Simulation Model to Aid Understanding into
Self-organised Collective Behaviour

Gregory Jones

Cognitive Science
2004/2005

The candidate confirms that the work submitted is their own and the appropriate credit has been given where reference has been made to the work of others.

I understand that failure to attribute material which is obtained from another source may be considered as plagiarism.

(Signature of student)_____

Summary

This report focuses on the conditions under which a population of agents initially able to survive independently will produce progressively more integrated behaviour, such that they become interdependent on one another, through the process of self-organisation. A simulation model was produced in order to understand these conditions. Many existing models focus on accurately modelling self-organised collective behaviour but not its emergence from independent behaviour. It was found that the transition was most clear when agents could benefit from each other's presence without having to make radical behavioural changes, and when their life-span was so short relative to their movement speed that their position within an aggregate effected the evolution of their behavioural strategy, causing spatial organisation of different strategies within aggregates.

Acknowledgements

Contents

Chapter 1

Introduction

There are many examples in nature of systems which are made up constituent parts which were once individual systems themselves. For example there is evidence that multicellular organisms evolved from aggregations of unicellular organisms [26], complex ecosystems are made up of simple transactions of energy and nutrients between organisms [25], and social systems are made up a collection of interactions between individuals. This project focuses on how and why a transition from a collection of independent entities to a single interdependent whole could occur.

The aim of this project was to produce a simulation model in order to gain an understanding into the conditions under which a group of agents initially behaving independently will self-organise to behave as a collective whole. It was hoped that producing such a model would allow conclusions to be drawn about self-organisation in natural systems. However, it was taken into account that extreme caution must be taken when using results obtained from computer simulations to make assertions about the natural world, and it was not be taken for granted that this is a valid approach.

The idea was to make the simulation as general as possible; not to model a specific example of a self-organising system in nature, but the general process of self-organisation so that the model can be interpreted in different ways, and general conclusions drawn.

Self-organisation and collective behaviour are widely applicable concepts, and relevant to cognitive science in a number of ways. It has been suggested that the nervous system is a self-organising system [39], artificial neural networks have been found to be self-organising [4], collective behaviour is a key concept in the field of swarm intelligence, and the techniques which are often used to simulate these, such as genetic algorithms and clustering algorithms, are commonly used in other areas of cognitive science [19, 27].

The project objectives were:

- To assess existing simulation models which demonstrate self-organised collective behaviour.

- To better understand the problem of altruism, and try to learn why co-operative behaviour occurs.

- To build a simulation model which can be analysed in different experimental conditions.

- To relate the results of studying the model to natural systems in the real world.

- To assess the validity of using this approach to draw conclusions about nature.

Chapter 2

Background reading

2.1 Introduction

The use of computers to simulate natural processes has been a common trend since their invention. The field of artificial intelligence set out to recreate properties of human intelligence [36]. It was based on the assumption that the driving principle behind human intelligence is our ability to reason. Since the enlightenment period western philosophy has been dominated by the view that reason is the source of knowledge [17]. When the computer was invented it appeared to be the ideal tool to emulate rational intelligence: centrally controlled and determined by logic alone. However it soon became apparent that there were limitations to this approach; tasks we commonly think of as simple, such as recognising letters, could not be expressed as a logical set of rules.

More recently, natural processes have been simulated in a completely different way. The emphasis has moved away from serial, centrally controlled computation of discrete facts and propositions, to parallel, distributed systems which aim to model complex, unpredictable phenomena. Research in swarm intelligence (where intelligence is an emergent property held by a 'swarm' of many unintelligent agents) aims to tackle similar problems to traditional AI, but takes a decentralised approach.

Artificial neural networks, originally based on the neuron model by McCulloch and Pitts [27], offer a bottom up approach to modelling the brain. The networks are able to generalise or cluster together data and are tolerant of noise. Intelligence is seen as an emergent property, brought about by many simple low-level interactions.

Many of these approaches have been developed not to accurately model nature but to provide useful computational tools. For example, neural networks have been used in letter recognition systems [16] and genetic algorithms [19] as a robust design tool. The simulations this project focuses on are developed to

help us better understand natural systems.

Many models of natural systems are based on linear equations. However, the new science of complexity has revealed problems with this approach [8]. By taking the interactions of many components of a system into account, circular feedback effects produce non-linear relations between them (for example in population dynamics [25]). Computer simulation provides a way modelling these systems without relying on linear equations. However, it remains unclear as to whether natural systems have such fundamental differences from computer simulations that simulation techniques teach us very little about nature.

2.2 The transition from independent to collective behaviour

The term 'collective behaviour' is typically used to describe groups of organisms behaving as a whole, for instance birds flying in a flock. Without wanting to get into difficult questions about what constitutes 'behaviour', it would be useful to use this term to describe any group of entities which in some way act together. Many of the major transitions in evolution discussed by Maynard-Smith [26] involve entities which initially behave independently becoming an inter-dependent whole. For instance one of the first of the major transitions is from independently replicating nucleic acid molecules to sets of linked molecules which replicate together. By being general about what is meant by collective behaviour it may be possible to identify some similar properties of the conditions under which these transitions occur.

These transitions, which often occur through the process of evolution, require the entities involved to become interdependent. In the case of groups of organisms this involves the production of social behaviour, either by organisms cooperating or exploiting one another. There has to be some selective advantage for these behaviours to evolve, and in the case of altruistic behaviour it is often not clear what the advantage is. Possible explanations are that altruistic social behaviour might be the only way a group can survive in a particularly harsh environment (as in the case of the slime mould [34]), or that reciprocal cooperation makes altruistic behaviours an evolutionary stable strategy within a population, or that altruistic behaviour tends to occur between genetically similar individuals (Hamiltonian kin selection [15, 38]). By trying to simulate the evolution of cooperative or exploitative behaviour, the later chapters of this report aim to better understand why they occur in nature.

While Maynard-Smith's major transitions provide milestones in evolution by which to classify living systems into levels, there are often grey areas where it is difficult to decide whether something should be described as a single entity or a group of entities. For example, Shapiro and Dworkin [33] note that the collective behaviour of bacteria is so integrated that it is difficult to perceive them as a group of individual organisms, and they might be better described as a multi-cellular organism.

A well documented example of this difficulty is the unusual life-cycle of the cellular slime mould Dictyostelium discoideum [34]. Under conditions where food is scarce a group of individual amoebae aggregate to form a mould which then behaves like an individual organism itself; it moves as a whole
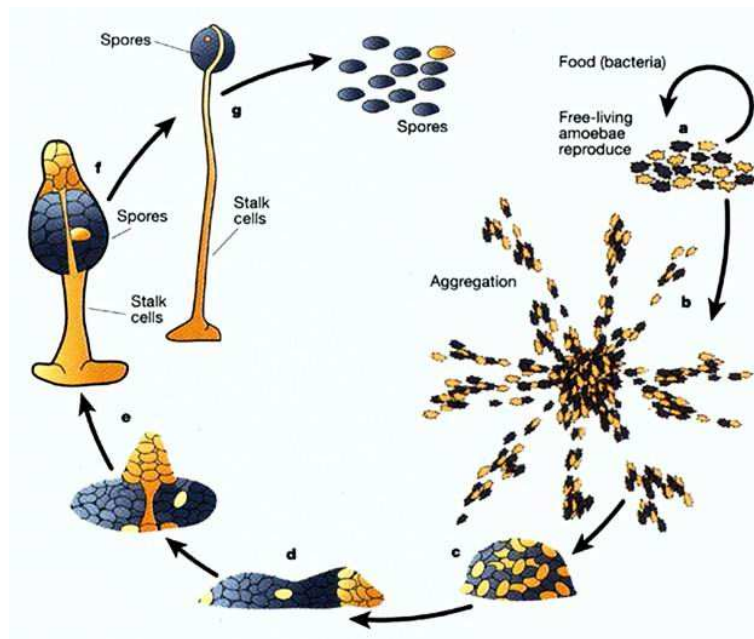
Figure 2.1: Life cycle of Dictyostelium discoideum (from [12])

slug-like aggregate and then its cells differentiate to form spores and stalk like a mushroom, and the spore cells are scattered from the top of the stalk (See figure 2.1). Since the stalk cells die to allow the spores to reach a better location, the mould acts as a whole, and the amoebae can no longer be seen as individuals. It is impossible to say whether the individual cells are acting altruistically or the whole aggregate is acting selfishly.

These difficulties with dividing life into different levels has caused many ecologists to see life as a single interconnected web [24], [8]. The transition from independent to collective behaviour forms the links between these levels in the network: from prokaryotes to eukaryotes, from uni to multi-cellularity, from individual organisms to societies of organisms. It is possible that the conditions under which collective behaviour arises hold some similar properties in all living systems. Also, it could be that these transitions occur through the process of self-organisation.

2.3 Self-organisation in living systems

According to Darwinian evolution, order in living organisms is brought about largely by chance through random mutations and natural selection. More recently it has been suggested that the presence of different kinds of feedback loops allow initially disordered conditions to settle into organised patterns [18]. The most basic of these feedback loops are the catalytic hypercycles described by Eigen [21]. In biochemical systems, networks of catalytic reactions form by one reaction causing the next. Often these networks form closed loops of reactions which are very stable structures capable of self-replication. These hypercycles form the fundamental metabolic processes in living cells. Another example of these

feedback loops is found in Robert May's [25] work on population dynamics where the population from one year effects the next. He found that when the rate of population growth reaches a critical point, populations from year to year begin to oscillate between two stable values from year to year. As it increases further, these become three and then four phase oscillations, and eventually the population values fluctuate chaotically. May called these critical points 'bifurcations'.

The Belgian chemist Ilya Prigogine developed a theory of self-organisation [31]. Traditionally, science has aimed to model the world mechanically, and studied its components in isolation. According to Prigogine only a small set of systems known as 'closed' systems can be understood mechanically. These systems are subject to the second law of thermodynamics; they have a tendency to become progressively more disorganised until they reach a state of maximum disorder (entropy). However, the vast majority of natural systems including biological and social systems are 'open' dynamic systems which cannot be described in sequential terms. In contrast to closed, mechanical systems, many natural systems become progressively more organised, through evolution for instance. To Prigogine, this progressive order in nature is brought about through a process of self-organisation, where many low level interactions between components of a system bring about high level order. All living systems and many non-living systems can be said to be self-organising systems. Self-organisation has been defined as the "irreversible process which results from the co-operative interaction of subsystems in complex structures of the whole system." (Ebeling and Feistel, 1986 [13])

The major defining characteristics of self-organising systems described by Prigogine are outlined as follows:

- They are 'open' systems in constant interaction with their environment. This makes it near impossible to study them in isolation from their environment and from other systems, as the components of the system are constantly changing (Prigogine coined the term 'dissipative structures' to describe these systems).

- They actively create and maintain their structure in 'far-from-equilibrium' conditions. Since they are a source of order they can be said to be 'creative'. This means that order arises from within rather than from some external controller or blueprint. By 'far-from-equilibrium', Prigogine means the environment is not static but constantly changing as it interacts with the system.

- They are complex in that:

  - They are composed of so many inter-dependent components that causal relationships between them cannot be established. Future states cannot be predicted from the current state without actually simulating the complete system. For example, in ecology it is impossible to predict that a slight change in temperature might wipe out a whole species.

  - Their components are interconnected, and feedback loops allow tiny changes in conditions to be amplified. The cyclic nature of feedback has been described as a fundamental property of living systems [8]. This is what makes recurrent neural networks (where each neuron is

6

connected to both itself and every other neuron, meaning activity can spread in any direction) behave in a life-like unpredictable way [4]. The instability of these networks has been seen as a disadvantage of them, as Carling suggests: "Perhaps the best way of avoiding instability is by not having feedback at all" [9]. But this instability could be seen as a life-like feature not normally present in computer models.

2.4   Computer simulation models as a method for studying nature

Computers have only been available as a tool to model nature for a relatively short period of time. Simulation models are advantageous in that they force the designer to be completely specific — it is impossible to produce a simulation of a half-baked idea; the process of implementing a simulation requires the consideration of every aspect of the system to be modelled, and the assumptions involved in doing this. Also, with the use of computer graphics the model can be visualised which not only makes it more interesting and easier to identify problems with, but provides something qualitative to compare to the natural system being simulated in a direct, intuitive way. Without the use of graphics, the strikingly life-like behaviour of John Conway's 'Life' [14] simulation might never have been discovered.

Despite these attractive qualities, the validity of using computer simulation models to gain understanding into the underlying properties of natural systems has to be questioned. It could be that natural systems have fundamentally different properties which cannot be expressed as formal programs. Early work in AI demonstrated some very convincing intelligence-like properties which turned out to be working in a fundamentally different way to human intelligence and were extremely limited in the range of 'intelligent' behaviours they could produce (Joseph Weizenbaum's 'Eliza' [37] program for instance). So new techniques should be used with caution and scepticism.

Computer simulations rely on the assignment of discrete numerical values to every aspect of the system modelled: continuous space and time are divided into discrete blocks or slices, the intensities of different physical conditions are assigned standardised values. There is a limit to how many factors can be modelled, but the range of factors affecting natural systems is virtually limitless; often complex physical factors like friction and air resistance are ignored, or levels of only a few chemicals are modelled, when in reality there are thousands [10]. There is also a limit to the size of the system modelled and how many systems which in turn affect the system are modelled. It could be argued that since natural ecosystems are all connected in some way, it is impossible to accurately model any one system without modelling the entire biosphere or even the entire universe. Clearly to model a natural system a great deal of approximation must take place. The question is whether studying these simulations can teach us anything about natural systems.

Di Paolo, Noble and Bullock [11] outline two extreme views on the epistemological value of simulation models: The 'strong A-life' view that simulations are in fact genuine instances of the phenomenon being simulated and so any observations made of them are just as valuable as observing the actual system, and the opposing view that since every aspect governing the simulation must be known beforehand

in order to create it, no new knowledge can be gained from running the simulation.

They argue that simulation models are similar to thought experiments, commonly used in philosophy, in that they involve only existing knowledge; no new empirical data is added during the experiment, but still provide a useful tool for understanding natural phenomena. Also, while simulation models have their limitations they do overcome some of the difficulties with the assumptions made in other modelling techniques (for example equational models).

Daniel Bruiger notes that simulation implies only selective functional equivalence and should not be confused with replication [5]. He claims that simulations appear functionally identical to natural systems only "when viewed through formalising spectacles.":

> "Science, as an intentional creation, may be a completely formalisable system, but this does not mean that nature is. That physical processes are rule-governed, and that these rules can be formally expressed in an exhaustive way, are mere assumptions. Physical processes may appear rule-governed for the same reason that they are recognised at all: because they correspond to categories which are intentional creations. Nature may be unfathomable because it is infinitely complex in one or both directions of scale. Alternately, thought itself may be interminable, condemned to essential incompleteness." (Bruiger, 2003)

J.S. Lansing [20] used a computer simulation to demonstrate the importance of bottom-up coordination in a real world problem. For centuries Balinese farmers have used their 'water temple' system of coordinating irrigation among hundreds of farming communities. Since water is scarce and there are many farmers, coordinating irrigation on the slopes is a difficult and complex task. The Indonesian government during the 'green revolution' of the 70's abolished the temples and encouraged farmers to grow three crops a year instead of two and use pesticides and fertilisers in an attempt to modernise farming and increase yields. However without the water temple system, there was not enough water, and without the stable ecosystems they provided, disease and pests rapidly became a problem and yields dropped. Lansing's simulation modelled the water temples and demonstrated that they were the most efficient way of coordinating irrigation. This could not have been shown by studying the properties of a single water temple in isolation - the system relies on low-level cooperation between adjacent farmers to coordinate irrigation over the whole mountain slope. Modern methods of doing this involve flying a plane over the mountain and taking photos to produce a topographical map so that irrigation can be directed; the opposite approach.

So while there are philosophical worries about the validity of generalising what is learnt from a simulation to the real world, Lansing's example shows it is possible to use them as a practical tool to solve real problems, and in a sense this gives them ecological validity.

2.5   Existing simulation models of collective behaviour

Before designing the simulation, a number of existing models were assessed in terms of their implications regarding the conditions that give rise to self-organised collective behaviour.

8

2.5.1   Stevens and Sosic — Emergent Behaviour in Slime Mould Environments

This model simulates the aggregation phase of the cellular slime mould *dictyostelium discoideum* life-cycle (when independent amoebae clump together to form an aggregate which then moves as a whole) [35]. Agents in the model (individual amoebae) move in continuous space, but chemical distribution is modelled as a grid of intensities. Amoebae respond to the signalling chemical cAMP by moving towards it and also releasing more cAMP themselves. So if a single amoeba in the centre releases a little cAMP they will all gradually move towards it, forming clusters which branch together, eventually forming a single connected shape. Fractal branching patterns are formed, similar to those found in real slime moulds.

