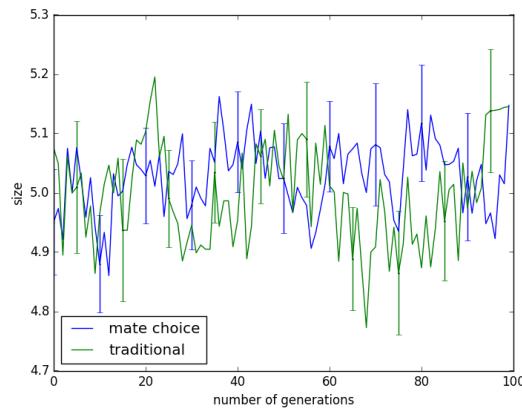
FIGURE 4.72: Expt 4.a.3 - Results of the tail's size

Also, the results from this experiment are very **similar** the the results from the **previous results** (GA). Except, for the results from **physical diversity** (figure 4.69), in which there are no significant differences. In these conditions, it seems that there are **no significant differences** between the results shown above, this within a **95% confidence interval**.

Analysis of the results

It seems to exist a great similarity among the various results. The increase in the pressure of the brightness caused different results when compared with the results of the second experiment. The greater pressure provokes that the values of **best fitness** and **average fitness** start with **lower** values that are increasing over time, at a slow pace. Also, it can be seen an **increase** in the **mortality** of the individuals which is demonstrated by a lower average fitness and a raise in the number of deaths.

That raise in the mortality, it will provoke a **increase** in the number of **immigrants** since the number of preys is highly unstable. This situation seems to cause the results in the **physical** and **sexual diversity**, wherein they are decreasing over the generations but these values are vastly **superior** when compared with the values of the **second experiment**, which is expected by the greater number of random individuals. As referred before, one part of the immigrants consists in random individuals and the other part consists in previously dead individuals that are brought back to the simulation.

Once again the results of the features of the tail are far from the expected values. Once more, the results of the size seem to not contribute for the evolution process. The values of the **brightness** are **decreasing** over the generations, as expected, but then they **don't stabilise around** the lowest possible value for the brightness (**zero**). This observation may suggest that even with a greater increase in the pressure of the brightness, the individuals are more concerned with the evolution of the motion rules and then when they have effective motion rules to describe their movements. They are able to bear the possible handicaps which they can have, ignoring then the evolution of its brightness.

4.4.b   Experiment 4.2

As the experiment above, this experiment aims to **explore** other **energy function** which may cause different conclusions. As been seen in the previous experiments, the brightness of the tail seems to only have slightly effects in the beginning, after that its values increase until they stabilise. This may suggest that the individuals are **not sufficiently constrained** by the **brightness**. For that reason in this experiment, the energy function was replaced by the following.

$$energy \mathrel{+}= collectedEnergy - 0.01 \times (1 + logsig(\frac{age - 100}{12}) + 2 \times logsig((energy \times tailBrigh + 0.5) * 9.5))$$

As can been seen, the values of the evaluated functions don't only depend on age of the individuals but also in brightness of their tails, which unlike the previous experiments its outcome **depends** on the value of the **current energy** of the individual. This with the purpose of the individuals don't always lose the same amount of energy but a quantity directly related with its own energy. Thereunto, a new sigmoid function was carefully chosen, wherein its parameters may vary from 0 to 1. In this experiment, the **rules** of the **sexual genotypes** were **slightly changed**. In the previous experiments, the values of the features of the tail could vary from 0 to 10. However,

in this experiment those **outcomes** were **limited from 0 to 1**, wherein those outcomes represent the results of operations between percentages. For example, the rule "**distanceToMatingSeason**' now represents the percentage of time until a mating season and the rule "**numPreysNeighbourhood**" now represents the percentage of preys among the neighbourhood of the individual.

Firstly, the results obtained through simulations using GA to represent the motion rules are analysed, followed by the analysis of simulation using GP for the representation and by the analysis of simulation using Push for the representation. Afterwards, a discussion covering general aspects of the results is presented, discussing in particular the evolution of the values obtained from the sexual genotype.

GA

The following graphics were used to compare the outcomes of seven metrics: best fitness, average fitness, number of deaths, diversity of the physical genotype, diversity of the sexual genotype, tail's brightness and tail's size. The results were obtained from two set of simulations, one using a traditional method for the parent selection and the other a method inspired on female mate choice, wherein both simulations use GA to represent the physical genotype.


FIGURE 4.73: Expt 4.b.1 - Results of the best fitness
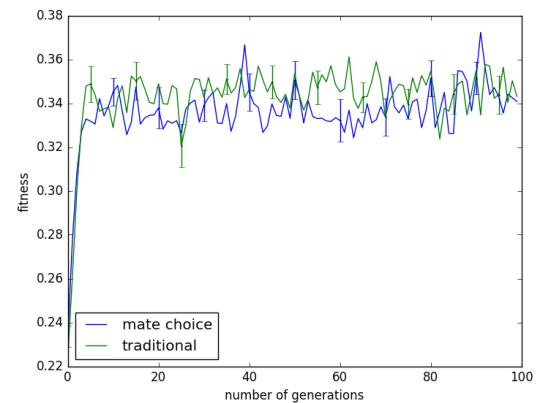
FIGURE 4.74: Expt 4.b.1 - Results of the average fitness
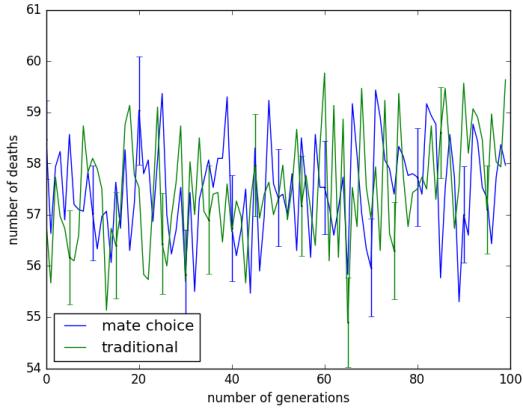
FIGURE 4.75: Expt 4.b.1 - Results of
the number of deaths
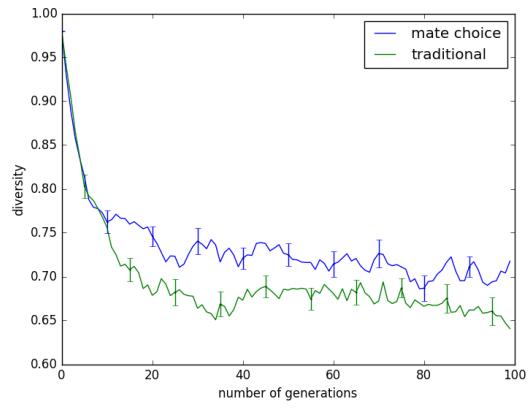
FIGURE 4.76: Expt 4.b.1 - Results of
the physical diversity
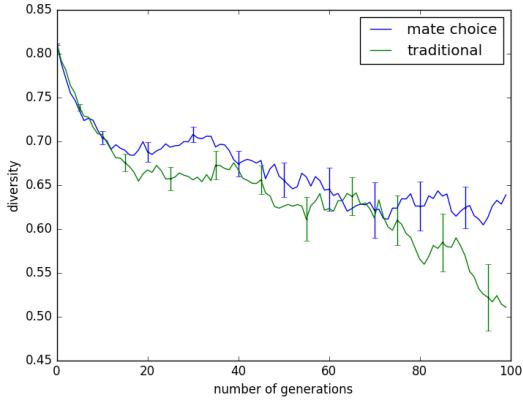
FIGURE 4.77: Expt 4.b.1 - Results of
the sexual diversity
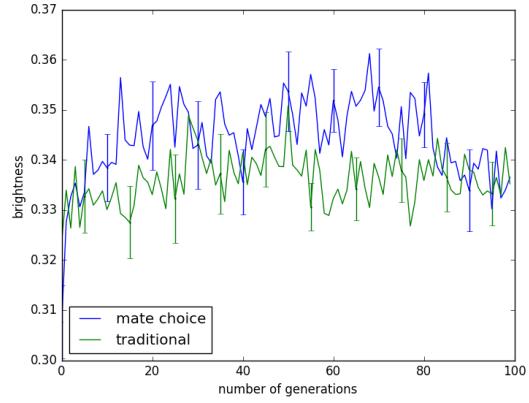
FIGURE 4.78: Expt 4.b.1 - Results of
the tail's brightness
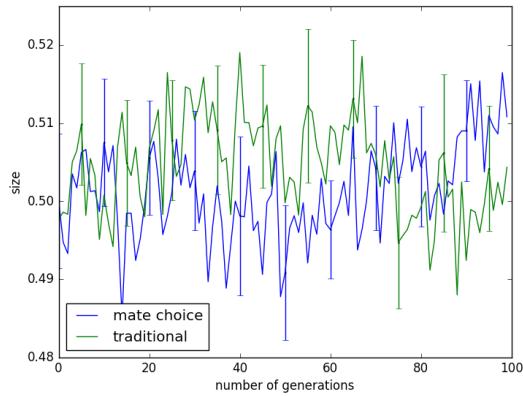
FIGURE 4.79: Expt 4.b.1 - Results of the tail's size

