E*PLORE-ING THE SIMULATION DESIGN SPACE

Luis Antunes Helder Coelho João Balsa

GUESS/Universidade de Lisboa, Portugal
{xarax, hcoelho, jbalsa}@di.fc.ul.pt

Abstract

One of the major puzzles in performing multi-agent-based simulations is the validity of their results. Optimisation of simulation parameters can lead to results that can be deceitful, optimistic, or plainly wrong. When the issue at stake is inherently complex, which is frequently the case with social phenomena, the search for emergent outcomes is closely related to macro effects deriving from micro behaviours, and the drawing of valid conclusions from the analysis of the observed results should be done with extra care.

Multi-agent-based social simulation is increasingly used not only to understand and explain phenomena, but also to predict outcomes and even to prescribe measures to be adopted by colective (public or private) entities. The notion that conclusions of simulation studies will be applied to real social settings brings an added responsibility to the researcher. Principled methodologies are needed that can minimise the *ad hoc* nature of experimentation.

In this paper, we present a set of methodological principles to explore the space of possible designs involved in simulation experiments. Principles are needed not only for the design of agents and the societies they are immersed in, but also for the design of models of simulations themselves. Several techniques are shown that can provide an increasingly broad covering of the space of possible *experiment* designs. We also explore some alternatives on how to progressively complexify particular mechanisms.

1 Introduction

In multiagent systems (MAS), the main concern has been to develop a sound principled recipe to develop and deploy a system from a more or less formal specification. Recent work by Wooldridge et al. [25] was preceded by other attempts such as ours [12], or those of Cohen at al. [13, 21]. A good overview can be consulted in [9] but we will not address all the methodologies therein because of space limitations. Early inspirations already realising the complexity of the task can be found in [24].

In a discipline such as Multi-Agent-Based Simulation (MABS), the idea is to bring together MAS and Social Sciences in a mutually fruitful cooperation. Concepts and techniques from the Social Sciences have been in the genesis of MAS, and social scientists now resort to MAS environments as an additional means with which to conduct experiments and validate theoretical work. The MABS endeavour is a fertile cross-cultural field, where some of the most exciting ideas from the several areas involved get assessed and tested. Methodologically speaking, MABS is a hard venture, because of the complexity of the systems involved, which is severely boosted up by desirable characteristics of MABS systems, such as agent autonomy, agent heterogeneity, and the sheer number of interactions among agents. Gilbert's methodology [17] is similar to Cohen's MAD (Modelling, Analysis and Design), but differs in a significant step, as we will show below. A note on the tackling of complexity at large, through MAS installations, is necessary to support the idea that *ad hoc* procedures are not advisable today.

When MABS is used to do Social Simulation, our aim is to have a deeper understanding of some selected social phenomena, overtaking some of the typical pitfalls met when using reductionist perspectives over an intrinsically complex problem. We conduct Social Simulation by bringing an holistic view into exploratory agent-based simulations. However, if the methodological stance towards MAS and MABS has already been addressed, the field of exploratory simulation is even more complex, and is still in need of some principles with which to guide the researchers in such a way that strengthens confidence in the obtained results, and their analysis.

So, in this paper we will propose a draft of a set of methodological principles with which to guide exploratory simulations in Social Science phenomena. This methodology builds up from other MAS and MABS methodologies to address all levels of complexity in such a simulation, namely, the agent cognitive level, the societal level, and the experimental (simulation) level itself. The *leitmotiv* of this methodology will be centred around complexity. We need to explore complex systems to get to know them, not to simplify them to a point we can easily know them. To this end, we build up on our vision of MAD methodology, (back and forth journeys in design proposed in [12]), complement it with more recent developments on individual decision in the BVG (Beliefs-Values-Goals) choice framework [6], and a schematic vision of exploratory simulation we addressed in [5].