The model gives a good example of how life-like collective behaviour can be produced from very simple conditions. However, the simplicity of the model makes the conditions seem somewhat engineered by the designers. The branching has to be 'set off' by setting one agent to release cAMP without detecting some cAMP themselves. In real slime moulds cAMP production occurs when food is scarce and is a survival technique; if the chances of survival as individuals are looking unlikely it is best to move far away and a fast way of doing this is to aggregate (and then form a stalk and spores to be dispersed). Since there is no notion of survival in this model (food is not modelled and the agents cannot die) the agents cannot be said to be cooperating.

2.5.2   Buzing, Eiben and Schut — Emerging Communication and Cooperation in Evolving Agent Societies

In this paper, a version of the 'sugarscape' world developed by Axtell and Epstein was modified to encourage cooperation [7]. Agents forage for 'sugar' in a grid world and are limited in the amount they can consume at one time. Large 'piles' of sugar can only be consumed by two or more agents, who share the total amount equally. This cooperation pressure can be varied by modifying the maximum amount of sugar a single agent can consume alone. Agents also have the ability to give signals to indicate the location of sugar that other agents can respond to. So if an agent encounters a pile of sugar which cannot be consumed alone it can communicate to encourage other agents to come to that location.

The model is evolutionary; the likelihood of agents giving or responding to signals is represented in their genome, and they are evolved with different levels of the cooperation pressure parameter. Unexpectedly, it is reported that a tendency to respond to signals evolves earlier than a tendency to give signals. This is presumably because responding to signals gives an immediate survival advantage, whereas the advantage of giving signals is indirect (through reciprocal cooperation or kin selection).

This model gives an insight into conditions that give rise to cooperative behaviour. Importantly, agents are able to survive as individuals (if they encounter enough small piles of sugar which can be consumed alone) but evolve cooperation as it gives a significant survival advantage. This is an example of a transition from independent to collective behaviour.

2.5.3 Marshall and Rowe — Viscous populations and Their Support for Reciprocal Cooperation

Marshall and Rowe's [23] model sets out to asses the role of 'viscosity' in a population of spatially distributed self-interested agents playing a cooperation game. Highly viscous populations are those which are limited to moving slowly and only by a small amount, as though they are moving through a viscous liquid, meaning that they tend to keep the same set of neighbours for a long period of time. This means they are more likely to re-encounter the same individual a number of times in the cooperation game than if they were able to move very quickly. Reciprocal cooperation ('returning the favour' by cooperating again) is a good strategy if the partners in the game remain the same for a number of interactions, but not if the partners are constantly changing. Marshall notes that a prime explanation for reciprocal cooperation in nature is that individuals are able to recognise one another, remember the cooperative act, and return it. It could be that reciprocal cooperation can occur without the need to remember a specific individual; it is just that when viscosity is high nearby pairs of individuals are likely to interact a number of times.

The model uses a wrapped grid world. Viscosity is represented by m, the probability that an agent will move into an adjacent cell at each time step. The higher m is, the lower the viscosity, since the more likely the agents are to move at each time step. Grid cells can accommodate any number of individuals, and only individuals who occupy the same cell may interact. At each time step, a random pair of individuals from each cell are selected to play a round of the iterated prisoner's dilemma game (See [3]). They look for the number of interactions that must take place for the 'tit-for-tat' strategy (responding with the same action the other player used in the last interaction, which can lead to reciprocal cooperation) to be a collectively stable strategy (using Axelrod's condition for defining a collectively stable strategy [3]).

They find that even with high viscosity (low values of m), the local populations size (number of agents that can interact as they occupy the same cell) is to too large to allow reciprocal cooperation to be a collectively stable strategy. They conclude that viscous populations do not provide complete support for reciprocal cooperation, and so reciprocal cooperation in nature must be caused at least partly by some other factor, such as the ability to recognise previous opponents.

This model concentrates on the effects a particular condition (population viscosity) on cooperative collective behaviour. It is especially relevant to this project, as a major aim was to try to understand the conditions that bring about collective behaviour. However, the use of a two-player cooperation game means the agents are pre-determined to interact with one another; they cannot survive without social behaviour and so cannot be said to be initially behaving independently. So while it helps us to understand the conditions for collective behaviour, it is not an example of a transition from independent to collective behaviour.

2.5.4 Di Paolo — Ecological Symmetry Breaking can Favour the Evolution of Altruism

In this model, agents in a continuous two dimensional world play an action-response cooperation game [29]. Unlike Marshall and Rowe's model, the tokens in the game (referred to as energy) are distributed spatially, such that there are different amounts of energy at stake when the game is played in different regions of the world.

The action response game involves each partner taking one move, which can be one of two types; 'even' or 'odd'. The game results in one of two states: coordination where each of their moves are different (one choosing even followed by the other choosing odd, or vice-versa) or non-coordination where each of their moves are the same. Coordination results in the energy at stake being shared equally between them. Non-coordination results in the partner who took the first move getting a proportion of the energy, and the other getting nothing. This proportion, c, can be modified and represents the level of conflict of interest between the two partners, since higher values of c make defecting a better strategy.

Since the only way to acquire energy is to win it in the game, agents who get many opportunities to interact with other agents tend to do better. This causes the agents to cluster together spatially to maximise the number of local interactions between them. However, since energy is used up in the game, there will be less energy at stake in the centre of a cluster where many games are taking place than there is at the edges. This causes a pressure for agents to spread out to get access to regions with more energy. The equilibrium between high player density causing the clusters to contract and low energy resources causing the clusters to expand makes the clusters dynamic but stable structures (see figure 2.2). Also, there is some spatial organisation of agents with different strategies within the clusters; responders tend to be located at the centre of clusters. Since agents are initially randomly distributed throughout the world, and the cluster is formed entirely by small scale interactions between agents, the model demonstrates the conditions under which stable structures can be produced through self-organisation.

The model provides a good example of spatial self-organisation which is what this project aimed to investigate. However there is not a clear transition from independent to collective behaviour. Even though the agents begin in random positions and then form spatial structures, they are unable to survive independently because the only way they can gain energy is through interaction with another agent. For the aims of this project, it is desirable to find a situation where these stable, self-organising structures are formed by a population whose members are initially able to survive as individuals, and if this is possible try to understand the necessary conditions for this to occur.

2.5.5 Pfeiffer and Bonhoeffer - Scenario for the Transition to Undifferentiated Multi-Cellularity

The simulation discussed in this paper [30] is of particular relevance to this project as it specifically focuses on the evolutionary advantages of the *transition to* collective behaviour, rather than just modelling an example of pre-existing collective behaviour. Pfeiffer and Bonhoeffer note that the transition
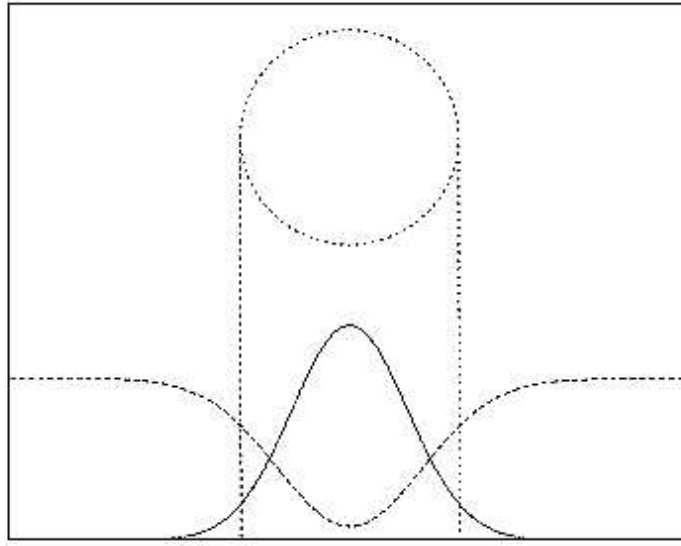
Figure 2.2: Energy and population distribution of a cluster (circle) (— Player density; - - - Energy) (taken from [29])

to multi-cellularity must have began with the clustering of undifferentiated cells, but that the main advantage of multi-cellularity comes from the division of labour between differentiated cells, so it is not immediately clear what the initial advantage of clustering was. They hypothesise that in a population of cooperating cells, clustering may be advantageous to avoid invasion from non-cooperators. They use a simulation to assess how advantageous clustering is in this situation.

Their simulation in based on a grid world, where only one cell (agent) can occupy a grid point at one time. Cells consume a food to produce a chemical energy, ATP. There are two types of cell: respirators which produce ATP slowly but consume the resource efficiently, and respiro-fermentors which consume quickly but inefficiently. The respiro-fermentors are seen as non-cooperators because only the individual benefits from the high rate of ATP production, but the cost of consuming the energy inefficiently effects other nearby individuals. When a cell produces a critical amount of ATP, if there is an empty neighbouring grid point, the cell splits in two. In the case of clustering cells, the child and parent are attached and not motile. They find that clustering respirators (cooperators) do better than non-clustering respirators, as by clustering they shield their food resource from invading non-cooperators, which supports their hypothesis that clustering is a strategy to avoid invasion.

This model gives a clear example of a condition which might give rise to the transition from independent to collective behaviour, although the authors note that evidence indicates that multi-cellularity evolved several times independently, and so this explanation is unlikely to be the only one. Agents are initially able to survive independently of each other which shows that cooperation is not pre-given and built in to the model. However, they only show that clustering would be advantageous *if it were* to evolve, they do not actually simulate the emergence of clustering.

Chapter 3

The basic model

Of the existing models studied, Di Paolo's model [29] came closest to simulating the behaviour that this project focuses on. The idea of localised energy in a pre-existing environment being 'used up' so that if clusters of agents get too big there will be less energy at stake in the centre appears to be a promising characteristic for the production of stable spatial structures through self-organisation. But the use of an action-response game as the method of accessing this energy is problematic as it forces agents to be mutually dependent on one another right from the start, so this dependence cannot be said to be an emergent property; it is explicitly built into the model. Also, it was considered that allowing for cooperative interactions between more than two agents at a time may produce more integrated collective behaviour than with Di Paolo's model. With these ideas taken on board, the initial idea for the model was formulated as follows.

3.1 Initial idea

A population of agents have to consume an energy source to survive in a spatial world. Individual agents can harvest a small proportion of the available energy at their current location. Two agents at adjacent locations can harvest more energy each than they could on their own. The proportion of the available energy that can be harvested raises exponentially with the number of adjacent agents cooperating, meaning a dense cluster of agents would be able to harvest a lot of energy per agent than a more spread out group of agents. However, because the energy can be used up in an area, they must spread out and keep moving to get a constant supply of energy. If clusters of agents are too small they won't be able to harvest enough energy to survive and if they get too big there will be not enough to go round. It was hoped that these opposing pressures would cause stable clusters to be formed, similar to those in Di

Paolo's model. Di Paolo reports that the equilibrium between high player density causing the clusters to contract and low energy resources causing the clusters to expand "determines the size of a cluster and it's stability as a discrete ecological entity." (Di Paolo, 2000 [29] p144.)

With reference to the aims of this project, the crucial factor in designing this model was to provide conditions which allow self-organised collective behaviour to emerge. In terms of agents moving around a spatial world, any kind of spatial organisation of the population, such as grouping into clusters, can be seen as collective behaviour provided it is caused by the agents interacting. The more stable and discrete the clusters are despite a changing environment (i.e. the longer they last, and the fewer bits break off and join other clusters), the more integrated the collective behaviour. So it is important to try to identify conditions which effect cluster stability.

At an early stage of implementation it was found that relying on the agents to be directly adjacent to each other to interact caused problems. With a fine grid (initially 50x50 grid points, and 50 agents) it was relatively rare that two randomly moving agents were adjacent to one another, and even when the simulation was modified so that they actively seek out other agents, they only took the four immediately adjacent points into account. It was considered that a better way cooperation could work would be instead of only adjacent agents being able to cooperate to gain a higher yield, they could release some chemical enzyme which makes harvesting more efficient. Clusters of agents would tend to have lots of the enzyme, and so would be able to harvest more energy than isolated agents. This meant agents could find each other by following the gradient of the enzyme, and so they did not need to be able to 'see' each other directly. This kind of enzyme production is similar to the cooperative production of siderophores (iron scavenging agents) in pathogenic bacteria described by Griffin, West and Buckling [2].

3.2   Design of the model

3.2.1   The world

A grid world was used, wrapped into a torus to eliminate edge effects. Only one agent per point on the grid was allowed, as this made it easier to see what was going on and was also more biologically realistic since agents were forced out of densely populated areas as they competed for space. However during the implementation it was realised that using collision detection to stop them occupying the same points slowed down the simulation a lot (which was a problem when clustering coefficients were being calculated and the simulation took four or five hours to run). Also, it was considered that allowing them to occupy the same point may give some interesting behaviour. So collision detection can be turned on or off depending on the experiment, and for the majority of experiments it was switched off.

3.2.2   The energy landscape

The energy (food) landscape was modelled by a two dimensional array of positive real values. There were three types of energy distribution: flat, random or normalised.

With the flat distribution energy at every point started at a constant initial value. With the random distribution energy at every point started with a random initial value. With the normalised distribution, energy was built up by adding energy using a random-normalised distribution with the mean value set to a number of random points to give a landscape which already has areas of high and low energy. This is useful as it provides gradients for energy-seeking agents to climb to find areas of high energy, since the 'blobs' of energy have low amounts of energy at the edges and high amounts at the centre. It was hypothesised that the rich initial environment created by these gradients would help to 'set off' the process of self-organisation.

At each timestep energy was incremented so that it never completely ran out. In the case of the flat landscape energy at every point was incremented by a constant which resulted in the same about of energy being added overall per timestep as with the other two distributions. In the case of the random landscape energy was incremented by one at a specified number of random points. In the case of the 'blobby' normalised landscape energy was incremented by one at a specified number of normalised random points where the mean value was the centre of the current 'blob'. When a blob gets to a critical size, the blob centre is moved to another random point and so a new blob is started.

3.2.3 The chemical enzyme

The enzyme was also modelled by an array of positive reals, between 0 and 100. Initially there was no enzyme present.

At each timestep, the enzyme level at each point was decremented by a constant as though it gradually evaporates. This was to ensure areas currently rich in enzyme are also currently rich in agents; otherwise eventually every point in the world would have maximum enzyme.

The enzyme diffuses to allow nearby agents to 'share' the enzyme they produce. Although this would occur to some extent without diffusion since the agents would move around and leave enzyme behind (so a region with many agents would have a lot of enzyme), diffusion allows the enzyme to be spread evenly and also creates a smooth gradient for the agents to follow in order to reach areas with high levels of the enzyme. It is also more realistic since chemical diffusion occurs in the real world.

So at each timestep, at each point in the world, the enzyme level is decremented by a constant diffusion coefficient (0.3 units of enzyme) and each of the four adjacent points are incremented by a quarter of this amount.

3.2.4 The agents

The agents were modelled by a data structure which held their coordinates on the world array, and their current energy level. Initially their coordinates were random and their energy level set to a constant.

At each time step they release enzyme, harvest the energy at their current location (depending on the level of enzyme and amount of energy), and move - either to a random adjacent point, or to the adjacent point which will result in them harvesting the maximum amount of energy at the next timestep (taxis),

At each timestep:

For each agent:

If they are dead, remove them

More energy added to landscape

Move to the adjacent square which will result in the most energy being consumed in the next timestep.

Some enzyme evapourates

Eat a proportion of the available energy depending on the amount of enzyme present

Enzyme diffuses

Release enzyme

Loose energy (metabolic cost)

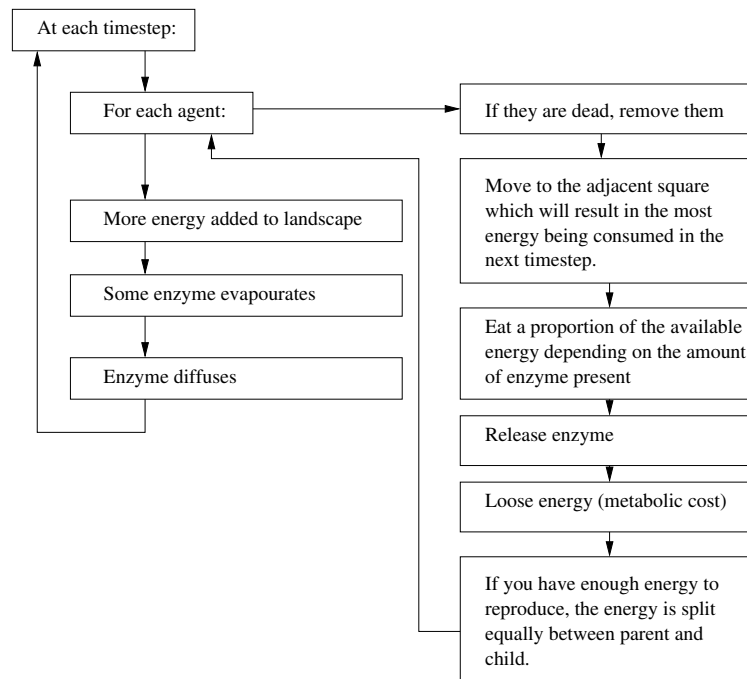If you have enough energy to reproduce, the energy is split equally between parent and child.