Comparing the above results with results from the second experiment, it is possible to
observe that the **best** (figure 4.73) and **average fitness** (figure 4.74) are smaller, also
these values seem to take slightly longer to stabilise. By the other hand, the **number
of deaths** (figure 4.75) seem to be higher and they are more unstable. The results from

the **physical** (figure 4.76) and **sexual diversity** (figure 4.77) seem to suggest that the results from mate choice are slightly higher. However, it was not verified statistically.

Regarding the results of the features of the tail, they don't seem different from the results from the second experiment, wherein the **brightness** (figure 4.78) seems to have only a slightly influence on the individuals, especially in the first generations. Once again the **size** (figure 4.79) seems to not influence the evolution process.

GP

The following graphics were used to compare the outcomes of seven metrics: best fitness, average fitness, number of deaths, diversity of the physical genotype, diversity of the sexual genotype, tail's brightness and tail's size. The results were obtained from two set of simulations, one using a traditional method for the parent selection and the other a method inspired on female mate choice, wherein both simulations use GP to represent the physical genotype.
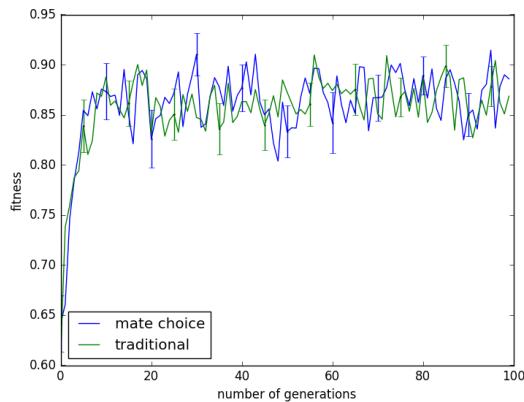
FIGURE 4.80: Expt 4.b.2 - Results of the best fitness
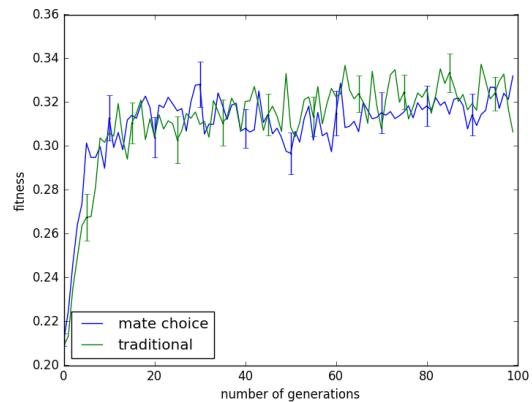
FIGURE 4.81: Expt 4.b.2 - Results of the average fitness
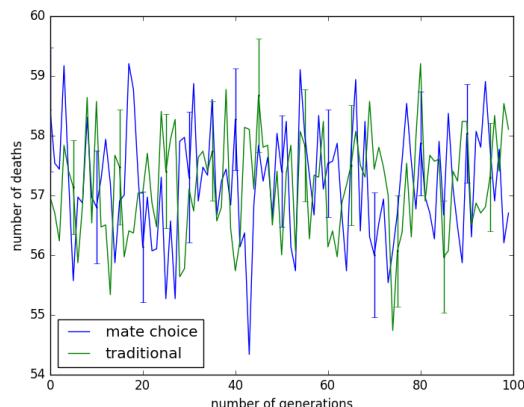
FIGURE 4.82: Expt 4.b.2 - Results of the number of deaths
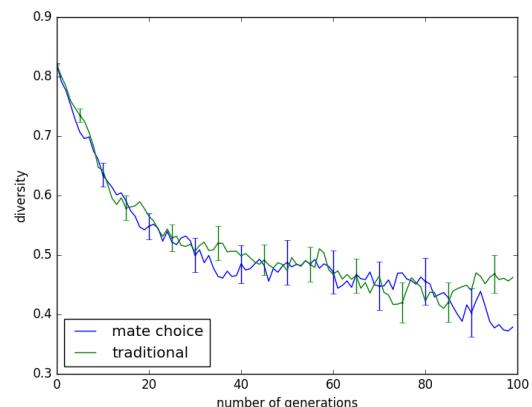
FIGURE 4.83: Expt 4.b.2 - Results of the physical diversity
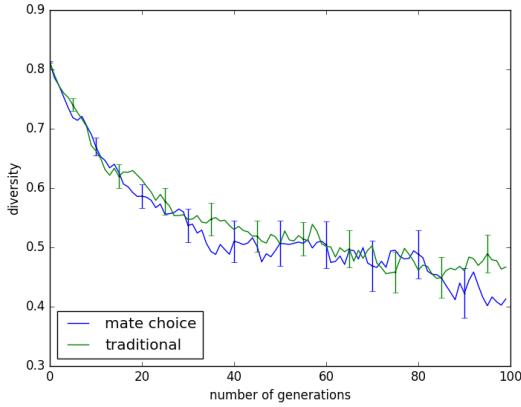
FIGURE 4.84: Expt 4.b.2 - Results of the sexual diversity
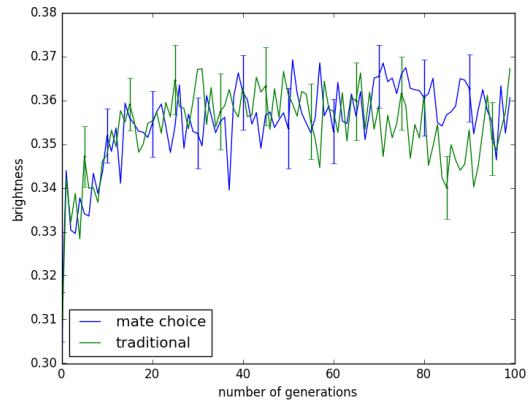
FIGURE 4.85: Expt 4.b.2 - Results of the tail's brightness
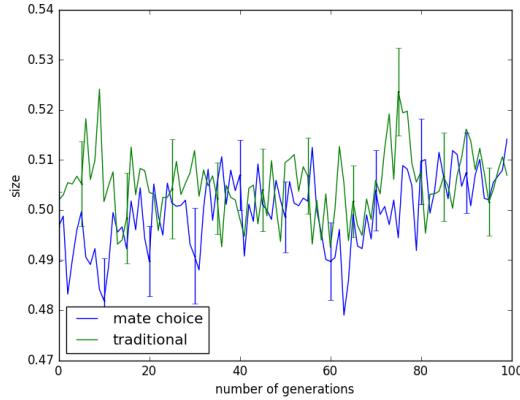
FIGURE 4.86: Expt 4.b.2 - Results of the tail's size

