Book Reviews

Brian Skyrms, **The Stag Hunt and the Evolution of Social Structure**, Cambridge University Press, Cambridge, UK., 2003, US$ 55, pp. 166, ISBN 0-521-82651-9, hardcover. Dimensions (in inches): $0.8 \times 5.5 \times 8.5$.

1. Overview of the book

When von Neumann and Morgenstern invented their theory of interacting rational decisionmaking, game theory promised to be the formal tool required to make the social sciences more precise and predictive and so to reduce the gap between the social and natural sciences. Although game theory has indeed received a central place within economics, in most other social sciences game theory never acquired such a relevant position. This is unfortunate, because game theory already proved to be useful for a better understanding of e.g. conventions and the social contract (cf. Lewis, 1969). But it is understandable as well, because if based on Bayesian decision theory, standard game theory has to make unnaturally strong presuppositions concerning rationality and common knowledge.

Brian Skyrms' 'The Stag Hunt and the Evolution of Social Structure' is the third book where the author argues that patterns of coordinated behavior can best be explained from an adaptive dynamic perspective where the agents are only boundedly rational. In the first book, 'The Dynamics of Rational Deliberation', he focussed on explaining the evolution and stability of equilibria in terms of dynamic deliberation. Here, rationality and common knowledge still play a role, although much less of it is required to explain game theoretical equilibria concepts than in standard motivations. In his more recent book 'Evolution of the Social Contract' as well as in the book reviewed here, Skyrms adopts a more radical position by making use of Evolutionary Game Theory (EGT). EGT was developed by the biologist John Maynard Smith as a formalization of the neo-Darwinian concept of evolution via natural selection where rationality and common knowledge are virtually irrelevant. More recently, EGT has become important in economics and other social sciences as well. It was realized that if we let the dynamic process that lies behind the notion of evolutionary stability guide by imitation or simple reinforcement learning, EGT provides a useful tool to study cultural evolution. For instance, for the study of the evolution of individual habits that lead to social contracts, conventions, and other social institutions that are useful, even if seen from a completely individualistic point of view.

The main problem to explain these institutions is to give a convincing story of why agents will coordinate their behavior so as to establish a collective good. The choice situation that is normally used to discuss this problem is the well-known prisoner's dilemma. In such a situation, the adoption by each agent of the strategy that seems individually best for all of them leads to an outcome that all agents agree is dispreferred to one specific other outcome. But perhaps the importance of the

Studia Logica (2007) 85: 133–138
DOI: 10.1007/s11225-007-9024-2

prisoner's dilemma to explain social institutions is overrated. To explain the concept of asymmetric institutions such as possession (some have it, others have not), for instance, Sugden (1986) proposed to focus on the game of Chicken, also known as the Hawk-Dove game. To give a neo-Hobbesian theory of the social contract, Binmore (1994) suggested to think of situations where the state where everybody cooperates and the one where nobody cooperates are both game-theoretical equilibria. Although the prisoner's dilemma is not, there are many situations of this type. The most interesting ones, however, are those symmetric games where, in some sense, the latter non-cooperative equilibrium is preferred by individualistic rationality. The Stag hunt game – described informally by philosophers as Rousseau and Hume – is exactly such a situation, and it is this type of game that Skyrms focusses on in his new work.

The Stag hunt game, as stated in the left-hand table below, is a simple two-player symmetric game with two strict equilibria: both hunting *S*tag, $\langle S, S\rangle$, or both hunting *H*are, $\langle H, H\rangle$. The first equilibrium gives the highest payoff to both, i.e., is payoff-dominant (or Pareto optimal), because it gives to both a utility of, let us say, 6, while the second equilibrium yields only one of 4. However, assume that if one hunts Stag but the other Hare, the payoff is (4,0) in 'favor' of the Hare-hunter. In that case, the payoff-dominated equilibrium where both are hunting Hare still has something in its favor: if one player is equally likely to play either strategy, the expected utility of hunting Hare for the other is optimal. Harsanyi and Selten (1988) call the equilibrium $\langle H, H\rangle$ *risk-dominant*, because in the absence of further information what the other player will do, it is safer to hunt Hare instead of Stag. The expected utilities of performing the acts in the Stag hunt game before deliberation are given in the right-hand table below:

Stag hunt	S	H
S	$6, 6$	$0, 4$
H	$4, 0$	$4, 4$

expected utilities	S	H
S	$3, 3$	$3, 4$
H	$4, 3$	$4, 4$

In the first chapter of the book, Skryms describes the Stag hunt and shows that the problem of reforming the social contract can be stated appropriately by this game. We can take the 'state of nature' to be the risk-dominant equilibrium and the social contract as the risky but rewarding Stag-hunt equilibrium. As another motivation for looking at the Stag hunt, Skyrms shows that when future plays of the game or group selection are taken into account, the problem of cooperation in the prisoner's dilemma is transformed into the problem of cooperation in the Stag hunt. From the perspective of rational choice-based game theory it is easy enough for any Stag hunt-type of game to calculate what each agent's expectations have to be concerning the other agent's behavior in order to predict that equilibrium $\langle S, S\rangle$ will in fact be selected. Skyrms, however, concentrates on the problem of how to move from the $\langle H, H\rangle$ equilibrium to the Stag hunt equilibrium, and here a dynamic, or evolutionary perspective on games is required.

Evolutionary game theory predicts that all and only all strict Nash equilibria will be evolutionary stable. This is an appealing result for those coordination games

where all strict equilibria have the same payoff, because then it is appropriately predicted that the equilibrium that in fact evolves is just a matter of chance. When some strict equilibria are better than others, however – as in the Stag hunt –, we want something more than pure chance, which is possible by (slightly) changing the evolutionary dynamics. Skyrms discusses four modifications of standard EGT as a result of which evolution favors some strict equilibria above others. The first way to modify standard EGT is to add to the process of reproduction a *stochastic* process of *mutation*, or imperfect learning.[1] In general, stochastic EGT favors those strict equilibria which have the greatest basin of attraction. For the simplest coordination games these are the payoff dominant equilibria, but not so for the Stag hunt game. This might be one reason why Skyrms is short on Stochastic EGT, on top of his valid complaint that it accounts for equilibrium selection only at the *very* long run. The other three modifications of standard EGT are discussed in much more detail, and all involve a form of *correlation*: the assumption that agents are not simply paired randomly with each other to play games as is assumed in standard EGT.

The first way to build correlation into the dynamic adaptive process is to assume that players are organized in a spatial structure and interact only with their immediate neighbors. This is discussed in chapters 2 and 3 of the book. Discussing a bargaining game, the second chapter suggests that interaction with neighbors enforces fair division of goods, a form of cooperation. But in the third chapter Skyrms shows by means of computer simulations that things are not that simple: whether cooperation comes about in dynamic systems using spatial structure depends on how many neighbors one has and on the exact sort of adaptive dynamics that is assumed. The main conclusion of this chapter is that local interaction tends to favor cooperative outcomes, but also that coordination is more easily established when agents are located on a two-dimensional lattice than when they are located on a circle: in the former case they each have 8 neighbors, in the latter only 2.

The most straightforward way to enforce cooperation in coordination games is by signaling before the actual play of the game which action one is going to perform. But of course, this works only when the sender has an incentive to cooperate, and when the signal sent is a reliable indicator of the sender's type. In the second part of Skyrms' book he discusses how signals can become informative, and under which circumstances the sending of signals might have an effect on the adaptive dynamic process. In the fourth chapter it is argued that taking an evolutionary perspective on the well-known Lewisian signaling games is useful, not only because the equilibrium concept used here fits better with the notion of equilibrium chosen by Lewis himself to single out linguistic conventions than the standard equilibrium notion of classical game theory, but mainly because it shows that one does not have to make strong assumptions concerning rationality and common knowledge to explain the stability, but also the emergence of linguistic conventions and a corresponding notion of inference. Sending costless signals is useful in the simplest kind of coordination games, but standard game theory predicts that signaling is useless and without

[1] Technically, this results if we give up the assumption of standard EGT that populations of players are infinite.

any effect – i.e., is just cheap talk – in games where the preferences are not so well allinged, but where coordination still seems useful: in the prisoner's dilemma, in assurance versions of the Stag hunt, and in bargaining games. Still, intuition and experiments suggest that agents also signal costless messages in these less ideal circumstances to enforce coordination. In the for me most surprising chapter of this book, Skyrms shows that this can be explained from an evolutionary point of view. In the prisoner's dilemma, signals can be used to destabilize the non-cooperative equilibrium; in the Stag hunt, signaling gives rise to a new type of equilibrium, while in bargaining games signaling extents the basin of attraction of fair division. Whereas according to standard game theory costless signaling can be effective only in as far as the preferences of the agents involved are aligned, Skyrms shows that our intuition that signaling is not limited to these circumstances can be explained when we take the evolutionary dynamics behind the notion of stability seriously.

