Chapter 3

Artificial Society Methodology

3.1 The Need for Methodological Clarity

The concept of artificial societies, and experimentation within the artificial domain, indicates a break with more traditional conceptions of simulation. Traditionally simulation relates to some pre-existing "target" system. A model is specified formally (either mathematically or computationally) based on what is known about the target, along with hunches concerning what is not known. The model is then executed with realistic initial conditions and the simulated results compared to the target in some way. The agreement between the model and the target is a measure of the accuracy of the model. This comparison of the model with the target can be seen as a form of validation which grounds the model [43]. Without this grounding the model can not be said to represent any particular target system or process.

The work done in the area of artificial societies is less "simulation" and more "construction". By analogy with artificial intelligence and artificial life, this synthetic approach does not attempt to model existing targets, but rather to study general processes which characterise many possible targets. In this sense, artificial systems and their study can be seen as form of theory construction in an abstract computational domain. Methodological confusion can arise due to the use of the word "simulation" with its traditional meaning and implication of a real world target. Scope for confusion is also produced due to the engineering stance required for construction of a computational model and its relationship to the traditional conceptions of social theory. For these reasons it is considered important to be clear about the methods used within artificial society work.

A set of methodologies, loosely based on the Popperian [136] concept of theory refutation through the falsification of hypotheses and consequent theory development, will now be presented. This process occurs within a deductive system (the artificial society) which, due to its complexity, requires empirical investigations to refute hypotheses and generate theory. Essentially then, the computational model itself is seen as a "target", and theories are constructed based on empirical investigation of the model via execution of the model.

3.2 Artificial Societies

Computational modelling and simulation of social systems has a history of almost forty years [78]. The exploratory "Artificial Societies" (ASoc) approach is a more recent trend [60]. ASoc models address "possible societies", their general processes, dynamics, and emergent properties [63]. In the same way that Artificial Intelligence is not limited to the accurate modelling of physiological brain processes, so ASoc research does not start from some given scenario or particular social system. The aim is to model features and processes which characterise societies in general: co-operation, specialisation, group formation, hierarchy etc. ASoc work does not strive for superficial realism or direct correspondence with existing societies, but for abstract logical relationships that characterise whole categories of phenomena. The aim then, is to model what might be termed "abstract social processes". Generally, ASocs consist of multiple interacting agents. Each agent minimally consists of: internal state; sets of possible actions; percepts (or perceptual inputs); a shared environment and some form of decision process informing action selection. This latter component of the agent "architecture" may vary considerably. It may consist minimally of fixed precoded rules (e.g. the Sugerscape [52]), deliberative, planning and goal directed AI systems (e.g. EOS [48]), inductive learning (e.g. via connectionist models [94]) or population level evolutionary methods (e.g. evolutionary game theory [84]).

3.3 A Special Methodology for Artificial Societies?

Many have recently noted that ASoc work does not fit neatly into existing experimental methodologies. Consider the following comments by key researchers in the field:

> "Simulation is a third way of doing science. Like deduction, it starts with a set of explicit assumptions. But unlike deduction, it does not prove theorems... induction can be used to find patterns in data, and deduction can be used to find consequences of assumptions, simulation modelling can be used as an aid to intuition." Axelrod [9].
>
> "Clearly, agent-based social science does not seem to be either deductive or inductive in the usual senses. But then what is it? We think generative is an appropriate term... We consider a given macrostructure to be 'explained' by a given microspecification..." Epstein & Axtell [52].
>
> "We can therefore hope to develop an abstract theory of multiple agent systems and then to transfer its insights to human social systems, without *a priori* commitment to existing particular social theory." Doran [46].
>
> "Our stress... is on a new experimental methodology consisting of observing theoretical models performing on some testbed. Such a new methodology could be defined as 'exploratory simulation'... " Gilbert & Conte [63].
>
> "In the past social scientists have tended to espouse either deduction... or induction. Simulation provides a third possibility, in which one starts with a set of assumptions, but then uses an experimental method to generate data which can be analysed inductively. Keep in mind the need to iterate between deductive and inductive strategy as one developes the model." Gilbert [64].

Below a set of methodologies is presented which attempt to untangle and incorporate these various strands of investigation.

3.4 Artificial Society Experimentation

ASoc work can be seen as comprising a set of assumptions (A) used to construct the society; a set of runs (R) produced by execution of a computer program which embodies (A); a set of measurements or observations (O) of the runs; a set of explanations (a theory) (E) which attempts to link (A) and (O) and a set of hypotheses (H) linked to (E) based on (A) and (O). (E) should be thought of as a (possibly incomplete and intuitive) theory of why the program (A) actually produces the (O) that it does. (E) is expressed in the terms of the the structures and relationships specified in the program.