The **results** from this experiment are very **similar with** the results from the **previous** results (GA). **Except**, in the results from **physical** (figure 4.83) and **sexual diversity** (figure 4.84), in which there are no significant differences between the selection methods. It seems that there are **no significant differences** between the results shown above, this within a **95% confidence interval**.

Push

The following graphics were used to compare the outcomes of seven metrics: best fitness, average fitness, number of deaths, diversity of the physical genotype, diversity of the sexual genotype, tail's brightness and tail's size. The results were obtained from two set of simulations, one using a traditional method for the parent selection and the other a method inspired on female mate choice, wherein both simulations use Push to represent the physical genotype.
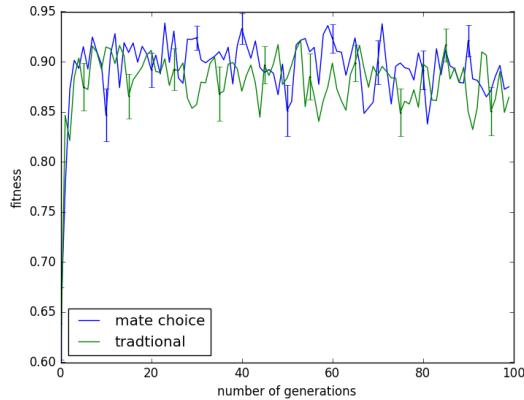
FIGURE 4.87: Expt 4.b.3 - Results of the best fitness
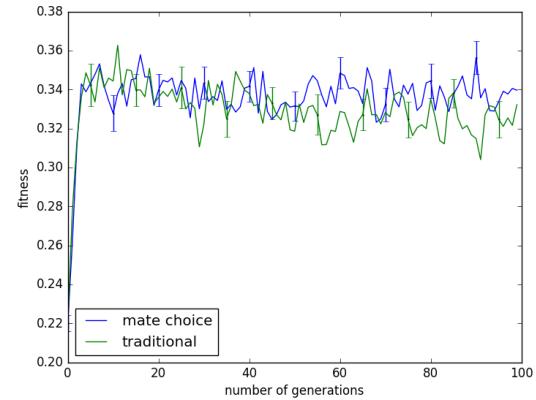

FIGURE 4.88: Expt 4.b.3 - Results of the average fitness
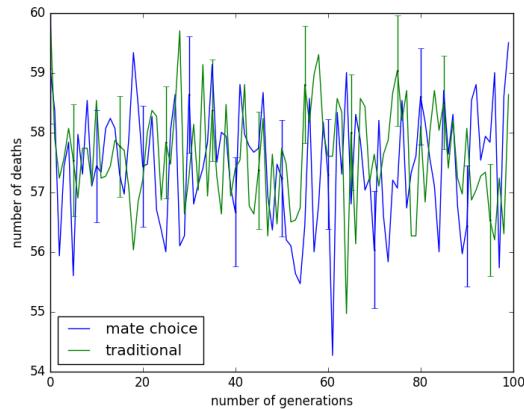
FIGURE 4.89: Expt 4.b.3 - Results of the number of deaths
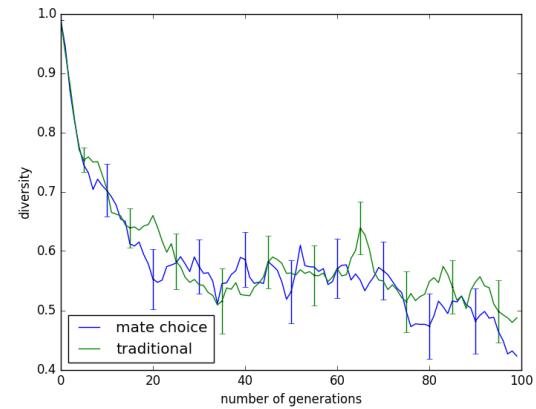
FIGURE 4.90: Expt 4.b.3 - Results of the physical diversity
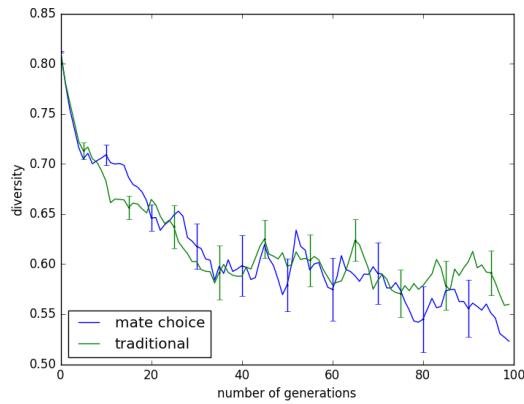
FIGURE 4.91: Expt 4.b.3 - Results of the sexual diversity
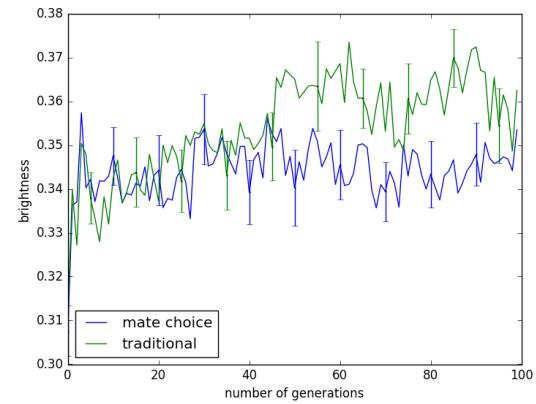
FIGURE 4.92: Expt 4.b.3 - Results of the tail's brightness
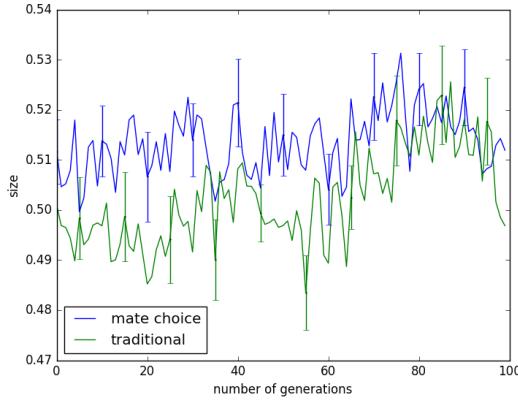
FIGURE 4.93: Expt 4.b.3 - Results of the tail's size

The **results** from this experiment are very **similar with** the results from the **previous** results (GP). It seems that there are **no significant differences** between the results shown above, this within a **95% confidence interval**.