Figure 3.1: Flowchart of the model

or not move at all, depending on the experiment. There is a metabolic cost to the agents at each timestep, meaning they have to constantly eat to maintain themselves; they cannot just eat for a while and then stop moving. This strategy was hard-wired in the basic model, but a later enhancement (discussed in the next chapter) was to have differing strategies which are passed onto offspring. An advantage of evolving strategies is that different strategies can prevail in different regions of the clusters, as Di Paolo reported.

Agents die if their energy reaches zero, and have a maximum energy level, producing offspring if a critical reproduction threshold is reached. When a new agent is born, the energy of the parent is split equally between child and parent. The child appears at a random point adjacent to the parent.

3.2.5 Other features, schedule and methodology

To allow for experimentation with different conditions whilst observing, aspects of the simulation could be changed in real time with a key-press — taxis could be turned on or off (with no taxis resulting in entirely random movement), the energy distribution could be toggled between flat, random and normalised distributions, and the effect of the chemical enzyme could be turned on or off. There was also a reset option so that after changing various conditions the simulation could be re-started to eliminate the effects of previous conditions.

The simulation also writes log files so that data about the simulation could be analysed. They contain the number of alive agents, the number of births, the total amount of energy consumed per timestep and the average distance to each agent's nearest neighbour. This is the inverse of the level of clustering: with

Date	Task
27th Sep	Formulate basic idea, begin general background reading.
8th Oct	Submit project proposal.
13th Oct	Look at existing models to get ideas for a new model.
25th Oct	Formulate ideas and design simulation.
1st Nov	Begin implementing simulation.
17th Nov	Complete a very basic working simulation for feedback from supervisor.
22nd Nov	Begin writing draft background reading chapter.
29th Nov	Begin series of experiments to gather data about the model in its basic form.
10th Dec	Complete and submit mid-term report.
11th Dec - 23rd Dec	Enhance the simulation to try to produce more stable clustering behaviour.
27th Dec - 9th Jan	Revise for exam.
11th Jan - 24th Jan	Finish exam, enhance simulation to introduce evolution.
24th Jan - 11th Feb	Complete write up of simulation design.
11th Feb - 1st Mar	Conduct experiments with evolutionary model.
1st Mar - 14th Mar	Write up results with evolutionary model and experiment with differentiated populations.
14th Mar - 27 Apr	Write discussion and conclusion and finish off anything that needs doing.

Table 3.1: Project schedule

an evenly distributed (spread out) population of agents the average distance to their nearest neighbour will be large compared to agents tightly packed into clusters. Unfortunately calculating this takes a long time so it is best to run the simulation overnight without graphical output to obtain this information.

The methodology of the project was to build the simplest possible model, undertake a series of experiments with it, make enhancements to the model in response to the results of these experiments, and then repeat the process with the enhanced model.

A weblog of progress with the project was maintained at http://ameba23.blogspot.com. Complete program listings are available as a link from this weblog. The project schedule is outlined in table 3.1.

3.3 Results with the basic model

Before making any enhancements to the model, a systematic series of experiments were conducted with the model in its basic form. This was to assess the extent to which the model exhibited the desired behaviour, identify problems from which enhancements could be made in response, and because there would be little to learn from an enhanced model if no data had been collected from the basic model to compare it to.
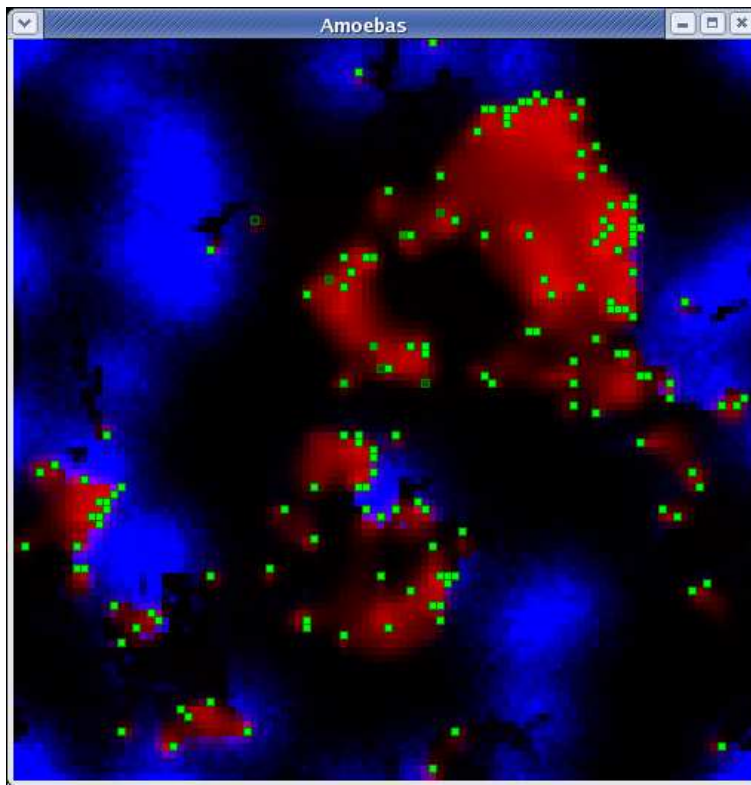
Figure 3.2: Screen-shot from experiment 1.1. Agents are represented by green squares, with brighter green for agents with more energy. The chemical enzyme is red and the food source is blue (colour intensities represent their levels).

3.3.1 Experiment 1.1: Blobby energy, Taxis on, Enzyme on, Collision on

When observing the simulation graphically for the first time, it was noticed that the agents appear to spread across the food source in wave-like patterns. They reproduce as they eat, causing some agents to get pushed out to the edges of the wave. Also, agents at the front of the wave oscillate back and forth as they move across the food. This is because they rely on enzyme they produce themselves to harvest the food, so they move forward to reach the small amount of enzyme that has diffused in front of them and then move back to the larger amount of enzyme they left behind. This is because the agents *have* to move at each timestep if it is possible to move, even if it's to a less desirable location than they currently occupy (this was changed in later versions of the model). Since the diffusion rate was initially set relatively low, the waves of agents leave a strong trail of enzyme over the food they've just eaten, and then move on to the next.

This spatial organisation into waves seems to be largely determined by the food distribution. Waves are far more visible with the normalised ('blobby') food distribution used in this experiment. In flat and random distributions the agents still oscillate back and forth but there are no discrete 'blobs' of food for them to spread across.
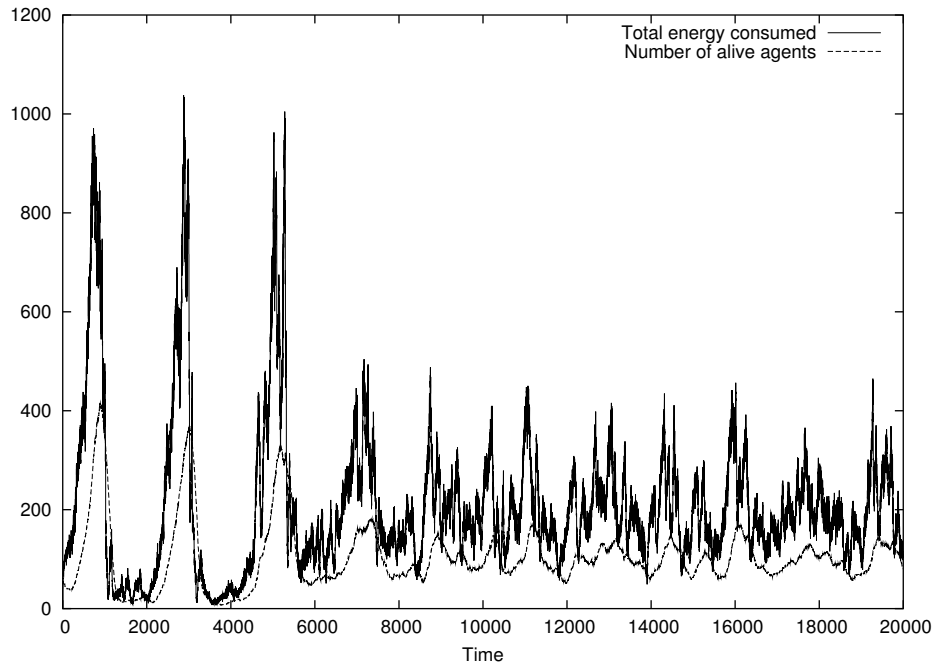
18

Figure 3.3: Experiment 1.1 - Normalised energy distribution

The energy consumed at each timestep fluctuates as agents reproduce and die. When a lot is consumed because of the initial food supply, the agents continually reproduce and consume more and more until a critical point is reached when there is no longer enough food to go round. Then many agents die and so consumption drops until it is replenished. Figure 3.3 shows that the effects of the initial large supply of food (large fluctuations in the energy consumed and population size) last for around 8000 timesteps.

3.3.2 Experiment 1.2: Random energy, Taxis on, Enzyme on, Collision on

When the simulation was run with random energy distribution, all the agents died at 1079 timesteps. This was because although there is a constant supply of food, there are no areas with very high food levels, and so the agents cannot move over the food fast enough to consume enough to survive. To allow data to be collected in this condition, the energy replenish rate was increased (from 200 to 1000 units) so that they would not all die out straight away. After the initial population explosion (which finishes around 2000 timesteps) the number of agents and food consumed did not fluctuate much compared to experiment 1.1 because there was a more constant supply of food rather than it appearing in discrete 'blobs', and there was less clustering. Because of the increased energy replenish rate they reproduced more than the agents in experiment 1.1, with on average five times more alive agents. Although the wave-like patterns were not as strong as with the normalised distribution, there was some organisation into clusters (see table 3.2).
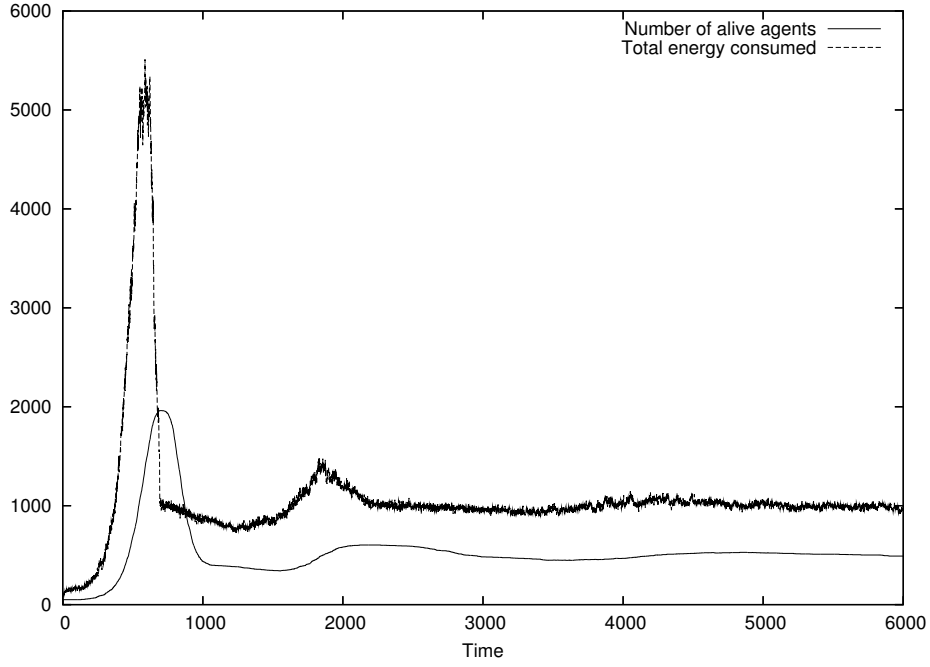
Figure 3.4: Experiment 1.2 - Random energy distribution

3.3.3   Experiment 1.3: Flat energy, Taxis on, Enzyme on, Collision on

With a flat energy distribution, the simulation behaved almost exactly as it did with the random one. Initially all agents died at 1025 timesteps, so the energy replenish rate was set to 1000. The average energy consumed, number of alive agents, and level of clustering were all almost the same as in experiment 1.2 (see table 3.2). It seems that the random energy distribution does not add any more 'richness' to the simulation when compared to the flat distribution.

3.3.4   Experiment 1.4: Normalised energy, Taxis off, Enzyme on, Collision on

Without taxis (food and enzyme gradient following behaviour), all agents died after only 112 timesteps so the energy replenish rate was increased from 200 to 1000 to allow the simulation to run long enough to analyse it. In this condition they died after 106 timesteps so it was increased further to 10,000. They still lasted only 128 timesteps. It was concluded that their difficulty in surviving was not due to not being able to find food, but not being able to find the enzyme they released which they need to consume the food (as their movement was random they would often move away from areas rich in enzyme). So the amount of enzyme they released per timestep was increased (from 20 to 50 units) and the enzyme diffusion rate (from 0.3 to 0.5, see section 3.2.3 for an explanation of enzyme diffusion), and the energy replenish rate at was left at 1000. So taxis appears to be crucial to survival of the agents.

   As with the other experiments, they had an initial population explosion around 1800 timesteps, and then the population dropped, and stayed relatively stable. They were distributed more randomly (less
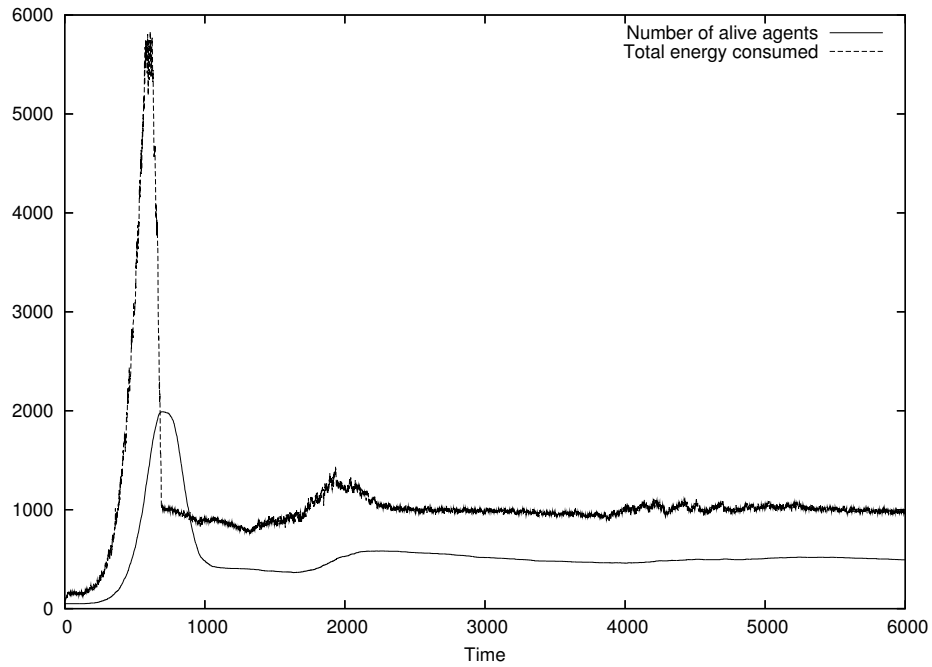
Figure 3.5: Experiment 1.3 - Flat energy distribution

clustering) than with taxis on, but were still to some extent clustered into areas rich in food and enzyme. Since they were moving randomly, this clustering must occur because agents in areas with little food tend to die and agents in areas with lots of food and enzyme tend to reproduce (and offspring appear adjacent to parents). This is further reinforced because they produce the enzyme themselves; areas rich in agents are consequently rich in enzyme. So while taxis considerably increases clustering, clustering will occur without it.

