has to confront in assessing public policy toward globalization and aid to developing countries."

**Peter Dietsch**
***Université de Montréal***

**REFERENCE**

Wolff, E. N. 2006. Changes in household wealth in the 1980s and 1990s in the U.S. In *International Perspectives on Household Wealth*, ed. Edward N. Wolff. Northampton, MA: Edward Elgar.

doi:10.1017/S0266267108002320

*The Structural Evolution of Morality*, Jason McKenzie Alexander. Cambridge University Press, 2007, ix + 300 pages.

This book draws together and expands upon Jason McKenzie Alexander's previously published work using evolutionary game theory. His central claim is that our moral principles act as heuristic devices which tend to maximize our expected utility over a lifetime. The models he presents aim to demonstrate that a range of simple moral principles often emerge as the victors over other strategies in a process of cultural evolution. What is relatively novel about Alexander's modelling, compared with the present philosophical literature, is that it takes seriously the fact that we interact with each other in a structured environment with non-random interactions. Another philosopher notable for making moves in this direction is Brian Skyrms and this book is certainly a must-read for anyone who has been enthused by his recent collections of evolutionary modelling (Skyrms, 2004). Skyrms' works have raised philosophical concerns for some, which *The Structural Evolution of Morality* makes efforts to address. Before considering these concerns I will outline the book's structure and highlight, by way of example, some results that I found particularly novel or thought provoking.

The majority of the text is devoted to modelling four game types using each of four forms of social network. Each game is taken to represent a canonical type of moral dilemma:

- Prisoner's Dilemma (PD): The problem of cooperation where interests partially diverge.
- The Stag Hunt: Trusting someone to play their part in a joint venture.
- Divide the Dollar/Cake: The problem of fair division of resources.
- The Ultimatum Bargaining Game: The phenomenon of moral retribution.

https://doi.org/10.1017/S0266267108002320 Published online by Cambridge University Press

Before reaching the modelling stage, however, Alexander provides two introductory chapters. Chapter 1 offers a strong defence of the claim that humans are boundedly rational and that classical game theory, based as it is on rational choice theory, is the wrong tool to use when modelling human behaviour. The bounded rationality thesis claims that humans are not cognitively sophisticated enough to meet the strong rationality strictures of classical game theory. In defending this view, Alexander provides expositions of the assumptions and foundations of both classical and evolutionary game theory. These accounts are both concise and accessible to newcomers to the field. Note that his preface grants mathematical sophistication as a prerequisite for the reader. I tend to disagree: those with little mathematical background (perhaps only elementary algebra) should be able to follow the main arguments and gain useful insights from both the introductory chapters and the modelling results that follow. Chapter 1 also offers a clear introduction to Gert Gigerenzer *et al*.'s "simple heuristics" approach to bounded rationality (Gigerenzer *et al*., 1999). Their work argues that humans employ fast and frugal rules-of-thumb to deal with decision making problems and the central claim of Alexander's work is that moral principles act in such a capacity.

Chapter 2 introduces the reader to the model types that feature in the remainder of the book. Firstly, we have the *replicator dynamics*, originally derived from the biological theory of natural selection. These dynamics model an unstructured population and have already been used to study the evolution of strategies in most of the games that Alexander considers. His point in including this model is to compare the results it generates with those from four agent-based models. Alexander argues convincingly that the replicator dynamics' assumptions of infinite populations and equiprobable pairwise interactions are unrealistic and that agent-based models move in the direction of increased realism. Happily for his programme, it turns out that, in most cases, these models yield results more favourable to the evolution of moral behaviour than do those using the replicator dynamics.

I now present the basics of the four agent-based models in some detail for two reasons: firstly, to give the reader a sense of their degree (or lack) of realism; secondly, I go on to discuss several results that depend on the details of the models in question.

In all agent-based models, each member of the finite population occupies a node in a network of connections that represent social relations which, in turn, allows us to define two neighbourhoods for each agent. The first is an *interaction neighbourhood*; the set of agents with whom the specified agent plays the game in question. The second is an *update neighbourhood*; the set of agents considered by the specified agent when she decides whether, and how, to update her strategy. As we will see, the

https://doi.org/10.1017/S0266267108002320 Published online by Cambridge University Press

outcome of a model can vary dramatically depending on the size of these two neighbourhoods.