Analysis of the results

It seems to exist a great similarity among the various results. The increase in the pressure of the brightness caused different results when compared with the results of the second experiment. The increase in pressure provokes that the values of **best fitness** and **average fitness** start with **lower** values that are increasing over time, also the values seem to take slightly longer to stabilise. Also, it can be seen an **increase** in the **mortality** of the individuals which is demonstrated by a lower average fitness and a raise in the number of deaths.

That raise in the mortality, it will provoke a **increase** in the number of **immigrants** since the number of preys is highly unstable. This situation seems to cause the results in the **physical** and **sexual diversity**, wherein they are decreasing over the generations but they values are vastly **superior** when compared with the values of the **second experiment**, which is expected by the greater number of random individuals. As referred before, one part of the immigrants consists in random individuals and the other part consists in previously dead individuals that are brought back to the simulation.

Once again the results of the features of the tail are far from the expected values. Once more the results of the **size** seem to not contribute for the evolution process. The values of the **brightness** start lower but over the generations their values increase until they stabilise. This fact may suggest that the brightness slightly influence the preys in the first generations, but when their motion rules become noticeably effective they start to ignore that handicap once they are able to bear it.

4.5   Analysis of the experimental results

The following section has the purpose to summarise the various conclusions of the experiences analysed above. An important issue is the fact that the above experiments were only analysed by the results of each metric. There are other important variables that could be studied, e.g. the **existence** of a **motion of flocks**, which it's difficult to analysed by the obtained data. Therefore, the existence of a motion of flocks was only confirmed by the **visual analysis** of the movement of the individuals. An intriguing fact is the results from **Push**, which don't seem to different from the others. However the movements generated by this representation are far **simpler** than the **movements** generated by the other representations, the main rule among those individuals seem to be the **target**, in other words, they seem to only pursue their energy sources.

The first experiment had the purpose to study possible differences between different representations, by the other hand, the **second** and the **third experiments** aim to **study** the **effects** of using a method inspired on **mate choice** instead of using a traditional method for the parent selection. The analysis of the results revealed some interesting facts, the results from the conducted experiments seem to **suggest** that there are **no** significant **difference between** the results from simulations that use **different representations** for the motion rules. This excluding the results of the sexual diversity for the GP representation that as been seen in the experiment I, it seems to produce individuals with sexual genotypes less dissimilar. That fact can be caused by the strong relationship between the physical and the sexual genotypes, for the GP, which are stored in the same data structure.

The **results** from the **second experiment** suggested that there are **no significant differences** between the results. The results from the **physical** and **sexual diversity** seem to suggest that there are a slightly difference between the results from mate choice and the traditional method, this for the GA and GP representation, wherein the results from mate choice seem to be higher. However, the statistical test did not validate that hypothesis.

The results of the **features** of the **tail** are **far from** the **expected** results. It was expected that their values will decrease over generations which didn't happen. The results from the brightness seem to start smaller and then, they increase until stabilise. Also, the values of the size seem to be random. It should be noted, that in each simulation there are four aspects that can constrain the evolutionary process, they are the motion rules, the fitness, the brightness of the tail and the size of the tail. The observed results suggest that in the beginning the **brightness** influences slightly the preys, but after

they have effective motion rules they seem to ignore that handicap, since they can bear it. Conversely, the results of the tail's **size** suggest that the velocity of each Boid isn't a main concern in the evolutionary process, where it can be random.

In order to confirm the previous conclusions, a **third experiment** was conducted. This experiment is very similar with the second experiment, except in the initial **conditions** which are **harsher**, the amount of food resource is lower and the distance between them is bigger. The initial conditions were changed with the purpose of studying if the conclusions reached in the second experiment remain true with other initial conditions. The change in the environment, as would be expected, provokes an increase in the mortality in each population. The obtained results suggest, once more, that there are **no significant differences** between using a method inspired on mate choice and the traditional selection method.

The **previous results**, especially of the features of the tail, are **far from the expected** values. The values of the handicaps should have a more relevant place in the evolutionary process which could provokes different conclusions when the results from the mate choice are compared with the values from the traditional method. In order to further explore the developed simulator **were conducted several experiments**, varying various aspects of the simulations. Among those experiments, **two** were chosen to be **shown** and **analysed** in this dissertation, wherein the **pressure** under the **brightness** of the tail was **increased** in both experiments.

Those two new experiments have **some differences between** the results of the **base case** (second experiment), but they both have in common the fact that **don't** seem to exist **significant differences** between using a method inspired on mate choice or a traditional method, for the selection of the parents. In the new experiments, the **pressure** on the **brightness** was **increased**, **causing** a large **increase** in the **mortality** of the preys.

The **results** from both experiments, especially from the fourth experiment, **suggest** that even with a greater increase in the pressure of the brightness, the **individuals** are **more concerned** with the evolution of the **motion rules**. Then, when they have effective motion rules to describe their movements, they are able to bear the possible handicaps which they can have. Ignoring then, the evolution of those values.

Chapter 5

Conclusions and Future Work

Darwin presented two mechanisms responsible for the evolution of species. The natural selection was widely accepted by the scientific community, but the same did not happen with the sexual selection which was widely criticised. It was only almost a century later, that the sexual selection came back to the spotlight. Nowadays, due the contribution of various authors, the sexual selection, like the natural selection, is accepted as one of the main mechanisms responsible to the evolution. However it has not established itself on evolutionary computation, field that is strongly influenced by the natural selection. Even, if some researchers suggest that there are advantages of using sexual selection instead of the traditional algorithms for the parent selection.

Another matter that has intrigued the scientific community, is the cooperative behaviours seen in flocks. The community has question themselves about the reasons that can lead several individuals to group together in a herd and how to describe mathematically the different movements seen in herds. *Reynolds* suggests that the behaviour seen in flocks may emerge through the application of simple movement rules in each individual. Recently, a new kind of algorithm was presented, combining the ideas from evolutionary computation with the algorithms of motion of flocks, wherein it is observed the evolution of the movement of each agent in order to produce elements more suited to the environment.

Therefore, it should be interesting to use a method inspired in female mate choice on the above algorithm, one idea that is not yet found in the literature and that is explored in this dissertation. A simulator was developed combining of motion of flocks with the ideas of a variant of evolutionary algorithms that uses sexual selection instead of the traditional selection methods.

The developed simulator is composed by several agents that can belong to one of three types: food resources, preys or predators, wherein the motion rules of the preys and predators are evolving over time. While the preys aim to better catch the food resources and to run away from predators, the predators have the purpose to be better effective at catching the preys. Cooperative movements between the agents of the same type are expected.

The evaluation of a simulator with such characteristics is not an easy task, wherein both species are evolving and the size of each population can vary a lot over the time. Reason why it was developed a separate system to evaluate the simulation. Each simulation only evaluates one specie at time, with some restrictions: the size of the population is limited and only the genotypes of the evaluated specie will evolve.

Through that system, it was possible to meet one of the objectives of this research, of analysing possible effects that may happen when the traditional method is replaced with the method inspired on mate choice. Several experiments were conducted, each one with their own objectives. At the end, it was concluded that there are no significant differences between the results.