The development of this methodology was based on the tax compliance scenario as inspiration and applicational support [3, 4, 8]. We should note that the kind of activities that e*plore involves, by no means eases up the task of the developer and simulator. What it does it to explicitly consider the structure of the development of the several models. The result of this exploration of the space of possible models could be compared in terms of complexity and effort with the usual process of sequential development, programming, and refinement of one model. However, instead of looking for *the* model, it does consider design options and lists alternative models. In these alternative scenarios, somewhat simplified visions of the problem are studied. Admittedly, this involves the risk that some necessary complexity is lost in the separation of characteristics. Still, no single one of this models is the absolute answer to the proposed problem. In the exploration of these individual models and their variability, we aim at getting deeper insight into the several facets of the target phenomenon, so that a unified view can be built, modelled and simulated.

The rest of the paper is organised as follows. In the next section we address some of the most representative methodologies for experimentation in MAS and MABS, and focus on their evolution. We then summarise the idea of exploratory simulation as proposed in the literature and enumerate and discuss the persistent methodological problems still to be found despite all systematisation efforts. We then present our first attempt at a unifying methodology for (exploratory, multi-agent-based) social simulation. Section 5 discusses the purpose of social simulation, and recommends prudence on the generalisation of its findings. The following section discusses the methodological steps in depth, focussing especially on evaluation. Section 6.1 takes on Sloman's idea of exploration of design space in this context, and proposes cumulative ways of covering design space by manipulating models design. Finally, section 7 enumerates the steps of the methodology, before we produce some concluding remarks.

2 Methodologies for Development of Multi-Agent Systems and Multi-Agent-Based Simulation

Recently, very serious efforts were produced on the issue of building up a solid methodology for deploying multi-agent systems (MAS). Perhaps the most achieved and influent of these efforts is Gaia, by Wooldridge et al. [25]. Gaia is involved in the MAS area coming of age, in what it attempts to establish a set of concepts and principles to build on a system and its components that is general and comprehensive, and apt to deal with the enormous development of agent systems we have watched.

In Gaia, the founding idea is that a MAS is a computational organisation consisting of several interacting roles. Gaia is proposed from an engineering standpoint, which is clear from the domain characteristics adopted. However, some of those characteristics are not adequate when we take on a more scientific stance. Gaia assumes that "the goal is to obtain a system that maximises some global quality measure (...) [and] is not intended for systems that admit the possibility of true conflict." [25, page 286]

In this light, we start our search for a more general methodology for social simulation, having Cohen's 1991 MAD (Modelling, Analysis and Design) [13] in mind. Cohen was worried about defining the general lines of an experimental method for artificial intelligence. Controlled experiments are designed to suggest or provide evidence for theories that can explain differences in the performance of systems. Acknowledging that empirical results are seldom general, Cohen insisted that nothing prevents the researcher from "inventing general theories as interpretations of results of studies in simulation testbeds, and nothing prevents (...) from designing additional studies to test predictions of these theories in several simulation testbeds" [21, page 39].

MAD (Modelling, Analysis and Design) involves seven activities [13]: (1) evaluate the environmental factors that affect behaviour; (2) model the causal relations between system design, its environment, and its behaviour; (3) design or redesign a system (or part of one); (4) predict how the system will behave; (5) run experiments to test predictions; (6) explain unexpected results and modify the models and design of the system; and (7) generalise models to classes of systems, environments and behaviours.

In [12] we have critically addressed this methodology from a systems development standpoint: to program is not only to code either formal or informal descriptions, so we have proposed to slide Cohen's ecology triangle along a line that could be travelled back and forth, as we depict in figure 1.

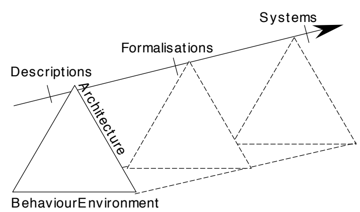

Figure 1: *Extended MAD: moving the ecology triangle along the design axis (adapted from [12]).*

In [5] we readdressed this methodology and confronted it with Gilbert's methodology for computational simulation [17]: (1) identify a "puzzle," a *question* whose answer is unknown; (2) *definition* of the target of modelling; (3) normally, some *observations* of the target are necessary, to provide the parameters and initial conditions of the model; (4) after developing the model (probably in the form of a computer program), the *simulation* is executed, and its results are registered; (5) *verification* assures the model is correctly developed; (6) *validation* ensures that the behaviour of the model corresponds to the behaviour of the target; and (7) finally, the *sensitivity analysis* tells how sensitive the model is to small changes in the parameters and initial conditions.