(A), (R) and (O) are semi-formalised since (A) represents a computer program, (R) some set of executions of the program and (O) some specified measures of (R). (E) and (H) may or may not be formalised. They are often given in a mixture of natural language using qualitative concepts and statistical or mathematical relationships. In either case the explanation (E) aims to illuminate the dynamic processes in (R) with reference to (A) and (O) and possibly via the identification of some emergent properties (see section 1.1.5 in chapter 1).

3.5 Pick & Mix Methodologies

Connecting the above components in different ways reveals several modes of inquiry, some of which are now detailed. An *existence proof* does not require an intial (E) at all. Here some (A) is demonstrated to be sufficient to produce some (O) which then may

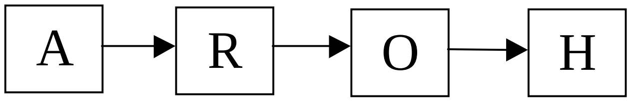

Figure 3.1: Existance Proof. Here a set of assumptions (A), represented by a program is executed to produce a set of runs (R). A set of observations (O) over the runs may be used to support a specific hypthesis (H), namely that (A) is *sufficient* to produce (O).

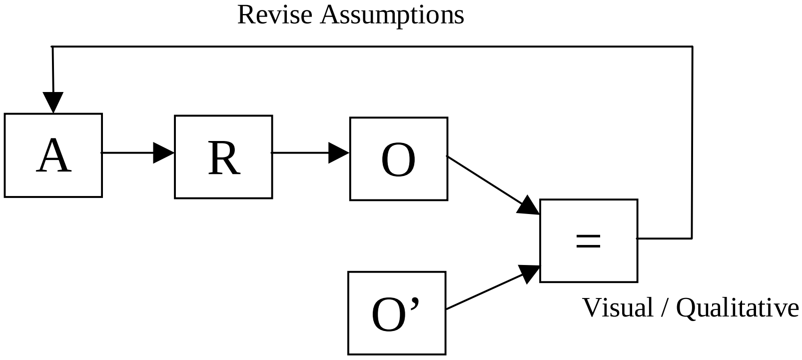

Figure 3.2: Behaviour Modelling or Reverse Engineering. Given some desired set of observations (O') the assumptions (A) represented by the program can be revised until some level of correspondance is produced.

be used to justify some (H) (see figure 3.1).

Behaviour modelling (or reverse engineering) does not require (E) or (H). Here some desirable observation is formulated (O') and compared (possibly visually / qualitatively) against actual observations (O) and, based on divergence, (A) is revised (see figure 3.2). This process is continued until a satisfactory level (however defined) of correspondence is observed.

Theory testing involves the translation / abstraction of some existing theory concerning real social processes (T) into (E), (A) and (H) and then the testing of (H) against (O) in order to either support or refute (H) and by implication (T) (see figure 3.3). In order to "translate" existing theory, often expressed in qualitative and narrative form, into a computational model generally requires additional assumptions to be made which are not

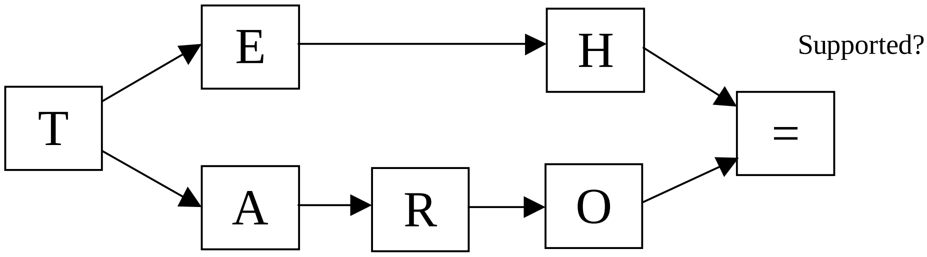

Figure 3.3: Theory Testing. Here some existing theory (T) which relates to some real social target system is translated into a program (A) that attempts to capture the assumptions of the theory and an explantion (E) which expresses the theory in the terms of the program (A). From (E) some set of hypotheses can be generated which can then be compared to observations (O) of actual runs (R) of the program (A).

part of the original theory. This requirement to augment a theory with other assumptions in order to reach a level at which a program can be constructed is the norm [49]. It is obviously a matter of debate as to how well the final implementation (A) captures the original theory. The process of implementing the theory in this way however is beneficial in its own right since it clarifies what is missing, what is vague, what is contentious etc.

What does this kind of "theory testing" tell us about the original theory (T)? It certainly does not "test" the theory empirically in the real world. The empirical "test" here is in an artificial and highly abstract world (often termed testing *in silico* [52], [81]). Also the translation of (T) into (A) forces a formal (algorithmic) representation of the salient entities and relationships. In order to produce an executable program, (A) must be consistent [124]. Any inconsistences in (T) would therefore be brought to light. This is a second form of "test". When (T) is not formally specified it is a matter of debate as to whether (A) captures the salient features of (T).