In summary, the analysis of the different experiments suggest, that for the proposed model, there are no significant differences between the available representations for the physical genotype. Also, the analysis of the results suggest that there are no significant differences between using mate choice and traditional method for the parent selection. An interesting observation that was seen in all experiments is the fact that the features of the tail seem to only contribute slightly for the evolutionary process, wherein the preys seem to be more concerned with the evolution of their motion rules.

Also, the observation of each simulation suggests the existence of cooperative behaviours between the agents, which is in agreement with the other objective of this research.

Future Work

For future work, the features of the tail should be further explored, trying to make them more relevant for the evolutionary process. Also, it should be interesting to test other values for the parameters and to test the simulations in other environments (different initial conditions), in order to, respectively, further adjust their values and to gather more evidences that support the obtained results.

Other aspect that should be improved in the future is the performance of each experiment, in particular the execution of simulations in a row wherein the framework is not releasing the allocated memory. Finally, other aspect that may be explored is the application of the motion rules of the most suited individuals in a more realistic simulation of flocks. In order to better observe the evolved movements, facilitating their comparison with the real movements seen in flocks.

Appendix A

Graphical User Interface

This section aims to group the several descriptions of objects that can be read over the chapter 3. These objects correspond to all the agents that can be observed in a running simulation. The following print screen, figure A.1, is used to assist in the description of the several objects.
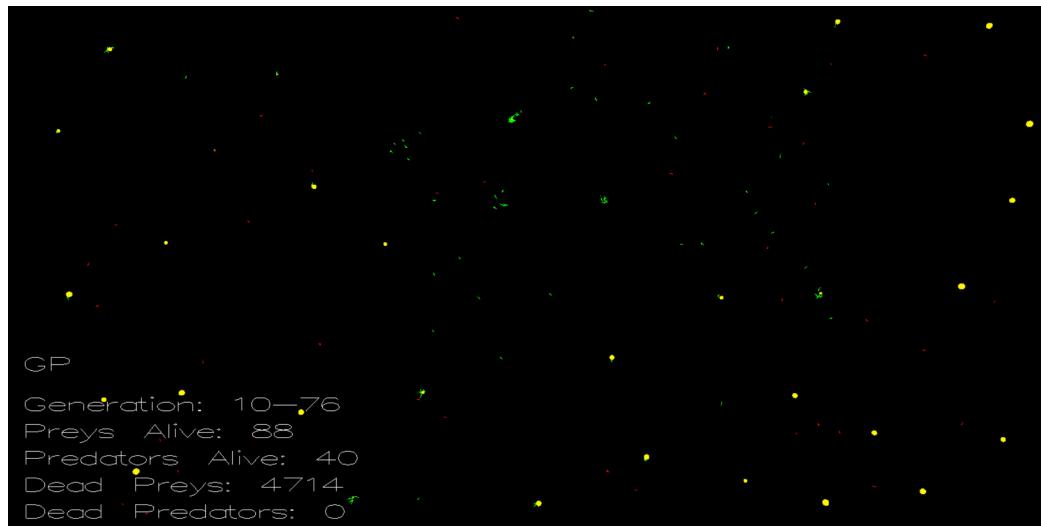
FIGURE A.1: Graphical user interface of the simulator

In the previous picture is possible to observe the environment of a running simulation. That environment has wrap boundaries, therefore when an agent overcomes the barriers it will reappear in the opposite side of the environment. The environment is composed by three types of agents, each one with its own graphical representation. The food resources are represented by *yellow circles*. The Boids (predators and preys) are represented by the same shape, *triangles*, wherein the predators are graphically represented by *red triangles*

and the preys by *green triangles*. As referred before, the preys are also characterised by features of its tails (size and brightness). However, due the small size of each Boid (triangle), those features of the tails are not graphically represented.

It is supposed that the movement of the Boids will evolve over generations, making the preys more efficient to catch the food resources and to run away from predators. While the predators will be more efficient to catch and steal energy from preys. The own energy of each agent will dictate their size, when they gain energy their size will become bigger but when they lose energy their size will decrease.

Also in the previous picture, it is possible to observe six captions in white text. The first two captions inform the user of the representation used to represent the motion rules and of the number of the current run and generation. The remaining captions provide information on the number of preys and predators alive at the moment and on the number of preys and predators that died during the current run.

Appendix B

Research Plan

The software development methodology used in this dissertation is a **lightweight variation of extreme programming** (XP), picture B.1, which is **intended** for only **one person**. This methodology, the XP, is intended to improve the software quality and responsive capacity to change of customer's requirements.

Planning/Feedback Loops
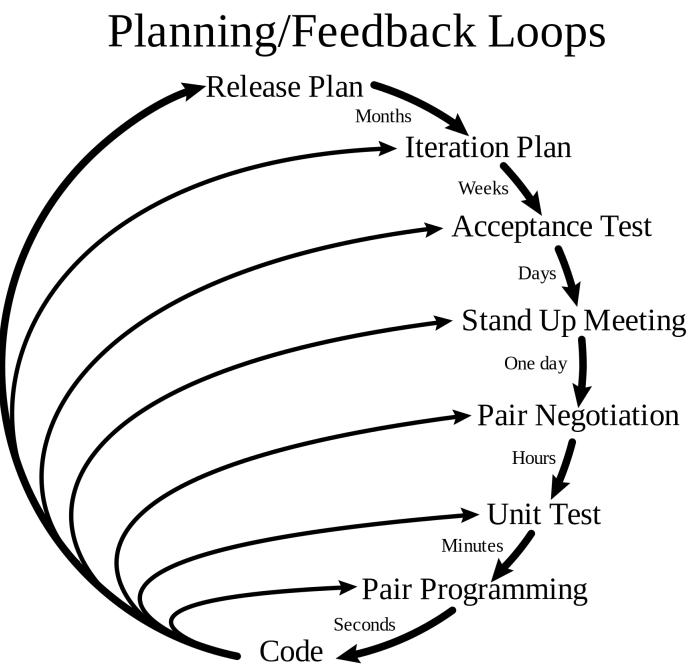

FIGURE B.1: Extreme Programming life cycle

The XP is an agile process and therefore it is a software development model open to future changes. It is an interactive process by having multiple short development cycles, rather than a long one. It considers that changes in the planning are natural and so the software process should be prepare to changes rather than trying to define a stable set

of requirements. To minimise the risk of incorrect decisions can be carried experiments (prototypes).

It aims to provide the highest value for a client in the fastest possible way, by prioritising the development of the most important features. XP was developed behind four core values: communication, simplicity, feedback and courage. Twelve values were derived from the core values: planning game, small releases, metaphor, simple design, testing, re-factoring, pair programming, collective code ownership, continuous integration, 40 hours a week, on-site customer and coding standards.

As described in some papers [69, 70], not all the core values and the derived values are suitable for a team of only one man. In lone programmer scenarios, the developer may be the own customer and since the team consists of only one element, he has the liberty to use only the characteristics of XP which he finds necessary. The practises that cannot be easily adopted by a single programmer are two: pair programming (since that the team is constituted by only one programmer) and collective code ownership (but even if there isn't a benefit of sharing code between the team, there are advantages in having a version control, such as Git or SVN, which facilitate the search and control of versions).