Regarding updating rules, there are many possibilities but these models concentrate on two. *Imitate-the-best* updates by following the strategy used by the most successful agent in the neighbourhood (including herself). *Best-response* moves in the direction of more sophisticated strategic behaviour and updates with the response that would provide the best aggregate result assuming that agents in the interaction neighbourhood will stick with their current strategy.

Finally, there are the four model social networks. Since in subsequent chapters each game type is modelled using all four networks, I introduce each one in turn.

*Lattice models*: A one-dimensional lattice is a world where each agent is connected to two neighbours to form a chain which is typically wrapped into a ring. The interaction neighbourhood is usually taken to have radius 1 but the update radius can vary. A two-dimensional lattice is a world on a chess board (which can wrap into a torus). For those familiar with them, the neighbourhood regions used are the Von Neumann (4) (N, E, S & W on the board), the Moore (8) (N, NE, . . ., W & NW) or the Moore (24) neighbours. These networks reflect the fact that our social relations are often highly correlated with spatial proximity.

*Small-world networks*: These acknowledge that, while we have relationships principally with those spatially close to us, there are some individuals who have long-distance connections. This results in the well-known phenomenon of strangers often finding they have a friend-of-a-friend-of-a-friend . . . in common. In subsequent chapters the small-worlds Alexander uses are basic ring lattices with a small number of "bridge" edges; dramatically reducing the mean number of connections between spatially distant individuals.

*Bounded-degree networks*: These are randomly connected networks apart from the fact that the number (degree) of connections for each agent is constrained to fall between a maximum and minimum value. It is plausible that some social structures exhibit such irregularity compared with lattices.

*Dynamic networks*: Each agent in this network is connected to all other agents but the "weight" of these connections can vary and determine the interaction probabilities between agents. These models incorporate the phenomenon of reinforcement learning as weights are increased depending on the size of payoff received from that interaction. Hence agents gradually tend to interact with individuals with whom they have had successful relationships in the past.

Having introduced these network types, Alexander moves on to presenting the results of his extensive modelling. It is impossible to do full justice here to their range and subtlety. Quite rightly, Alexander

https://doi.org/10.1017/S0266267108002320 Published online by Cambridge University Press

demonstrates the robustness of his results by varying the free parameters in the models in multiple ways. At times this leads to a dizzying plethora of variations but he is absolutely right to take the reader through them. We need to be reassured that the results are not merely artefacts of his choice of parameter values in a particular case. I will limit myself to discussing a few results which stood out as being of particular interest and also provide a taste of the kind of results Alexander claims to have deep connections with our moral principles. No doubt others would be able to make an equally long, and different, list. Alexander's website at http://evolve.lse.ac.uk/compass/ allows one to play for oneself with some of these structured environments.

Chapter 3: *Cooperation* (The prisoner's dilemma, PD). As this game is the most renowned one in game theoretic literature I take it that the readers of this journal are familiar with the problem of the non-equilibrium nature of mutual cooperation. It has previously been demonstrated that playing the one-shot PD on a lattice can help the evolution of cooperation and Alexander's results in Chapter 3 reinforce this conclusion. What I take to be his most important contribution here is found in one brief section. It is well known that in the *repeated* PD cooperative behaviour can be both rational and the product of an evolutionary process. This fact was popularized by Robert Axelrod's important work that hailed tit-for-tat (TFT) as the victor in two evolutionary "tournaments" (Axelrod, 1984). Alexander contributes, in this chapter, to what Ken Binmore has called "breaking the tit-for-tat bubble" (Binmore, 1994: sec. 3.2.5). It is important to stress that TFT is one of many possible successful strategies in the indefinitely repeated PD and not all of those share Axelrod's features of being "nice, retaliatory, forgiving and clear" (Axelrod, 1984: 54). Chapter 3 presents a model where "*none* of the surviving strategies are even 'tit-for-tattish'" (p. 63, Italics in the original).