3.3.5 Experiment 1.5: Normalised energy, Taxis off, Enzyme on, Collision on, Reproduction off

This was to test whether the clustering found in experiment 1.4 is dependent on agents in rich areas being able to reproduce. Parameters were the same as with experiment 1.4 except reproduction was off. There was less clustering, but this could be because there were less agents due to no reproduction (clustering is taken to be the average distance to each agent's nearest neighbour, so a smaller number of agents will be less clustered). To get round this, the clustering measure could be taken to be the average distance to nearest neighbour divided by the number of alive agents. Since they cannot reproduce, there ends up being a small number of agents with much more energy than they can consume.
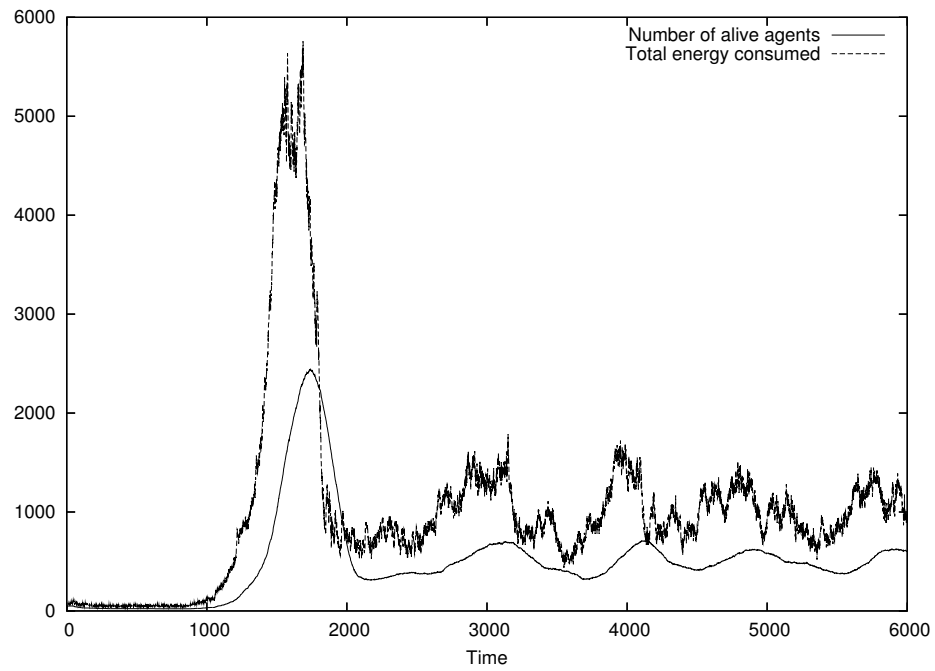
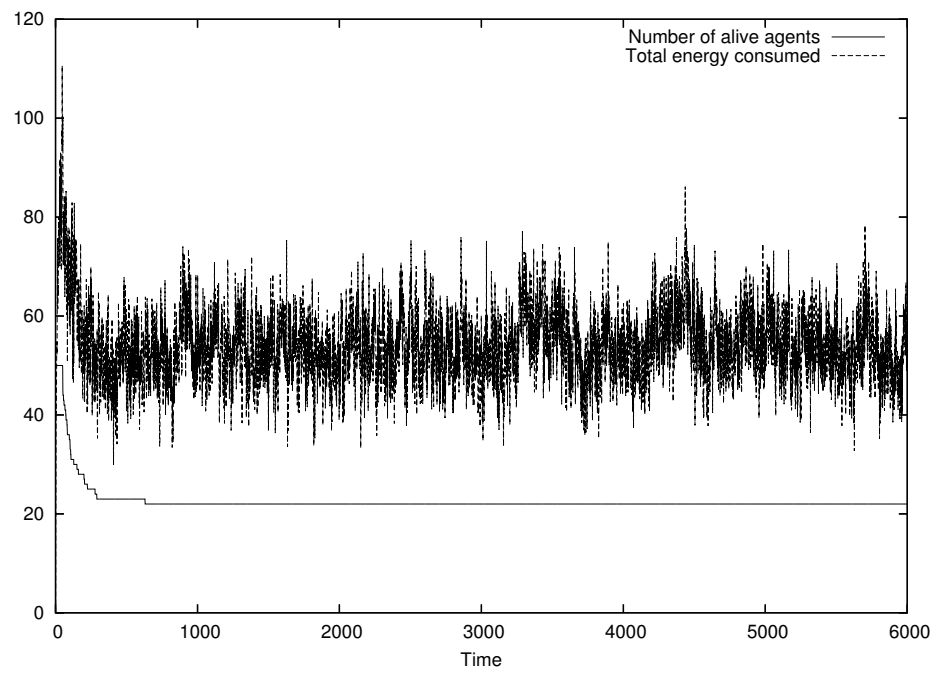Figure 3.6: Experiment 1.4 - Without taxis

Figure 3.7: Experiment 1.5 - Without taxis or reproduction

Exp	Mean alive agents	Mean energy consumed	Mean nearest neighbour	Births	Deaths
1.1	104.02	208.11	4.30	1650	1594
1.2	530.49	1076.82	2.51	2353	1864
1.3	530.43	1077.91	2.50	2345	1852
1.4	500.15	1003.19	2.69	7534	6936
1.5	22.61	54.01	12.52	50	28

Table 3.2: Summary of results with the basic model

3.4 Conclusions drawn from experiments with the basic model

The spatial organisation observed seemed to be partly determined by the energy landscape; a normalised ('blobby') energy landscape gave richer organisation into wave-like patterns. The different energy landscapes are explicitly designed into the simulation (in the case of the normalised landscape by creating a number of normalised 'blobs' at random locations). This is not desirable as this project aims to study self-organisation rather than organisation determined by a pre-given environment. It can be difficult to decide whether a certain behaviour is self-organising or entirely environmentally determined, since without some sort of environment no behaviour can occur at all.

However, a good indication that a behaviour is self-organising is its persistence despite environmental change. The clusters of agents and wave-like patterns produced as they spread across areas rich in food disappeared as soon as the food disappeared. One of the defining characteristics of self-organising systems (outlined in section 2.3) based on Prigogine's work [31] was that they create and maintain their structure in 'far-from-equilibrium' conditions (in a constant state of change). So if the spatial organisation of the agent population was to remain a relatively stable structure while the environment changed, I could conclude that their behaviour was not environmentally determined and so would be in a much better position to describe it as self-organising. Because of this, subsequent modifications to the model were made in an attempt to achieve more stable clustering.

3.5 Enhancements made in response to these experiments

3.5.1 Efficiency of food consumption

When looking at the data from the experiments, a problem was noticed with the way the agents use the enzyme. The enzyme increases how much of the available food they can consume, so if an agent is in a position where there is a lot of enzyme it will consume a large proportion of the food at that position. This means in areas rich in agents (and therefore rich in the enzyme they produce) the food is used up much faster than in areas with few agents. The wave-like patterns they produce as they spread across food areas occur because there is so much enzyme that they can consume all the food at their location in just a few timesteps and move on. There does not appear to be anything intrinsically wrong with this

method of enzyme use, but it does not allow clusters of agents to stay in one location for very long (as all the food gets used up), and so does not allow stable clusters to be produced. As was noted above, with respect to the aims of this project the emergence of stable clusters is highly desirable as they are a form of self-organised spatial structure.

A different way of using the enzyme would be to make it increase the *efficiency* of the food consumed rather than increase the amount of food consumed. So each timestep, the same amount of food is consumed regardless of the amount of enzyme present, but the energy gained from consuming that unit of food varies in proportion to the amount of enzyme. Agents in a cluster would gain more energy from the food they consume than agents on their own, but the food will not get used up any faster, making it much more advantageous to survival to be in a cluster, and also making the food a cluster feeds off last longer so the cluster will be able to remain stable for longer.

3.5.2   Sigmoidal food consumption function

Initially this new relationship between amount of enzyme present and efficiency of food consumption was linear, but it was considered that making a small increase in enzyme give a large increase in efficiency would make grouping into clusters even more advantageous and so more likely to occur. It was hypothesised that when an agent moved onto enzyme produced by another agent, the small increase in enzyme would make them considerably more likely to gain enough energy to reach their reproduction threshold and so a cluster would be started. So a sigmoid function between enzyme level and energy gained per unit of food consumed was implemented (see equation 3.1, where g is energy units gained by the agent per unit of food consumed, x is the level of enzyme at the agents' position, and c is the centre of the sigmoid curve. A value of $c = 5$ was used).

$$g = \frac{1}{1 + e^{c-x}} \qquad (3.1)$$

This new way of using the enzyme was implemented, and another experiment run so that this method could be compared with the original method (see figures 3.9 and 3.10). Higher levels of clustering were found (average distance to nearest neighbour was 1.83, compared with 2.50 for the basic model with a flat energy distribution (experiment 1.3)), and the population size and amount of energy consumed appeared to oscillate periodically (figure 3.9). This happened because the clusters grew while a lot of food was consumed, until almost the entire world was occupied by a single cluster, and then as food became scarce some agents died and the cluster repeatedly split into smaller clusters until the food was replenished, and the cycle repeated.

3.5.3   The saturation constant

In the basic model unconsumed food builds up and up indefinitely as more food is added, so moving to areas where no other agents have been for a long time gives a great survival advantage. Because of this,
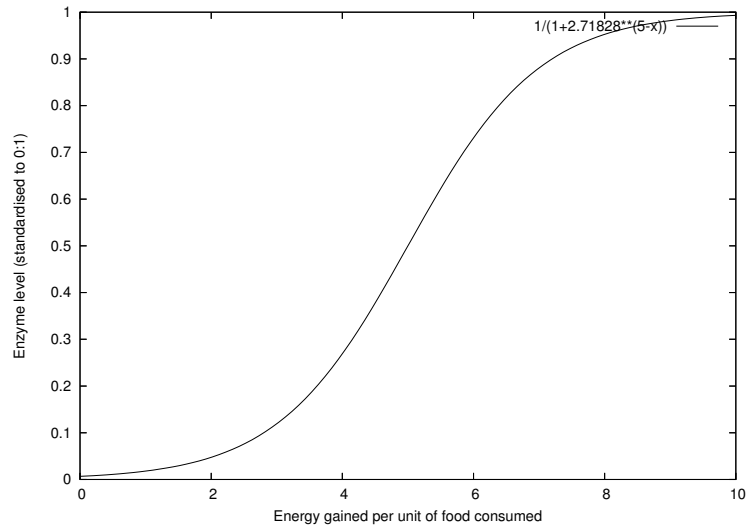
Figure 3.8: Sigmoidal relationship between energy gained and enzyme level
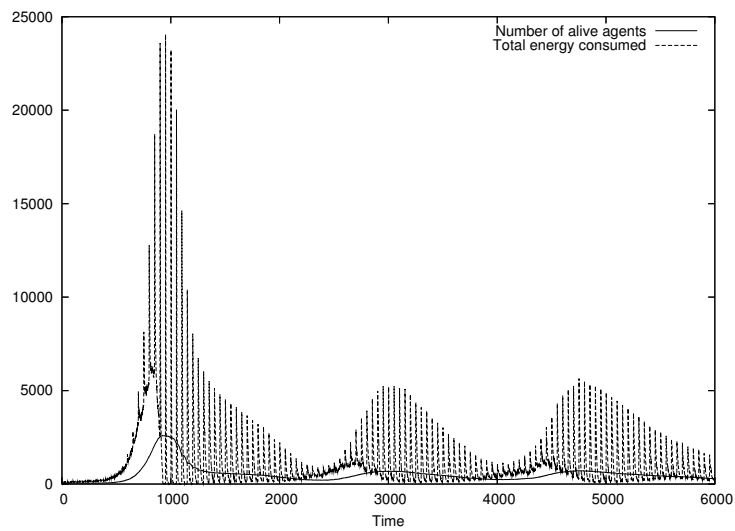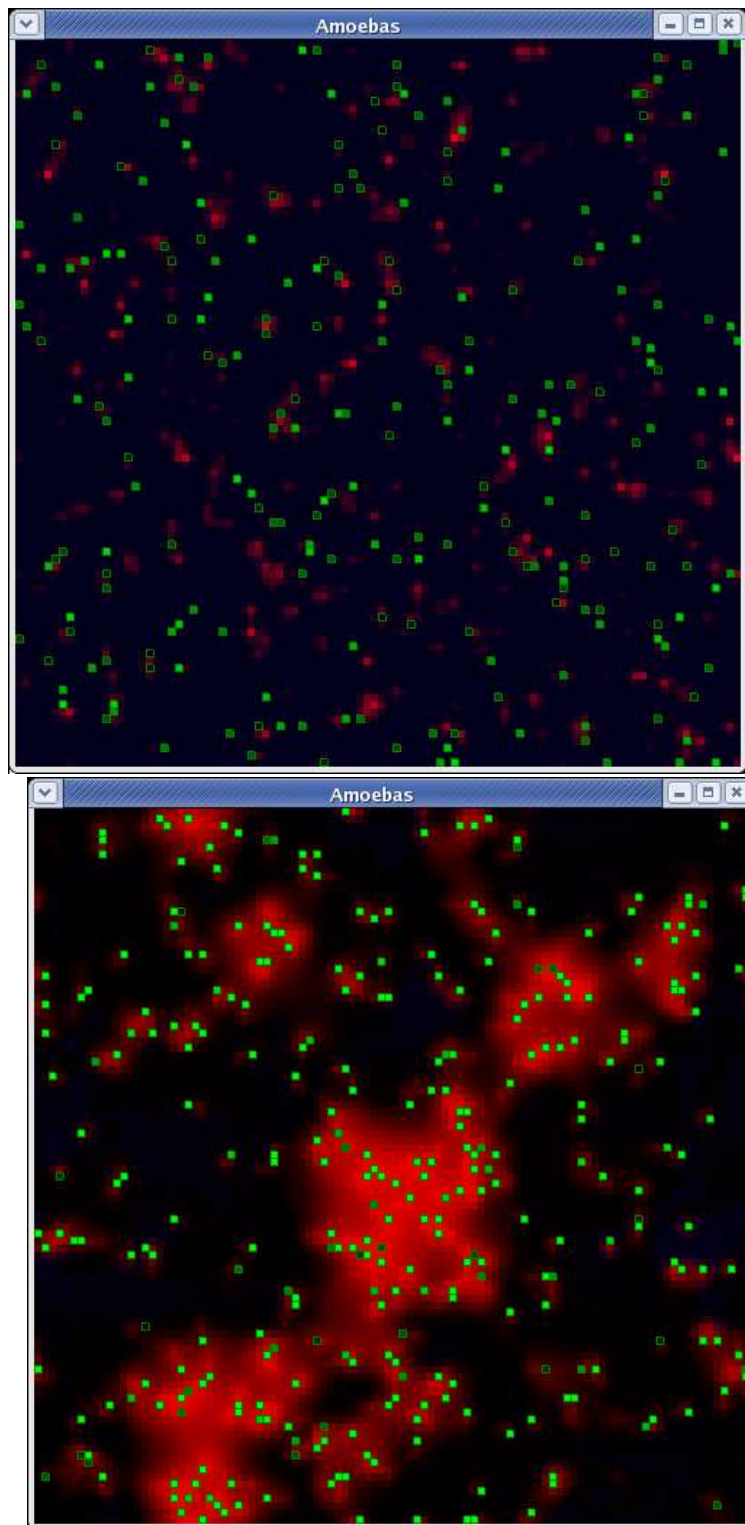

Figure 3.9: New method of enzyme use

25

Figure 3.10: Enhanced model (bottom) where enzyme is used to increase efficiency of food rather than food consumption, in comparison to basic model (top). Both shown with a flat energy distribution.

there was a tendency for individuals who brake away from a cluster to gain a lot of energy, reproduce and start a new cluster, so they eventually end up spread out.

To stop this, a method similar to that used in Di Paolo's model [29] was used. An upper limit on food level was imposed by 'evaporating' a constant proportion of the food at each grid point at each timestep, so unvisited areas of the world would only have a little more food than recently visited areas. Di Paolo refers to this as the saturation constant, as it is as if the environment can only retain a limited amount of food.

This makes the food landscape less variable; generally there is a low level of food all over rather than a few unvisited areas with high food levels and a lot of recently visited areas with no food. With less variation in food, taxis tends to make agents follow the enzyme gradient as there is less of a food gradient to follow (the taxis function is based on the amount of energy that will be gained by moving to a particular point, so is effected by both food and enzyme).

Making the enzyme level more likely to determine where an agent moves and whether an agent reproduces causes the Allee effect to take place (see [1]). At a low population density the per capita birth rate declines. In nature this is typically because of the increased difficulty in finding a mate but in the case of this simulation it is due to lower enzyme level. With higher population densities there is a higher birth rate. This positive relationship causes clustering to increase.

Chapter 4

The evolutionary model

4.1   Motives for introducing evolution

While the basic model produced some interesting results, it did not completely fulfil the aims of this project. In particular, I was sceptical as to whether the spatial organisation observed was genuinely self-organising in the way Prigogine [31] describes (order emerging as a result of many low-level interactions between components of the system - the agents in this case), as it appeared to be at least partly determined externally by the environment (the energy distribution). It would be desirable to see the emergence of more stable structures which persist despite changes in the energy distribution. Not only would this fit better with Prigogine's characteristics of self-organising systems, it would show that the agents are genuinely dependent on each other, since remaining clustered into a stable structure regardless of environmental changes would require them to constantly maintain proximity to each other, which will only occur under conditions where it is advantageous for their survival. So the production of more stable spatial structures would better fulfil the aims of this project by providing an example of self-organised collective behaviour which can be analysed to try to gain an understanding into the conditions which produce this behaviour.