Both methodologies are quite similar, but in MAD there is no return to the original phenomenon. While Cohen's emphasis is on the system, Gilbert is more concerned with the original phenomenon to be modelled and simulated. In [5], we proposed some methodological principles with which to confront the results of simulations, and proposed a merge between extended MAD and a description of exploratory simulation, crossed with the the idea of heterogeneous agents with an individual choice framework, that took the experiment designer inside the whole methodological scheme. The key idea is not to mask complexity away from experimentation with complex models and systems. The existing methodologies are not capable of dealing with the complexity contained in today's exploratory simulations (ES) with agent-based social systems. This concern (see also [10]) comes from the best of reasons: today's agent technology, together with the increased computational power available, brought the social scientists to tackle new problems (or scaled up old problems), through computational simulations, that they would not dream of until recently. The existing methodologies are too focussed on realising a system tuned for a given purpose, whereas in ES that purpose is too vague and complex to be defined from the start.

3 Exploratory Simulation

The notion of agent and computational simulation are the master beams of the new complexity science [15]. Computational simulation is methodologically appropriate when a social phenomenon is not directly accessible [19]. One of the reasons for this inaccessibility is the target phenomenon being so complex that the researcher cannot grasp its relevant elements. Simulation is based in a more observable phenomenon than the target one. Often, the study of the model is as interesting as the study of the phenomenon itself, and the model becomes a legitimate object of research [14]. There is a shift from the focus of research of natural societies (the behaviour of a society model can be observed "in vitro" to test the underlying theory) to the artificial societies themselves (study

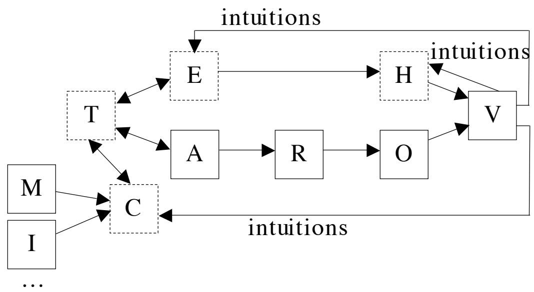

Figure 2: *Exploratory simulation. A theory (T) is being built from a set of conjectures (C), and in terms of the explanations (E) that it can generate, and hypotheses (H) it can produce. Conjectures (C) come out of the current state of the theory (T), and also out of metaphors (M) and intuitions (I) used by the designer. Results (V) of evaluating observations (O) of runs (R) of the program that represents assumptions (A) are used to generate new explanations (E), reformulate the conjectures (C) and hypotheses (H), thus allowing the reformulation of the theory (T) (from [5]).*

of possible societies). The questions to be answered cease to be "what happened?" and "what may have happened?" and become "what are the necessary conditions for a given result to be obtained?," and cease to have a purely descriptive character to acquire a prescriptive one. A new stance can be synthesised, and designated "exploratory simulation" [14]. The prescriptive character (exploration) cannot be simplistically resumed to a optimisation, such as the descriptive character is not a simple reproduction of the real social phenomena.

In this methodological stance, the site of the experimenter becomes central, which reinforces the need of defining common ground between him/her and the mental content of the agents in the simulation (see figure 2). Hales [20] claims that experimentation in artificial societies demands for new methods, different from traditional induction and deduction. Like Axelrod says: "Simulation is a third form of making science. (...) While induction can be used to discover patterns in data, and deduction can be used to find consequences of assumptions, the modelling of simulations can be used as an aid to intuition" [7, page 24].

However, as Casti stresses [11], there are difficulties in concretising the verification process: the goal of these simulation models is not to make predictions, but to obtain more knowledge and insight. In [5], we emphasised the fact that theories, explanations and hypotheses are being constructed, not only given and tested. Simulation is precisely the search for theories and hypotheses. These come from conjectures, through metaphors, intuitions, etc. Even evaluation needs intuitions from the designer to lead to new hypotheses and explanations. This process allows the agent's choices to approximate the model that is provided as reference. Perhaps this model is not as accurate as it should be, but it can always be replaced by another, and the whole process of simulation can provide insights about this other model.

