Chapter 4

Experimental Overview and Progression

4.1 Overview

In chapter 2 the problem of altruism and co-operation in MAS was discussed. In chapter 3 a number of possible methodologies were presented in relation to artificial societies. In this chapter an overview of the experimental progression of the rest of the thesis are given, drawing on these two previous chapters. The aim of this brief chapter is to give an overview of the logical progression of the work and the methodology applied at each stage. Very brief details are given of results obtained from experiments, if this is required to understand the progression of the work, but the reader is referred to the relevant chapters for detailed results.

4.2 The SwapShop

In chapter 2, kin altruism is discussed in the context of cultural evolution. The idea that a form of "memetic kin" altruism may emerge from a cultural evolutionary process is advanced. In chapter 5 an artificial society is presented (the SwapShop) which implements a resource sharing scenario in which agents learn via a simple cultural learning rule. Agents are based on a satisficing model in which the underlying process of meme propagation and transmission between spatially adjacent agents is modelled. The aim is to experimentally determine if such a model is sufficient to produce altruisitc behaviour consistent with a "memetic kin" hypothesis. Additionally the cultural learning rule is compared to a more traditional optimising "genetic" learning rule, in order to compare genetic kin altruism with memetic kin altruism.

Methodologically this is an application of *theory testing* (see figure 3.3 in chapter 3). We start with a minimal implementation of an existing theory of altruism based on "memetic kin" (as given by Allison [2]), and test the hypothesis that altruistic behaviour will be selected for. Additionally, Allison speculated that memetic kin altruism may produce higher levels of altruism than genetic kin processes. This hypothesis is also tested.

Therefore, two qualitative hypotheses concering memetic kin altruism are tested in the SwapShop experiments. Both are found to be consistent with the results obtained. Consequently Allison's informally stated theory has been augmented with a concrete computational model which produces results consistent with two qualitative hypotheses derrived from the theory.

The SwapShop demonstrates a form of "cultural kin" selection in which "kin" (those sharing the same memes) are able to recognise one another. However, the ability of agents to recognise cultural kin is built into the model. This assumption that cultural

kin can recognise one another is a kind of "fix" since the SwapShop *does not model any underlying mechanisms that produce cultural kin recognition*, it is simply assumed. The StereoLab attempts to model that inderlying process, minimally, via the mechanim of a cultural "stereotype".

4.3 The StereoLab

In chapter 6 an artificial society is outlined which consists of agents with simple stereotyping abilities (the StereoLab). Chapter 8 describes experiments carried out in the StereoLab. The methodology employed here is theory construction (see figure 3.4 in chapter 3). We start with the problem of how agents can come to recognise cultural kin based on observable markers *when there is no essential linkage between the two*. In the SwapShop (described above) agents have the ability to recognise those who share all the same memes as themselves. That kind of ability does not seem reasonable in human societies or practical in open MAS. Allison speculates that what is required are mechanisms which he terms "cultural packaging" techniques. Such mechanisms would produce correlations between observable cues (i.e. surface memes) and agent behaviours (i.e. hidden memes)[1]. However, we have no *a priori* theory of how such packaging techniques may operate. So on what basis can an artificial society be designed to aid construction of such a theory?

In the StereoLab, agents are specified with cognitively plausable (yet limited) stereotyping characteristics in which observable markers (tags) are mapped to interaction strategies. Agents interact by playing pairwise games of the Prisoner's Dilemma and culturally learn in a similar way to agents in the SwapShop.

Here then, we implement agents based on a simple theory of stereotyping. The

[1]See section 2.6.2 in chapter 2 for a discussion of surface and hidden memes.

scenario produced captures the ability of agents to observe, copy and recognise cultural markers (tags) and to select differential behaviours towards others based on these. Agents culturally learn the differential behaviours (stereotypes) from each other. So although we have no theory of how co-operation might emerge from a cultural process involving tags and behaviours, we can still construct an artificial society based on a simple theory of how individual agents behave.

Since we start with no *a priori* theory of how altruistic and co-operative behaviours may emerge from such a stereotyping process, many of the assumptions in the StereoLab concerning agent behaviour and interaction are parameterised. Given this parameterisation the StereoLab expresses a large space of possible artificial societies. This space is then searched in a semi-automated way (see chapter 7) in order to locate those societies from the space which produce high co-operation. This is a form of *reverse engineering* (see figure 3.2 in chapter 3) in which a desired behaviour (in this case high co-operation between agents) is searched for in a space of assumptions.

The experimental objective of chapter 8 is therefore to locate and characterise regions of the StereoLab parameter space which produce high co-operation. An additional methodological and technical objective is to apply techniques outlined in chapter 7 to achieve that goal.

The experimental work condenses results, from tens of thousands of individual experimental runs, into a set of regions in the parameter space of the model which result in high co-operation between agents. From within these regions individual runs are observed and explanations are produced as to why high co-operation is occuring. This is a form of *explanation finding* (see figure 3.5 in chapter 3). In different regions different co-operation forming processes are identified.

One of the co-operation forming processes identified involves the formation of groups sharing the same tags and biased game interaction within groups. An explantion of how this process promotes co-operation is produced, and this is fed back into a tentative theory of co-operation based on tags and interaction biasing. In order to further test the new theory, and produce more detailed analysis of the process, a further simplified artificial society is constructed (TagWorldII). This society expresses the theory as simply as possible and also applies it to an optimising agent scenario.

4.4 TagWorldII

In chapter 9 the theory produced from experimentation in the StereoLab (chapter 8) is tested in a different scenario (TagWorldII). The experimentation is therefore an application of the *theory testing methodology* (see figure 3.3 in chapter 3).

TagWorldII is a simplified version of the StereoLab which minimally captures the theory. It is also different from the StereoLab in that agents are represented as optimisers. The optimising assumption is applied in order to test if the theory is robust under both the satisficing and optimising assumptions. The simplification allows for the process to be visualised and more precisly characterised mathematically. Several of the key assumptions of the TagWorldII society are parameterised and the co-operation promoting process is shown to be present over a large part of the space.

4.5 Summary

Overall the experimental work follows the following pattern:

- Initially an artificial society is constructed to test an existing theory (chapter 5).

- Questions arising from this lead to the construction of an artificial society to aid the construction of new theory (chapters 6 and 8).

- The newly constructed theory is then tested and refined in a further simplified artificial society (chapter 9).

The possible implications and applications of the new theory are discussed in chapter 10.