In order to get this result, it was necessary to change the simulation in some way. Since there are a lot of parameters effecting the way the agents behave, it was decided that it would be appropriate to introduce a genetic algorithm so that the agent parameters evolve. While there is no guarantee that making the agents evolve better survival strategies will make them more likely to form stable, self-organised structures, it seemed likely that they would, given that mutual enzyme production between neighbours would greatly increase their energy consumption, and so there would be a selection pressure for them to cluster together. This selection pressure was present in the basic, non-genetic model but the
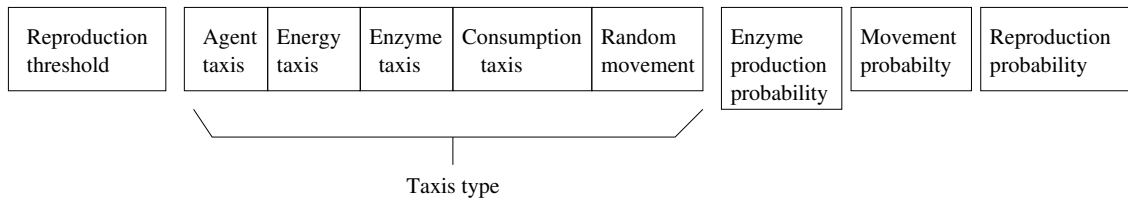
Figure 4.1: The genome

agent population had no way of modifying behaviour to react to this pressure.

Also, genetic variation may cause the population to differentiate into groups of agents with distinct strategies (like different species), and it is possible that different types may become more prevalent in different regions of the world. Di Paolo [29] reports that in his model different agent strategies formed different parts of the spatial structure, with one strategy being more prevalent around the edges of clusters and another being prevalent in the centre. So it is possible that introducing evolution may not only allow parameter values to be found which result in the formation of more stable structures, genetic variation itself could form part of the spatial organisation.

4.2 Method

Evolution can be implemented as generational or steady state. With generational evolution, a population is simulated for a specified amount of time, their fitness calculated, and then offspring are produced for simulation in another distinct 'run', in the absence of their parents. This keeps the population at a stable size which has advantages: there is always a large enough population to analyse in detail, and analysis is not biased by variance in the population size (it is difficult to compare populations of very different sizes). Also, the population cannot get so large as to be computationally very demanding. With steady state evolution, population members can be born or die at any time in the simulation, so evolution occurs continuously rather than in discrete stages, which is more biologically realistic [28]. Since the model already allows agents to reproduce or die at any time, steady state evolution was used.

4.2.1 The genome

The choice of parameters to be included in the genome has a great effect on evolution [28]. The factors which effect the agents' behaviour directly are of most importance. The genome parameters used are outlined as follows (also see figure 4.1).

- Reproduction threshold - The amount of energy that must be gained for an agent to reproduce.

- Taxis - Different types of taxis were represented as probabilities that the type of taxis will be used at each timestep. The probabilities that the different types of taxis will be used all sum to one. So agents can use a mixture of taxis types, different at each timestep. There are five types of taxis, as follows:

- Agent taxis - movement towards other agents, based on the number of agents directly adjacent to each of the four possible spaces to move into.

- Energy taxis - movement to the space with most energy.

- Enzyme taxis - movement to the space with most enzyme.

- Consumption taxis - movement to the space which will result in the consumption of the most energy in next the timestep (as was used in the basic model).

- Random movement (no taxis) - movement to a random adjacent space.

- Probability of enzyme production at each timestep. This is particularly likely be effected by a metabolic cost for enzyme production.

- Probability of movement at each timestep. Similarly this is likely to be effected by a cost for movement.

- Probability of reproduction at each timestep, given that the reproduction threshold has been reached. It was expected that this probability would always evolve to one (certain), since it is generally advantageous to reproduce, given that enough energy is gained to sustain the parent and child. Reproduction can be controlled by mutations in the reproduction threshold but it was decided that allowing a probability for reproduction to evolve as well would allow the level of reproduction to evolve independently of the amount of energy gained.

4.2.2 Survival costs

How the agents evolve is greatly effected by the costs to survival imposed on them. In the basic model the only metabolic cost was made at every timestep regardless of the agent's activity (each timestep each agent's energy was decremented by a constant). This was not very biologically realistic and does not provide a very complex selection pressure. In this evolving version of the model, metabolic costs were introduced for movement, enzyme production, and reproduction. The reproduction cost is made to the parent before reproduction.

Since the relative size of these costs will drastically effect evolution, a series of experiments were run to see how the agents would evolve with different survival costs, and to see which arrangement of cost values gives rise to the most spatial organisation and collective behaviour. There was a difficulty in selecting survival costs for the experiments because if the costs are set too low the agents reproduce so much that they occupy every point in the grid world and so spatial organisation is impossible. Setting the costs too high resulted in all the agents dying.

4.2.3 Mutations

Each time there is a birth, there is a probability that each parameter in the genome will be mutated. Mutations were carried out by adding a normalised random number with a mean of zero to the old

value for the parameter. A difficulty in making mutations in genetic algorithms is how to interpret mutated values which fall outside of the limits of acceptable values for the parameter. For example if the parameter is a probability (which most of the parameters in this simulation were), how should mutated values which are below zero or above one be interpreted? While this might appear to be a trivial technical detail, if mutation operators are the only source of novelty in the simulation (there was no sexual recombination of genotypes in this model) then any bias in the way the mutation is carried out will have a dramatic effect on evolution.

Bullock [6] assessed different methods of interpreting these values to see which gives the least evolutionary bias. Illegal values can be: Absorbed (truncated to nearest legal value), repeated (keep generating new mutations until a legal value is produced), reflected (the amount of error is negated and then added to the nearest legal boundary), wrapped (the amount of error is added to the opposite legal boundary), or ignored (illegal values default to the unmutated parental value). It was reported that the method resulting in the least bias was to ignore out of limits mutations, so this is the method that was used in this project.

The probability that each different type of taxis will be used must sum to one to prevent bias, so after mutating, each taxis probability value was divided by the sum of all taxis values so that they all sum to one.

4.3 Results with the evolutionary model

4.3.1 Experiment 2.1 - Normal costs

This arrangement of costs is the same as for the basic model: a standard cost is made at each timestep regardless of the agent's activity and no other costs are made. This gives a control condition for comparison with the other experiments.

From the initial random values, the probability of movement increased, as did the probability of enzyme production (with a little fluctuation) which is what was expected since these are behaviours which aid survival (see figure 6.1 in appendix B). The most popular type of taxis by far was energy taxis (moving towards food) which was not expected as consumption taxis is a more accurate indicator of the amount of energy that will be gained at the next timestep as it takes both food and enzyme levels into account. A likely explanation for this is that as the world becomes very populated there is a lot of enzyme and not much food, so enzyme level is not a major factor in determining the best position to move to; it is those agents which move towards food which have the biggest survival advantage.

4.3.2 Experiment 2.2 - Cost for movement

Even with a large cost for movement (20 energy units, compared with the basic cost in experiment 2.1 of 2 units) the agents were still likely to move after 6000 timesteps, so the survival advantage of moving must be far greater than this cost.
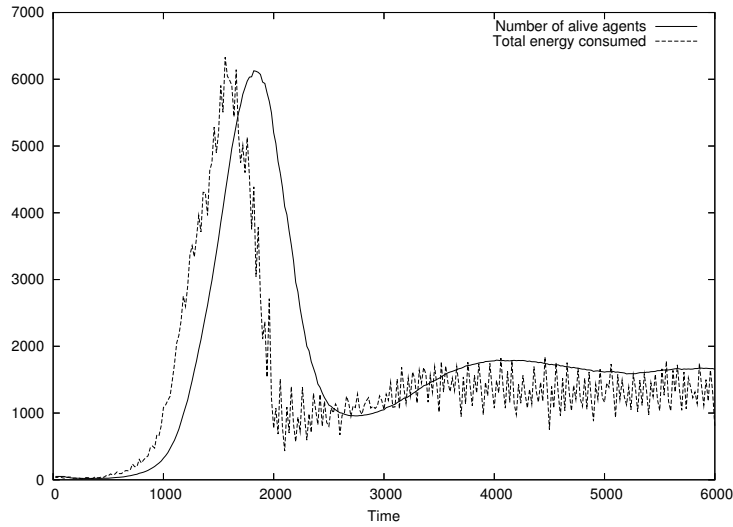
Figure 4.2: Experiment 2.1 - Normal costs

It was hypothesised that introducing a movement cost would increase cluster stability as agents would be less likely to move and so offspring would remain nearby their parents in a cluster. However because the agents still evolve a high movement probability, there was only a small drop in average distance to nearest neighbour, so only a small increase in clustering (see table 4.1).

The genome appears to fluctuate more than in the control experiment (see figure 6.2, appendix B) because the population was very low for the first 3000 timesteps, so the genetic mutations of each individual greatly effects the average genome shown in the graph.

Although energy taxis was again the most popular type of taxis at the end of the simulation run, it was only slightly more popular than consumption taxis. This is because with less movement the enzyme was not spread about as much by agents constantly moving around and leaving enzyme behind, which creates greater competition for enzyme, making it more important to survival for agents to follow the enzyme gradient.

4.3.3   Experiment 2.3 - Cost for enzyme production

This was probably the most important experiment in terms of assessing the extent to which collective behaviour is taking place as introducing a cost for enzyme production and representing the probability of enzyme production in the genome allowed the extent to which the agents were cooperating to be assessed. Since agents in clusters rely on enzyme produced by others to increase their efficiency of food consumption, if they are very likely produce it despite it being costly they can be said to be cooperating to some extent. A non-cooperative agent (one who is unlikely to produce enzyme) is able to exploit a cluster of cooperators by gaining an advantage from enzyme produced by others whilst not contributing to the cluster themselves. It was found that very little enzyme was produced overall in comparison to other experiments, but that pockets of cooperators appeared from time to time (which can be seen in
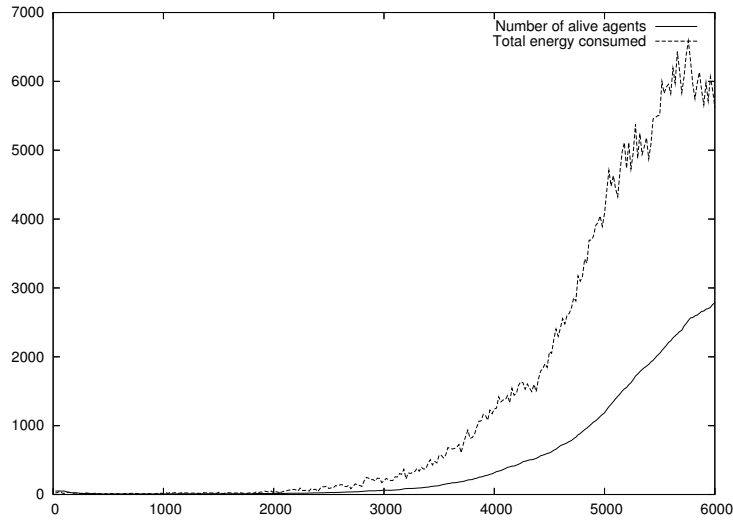
Figure 4.3: Experiment 2.2 - Cost for movement

graphics mode as a blob of enzyme appears around them) which die out when they come into contact with non-cooperators as though they are invaded. In this sense, the presence of the enzyme acts as a marker for cooperation.

4.3.4 Experiment 2.4 - Cost for reproduction

With an increased cost of reproduction many of the initial agents die and the few which happen to have started with a high reproduction threshold reproduce, but more slowly as they need to gain more energy before they can reproduce. Once these agents have become prevalent, and so the average reproduction threshold increased to account for the cost, the simulation appears much the same as experiment 2.1.

4.3.5 Experiment 2.5 - Costs for movement, enzyme production and reproduction

The final experiment was with a mixture of costs to see what cumulative effect they had. As figure 4.7 shows, the population drops very low and then begins to rise, and is still rising steadily after 6000 timesteps. After this time there is a population explosion, and a single cluster grows (see figure 4.6). As the single cluster continues to get bigger, the food in the centre gets used up, resulting in agents arranged in a doughnut shape. As the food builds up again, some agents move back inwards. Eventually, once the cluster gets so big as to have used up all of the initial supply of food the agents become much more evenly distributed across the world and clustering decreases as they have to keep moving to find food.
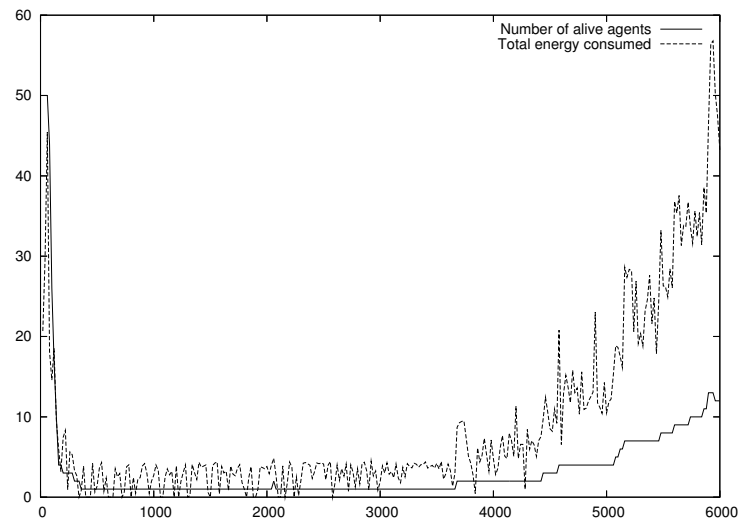
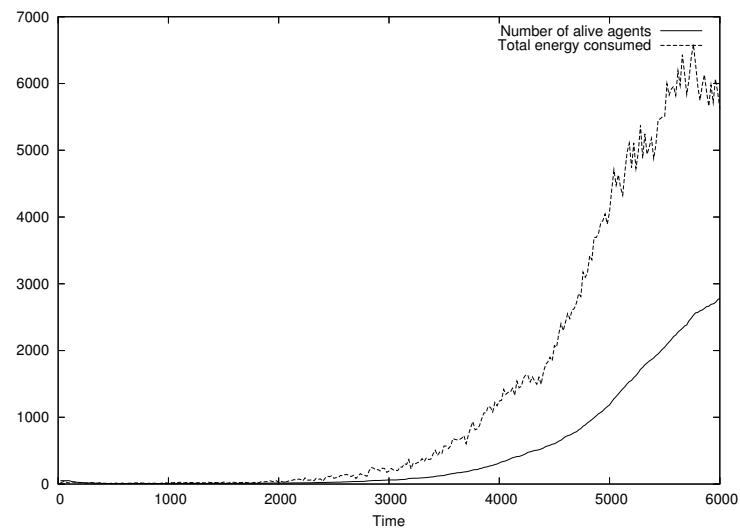Figure 4.4: Experiment 2.3 - Cost for enzyme production

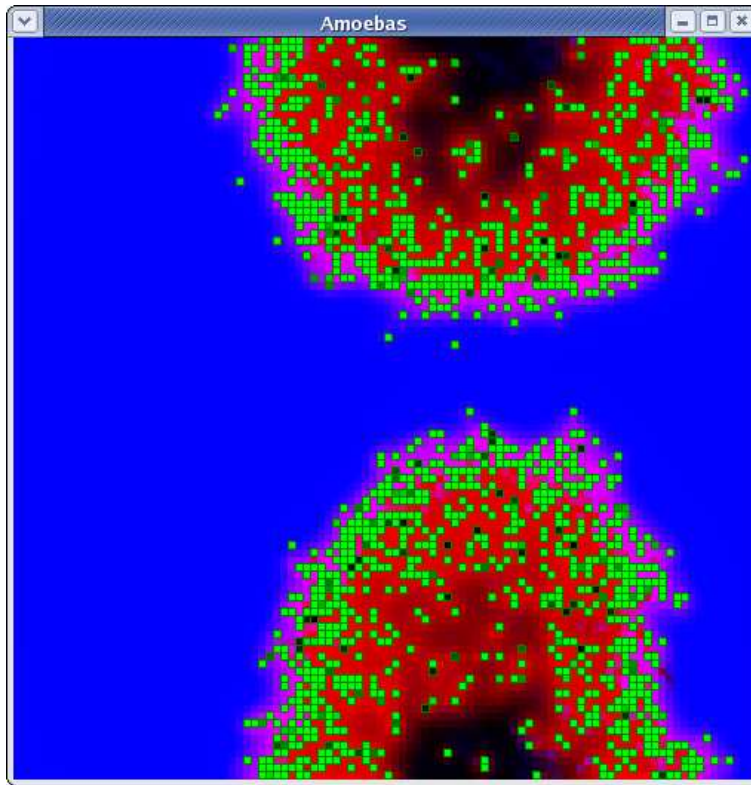Figure 4.5: Experiment 2.4 - Cost for reproduction

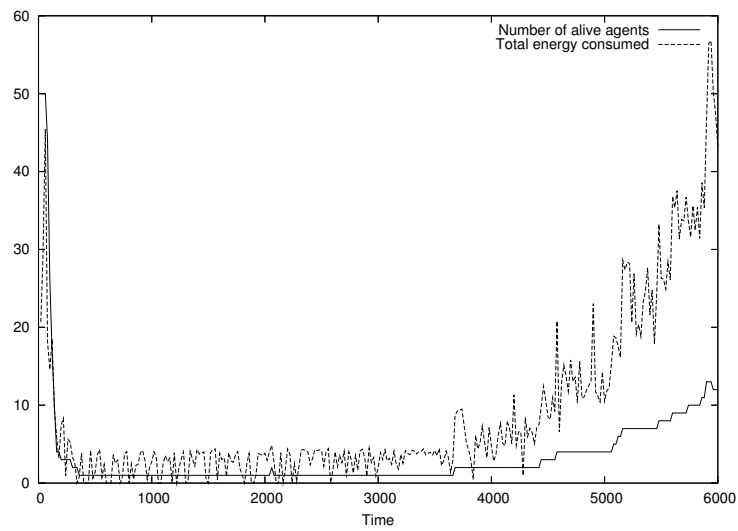Figure 4.6: Experiment 2.5 - Screen-shot. Note that the world is wrapped

Figure 4.7: Experiment 2.5 - Costs for movement enzyme production and reproduction

Parameter	Exp 2.1	Exp 2.2	Exp 2.3	Exp 2.4	Exp 2.5
Mean alive agents	842.08	99.42	16.26	103.71	16.35
No. births	10121	4843	65	4832	65
No. deaths	8465	2959	53	2059	53
Mean Energy consumed	1470.17	776.33	509.14	632.54	509.14
Mean nearest neighbour	4.24	3.43	13.09	1.29	16.18
Reprod threshold	500.45	500.21	499.47	1000.18	999.47
Agent taxis	0.13	0.17	0.30	0.17	0.30
Energy taxis	0.53	0.27	0.18	0.27	0.18
Enzyme taxis	0.03	0.11	0.16	0.11	0.16
Consume taxis	0.18	0.25	0.09	0.25	0.09
Random Taxis	0.13	0.21	0.27	0.21	0.27
Enzyme Production	0.74	0.83	0.89	0.83	0.89
Movement prob	0.92	0.72	0.90	0.72	0.89
Reproduction Prob	0.46	0.37	0.36	0.37	0.36

Table 4.1: Summary of results with the evolutionary model, with the average genome after 6000 timesteps.