4 Persistent Methodological Problems

In this section we summarise the problems that persist after all these methodological undertakings that have crossed the last decade or so whilst this multi-disciplinary area of multi-agent-based exploratory social simulation was being delineated, and its goals and possibilities were better understood. Next, we will claim that the area as a whole is ready to go further and propose solutions for real world (target system) problems and questions.

4.1 Validity and Significance of Results

All modellers, simulators and experiments are worried about the validity and significance of the models they build and use. Unfortunately, as we have seen from the comparison between the two methodologies above, once the models are built, tested and deployed, the experimenter may tend to look at them as being the real system, and forget they are still only models. And so, outcomes of the MABS are still outcomes of a simulation, not necessarily similar or representative of how the world would react in the same conditions. This was the criticism behind the proposal of Extended MAD [12], but as more and more models and simulations are being created and explored, we notice that this basically flawed stance should still be stressed and fought against. Promises can kill a research program, and social simulation is still at its infancy and needs to be protected.

4.2 The Role of the Observer/Experimenter

Another persistent issue is the place and role of the experiment designer. Discrepancies between the notions of causality and correlation may lead to poor interpretations of the modelling efforts. Since a recurrent issue of exploratory simulation is emergence, and this concept depends on what the observer is expecting (or, more formally, can demonstrate to be derivable) from the system design, there are several issues to be addressed. In truth, they have been mentioned by several authors in the literature and public addresses, perhaps only not systematically. We will provide some illustrations of the importance of this issue:

- Axelrod defended in [7] that models and simulations should be described in such a way so as to be reproducible and indeed reproduced by different people, in an effort to ensure validation of experiment designs and their outcomes;
- Gilbert described [18] several varieties of emergence, including 'second order emergence,' in which agents themselves recognised emergent features of the society and this influenced their behaviour, while Antunes et al. [3] introduced a micro-level 'perception' of a macro-level measure as influencing individual agent's behaviour;
- Campos et al [10] enumerate seven roles for experimenters in a multi-agent simulation. Many before have argued the necessity of the 'tester' role being played by a different individual from the 'designer' or 'developer.' This set of roles does not stress this necessity, but goes far beyond in specialising the roles involved in experimentation.

4.3 Exploring Design Spaces

The notion of exploration of the design space against the niche space was introduced in MAS by Aaron Sloman [22, 23] to clarify how one can find a solution (architecture) for a particular problem. Stemming from broad but shallow agent architectures, designs are proposed and tested against original specifications, and finally, some *variations* introduced to check how the specific architecture adapted to the niche space it was developed for. In most MABS simulations reported in the literature, this last step is not performed, and again the reader is left with the notion that the way models were built was either the *only* or the *best* possible design. This brings us back to the concern about exemplification instead of demonstration.

However, the picture gets even darker when we consider not only agent design, but also experiment design. It could be said that we are exploring a multi-dimensional region using only two-dimensional tools. Any kind of variation could be introduced by considering any other relevant dimension, and we must possess the means with which to assess relevance of the features under exam and their consequences for the outcome of experiments.

5 The Purpose of Agent-Based Exploratory Simulation

The dramatic effect of considering ill, biased, or flawed methodological principles for complex simulations becomes apparent when we consider its possible purposes. Many of these are often only implicitly considered, so it is important to stress all of them here.