Chapter 4: *Trust* (The Stag Hunt). Here we find a welcome continuation of a trend that moves away from the PD being the dominant model for the state of nature and the development of the social contract (Skyrms, 2004). In the stag hunt there are two equilibria, both hunt hare and (the optimal) both hunt stag and so we face an equilibrium selection problem. Here the clearly introduced concept of risk dominance plays a central role. Some interesting results undermine the practical importance of *stochastic stability* proofs. Where there is a unique stochastically stable equilibrium in an evolutionary game played with random mutations then the population will, in the *ultra long-run*, spend approximately all of its time at that equilibrium. With the right payoff values and best-response learning on a 2-D lattice, All Hunt Stag is such an equilibrium. However, section 4.2 demonstrates that in such a case even 100 000 000 generations will not shift the population significantly away from a sub-optimal initial state of All Hunt Hare (p. 126).

https://doi.org/10.1017/S0266267108002320 Published online by Cambridge University Press

The more general and slightly unsettling theme of this chapter is that the relative sizes of the payoffs, whilst holding their ordering constant, can dramatically affect the outcome of the evolutionary process. I say "unsettling" because a result that holds for all stag hunts would make life simpler in the state of nature. Alexander's heuristics story now requires us to have reasonable sensitivity not just to the ordering of payoffs, but to their actual values and hence be in a position to conditionalize our deployment of a "trust" principle.

Chapter 5: *Fairness* (Divide the dollar). Also referred to as the Nash bargaining game, two players are presented with a dollar or cake and can demand a proportion of it. If their demands add up to less than the full cake then both get what they want and the rest is wasted. If their demands exceed the full cake then no one gets anything. Brian Skyrms uses this game in his own investigation of fairness norms (Skyrms, 2004: Ch. 2). Although its representativeness as a model of a typical distribution problem has been questioned I do not want to enter that debate here but instead highlight the effect of structuring the population (Kitcher, 1999: 223).

The typically acknowledged fair outcome is a 50–50 split and this is an *evolutionarily stable strategy* (ESS), that is, if everyone demands half then no other strategy can gain a foothold in the population. Unfortunately, under the replicator dynamics, there are other ESSs. These are polymorphisms where two demands are present in the population that, when paired, sum to the whole cake. In contrast, Alexander demonstrates that these polymorphisms can occur in regions of a structured world but that there is a strong tendency for them to be driven out by the fair demand (sec. 5.2). The root of the problem for the polymorphism is that, where "greedy" players are adjacent to each other they do badly when paired together.

Such a problem points the way towards a more complicated model that Alexander does not pursue but which I note here, hopefully, to demonstrate the rich possibilities for future work opened up by *The Structural Evolution of Morality*. If a game is asymmetric in the sense that the two players are assigned roles, then a famous result from Reinhard Selten demonstrates that polymorphisms cannot form an ESS (Selten, 1980). It would be interesting to know the effect on a structured model if players are randomly assigned one of two roles and can conditionalize their demand based on that assignment. A player that is greedy in one role but modest in the other will not suffer when paired with itself and the mean outcome will be a 50–50 split.

Another interesting result in Chapter 5 lends support to the bounded rationality thesis. If the game is played on a lattice with the best response learning rule then fair division can fail to overtake the population (p. 181). As in the PD, we have a case where more strategically sophisticated

https://doi.org/10.1017/S0266267108002320 Published online by Cambridge University Press

behaviour yields a sub-optimal outcome (cooperation does not evolve for the PD on a 2-D lattice with best response dynamics). Such results provide a positive reason for the use of simple imitation heuristics rather than the negative claim that features in Chapter 1 of *The Structural Evolution of Morality* (Gigerenzer *et al*., 1999: Ch. 1) stress the advantages of heuristics over perfect deliberation in their work).

Chapter 6: *Retribution* (The Ultimatum Game). Once again, a cake is divided but this time player one makes an offer to player two who can accept it and let player one have the rest, or reject it so that both receive nothing. Rational choice theory recommends as an equilibrium a minimum offer and an acceptance of any non-zero offer. Such behaviour seems to contradict a fairness norm of offering half and the retribution of rejecting offers of less than half. However, Alexander acknowledges that results from experimental economics find no conclusive cross-cultural norm (p. 236). He makes a virtue of such inconclusive findings by claiming that they are, to some extent, in keeping with his own inconclusive modelling results. Structured environments do less to promote fairness here than in the case of divide the dollar.