4.4   Conclusions from the evolutionary model

In general, the most popular type of taxis was energy taxis (moving towards food), which was not expected. It seemed to make more sense for consumption taxis to be the most widely used since this is the most direct indicator of how much energy will be gained by moving to a particular point. However, since the population was clustered, and the agents produce enzyme, agents living within a cluster are likely to be constantly surrounded by enzyme, making them indifferent to it. So the strategy they evolved was particularly suited to living with other agents.

Chapter 5

Differentiated populations

5.1 Motives for introducing differentiated populations

In experiment 2.3 it was difficult to assess the extent to which cooperation was taking place, because there was no discrete divide between cooperators and non-cooperators; some agents were just more likely to produce enzyme than others. The pockets of enzyme observed show that there was some differentiation in the population going on but it was very difficult to quantify. In order to look at the behaviour of differentiated populations, the simulation was changed so that the agents' enzyme production was represented in the genome by a binary on/off value instead of a probability. So the population was divided into enzyme producers and non-enzyme producers. Since there was a metabolic cost for enzyme production, the non-enzyme producers can be seen as 'cheaters'; relying on the enzyme produced by others to survive. So that this could be observed graphically, enzyme producers were coloured green and cheaters yellow.

Having a discrete set of cooperators and non-cooperators makes the simulation somewhat similar to Pfeiffel and Bonhoeffer's [30] simulation where cooperators cluster together to shield their resource from non-cooperators.

5.2 Results with differentiated populations

The simulation was run for 40,000 steps (much longer than previous experiments of 6000 timesteps) because it often took a long time for a large cluster to appear.

Initially, the agents were not clustered at all and moved around using only their own enzyme to consume food. As the population started with a mixture of enzyme producers and non-enzyme produc-
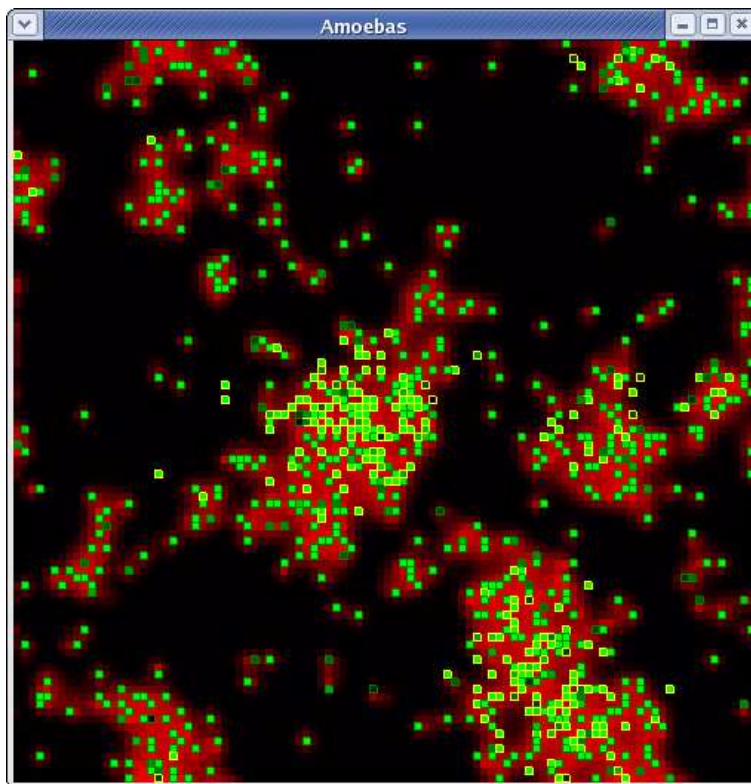
37

Figure 5.1: Populations of enzyme producers and non-enzyme producers. Note that non-enzyme producers (yellow) are only present inside clusters.
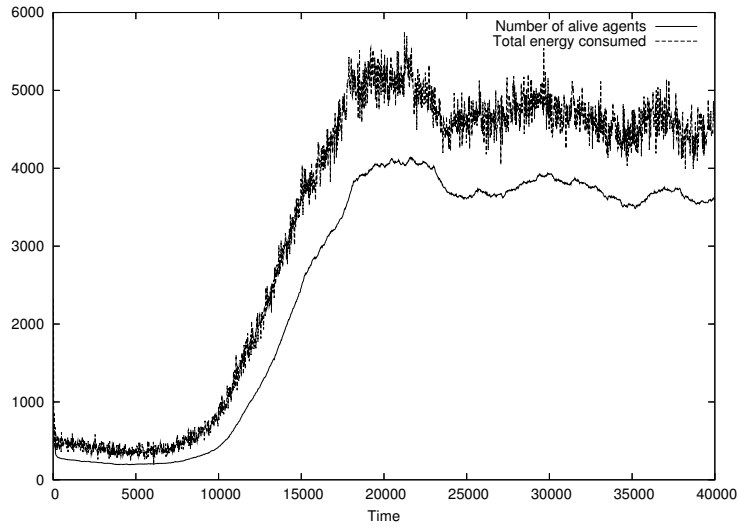
Figure 5.2: Differentiated populations

ers (the genome of each agent was initialised randomly), there was a sudden drop in the population as the non-enzyme producers all died out (see figure 5.3, the percentage of enzyme producers is around 100 at 1000-8000 timesteps). As clustering gradually increased (average distance to nearest neighbour decreased), non-enzyme producers were more likely to survive as there was a lot more enzyme around, and since they did not have the metabolic cost of producing enzyme they did better than the enzyme producers and reproduced more. Eventually the population of non-enzyme producers stabilised at around 75%, as this was the biggest proportion that can survive solely on the enzyme produced by others.

As the proportion of cheaters rose above 50%, consumption taxis (which is determined by both food and enzyme levels) became significantly more popular than the other types of taxis (see figure 6.6, appendix B), and the popularity of energy taxis dropped (which was the most popular type in the control evolutionary experiment (experiment 2.1)). This was because it is critical for survival that non-enzyme producers seek out enzyme, whereas enzyme producers are likely to be in high enzyme areas whatever their taxis behaviour, as they produce it themselves, so moving towards food is far more important. Also movement probability increased significantly around this point (10,000 to 15,000 timesteps), as non-enzyme producers have to keep moving to actively seek out enzyme.

5.3 Conclusion

Non-enzyme producers (cheaters) are effectively parasites on clusters of enzyme producers. Since they only survive because of the enzyme produced by others, they are dependent on other agents; meaning their behaviour can be described as social (although it is exploitative rather than cooperative). Being a cheater is only an evolutionarily stable strategy when there is a large population of non-cheaters around. Figure 5.1 shows that non-enzyme producers only survive within clusters.
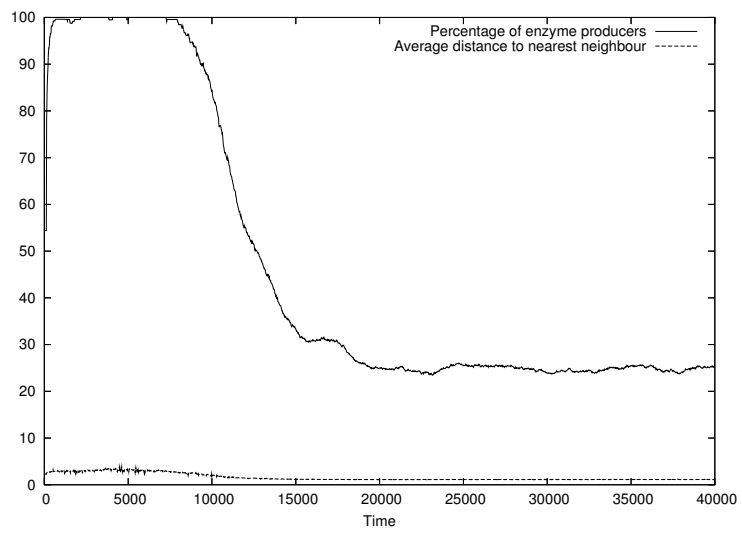
Figure 5.3: Percentage of enzyme producers over time with differentiated populations

Chapter 6

Discussion

6.1 Criteria for defining self-organised collective behaviour

To decide whether the behaviour reported can be described as self-organised or collective, it is necessary to return to the criteria used to define these concepts.

A problem with looking for self-organised collective behaviour in this kind of model is that it is easy to jump to conclusions when some kind of spatial pattern is produced, but it is not always clear whether it constitutes self-organisation. As this is a graphical model, there is a great temptation to draw conclusions from what is observed without backing it up with an objective measure. The main numerical measure that was used to indicate this sort of organisation in this project was the clustering coefficient based on the average distance to nearest neighbour. This rested on the assumptions that benefiting from mutual enzyme production constituted collective behaviour and that greater proximity to one another indicated greater use of each other's enzyme. So even when objective measurements are used, conclusions drawn from them will only be as valid as the underlying assumptions. The problem is that self-organisation and collective behaviour are vague concepts which lack a standardised way of measuring them.

Prigogine's defining characteristics of self-organisation, discussed in section 2.3, provide a possible way of evaluating the extent to which the reported behaviour if self-organising.

Firstly, the system must actively create and maintain it's structure in 'far from equilibrium' conditions. The structure, in the case of this project, is the clusters of agents and the surrounding enzyme they use. The fact that these clusters are not present initially, and exist for some time (as shown by changes in the clustering coefficient) makes it clear that they are somehow created and maintained, but it is not immediately clear what the source of this organisation is. As discussed in section 3.4, organisation in the

41

basic model appeared to be largely determined by the layout of the environment rather than the system of interacting agents. In reaction to this it was proposed that the structure must persist despite environmental changes in order to be considered self-organising. But it is not clear how radically changing the environment must be to be able to rule out environmental factors as being the source of order; i.e. what constitutes 'far from equilibrium' conditions. The only environmental factors outside the control of the agents' evolutionary strategy were the initial energy distribution, the way energy was replenished (and 'evaporated', with the introduction of the saturation constant in the enhanced model), and the diffusion and evaporation of the enzyme. So a possible criteria could be that if clusters are able to persist despite radical fluctuations in the energy distribution, they can be considered self-organising. The problem with this criteria is that fluctuations in food are caused by agents consuming it and it being replaced, so it is impossible to separate environment conditions from conditions created by the agents themselves.

However, another way of showing that the organisation was due to interaction between agents rather than environmental factors is to make the environment simple and unchanging, and assume that because the environment is static any rich organisation which occurs must be determined by other factors. This was the purpose of using a flat energy distribution for all experiments in chapters 4 and 5. Since there was a high degree of spatial organisation in these experiments (reflected by high levels of clustering), it was concluded that self-organisation must be taking place.

Another characteristic is that self-organisation, by its very nature, is supposed to be order which emerges spontaneously rather as a result of some explicit design process. This is a potentially serious problem with the methodology of this project; systematically designing experiments in order to find self-organisation may be a fundamentally flawed method. For instance, changes were made to the simulation in reaction to the unsatisfactory results with the basic model (changes to the way enzyme was used and food was replenished) and as a result the simulation became 'more self-organised'. Making changes in this way is a design process, which again prompts questions as to whether the organisation is designed and built-in or emergent.

However, the purpose of these design changes was to find the conditions which give rise to self-organisation rather than to actually create it. This is a subtle difference and open to debate; it is difficult to say whether the results observed were discovered or created.

The criteria for defining collective behaviour was that which was produced not just by an aggregation of many independently acting agents in some kind of formation, but as the result of behaviour which evolved only because of the interaction with other agents; i.e. it is truly social behaviour, whether cooperative or exploitative, and so would not have evolved in an agent in isolation.

With distinct populations of enzyme producers and non-enzyme producers, there was a shift from independent to collective behaviour, as non-enzyme producers were unable to survive at the beginning of the simulation run, but as the enzyme producing population became organised into clusters, conditions were created which allowed non-enzyme producers to survive. So non-enzyme producers were dependent on enzyme producers for survival, meaning their strategy was collective.

A more objective way of testing if an evolved trait is 'social' would be to see if it is present in agents

evolving in isolation from each other. This could be done by using generational evolution, where a series of trials for each agent alone are run, which could be compared to steady state evolution where they all evolve together. Traits which become prevalent in isolated agents cannot be in any way social since there is no interaction going on. This criteria would take into account more indirect social behaviours which can otherwise be difficult to identify. For example stigmergic communication, where agents effect one another's behaviour by one agent modifying the environment in some way, which will later cause another agent to act differently. In the case of this project an example of this could be that areas high in enzyme are likely to be recently visited by agents, and so less likely to be high in food. Since the enzyme diffuses but food does not, the enzyme gradient might often lead to areas low in food, which could be why consumption taxis (based on both food and enzyme levels) was not always the most popular type of taxis. Unfortunately, due to time restrictions generational evolution was not implemented.

6.2   Conditions which cause the transition

Having assessed the criteria for self-organised collective behaviour, it seems that in some of the experiments performed, the population of agents in the simulation can be described as making a transition from independent to collective behaviour through the process of self-organisation; so the model succeeded in simulating what it aimed to simulate. Since this was not the case with every experiment, and different experiments showed varying degrees of this behaviour, by looking at the results it must be possible to identify the conditions which cause this transition.

The presence of the chemical enzyme appeared fundamental to the production of co-operative behaviour, as without enzyme the only interaction between agents was the stigmergic effect of their competition for food (areas recently visited by other agents would be low in food, so by following the food gradient alone, agents tended to avoid each other), or in cases where collision detection was switched on, there was direct interaction as they compete for space. When the simulation was modified so that agents could consume food regardless of enzyme level, and did not produce enzyme, (i.e. enzyme was not modelled at all) the clustering level was significantly lower (agents were more spread out).

Since this project aimed to draw general conclusions which could aid understanding into any self-organising system, not just the simulation discussed, it is necessary to identify exactly what it is about the way the enzyme allows the agents to interact that causes this transition to collective behaviour. Specifically, if they are able to survive independently to begin with, how could a cooperative relationship whereby they become dependent on each other be initiated?

A possible explanation is that the enzyme allows cooperation to be introduced gradually and passively, without any sudden change in the agent's strategy. This is because the enzyme is produced by independent agents since they need it to survive, and all that is required for two agents to benefit from mutual enzyme production is their close proximity to one another. Proximity could occur by chance in the case of random movement, by both agents following a common food gradient in the case of energy taxis, or by them actually following each other's enzyme gradient in the case of enzyme or consumption

taxis. This mutual benefit means both agents will be more likely to survive and reproduce, and since offspring appear directly adjacent to parents, they too will benefit from the cluster's enzyme and be more likely to survive themselves. In the case of the evolutionary model, the constant exposure to the surrounding cluster of agents and enzyme will effect the genome over many generations, making social behaviours more likely to become stable strategies.