1. By building computational models, scientists are forced to *operationalise* the concepts and mechanisms they use for their formulations. This point is very important as we are in cross-cultural field, and terminology and approaches can differ a lot from one area to another;
2. The first and many times only purpose of many simulations is to get to *understand* better some complex phenomenon. In MABS, 'understand' means to describe, to model, to program, to manipulate, to explore, to have a hands-in approach to the definition of a phenomenon or process;

3. Another purpose of exploratory simulation is to *experiment* with the models, formulate conjectures, test theories, explore alternatives of design but also of definitions, rehearse different approaches to design, development, carry out explorations of different relevance of perceived features, compare consequences of possible designs, test different initial conditions and simulation parameters, explore 'what-if' alternatives. In sum, go beyond observed phenomena and established models, and play with the simulation while letting imagination run free;
4. With MABS, we ultimately aim to *explain* a given phenomenon, usually from the real social world. The sense of explaining is linked to causality more than to correlation. As Gilbert [18] says, we need explanation not only at the macro level, but also at the individual level. Our explanation of the phenomena we observe in simulation is solid because we must make the effort of creating and validating the mechanisms at the micro level, by providing solid and valid reasons for individual behaviours;
5. When we achieve such a level of understanding, we are able to *predict* how our models react to change, and this prediction is verifiable in the real phenomenon, through empirical observations. It is important to stress that even empirical observations presuppose a model (which data were collected, which questionnaires were used, etc.). A recent effort that may prove very useful in understanding the complexities of this process is the Model to Model workshop series [1, 2];
6. Finally, we have such confidence in the validity and prediction capability of our simulation system, that we are ready to help rehearse new policies and *prescribe* measures to be applied to the real phenomenon with real actors. It is obvious that no rigour can be spared when a simulation program achieves this point, and initial restrained application is highly recommended.

6 How to Conduct Agent-Based Exploratory Simulation

In the most interesting social simulations, agents are autonomous, in what individual agents have their own reasons for the choices they make and the behaviours they display. Simulations are hence run with a heterogeneous set of agents, closely resembling what happens in real social systems, where individuality and heterogeneity are key features. So, individual action is situated, adaptive, multi-dimensional, complex. If individual autonomy produces additional complexity in MAS, emergent, collective and global behaviour derived from the interactions of dozens of agents renders the whole outcome of simulations even more complex and unpredictable.

An important feature of social simulation is that usually researchers are not only concerned with the overall trajectories of the system, much less their aggregated evaluation (in terms of averages or other statistical measures). Equilibria, non-equilibria, phase transitions, attractors, etc. are as important as observing the individual trajectory of given agents, and examining its reasons and causes. This is important both to validate the model at individual and global levels, but also because the whole dynamics of the system and its components is influenced by the micro-macro link.

In e*plore, important phases are to determine what characteristics are important and what measures (values) are to be taken at both those levels, what is the appropriate design of individual cognitive apparatus and of inter-personal relationship channels (other methodologies such as Extended MAD or Gaia might prove useful for this), what roles the experiment designer will play and how his/her beliefs are represented inside the simulation, how to perform translation (specification, coding, validation, etc.) along the lines of a new (hyper-)triangle (much more complex than the one in figure 1) and complement it with complex dynamic evaluations, how to design models, agents, systems, experiments, simulations, in order to travel alongside the space of models to cover problem characteristics and to evaluate truthfulness of a certain agent design. All this while keeping in mind that we are looking for a solution for a problem in the real world.

6.1 Systematically Transversing Design Space

According to Gilbert [18], it was Epstein and Axtell [16] who pioneered the technique of starting a simple model and refining it. This can be considered an adaptation of Sloman's increasing depth in his broad but shallow agent models, but this time applied to the hole MAS and not only the individual agent. In this section we propose that when we need to explore the space of possible designs, several techniques can be used to ensure complete and comprehensive covering. We have been used these ideas in the tax compliance scenario [3, 4, 8], where we envisage to get a deeper insight into individual and collective behaviour involved in tax evasion and better support and confidence for our exploratory ideas.

While we propose the following techniques as a way of consecutively enriching and rehearsing new agent and societal models, we offer their application to the exploration of the design space of *experiments* themselves. Variations of the models involved in experiments depend on an amazing number of features to be repeatedly fixed and spanned over their domain. These include initial conditions, parameters, realist estimations of lacking numbers, etc., but we have to consider *higher order* decisions, such as mechanisms that can change/update/vary those parameters, and even interconnections among those mechanisms. All of these are design options to be made, and their validity must be strengthened by convenient exploration around them.