The reasons which led the author of this dissertation to choose this software development process are hereafter enumerated:

- Changes in the project are regarded as natural

- Prioritisation of feature at the moment

- Frequent releases of new versions

- Uses systems of version control

- Test driven development

- Acceptance Tests

- Quick feedback

Since this dissertation is a research project changes can easily occur, distorting the prioritisation of functionalities. Flaws in the logic of some functionalities can also occur and they usually take some time until they are discovered. So, since various versions are released it's important to have a system to control the several versions, in order to find the source of those errors.

Planning

In this section, the planning for this dissertation is presented. The planning was divided in two parts, according to the semesters. It should be noted, that for the first semester a week of work corresponds to 12 hours, but in the second semester, it corresponds to 30 hours of work.

The software development process used is an agile process and so the requirements can change, it is likely that the proposed plan won't be strictly equal to the planning that will be followed. Also, the comparison of the planning of the two semesters may suggest that the planning for the first half is heavier, but that is far from the reality. Since that in the first semester is proposed the development of simple versions of the simulator. In the second semester, the simulator will be completed and validated, also, in this semester a considerable number of experiments will be held.

Another important aspect in the second half is the inclusion of tasks for analysis of the results. These tasks are intended to find and correct possible errors in the simulator, and so, it is intended that the beginning of the next tasks is practically free of bugs. Those tasks are not a replacement but rather an addition to the tests at the end of each functionality.
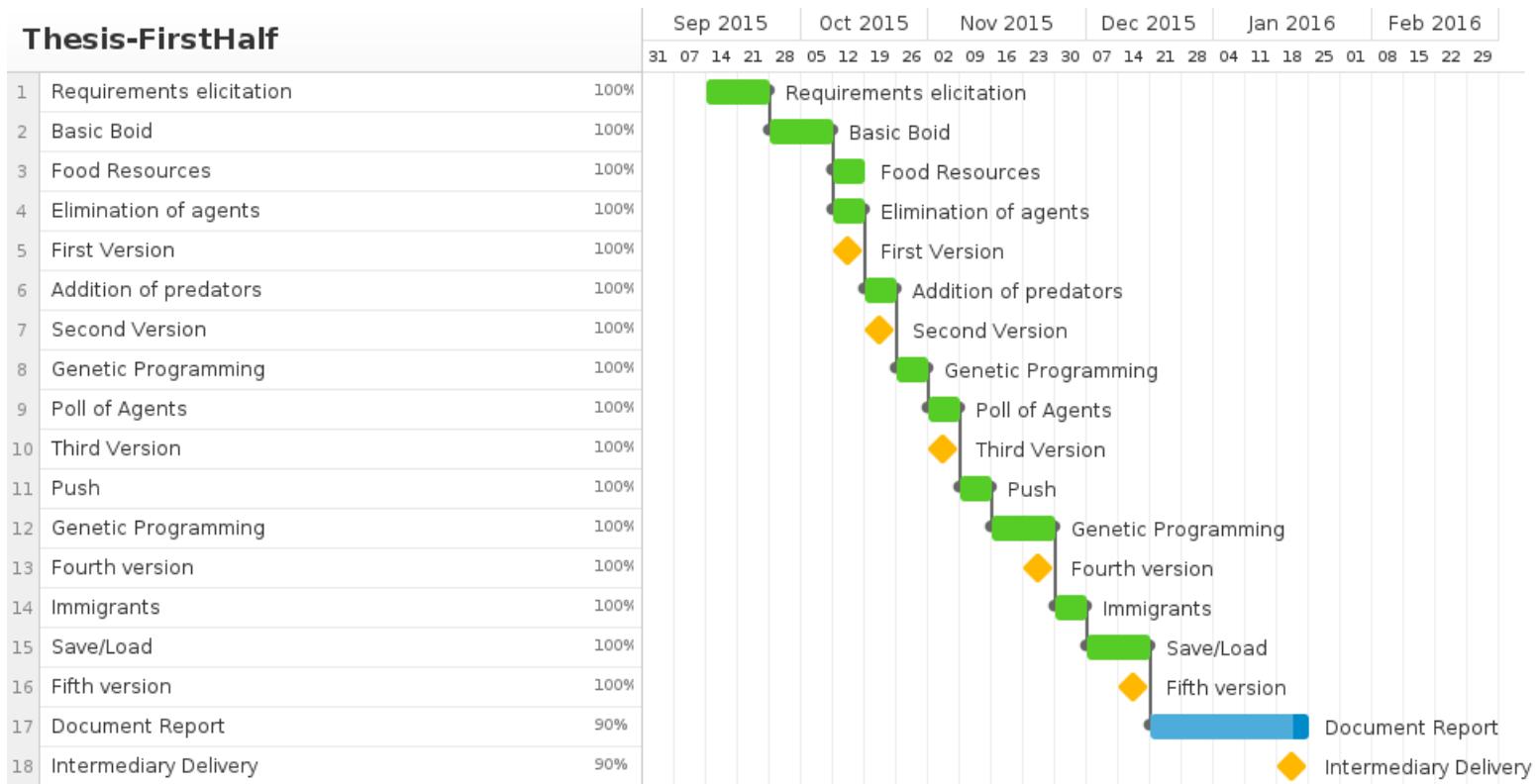
FIGURE B.2: Gantt chart representing the allocated time for first semester
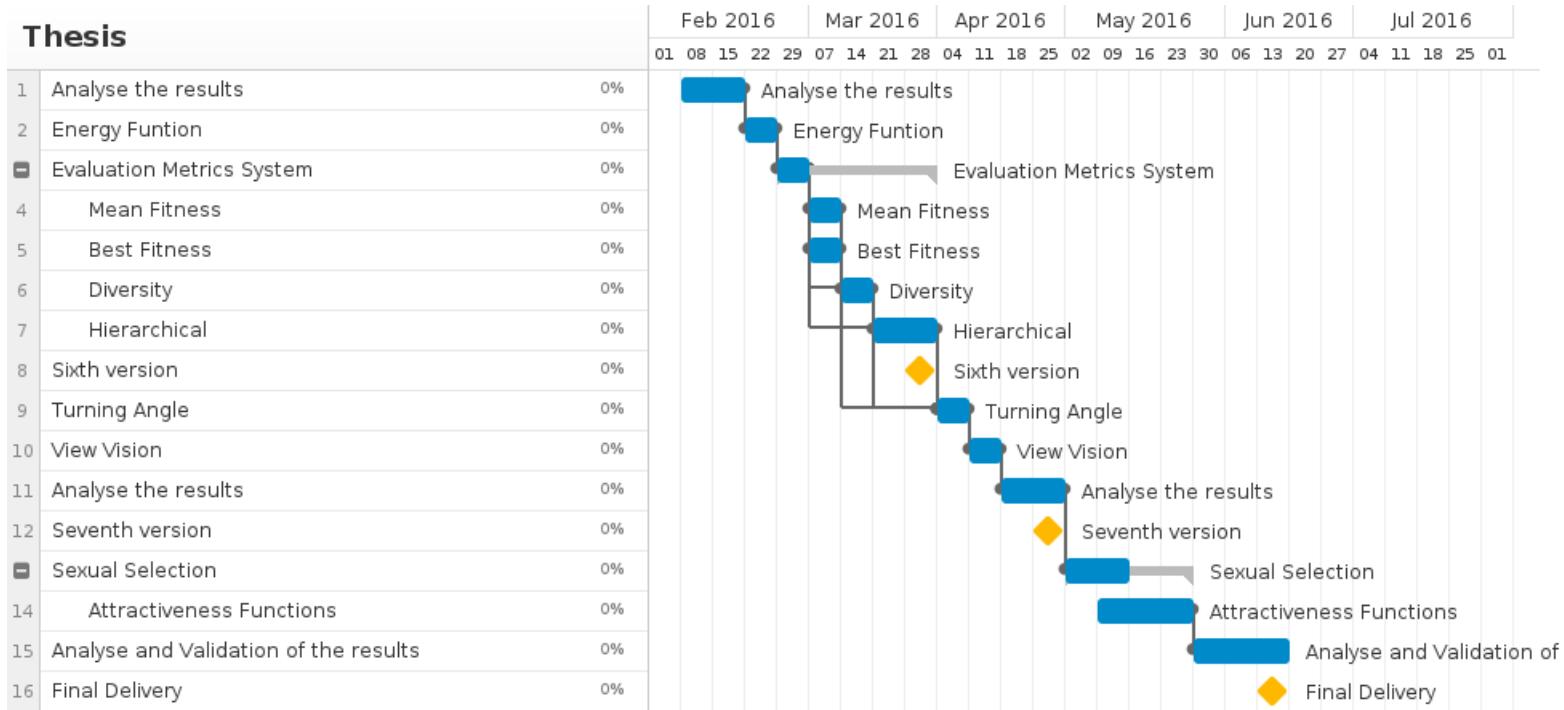
FIGURE B.3: Gantt chart representing the allocated time for second semester