So, in more general terms, one of the conditions for the transition to collective behaviour could be that benefiting from an interaction between components does not require any discrete, active change in behaviour.

Taxis was also a factor which helped to produce collective behaviour, as shown by the lower clustering reported in experiment 1.4 (basic model, without taxis) compared to experiment 1.1 (with taxis). In particular, the evolution of different types of taxis produced a strategy which was particularly suited to living within a cluster. This was the greater use of energy taxis over enzyme taxis (in experiment 2.1), because due to the high enzyme levels inside clusters, being able to find energy is a greater determinant of survival.

So in a scenario where social behaviour (mutual enzyme production in the case of this project) gives a survival advantage, evolution causes the production of behavioural strategies suited to social life (collective behaviour). By allowing novel behaviour to be produced as a result of many small interactions and chance mutations, evolution can also play a key role in self-organisation. When the population diverged into enzyme producers and non-enzyme producers, spatial organisation was produced as a result of both the nature of the enzyme encouraging cooperation, and evolution making different social behaviours prevalent in different regions. This occurred because the life-span of the agents was short relative to the speed that they move, so an agent's position within a cluster effected its chances of survival, meaning evolution was happening very fast in relation to the length of time a cluster remained stable for. Marshall and Rowe [23] looked at the effect of this population viscosity (relatively low movement speed of agents) in relation to reciprocal cooperation, but here it appears to effect spatial organisation of agents with different strategies.

The main explanation for the evolution of undifferentiated multi-cellularity proposed by Pfeiffer and Bonhoeffer [30] from the results of their simulation was that clustering is a shielding technique to protect organisms from invaders. So it would be expected that a condition for the transition was the presence of uncooperative invaders. However, although uncooperative agents are present in my model, it is unlikely that they cause the transition to collective behaviour for two reasons. Firstly, since uncooperative individuals are created by genetic mutation they can be born into a cluster of cooperators, rather than trying to invade from outside. This means it is impossible for cooperators to shield their food resource by clustering, as non-cooperators have access to their local resource regardless of their spatial formation. Also, clustering occurs in the absence of non-cooperators so they cannot be a necessary condition for the production of clustering.

In conclusion, the nature of the production and effect of the chemical enzyme created a scenario where cooperative relationships could emerge gradually, without discrete behavioural changes, which

accounts for the production of collective behaviour in the basic model. The high population viscosity relative to agents' life-span meant that evolution allowed different social strategies to prevail in different regions, creating spatial organisation within the aggregate of agents.

6.3   Problems and possible extensions for future work

As discussed in section 3.4, organisation in the basic version of the model appeared to be largely determined by environmental factors. On reflection it seems that starting with a rich environment was a mistake; in order to really understand what is going on in the simulation it is necessary to start with the simplest conditions possible. The normalised energy distribution was implemented because it was hypothesised that a rich environment might 'set off' self-organisation, but it actually almost completely determined the organisation to the extent that it was impossible to distinguish what was a result of the nature of the way the agents behaved and what was determined by the pre-existing environment.

Often graphical output alone was used to draw conclusions despite the availability of statistical tests which provide a more objective way of analysing the simulation. For example, figure 5.1 was used to show that non-cooperators only survive in clusters, but a statistical measure could have shown this, such as taking the average distance to nearest neighbour for non-cooperators alone and comparing it to that of cooperators. Similarly, sometimes conclusions were drawn from patterns in graphs produced from the log files. For example, it was noted that there were often oscillations in the amount of energy consumed over time, particularly when the enhanced method of enzyme use was implemented (figure 3.9). A Fourier transform could have been used to test for the degree of periodicity in this data, but wasn't due to time restrictions. So an enhancement would be to spend more time producing statistical data to support conclusions.

Another problem with the results is that often the data produced by the simulation still hasn't stabilised after a complete run, particularly with the evolutionary model where it takes a relatively long time for an evolutionary stable strategy to emerge. For example in experiment 2.2 (see figure 4.3) the population is still growing steadily after the run of 6000 timesteps. Clearly these experiments should have been run for longer, until either roughly constant or periodic values are produced.

As discussed briefly in section 6.1, another possible extension would be to implement generational evolution in order to compare agents evolving in isolation to agents evolving as a group, so that social traits can be objectively identified.

6.4   Interpreting the model in relation to natural systems

A central idea of the project was to make the simulation as widely applicable as possible, so that it would be possible to draw conclusions about different instances of the emergence of self-organised collective behaviour in natural living systems, rather than modelling a particular species. This makes it difficult to assess how biologically realistic the findings were.

6.4.1 Bacterial aggregates

Agents in the model could be compared to bacteria using enzymes to break down food. As discussed briefly in section 2.2, Shapiro and Dworkin [33] argue that the social behaviour of bacteria is often so integrated that an aggregate of bacteria has all the characteristics of a single multicellular organism. The findings of this project indicate that this integration of social behaviour comes about because of the ability to benefit from interaction between agents (mutual enzyme production between bacteria in this case) without having to make radical behavioural changes from living independently. Also, the findings of this report would suggest that because of the short life-span of bacteria (relative to viscosity; the speed and likelihood that bacteria will move), divergence in the population produced by evolution could cause spatial organisation of different strains of bacteria within an aggregate; different regions of the aggregate evolve different strategies.

Griffin, West and Buckling [2] studied the cooperative production of siderophores (iron scavenging agents) in the pathogenic bacterium *Pseudomonas aeruginosa*. They varied the degree of relatedness (whether the bacteria were all a single clone, or from two different strains) and the scale of the competition (global competition was created mixing up the bacteria so relatives were not nearby, and local competition by not mixing). They found that the proportion of cooperators (bacteria who did produce the siderophores) was greater in the case of high relatedness (because of kin selection [15]) and global competition. In my model, the agents are in local competition, since relatives were nearby spatially, and in the case of the differentiated populations there was low relatedness, as there were two distinct 'strains' of agent. Griffin and colleagues found a low proportion of cooperators (siderophore producers) in these conditions (eventually reaching around 5%) which is consistent with my finding of a low proportion of enzyme producers (around 25% after 20,000 timesteps).

6.4.2 Groups of larger organisms

Another way the simulation could be interpreted is that the agents are larger animals which can access richer food sources by cooperating, for example by hunting or foraging in groups (the bigger the group the bigger the prey that can be taken on, and so the richer the food). However this sort of cooperation requires a more active change in behaviour; two animals must work together to cooperate, rather than just continuing their non-cooperative behaviour whilst happening to be spatially close each other, although since individuals would be hunting small prey anyway, it is easy to see how cooperation could initially occur with very little behavioural change by two individuals simultaneously hunting a single prey. Here, the population is much less viscous relative to life-span (individuals move around a lot during their lifetime), so it is unlikely that the hunting group will form a stable aggregate for long enough for positions within the aggregate to effect survival chances.

6.4.3 Human agriculture

In agricultural human society an individual is able to grow a small amount of food on their own, but with more cooperating farmers the yield increases exponentially. When groups get too big they have to spread out to get more access to land and sometimes split into smaller groups. This dynamic balance between the size of the cooperating population and their access to a spatially distributed resource is similar to the simulation in this project. Again, cooperative farming requires a more active behavioural change than the bacteria case, but Marceau and Gordon [22] model the gradual transition from foraging communities to agricultural communities, which could be seen as another major transition from independent to collective behaviour. Also like the hunting groups, farming communities are likely to move around a lot during their member's lifetimes, meaning that spatial organisation of different evolutionary strategies within the cluster of farmers is very unlikely. Spatial organisation is more likely to occur with learnt social roles and behaviours, as learning occurs considerably faster than evolution.

6.4.4 Validity of the simulation approach

Having made several interpretations, it is necessary to return to the discussion of whether computer simulation is a good enough approximation of a real living system for the interpretations to have any value, in comparison to other approaches to learning about self-organisation in the real world.

As discussed in section 2.4, mathematical models as opposed to simulation are good for investigating the relationships between several variables in a natural system, but with a large number of interdependent variables, the system becomes too complex for equational relationships to be established.

Another alternative approach to studying self-organisation is by experimentation with real biological systems. This gives empirical data, and does not have the problems associated with producing a model (only taking a limited number of variables into account, approximating variables to discrete values, and being open to bias from the designer's assumptions [10]). However, data can be harder to obtain than with models, and the problem of bias shifts from the method of designing the model to the method of measuring the variables. The consistency between the simulation described in this report and Griffin, West and Buckling's [2] experiments with bacteria described above show that the simulation approach does to some extent have ecological validity, and that the two approaches can be complementary. Biological experimentation was deemed unsuitable for this project as the aim was to gain understanding into self-organisation in general, not a specific biological example.

Another alternative method of achieving this aim is through philosophical enquiry. This approach would be useful to clarify the concept of self-organisation, or what it is for a system to be self-organising, but identifying the conditions which give rise to self-organisation might require something to compare the concept to, such as a thought experiment or model.

Given the range of advantages and disadvantages of the different approaches, ideally a mixture of approaches should be used (for example Di Paolo [29] used both mathematical and simulation models in combination). Due to time and resource restrictions it was not possible to use a combination of ap-

proaches in this project, but the findings were related to research from different areas and methodologies (such as [22], [2] and [1]).

6.5 Conclusions

Despite the initial problem of the rich environment in the basic model appearing to be the main source of organisation, later experiments clearly exhibited self-organised collective behaviour and in the case of differentiated populations, the transition from independent to collective behaviour could be identified by changes in the proportion of non-cooperators who could only survive in a social context.

Two conditions for the transition from independent to self-organised collective behaviour were identified. The ability for social dependence to increase gradually due to the way the enzyme effects efficiency of food consumption, meaning that the transition required only changes in existing behaviours rather than the production of new behaviours. Also, spatial self-organisation of different behavioural strategies within aggregates was able to emerge through evolution because of the high population viscosity relative to life-span; agent's lives were so short that their position within a cluster effected the evolution of their behaviour.

These conditions help to explain the integration of social behaviour in bacteria, but while there are common characteristics in cooperation and self-organisation in groups of larger organisms and human social systems, the findings of this project are less applicable to them due to the partial absence of the conditions identified.

It appears there are similarities in the nature of the transition from independent to collective behaviour in different types of living system, and that the approach used can increase understanding into these conditions, but further work is needed to clearly identify conditions for the transition which are common to all living systems.

Bibliography

[1] W.C. Allee. *Animal Aggregations: A case study in general sociology*. University of Chicago Press, 1931.

[2] Stuart West Ashleigh Griffin and Angus Buckling. Cooperation and competition in pathogenic bacteria. *Nature*, 430:1024, 2004.

[3] R. Axelrod. *The Evolution of Cooperation*. New York: Basic Books, 1984.

[4] Randall D. Beer and John C Gallagher. Evolving dynamical neural networks for adaptive behavior. *Adaptive Behaviour*, 1:91, 1992.

[5] Daniel J Bruiger. The rise and fall of reality: Deliberations on the mind-body problem. http://www.island.net/ dbruiger/toc.html, 2003.

[6] S. Bullock. Are artificial mutation biases unnatural? *Advances in Artificial Life - Fifth European Conference (ECAL99)*, pages 64–73, 1999.

[7] A.E. Buzing, P.C. Eiben and M.C. Schut. Emerging communication and cooperation in evolving agent societies. *Journal of Artificial Societies and Social Simulation*, 2004.

[8] Fritjof Capra. *The Web of Life*. London: HarperCollins, 1996.

[9] Alison Carling. *Introducing Neural Networks*. Sigma Press, 1992.

[10] Christopher A. Chung. *Simulation modeling handbook : a practical approach*. Boca Raton : CRC Press, 2004.

[11] J. Di Paolo, E. Noble and S. Bullock. Simulation models as opaque thought experiments. *Artificial life*, 7, 2000.

[12] T. Farnham. Social evolution: Cell lineage conflicts in the social amoebae. http://www.ruf.rice.edu/ evolve/, 2005.

[13] W. Feistel.R., Ebeling. Evolution of biological organization and information processing. In N. Bakardjieva, editor, *Self-regulation of Plant Metabolism*. Publ.House Bulg.Acad.Sci. Sofia, 1986.

[14] Martin Gardner. "the fantastic combinations of john conway's new solitaire game 'life'". *Scientific American*, 233:120–123, 1970.

[15] W.D. Hamilton. The evolution of altruistic behaviour. *Am. Nat.*, 97:354–356, 1963.

[16] Hermann Hild and Alex Waibel. Connected letter recognition with a multi-state time delay neural network. In Stephen José Hanson, Jack D. Cowan, and C. Lee Giles, editors, *Advances in Neural Information Processing Systems*, volume 5, pages 712–719. Morgan Kaufmann, San Mateo, CA, 1993.

[17] Immanuel Kant. *Foundations of the metaphysics of morals, and What is enlightenment?* London : Collier Macmillan, 1959.

[18] S.A. Kauffman. *The Origin of Order: Self-Organisation and Selection in Evolution*. Oxford University Press, 1993.

[19] P.G. Korning. Training of neural networks by means of genetic algorithm working on very long chromosomes. *International Journal of Neural Systems*, 6(3):299–316, 1995.

[20] J.S. Lansing. *Priests and Programmers*. Princeton University Press, 1991.

[21] Eigen M. and Schuster P. *The Hypercycle: A principle of natural self-organization*. Berlin: Springer, 1979.

[22] Nicolas Marceau and Gordon M. Myers. From foraging to agriculture. Technical Report 103, CREFE, Universit du Qubec Montral, February 2000. available at http://ideas.repec.org/p/cre/crefwp/103.html.

[23] James A Marshall and Johnathan E. Rowe. Viscous populations and their support for reciprocal cooperation. *Artificial Life*, 9:327, 2003.

[24] Peter Marshall. *Nature's Web*. London: Simon and Shuster, 1992.

[25] Robert May. Bifurcations and dynamic complexity in simple ecological models. *American Naturalist*, 110:573–578, 1976.

[26] John Maynard Smith and Eors Szathmary. *The Major Transitions in Evolution*. Oxford University Press, 1997.

[27] W.S. McCulloch and W.H. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–133, 1943.

[28] Melanie Mitchell. *An Introduction to Genetic Algorithms*. MIT Press, 1996.

[29] Ezequiel A. Di Paolo. Ecological symmetry breaking can favour the evolution of altruism in an action-response game. *Journal of Theoretical Biology*, 203:135, 2000.

[30] T. Pfeiffer and S. Bonhoeffer. An evolutionary scenario for the transition to undifferentiated multicellularity. *Proceedings of the National Academy of Sciences of the USA*, 100(3):1095–1098, 2003.

[31] Ilya Prigogine and Isabelle Stengers. *Order out of Chaos - Man's New Dialogue with Nature*. London : Heinemann, 1984.

[32] Elisabet Sahtouris. *Earthdance : Living Systems in Evolution*. iUniverse, 1999.

[33] James Shapiro and Martin Dworkin. *Bacteria as Multi Cellular Organisms*. New York: Oxford University Press, 1997.

[34] James A. Spudich. *Dictyostelium discoideum : molecular approaches to cell biology*. London: Academic Press, 1987.

[35] R Stevens and R Sosic. Emergent behaviour in slime mould environments. *ICEC*, 274:684–688, 1995.

[36] A Turing. Computing machinery and intelligence. *Mind*, 59:433–460, 1950.

[37] Joseph Weizenbaum. Eliza: A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):35–36, 1966.

[38] G.C. Williams. *Group Selection*. Chicago : Aldine-Atherton, 1971.

[39] David WillShaw. Self-organisation in the nervous system. Foresight Cognitive Systems Project, http://www.foresight.gov.uk, 2003.

Appendix A - Personal reflection

I really enjoyed writing and experimenting with the simulation, it was easy to spend a lot of time experimenting with different conditions, but to be systematic and spend time analysing the results before trying out a new idea, as well as keeping to the project schedule (table 3.1) was quite difficult, and not something I have had much experience of doing.

I have not had the opportunity to devote this much time to a single piece of work before and found myself getting very involved in it. Writing a report of this length was also undoubtedly a useful experience, and allowed me to develop a skill which I will hopefully be using again.

There were times when I was very sceptical about the idea that the simulation could really be compared to a natural system, or that the behaviour produced was anything more than that which I had explicitly designed into it. Because I had a vague idea in my head from the start of what I wanted to see, it seemed impossible to argue it had been produced naturally and spontaneously. But every now and then I would get a result I had not expected and my interest would return.