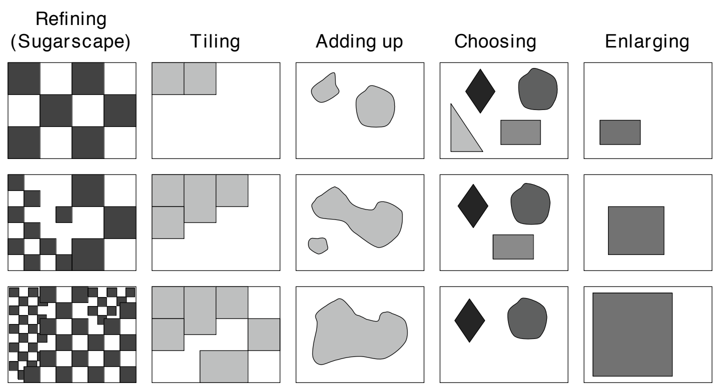

Figure 3: *Some techniques to cover design space.*

Figure 3 illustrates how a set of models can be designed and composed to comprehensively cover the space of possible designs. Models evolve from models by means of several different techniques and their combinations: refining, tiling, adding up, choosing, enlarging, etc. These are all standard techniques used in the development of models and systems. Exploration through these techniques involves moving from one model to another through introducing variability in the models characteristics, be them either parameters and variables, or objects, agents and environments, or social mechanisms (for interaction, protocols, dynamic structures), or even experiment design-related.

In a short explanation of these techniques, we will refer to the object of variation as a "mechanism." A mechanism can be simply seen as a variable that represents some concept, or a complex set of social rules that the model includes. So, a mechanism is not necessarily individual, and the variability we propose must be applied to all parts of design (individual agent, environment, interactions between agents, societal rules and even experiment design). So, *refining* involves substituting some simple mechanism for a slightly more complex one. *Tiling* means to explore some design alternative by covering the whole space of possibilities for a given mechanism. *Adding up* involves the summation of two or more models developped in parallel, and addressing different aspects of the target phenomenon. *Choosing* is the inverse of adding up, to give up some model or some characteristics of a model that do not seem promising for the overall solution. Enlarging means to augment a model by adding new features, and relating them to the existing ones.

The idea behind this strategic exploration of the experiment design space is to build up theory from the exploration of models. As an example, consider our experiments on tax compliance [3, 4, 8]. Existing theoretical models were plainly unsatisfactory. On the other hand, we had no solid empirical data with which to calibrate and ultimately validate our models. So, we opted for a strategy of mimicking the standard mainstream model (which we called Ec_0) with which we recorded a set of base data against which to compare the outcome of subsequent models. Then, we successively introduced new models with specific characteristics, either at the micro (individual) or at the macro (societal) levels, with some reasons, conjectures or intuitions.So, Ec_0^τ introduced expanded history in the individual decision; Ec_1 proposed agent *individuality*, whereas Ec_2 postulated individual *adaptivity*; Ec_3^* introduced sociality, it is the first model where the individual decision depends on a social perception; Ec_3^{*i} explored one particular type of interaction, imitation; and finally Ec_4^* postulated social heterogeneity, different agent breeds in a conflictual relation. Other models are still being shaped, such as $Ec_?^{*k}$ a model where perception is limited to a k-sized neighbourhood. This tentative coverage of our problem and model space uses several combined techniques of figure 3.

6.2 Deepening the design

When building up experimental designs, it is usual to defend and adopt the so-called KISS ("keep it simple, stupid!") principle [7]. In some sense, Sloman's "broad but shallow" design principle starts off from this principle. Still, models must never be simpler than they should. The solution for this tension is to take the shallow design and increasingly deepen (or thicken, as we proposed in Kyoto for WCSS'06) it while gaining insight and understanding about the problem at hand. The idea is to explore the design of agents, (interactions), (institutions), societies and finally experiments (including simulations and analysis of their outcomes) by making the initially simple (and simplistic) particular notion used increasingly more complex, dynamic, and rooted in consubstantiated facts. As Moss argued in his WCSS'06 plenary presentation, "Arbitrary assumptions must be relaxed in a way that reflects some evidence." This complex movement involves the experimenter him/herself, and according to Moss includes "qualitative micro validation and verification (V&V), numerical macro V&V, top-down verification, bottom-up validation," all of this whereas facing that "equation models are not possible, due to finite precision of computers."