Theory building involves the abstraction from (T) into (E), (A) and (H). (E) can then be used to produced (H). If the comparison between (O) and (H) leads to refutation of (H) then (E) and / or (A) can be revised. The revision of either of these means that new

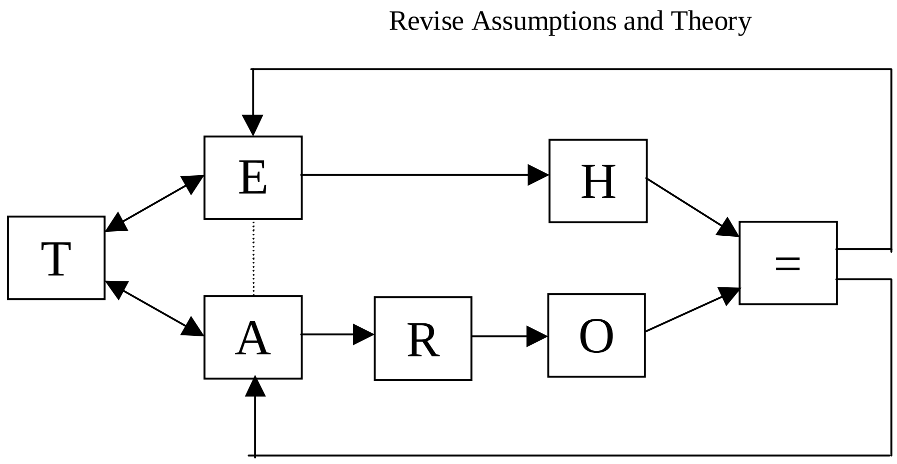

Figure 3.4: Theory Building. An existing theory (T) is translated into a set of assumptions (A) represented by a program and an explanation (E) which expresses the theory in terms of the program. The generation of hypotheses (H) from (E) and comparison against observations (O) of runs (R) of the program allows for both (A) and (E) to be revised. If eventually (H) and (O) correspond then (A), (E) and (H) may be fed back into a new revised theory (T) which may be applicable to a real target.

theory is being created. If a state is reached in which (E), (H) and (O) correspond, (E) and (A) can then possibly be "de-abstracted" into (T) producing a theory testable against real social processes (see figure 3.4).

Explanation finding involves iterative refinement of (E) based on comparison of (H) with (O) without changing (A) (see figure 3.5). This mode of investigation mirrors traditional empirical methods employed in the natural sciences. The difference is that here the computational model, rather than some aspect of the "natural" world, becomes the object of study for which explanations and hypotheses are generated and tested.

Many of these modes combine deduction and induction often in an iterative way. Elsewhere I have called this "Ceduction" [69], [74]. The inductive process here is viewed as iterative observation and revision of (E) and (A). The (O) is produced deductively (computationally) from (A) but the revision of (E) and (A) is an inductive process based on observations (O) guided by (H).

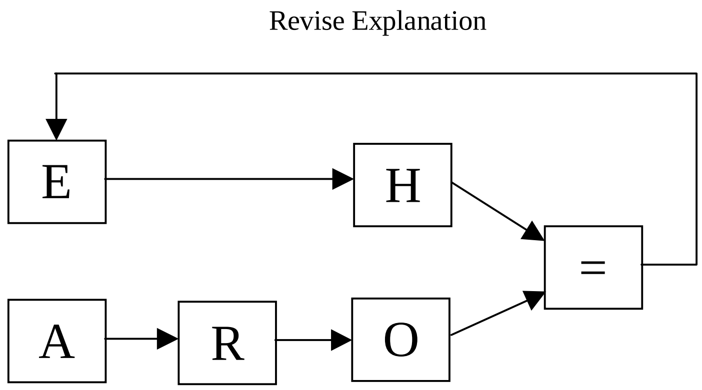

Figure 3.5: Explanation Finding involves iterative refinement of the explantion (E) based on comparison of hypotheses (H) with observations (O) without changing the assumptions expressed by the program (A).

3.6 Conclusion

Artificial societies do not aim to model real societies in any direct sense. They can be seen as an aid to intuition in which the researcher formalises abstract and logical relationships between entities. When constructed computationally they can be investigated empirically by producing explanations and hypotheses which can be refuted by observation of the behaviour of the computational model. More interestingly, since the assumptions upon which the society is constructed can be changed, there is no need to simply observe. Other modes of investigation are possible in which assumptions are changed to produce desired behaviour. Whatever the relationship between real societies and artificial ones (a problematic area) it is argued that such systems can be used as an aid to theory construction by sharpening intuition in the formulation of hypotheses which can be testable against real data. Such claims may sound far-fetched but within evolutionary biology a recent example of this kind of progress [161] has been demonstrated.

Artificial societies facilitate experimentation, discourse and theory formulation concerning abstract social processes. In that context they offer a way of moving social theory into an experimental and semi-formal realm. In this realm hypotheses can be tested, experimental results independently reproduced and models unambiguously communicated (via computer programs). On this basis abstract social theory can be developed in a more objective way than traditional social theory, where intuition and rhetoric are often used instead of experimentation.