Bibliography

[1] Charles Darwin. On the origin of species, 1859.

[2] Ronald Aylmer Fisher. The evolution of sexual preference. *The Eugenics Review*, 7(3):184, 1915.

[3] Ronald Aylmer Fisher. *The genetical theory of natural selection: a complete variorum edition*. Oxford University Press, 1930.

[4] Amotz Zahavi. Mate selection—a selection for a handicap. *Journal of theoretical Biology*, 53(1):205–214, 1975.

[5] Geoffrey F Miller and Peter M Todd. Evolutionary wanderlust: Sexual selection with directional mate preferences. *From Animals to Animats 2: Proceedings of the Second International Conference on Simulation of Adaptive Behavior*, 2:21, 1993.

[6] Charles Darwin. The descent of man and selection in relation to sex, 1871.

[7] CJ Barnard and DBA Thompson. Why feed in flocks. *Gulls and Plovers*, pages 1–45, 1985.

[8] Craig W Reynolds. Flocks, herds and schools: A distributed behavioral model. *ACM Siggraph Computer Graphics*, 21(4):25–34, 1987.

[9] Lee Spector, Jon Klein, Chris Perry, and Mark Feinstein. Emergence of collective behavior in evolving populations of flying agents. In *Genetic and Evolutionary Computation—GECCO 2003*, pages 61–73. Springer, 2003.

[10] Adam Birt and Samuel Shaw. Evolving boids: Incorporating machine learning into artificial life. *Journal of Computing Sciences in Colleges*, 18(5):290–291, 2003.

[11] António Leitão and Penousal Machado. Mate choice in evolutionary computation. In Amir H. Gandomi, Amir H. Alavi, and Conor Ryan, editors, *Handbook of Genetic Programming Applications*, pages 155–177. Springer International Publishing, 2015. ISBN 978-3-319-20882-4. doi: 10.1007/978-3-319-20883-1_7. URL `http://dx.doi.org/10.1007/978-3-319-20883-1_7`.

[12] Thomas Bäck, D.B. Fogel, and Z. Michalewicz. *Evolutionary Computation 1: Basic Algorithms and Operators*. IOP press, 2000.

[13] T. Bäck, D.B. Fogel, and Z. Michalewicz. *Evolutionary Computation 2: Advanced algorithms and operators*. Institute of Physics Pub., 2000.

[14] Agoston E Eiben and James E Smith. *Introduction to evolutionary computing*. Springer Science & Business Media, 2003.

[15] William M Spears, Kenneth A De Jong, Thomas Bäck, David B Fogel, and Hugo De Garis. An overview of evolutionary computation. In *Machine Learning: ECML-93*, pages 442–459. Springer, 1993.

[16] John Henry Holland. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press, 1992.

[17] Hans-Paul Paul Schwefel. *Evolution and optimum seeking: the sixth generation*. John Wiley & Sons, Inc., 1993.

[18] I Rechenberg. Evolutionstrategie 1 lab course, 1978.

[19] L. J. Fogel, A. J. Owens, and M. J. Walsh. *Artificial Intelligence through Simulated Evolution*. John Wiley, New York, USA, 1966.

[20] John R Koza. *Genetic programming: on the programming of computers by means of natural selection*, volume 1. MIT press, 1992.

[21] Darrell Whitley. An overview of evolutionary algorithms: practical issues and common pitfalls. *Information and software technology*, 43(14):817–831, 2001.

[22] Darrell Whitley, Soraya Rana, John Dzubera, and Keith E. Mathias. Evaluating evolutionary algorithms. *Artificial Intelligence*, 85(1–2):245 – 276, 1996. ISSN 0004-3702. doi: http://dx.doi.org/10.1016/0004-3702(95)00124-7. URL http://www.sciencedirect.com/science/article/pii/0004370295001247.

[23] Marek Obitko. Genetic algorithms, dec 2015. URL http://www.obitko.com/tutorials/genetic-algorithms.

[24] Stuart J. Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson Education, 2 edition, 2003. ISBN 0137903952.

[25] Fred Glover. Future paths for integer programming and links to artificial intelligence. *Computers & operations research*, 13(5):533–549, 1986.

[26] Scott Kirkpatrick, C Daniel Gelatt, Mario P Vecchi, et al. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.

[27] Matej Črepinšek, Shih-Hsi Liu, and Marjan Mernik. Exploration and exploitation in evolutionary algorithms: a survey. *ACM Computing Surveys (CSUR)*, 45(3):35, 2013.

[28] Jaroslaw Arabas, Zbigniew Michalewicz, and Jan Mulawka. Gavaps-a genetic algorithm with varying population size. In *Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the First IEEE Conference on*, pages 73–78. IEEE, 1994.

[29] Jürgen Branke and Hartmut Schmeck. Designing evolutionary algorithms for dynamic optimization problems. In *Advances in evolutionary computing*, pages 239–262. Springer, 2003.

[30] Franz Oppacher, Mark Wineberg, et al. The shifting balance genetic algorithm: Improving the ga in a dynamic environment. In *Proceedings of the genetic and evolutionary computation conference*, volume 1, pages 504–510, 1999.

[31] Andrea Grosso, Marco Locatelli, and Fabio Schoen. A population-based approach for hard global optimization problems based on dissimilarity measures. *Mathematical Programming*, 110(2):373–404, 2007.

[32] El-Ghazali Talbi. *Metaheuristics: from design to implementation*, volume 74. John Wiley & Sons, 2009.

[33] K Dejong. An analysis of the behaviour of a class of genetic adaptive systems. doctorat dissertation, dept. of computer and communication sciences, university of michigan. *Ann Arbor*, 1975.