A possible sequence of deepening a concept representing some agent feature (say parameter c, standing for honesty, income, or whatever) could be to consider it initially a constant, then a variable, then assign it some random distribution, then some empirically validated random distribution, then include a dedicated mechanism for calculating c, then an adaptive mechanism for calculating c, then to substitute c altogether for a mechanism, and so on and so forth. These sequence illustrates some of the combination of techniques depicted in figure 3.

7 e*plore v.0

We can synthesize the steps of e*plore methodology:

i. *identify the subject* to be investigated, by stating specific items, features or marks;

ii. *unveil state-of-the-art* across the several scientific areas involved to provide context. The idea is to enlarge coverage before narrowing the focus, to focus prematurely on solutions may prevent the in-depth understanding of problems;

iii. *propose definition* of the target phenomenon. Pay attention to its operationality;

iv. *identify relevant aspects* in the target phenomenon, in particular, *list individual and collective measures* with which to characterise it;

v. if available, *collect observations* of the relevant features and measures;

vi. *develop the appropriate models* to simulate the phenomenon. Use the features you uncovered and program adequate mechanisms for individual agents, for interactions among agents, for probing and observing the simulation. Be careful to base behaviours in reasons that can be supported on appropriate individual motivations. Develop visualisation and data recording tools. Document every design option thoroughly. *Run the simulations*, collect results, compute selected measures;

vii. return to step iii, and *calibrate everything*: your definition of the target, of adequate measures, of all the models, verify your designs, validate your models by using the selected measures. Watch individual trajectories of selected agents, as well as collective behaviours;

viii. *introduce variation* in your models: in initial conditions and parameters, in individual and collective mechanisms, in measures. Return to step v;

ix. After enough exploration of design space is performed, use your best models to *propose predictions*. Confirm it with past data, or collect data and validate predictions. Go back to the appropriate step to ensure rigour;

x. Make a generalisation effort and *propose theories and/or policies*. Apply to the target phenomenon. Watch global and individual behaviours. Recalibrate.

8 Concluding Remarks

When embracing a new project on the dynamics of tax evasion, we were struck by the difficulty in adequately designing the MAS models and simulation experiments in such a way that the results of our investigation could be reliable enough as to provide solid cues on how to act in the real world side of the problem. We crossed this concern with our old approaches to methodological principia to the design and deployment of MAS, to outline a set of steps that allow to think holistically and in a complex way about the carrying out of social simulation experiments.

The e*plore methodology goes beyond other proposals in MAS, because it takes a step back from the core of action, and looks at the experimentation process as a whole where the researcher has a role and intents. This is the reason why it starts from a broad, multi-disciplinary research on the issue to take on, and proposes a lot of cycles in the development process, to ensure not only verification and validation, but also comprehensive coverage of the experiment design space. This is accomplished through the use of several variation techniques, but its foundations lay on the researcher's experience, rigour and honesty, but also intuition and creativity. At this stage of our proposal, we cannot offer better guidance to transverse that space, since its cartography is not available, and its topology is too complex.

References

[1] Model to Model Workshop, March 31-April 1, 2003, Marseille, France. http://cfpm.org/m2m/.

[2] Second Model to Model Workshop, September 16-19, 2004, Valladolid, Spain. www.insisoc.org/ESSA04/M2M2.htm.

[3] Luis Antunes, João Balsa, Luis Moniz, Paulo Urbano, and Catarina Roseta Palma. Tax compliance in a simulated heterogeneous multi-agent society. In Jaime Simão Sichman and Luis Antunes, editors, *Multi-Agent-Based Simulation VI*, volume 3891 of *LNAI*. Springer-Verlag, 2006.

[4] Luis Antunes, João Balsa, Ana Respício, and Helder Coelho. Tactical exploration of tax compliance decisions in multi-agent based simulation. In Luis Antunes and Keiki Takadama, editors, *Proc. MABS 2006*, 2006.