Self-organisation and collective behaviour are such widely applicable topics that it was easy to integrate a lot of the material I had already studied for my Cognitive Science degree into this project, particularly from bio-inspired computing but also from cognitive neuroscience, where I have studied self-organisation in the nervous system, and from various philosophy modules which have touched on holistic perspectives of nature. I was lucky to have found an area which allowed me to do a programming project whilst pursuing my interest in philosophy.

Whilst working on this project I was also working on a 10 credit 'independent study' in philosophy, where I argued that human society can be described as a self-organising living system, based on the characteristics of self-organisation described by Prigogine [31]. I found that the two projects complemented each other well, as when some philosophical issue arose in this project I had an avenue by which to explore it.

If I did the project again I would definitely start with an even simpler model and add complexity more gradually and systematically, so that it was clear which aspect of the model was producing which

behaviour. Also, I would try to run the simulation without graphics first and draw conclusions from the statistical results before watching what is going on, as I think that seeing what was happening effected my interpretation of the statistics. So I would advise future students who are considering attempting a project in this area to try to be systematic and scientific from the start, rather than having a good mess around with the model and worrying about analysing what is going on later. Having said that, outside of the context of a computing project I like the idea of the simulation being open to any interpretation, and if I was to pursue it outside of my studies I'd put more effort into improving the graphical output rather than logging data for analysis. It is the lifelike movements which can only be captured visually and subjectively that bear the most striking resemblance to a living system. These movements, periodic but never quite repeating exactly, reflect the links between life and dance described by Sahtouris [32], and to me represent spontaneity, anarchy and partying.

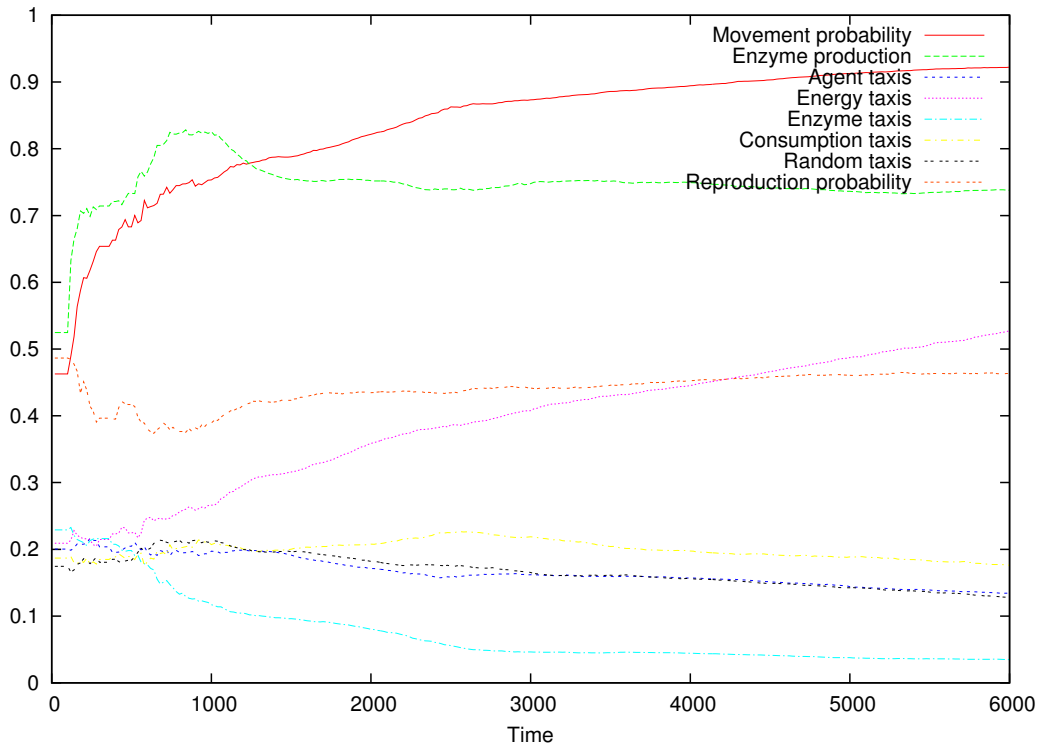# Appendix B - Changes in average genome over time for the evolutionary model

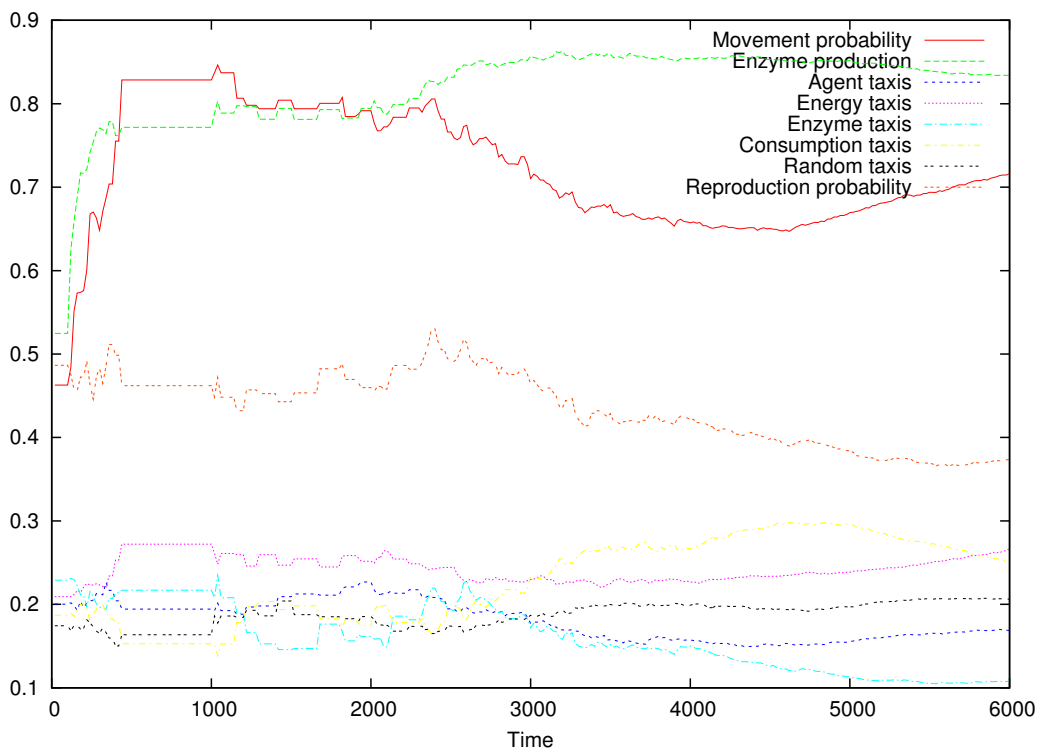Figure 6.1: Changes in genome in experiment 2.1
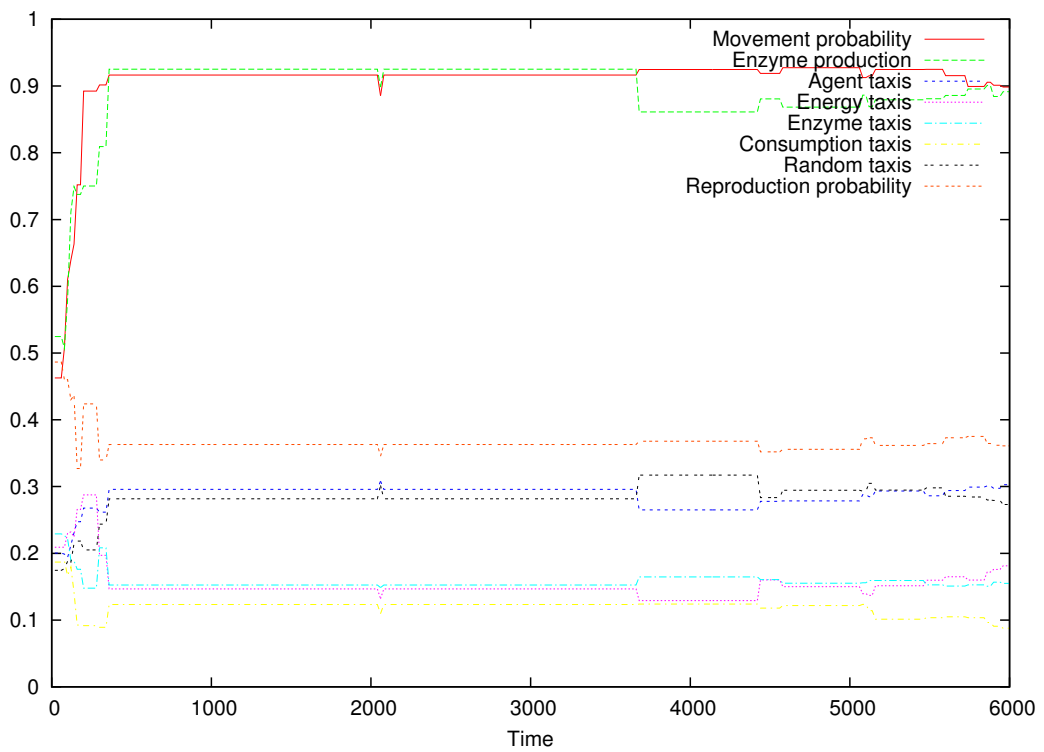
Figure 6.2: Changes in genome in experiment 2.2

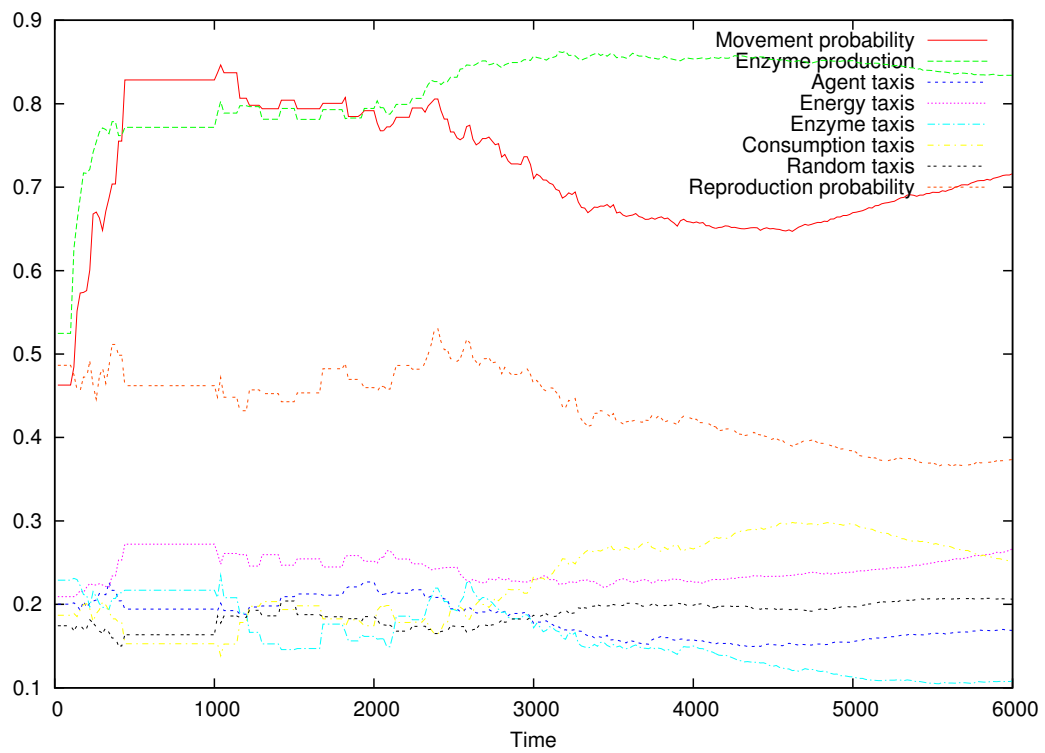Figure 6.3: Changes in genome for experiment 2.3

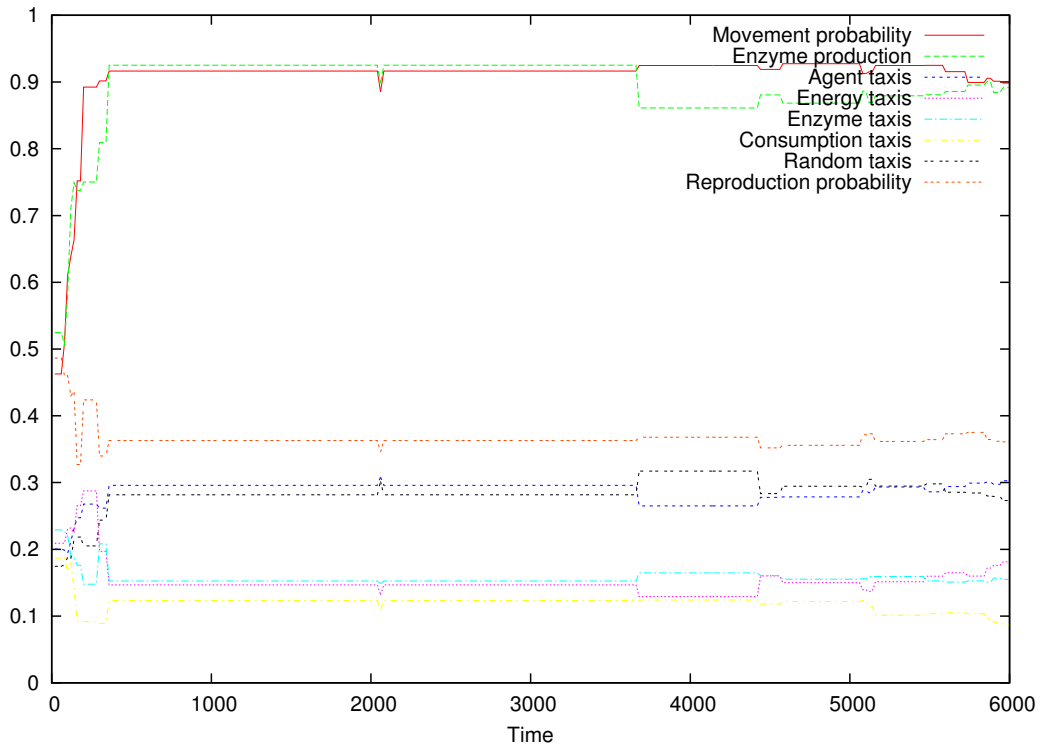Figure 6.4: Changes in genome for experiment 2.4
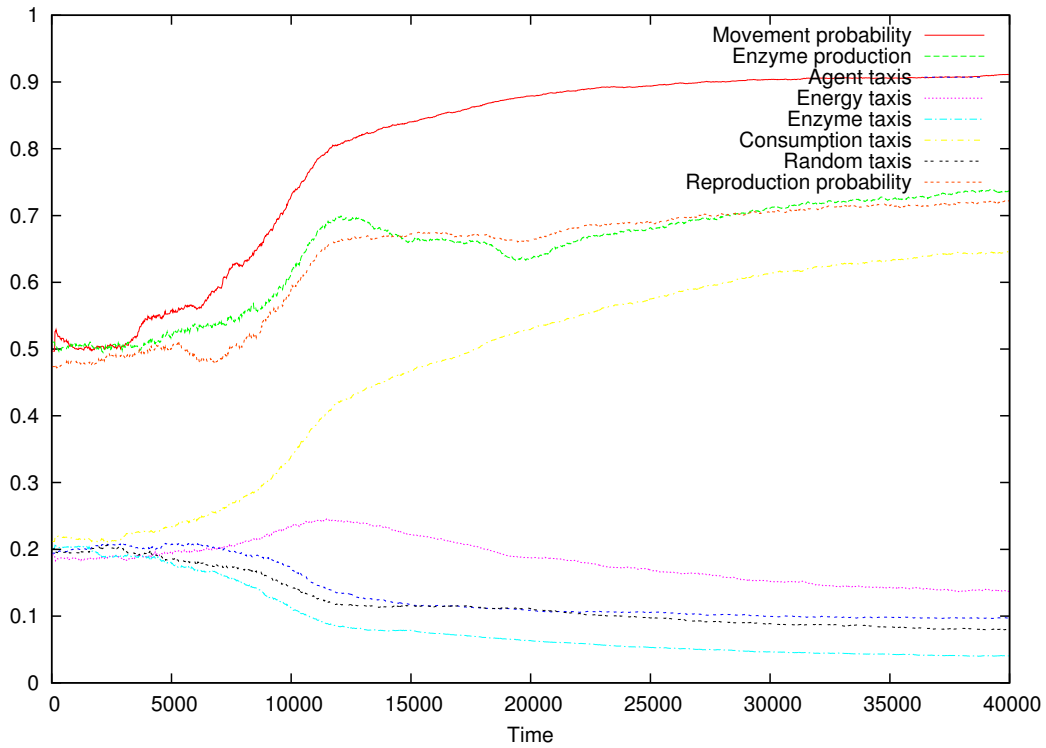
Figure 6.5: Changes in genome for experiment 2.5

Figure 6.6: Average changes in genome for differentiated populations