In standard dynamic systems, as well as in ones where players are spatially organized, it is assumed that it is fixed in advance who plays with who, and it is studied how the strategies chosen by the agents might change. In chapter 5 Skyrms shows that signaling one's future play of the game might change the interaction structure. In the third and final section of the book, Skyrms focusses his attention more radically to this latter type of evolution. Based on joint work with Robin Pemantle, he shows in chapter 6 that already by simple reinforcement learning one can end up playing only with partners such that cooperative behavior is beneficial. It is known already that reinforcement learning – or learning by association – and evolutionary game theory have a great deal in common: they both can explain self-emerging complex structure, and the paths through which these stable structures are reached are very similar. Even so, Skyrms' discussion is very illuminating. He discusses several surprising consequences that follow when contrasting forms of reinforcement are considered. For instance, one-sided versus interactive reinforcement; positive versus negative reinforcement; reinforcement with perfect versus imperfect memory, and payoff-neutral versus payoff-dependent reinforcent. Especially payoff-dependent reinforcement is elaborated on, concentrating most on the Stag hunt game. It is shown, for example, that when the reinforcement dynamics is driven by the payoffs of the Stag hunt game, this greatly improves the chances that cooperative Stag hunters pair only with each other, and will thus flourish. In the final chapter of the book, Skyrms discusses the interplay between the evolution of strategies and of interaction structure, and even examines the influence of various ways one might change one's strategy (by imitation or by best response) on this interplay. He does so by looking at various well-known type of games (including the Stag hunt and the prisoner's dilemma), and discusses various surprising consequences.

2. Evaluation

To those who are a bit like me and love books in which technical concepts from disciplines as diverse as evolutionary biology, philosophy, and economics are combined but explained in a very lucid way, without ever giving up the required nuance, I

can unreservedly recommend this delightful book. It is an absolute pleasure to read and a more stimulating introductory text on evolutionary thinking of social institutions is hard to imagine. Just like in its successful predecessor and award-winning 'Evolution of the Social Contract', Brian Skyrms argues in this book forcefully in favor of an evolutionary perspective on social institutions. He does so by combining theoretical insights with easy to imitate computer simulations of evolutionary games. Although the book deals with well-known problems and makes use of sometimes common techniques, Skyrms always manages to come up with illuminating new ways to think of things, and of fascinating new insights. Perhaps the last three chapters of this book are most surprising in this regard where he shows (i) that in an evolutionary setting effective costless signaling is not limited to pure coordination games, and (ii) that adaptive dynamic models can be used not only to explain the evolution of strategies, but also the evolution of interaction structures. As such, I have no doubt that the book is very useful not only for the layman, but also for social philosophers, economists, and evolutionary biologists.

The book is also very useful for anyone who just like me is interested in the evolution and change of language. Skyrms discusses several techniques that all seem to play a role here: the dynamic perspective on Lewisian signaling games can explain the evolution of linguistic conventions; stochastic EGT might be a fruitful tool to account for linguistic universals; the effect of spatial clustering and local interaction can help to clarify linguistic diversification and the emergence of dialects, while the evolutionary perspective on reinforcement might illuminate the process of grammaticalization: the evolutionary path where frequently used lexical items (like nouns and verbs) change into grammatical features (like auxiliary verbs or pronouns). Finally, I believe that the focus on the Stag hunt game can not only function as a useful antidote to the usual emphasis on the prisoner's dilemma in social philosophy, but also might help some to realize that it describes a situation very relevant to linguistic concerns as well. Brian Skyrms has done us a great service, and I am looking forward to his next book.

References

[1] Binmore, K. (1994), *Playing fair*, MIT Press, Cambridge, MA.

[2] Harsanyi, J. C. and R. Selten (1988), *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, MA.

[3] Lewis, D. (1969), *Convention* Harvard University Press, Cambridge, MA.

[4] Maynard Smith, J. (1982), *Evolution and the Theory of Games*, Cambridge University Press, New York.

[5] Neuman, J. von, and O. Morgenstern (1947), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.

[6] Skyrms, B. (1990), *The Dynamics of Rational Deliberation*, Harvard University Press, Cambridge, MA.

[7] Skyrms, B. (1996), *Evolution of the Social Contract*, Cambridge University Press, Cambridge.

[8] Skyrms, B.(2004). *The Stag Hunt and the Evolution of Social Structure*, Cambridge University Press, Cambridge.

[9] Sugden, R. (1986), *The Economics of Rights, Co-operation and Welfare*, Basil Blackwell, Oxford/New York.

Robert van Rooij
Institute for Logic, Language and Computation
University of Amsterdam
R.A.M.vanRooij@uva.nl