[5] Luis Antunes and Helder Coelho. On how to conduct experiments with self-motivated agents. In Gabriela Lindemann, Daniel Moldt, and Mario Paolucci, editors, *Regulated Agent-Based Social Systems: First International Workshop, RASTA 2002*, volume 2934 of *LNAI*. Springer-Verlag, 2004.

[6] Luis Antunes, João Faria, and Helder Coelho. Improving choice mechanisms within the BVG architecture. In *Intelligent Agents VII, Proc. of ATAL 2000*, volume 1986 of *LNAI*. Springer-Verlag, 2001.

[7] Robert Axelrod. Advancing the art of simulation in the social sciences. In Rosaria Conte, Rainer Hegselmann, and Pietro Terna, editors, *Simulating Social Phenomena*, volume 456 of *LNEMS*. Springer, 1997.

[8] João Balsa, Luis Antunes, Ana Respício, and Helder Coelho. Autonomous inspectors in tax compliance simulation. In *Proc. 18th European Meeting on Cybernetics and Systems Research*, 2006.

[9] Federico Bergenti, Marie-Pierre Gleizes, and Franco Zambonelli, editors. *Methodologies and Software Engineering for Agent Systems: The Agent-Oriented Software Engineering Handbook*. Kluwer Ac. Press, 2004.

[10] André M. C. Campos, Anne M. P. Canuto, and Jorge H. C. Fernandes. Towards a methodology for developing agent-based simulations: The masim methodology. In *Proc. AAMAS 2004*, pages 1494–1495, 2004.

[11] John L. Casti. Would-be business worlds. *Complexity*, 6(2), 2001.

[12] Helder Coelho, Luis Antunes, and Luis Moniz. On agent design rationale. In *Proc. XI Brazilian Symposium on AI*. SBC and LIA, 1994.

[13] Paul R. Cohen. A Survey of the Eighth National Conference on AI: Pulling together or pulling apart? *AI Magazine*, 12(1):16–41, 1991.

[14] Rosaria Conte and Nigel Gilbert. Introduction: computer simulation for social theory. In *Artificial Societies: the computer simulation of social life*. UCL Press, 1995.

[15] Rosaria Conte, Rainer Hegselmann, and Pietro Terna. Introduction: Social simulation – a new disciplinary synthesis. In *Simulating Social Phenomena*, volume 456 of *LNEMS*. Springer, 1997.

[16] Joshua M. Epstein and Robert Axtell. *Growing artificial societies*. The Brookings Institution and The MIT Press, Washington, D.C. and Cambridge, MA (resp.), 1996.

[17] Nigel Gilbert. Models, processes and algorithms: Towards a simulation toolkit. In Ramzi Suleiman, Klaus G. Troitzsch, and Nigel Gilbert, editors, *Tools and Techniques for Social Science Simulation*. Physica-Verlag, Heidelberg, 2000.

[18] Nigel Gilbert. Varieties of emergence. In *Proc. Agent 2002: Social agents: ecology, exchange, and evolution*, Chicago, 2002.

[19] Nigel Gilbert and Jim Doran, editors. *Simulating Societies: the computer simulation of social phenomena*. UCL Press, London, 1994.

[20] David Hales. *Tag Based Co-operation in Artificial Societies*. PhD thesis, Univ. Essex, 2001.

[21] Steve Hanks, Martha E. Pollack, and Paul R. Cohen. Benchmarks, test beds, controlled experimentation, and the design of agent architectures. *AI Magazine*, 14(4), Winter 1993.

[22] Aaron Sloman. Prospects for AI as the general science of intelligence. In *Proc. of AISB'93*. IOS Press, 1993.

[23] Aaron Sloman. Explorations in design space. In *Proc. of the 11th European Conference on Artificial Intelligence*, 1994.

[24] William R. Swartout and Robert Balzer. On the inevitable interwining of specification and implementation. *Communications of ACM*, 25(7):438–440, 1982.

[25] Michael Wooldridge, Nicholas R. Jennings, and David Kinny. The GAIA methodology for agent-oriented analysis and design. *Journal of Autonomous Agents and Multi-Agent Systems*, 3(3), 2000.