[34] Nils Aall Barricelli et al. Esempi numerici di processi di evoluzione. *Methodos*, 6 (21-22):45–68, 1954.

[35] Trivers Robert. Parental investment and sexual selection. *Sexual Selection & the Descent of Man, Aldine de Gruyter, New York*, pages 136–179, 1972.

[36] Alfred Russel Wallace. *Tropical nature, and other essays*. Macmillan and Company, 1878.

[37] Richard Dawkin. The selfish gene. *Oxford ni versity Press*, 1:976, 1976.

[38] Penousal Machado and António Leitão. Evolving fitness functions for mating selection. In *Genetic Programming*, pages 227–238. Springer, 2011.

[39] António Leitao, Joao C Neves, and Penousal Machado. A self-adaptive mate choice model for symbolic regression. In *Evolutionary Computation (CEC), 2013 IEEE Congress on*, pages 8–15. IEEE, 2013.

[40] Ekaterina A Holdener and Daniel R Tauritz. Learning offspring optimizing mate selection. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation*, pages 1109–1110. ACM, 2008.

[41] Susmita De, Sankar K Pal, and Ashish Ghosh. Genotypic and phenotypic assortative mating in genetic algorithm. *Information Sciences*, 105(1):209–226, 1998.

[42] Michael Ratford, Andrew Tuson, and Henry Thompson. The single chromosome's guide to dating. In *Artificial Neural Nets and Genetic Algorithms*, pages 171–174. Springer, 1998.

[43] Severino F Galán, Ole J Mengshoel, and Rafael Pinter. A novel mating approach for genetic algorithms. *Evolutionary computation*, 21(2):197–229, 2013.

[44] Yuexin Jiang, Daniel I Bolnick, and Mark Kirkpatrick. Assortative mating in animals. *The American Naturalist*, 181(6):E125–E138, 2013.

[45] Michael Ratford, A Tuson, and Henry Thompson. Applying sexual selection as a mechanism for obtaining multiple distinct solutions. *DAI RESEARCH PAPER*, 1997.

[46] Rodney Fry, Stephen L Smith, and Andy M Tyrrell. A self-adaptive mate selection model for genetic programming. In *Evolutionary Computation, 2005. The 2005 IEEE Congress on*, volume 3, pages 2707–2714. IEEE, 2005.

[47] Jose Sánchez-Velazco and John A Bullinaria. Gendered selection strategies in genetic algorithms for optimization. *survival*, 8(6):11, 2003.

[48] Kai Song Goh, Andrew Lim, and Brian Rodrigues. Sexual selection for genetic algorithms. *Artificial Intelligence Review*, 19(2):123–152, 2003.

[49] Mohammad Jalali Varnamkhasti and Lai Soon Lee. A genetic algorithm based on sexual selection for the multidimensional 0/1 knapsack problems. In *International Journal of Modern Physics: Conference Series*, volume 9, pages 422–431. World Scientific Publishing Company, 2012.

[50] Iztok Lebar Bajec, Nikolaj Zimic, and Miha Mraz. Simulating flocks on the wing: the fuzzy approach. *Journal of Theoretical Biology*, 233(2):199–220, 2005.

[51] Frank Heppner and Ulf Grenander. A stochastic nonlinear model for coordinated bird flocks. *AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, WASHINGTON, DC(USA). 1990.*, 1990.

[52] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Neural Networks, 1995. Proceedings., IEEE International Conference on*, volume 4, pages 1942–1948 vol.4, Nov 1995. doi: 10.1109/ICNN.1995.488968.

[53] Alberto Colorni, Marco Dorigo, Vittorio Maniezzo, et al. Distributed optimization by ant colonies. In *Proceedings of the first European conference on artificial life*, volume 142, pages 134–142. Paris, France, 1991.

[54] DT Pham, Afshin Ghanbarzadeh, E Koc, S Otri, S Rahim, and Mb Zaidi. The bees algorithm. technical note. *Manufacturing Engineering Centre, Cardiff University, UK*, pages 1–57, 2005.

[55] Akira Okubo. Dynamical aspects of animal grouping: swarms, schools, flocks, and herds. *Advances in biophysics*, 22:1–94, 1986.

[56] Dario Floreano and Claudio Mattiussi. *Bio-inspired artificial intelligence: theories, methods, and technologies.* MIT press, 2008.

[57] Iain D Couzin, Jens Krause, Richard James, Graeme D Ruxton, and Nigel R Franks. Collective memory and spatial sorting in animal groups. *Journal of theoretical biology*, 218(1):1–11, 2002.

[58] Lukas Felzmann. How the science of swarms can help us fight cancer and predict the future, dec 2015. URL `http://www.wired.com/2013/03/powers-of-swarms`.

[59] Michael Girard and Susan Amkraut. Eurhythmy: Concept and process. *The journal of Visualization and computer animation*, 1(1):15–17, 1990.

[60] Russ C Eberhart and James Kennedy. A new optimizer using particle swarm theory. In *Proceedings of the sixth international symposium on micro machine and human science*, volume 1, pages 39–43, 1995.

[61] Morten Lovbjerg, Thomas Kiel Rasmussen, and Thiemo Krink. Hybrid particle swarm optimiser with breeding and subpopulations. In *Proceedings of the Genetic and Evolutionary Computation Conference*, volume 2001, pages 469–476. Citeseer, 2001.

[62] Vladimiro Miranda and Nuno Fonseca. Epso-evolutionary particle swarm optimization, a new algorithm with applications in power systems. In *Proc. of the Asia Pacific IEEE/PES Transmission and Distribution Conference and Exhibition*, volume 2, pages 745–750. Citeseer, 2002.

[63] Crina Grosan, Ajith Abraham, Sangyong Han, and Alexander Gelbukh. Hybrid particle swarm–evolutionary algorithm for search and optimization. In *MICAI 2005: Advances in Artificial Intelligence*, pages 623–632. Springer, 2005.

[64] Yen-Wei Chen, Kanami Kobayashi, Hitoshi Kawabayashi, and Xinyin Huang. Application of interactive genetic algorithms to boid model based artificial fish schools. In *Knowledge-Based Intelligent Information and Engineering Systems*, pages 141–148. Springer, 2008.

[65] Lee Spector, Jon Klein, and Maarten Keijzer. The push3 execution stack and the evolution of control. In *Proceedings of the 2005 conference on Genetic and evolutionary computation*, pages 1689–1696. ACM, 2005.

[66] Jon Klein. Breve: a 3d environment for the simulation of decentralized systems and artificial life. In *Proceedings of the eighth international conference on Artificial life*, pages 329–334, 2003.

[67] Nelson Minar, Roger Burkhart, Chris Langton, and Manor Askenazi. The swarm simulation system: A toolkit for building multi-agent simulations, 1996.

[68] Mitchel Resnick. Starlogo: An environment for decentralized modeling and decentralized thinking. In *Conference companion on Human factors in computing systems*, pages 11–12. ACM, 1996.

[69] Edward Akpata and Karel Riha. Can extreme programming be used by a lone programmer? *Systems Integration*, page 167, 2004.

[70] Ravikant Agarwal and David Umphress. Extreme programming for a single person team. In *Proceedings of the 46th Annual Southeast Regional Conference on XX*, ACM-SE 46, pages 82–87, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-105-7. doi: 10.1145/1593105.1593127. URL `http://doi.acm.org/10.1145/1593105.1593127`.