A specific way in which this game bucks trends from previous chapters is the effect of increasing the update neighbourhood. In the ultimatum game such an increase can make fair offers less attractive (p. 210). In contrast, in the PD and stag hunt, a larger region for imitation aids the evolution of cooperation and hunt stag (pp. 72, 118). Alexander's aim in this book is not to draw practical conclusions but the update radius effect jumps out as one that others should explore further. In a modern mass media society, interaction neighbourhoods may have widened but the range of individuals, real or fictional, that we observe and could potentially imitate is even larger. Whether these models' early results have any consequence for public policy is difficult to tell at this stage but presumably a long-term goal of any study of social structure might be the manipulation of that structure in order to promote prosocial behaviour.

The final chapter of modelling is admitted by Alexander to be an underdeveloped survey of some multi-player interactions in a structured environment. His tentative conclusion is that such interactions are less conducive to the evolution of moral behaviour than the two-player games. The biggest contribution of this chapter is clearly its provision of a rich supply of inspiration for future research.

"Philosophical Reflections" is the title of Alexander's final chapter. Someone critical of what we have been presented with up to this point would be justified in asking; all very interesting, but where's the morality? There is a normative dimension to the story: we should follow these principles because then we do well for ourselves but where is the account of the *moral* "should"? Alexander anticipates these criticisms and frames the debate via the concepts of *thin* moral behaviour and *thick* moral

https://doi.org/10.1017/S0266267108002320 Published online by Cambridge University Press

action. Thinly moral behaviour is in accordance with moral principles, but the agent lacks "sufficiently many of the right beliefs, intentions, preferences and desires" (p. 268). Thick moral action occurs where the appropriate behaviour is present and the agent does not suffer from those deficits. Alexander's immediate project is an account of the evolution of thin morality and there will be some who feel that this is sufficient. Alexander does not subscribe to that view but he is happy here to limit himself to brief "gestures towards a solution" (sec. 8.2). He acknowledges that he does not provide an account of our rich moral psychology but points towards evolutionary psychology as the place to look to enrich a "science of morality" (p. 281). Alexander looks forward to an account of our moral sentiments by an evolutionarily (both cultural and biological) informed experimental psychology that will investigate the "morally relevant emotions, moods, interpersonal affective stances and attitudes (that) motivate us to act" (p. 275).

How convinced the reader is by this promissory note will depend on their attitude to evolutionary psychology, and *The Structural Evolution of Morality* would need to be twice as long to engage fully with that debate. In broad terms, those attracted to a Humean moral theory are likely to be sympathetic to Alexander's philosophical reflections. As I noted at the beginning of the review, this work is certain to be invaluable to anyone following the road laid by Brian Skyrms. Additionally, I recommend this book to many of those unconvinced about the possibility of making the leap from an explanation of thin morality to a thick account. Even with its more speculative moral musings removed, and a retitling as *The Structural Evolution of Social Behaviour*, this work would still make an important contribution to political philosophy and to the social sciences more broadly.

**Jonathan Grose**
***University of Bristol and University of Gloucestershire***

**REFERENCES**

Axelrod, R. 1984. *The evolution of co-operation*. New York: Basic Books.
Binmore, K. 1994. *Game theory and the social contract. Volume I: Playing fair.* Cambridge, MA: MIT Press.
Gigerenzer, G., P. Todd and the ABC Research Group. 1999. *Simple heuristics that make us smart*. Oxford: Oxford University Press.
Kitcher, P. 1999. Games social animals play: Commentary on Brian Skyrms's Evolution of the Social Contract. *Philosophy and Phenomenological Research* 59: 221–8.
Selten, R. 1980. A note on evolutionarily stable strategies in asymmetric animal conflicts. *Journal of Theoretical Biology* 84: 93–101.
Skyrms, B. 2004. *The stag hunt and the evolution of social structure*. Cambridge: Cambridge University Press.

https://doi.org/10.1017/S0266267108002320 Published online by Cambridge University Press