Interim Report **IR-02-019**

William D. Hamilton's Work in Evolutionary Game Theory

Karl Sigmund (karl.sigmund@univie.ac.at)

Approved by

Abstract

William D. Hamilton's impact on contemporary views of evolution has been nothing short of revolutionary, and his contributions to our reasoning on kin selection, genomic conflicts, parasitism and costs of sex dominate these fields. He also had an important, if somewhat less obvious, influence on evolutionary game theory. This field, which by now is well-established both in biology and economics, is usually (and rightly) attributed to John Maynard Smith: the canonical references are his brief joint paper with George Price, and his book 'Evolution and the Theory of Games'. But W.D. Hamilton played an important, and indeed pioneering role in the development of this field.

About the Author

Karl Sigmund

Institute of Mathematics, University of Vienna

Strudlhofgasse 4, A-1090 Vienna, Austria

and

Adaptive Dynamics Network

International Institute for Applied Systems Analysis

A-2361 Laxenburg, Austria

Wiliam D. Hamilton's Work in Evolutionary Game Theory

Karl Sigmund

William D. Hamilton's impact on contemporary views of evolution has been nothing short of revolutionary, and his contributions to our reasoning on kin selection, genomic conflicts, parasitism and costs of sex dominate these fields. He also had an important, if somewhat less obvious, influence on evolutionary game theory. This field, which by now is well-established both in biology and economics, is usually (and rightly) attributed to John Maynard Smith: the canonical references are his brief joint paper with George Price (Maynard Smith and Price, 1973), and his book 'Evolution and the Theory of Games'(Maynard Smith, 1982). But W.D. Hamilton played an important, and indeed pioneering role in the development of this field.

In the early sixties, game theory was looking back on more than a decade of heady progress. The seminal book on 'The theory of Games and Economic Behaviour' (von Neumann and Morgenstern, 1944) had inspired some of the best mathematicians of the time. It had also fired public imagination and been subject to some of the hype later lavished on catastrophes and chaos. The response of economists, to whom game theory was offered as a tool, turned out to be somewhat more guarded. Bill Hamilton had encountered the book 'in idle reading' (as he later wrote, Hamilton 1996) while he was a student of genetics, and returned to it 'almost as a leisure distraction' when he had landed at the London School of Economics for his doctoral thesis (officially in human demography). He later wrote that 'the idea of fitness being treated as equivalent to 'payoff' in a biological version of von Neumann's game theory had crossed my mind as soon as I read his earliest account'. Hamilton had read also the standard textbook of the time (Luce and Raiffa 1957), which included the concept of a Nash equilibrium and a discussion of the Prisoner's Dilemma.

At that time, there had been several attempts to use game theory in evolutionary biology, most notably a paper by Lewontin (1960) which interpreted life-history traits as strategies and reproductive success as payoff function, and which showed that the optimal solution of a certain game 'against nature' was a mixed strategy. Another forerunner was Verner (1965), who treated sex-ratios as strategies and aimed at showing that in the presence of fluctuations in the overall sex-ratio, an individual's 1:1 sex-ratio was optimal. These early papers did not yet apply game theory to model competition within a population, and did not even implicitly consider population dynamics.

Decades earlier, both these aspects had been used by R.A. Fisher, Bill Hamilton's 'hero of twentieth century evolution' (Hamilton 1996). Fisher's argument on sex-ratio selection (Fisher, 1930) was later couched in terms which made it appear as a paragon of evolutionary game theory (see e.g. Maynard Smith 1982). But when Fisher first conceived his argument, he knew nothing about game theory: in fact, in (Fisher, 1934), he proved from scratch a result about the optimal play in a defunct card-game, obviously unaware of John von Neumann's minimax theorem from 1930. Robert

MacArthur (1965) used an approach similar to Fisher's in discussing sex-ratios, again without using concepts from game theory.

Bill Hamilton applied game theory to reveal the essence and the limitations of Fisher's argument, and to explain the wide-spread occurrence of extraordinary sex ratios. This article, which many (including himself) consider as his best paper, appeared 1967 in *Science*, but it was based on ideas he had already developed by 1963. Hamilton stressed the 'unexpectedly close similarity to certain types of situations considered in the theory of games' (Hamilton, 1967), first 'in the sense of a play by the individual against the population' (what nowadays is called 'playing the field'), and then in the pairwise competition of two parasitoid females laying their eggs in the same host ('a refined version [of the model] which is very realistically game-like'). Technically, the main tool is that of an 'unbeatable' strategy. Hamilton did not define it explicitly — in fact, he only spoke of sex-ratios as being unbeatable, and always in the context of a specific genetic or phenotypic model. As Hamilton (1996) later stated, the notion required that no alternative strategy, 'no matter what starting frequency in a mixture, would be able to increase.

A few years later, John Maynard Smith introduced the definition of an evolutionarily stable strategy (or ESS), stressing that it had `been derived in part from the theory of games, and in part from the work of MacArthur and Hamilton on the evolution of the 'sex ratio' (Maynard Smith and Price, 1973). The concept met immediately with resounding success. In contrast to the 'unbeatable' strategy, it only required that an alternative strategy would not be able to increase when its frequency was low. In fact, it would be selected against: a rare mutant invading a resident monomorphic population of ESS-players would be eliminated under the influence of natural selection. This notion was stronger than that of a Nash equilibrium, but weaker than that of an unbeatable strategy (and of a strict Nash equilibrium). Roughly speaking, two strategies exposed to selection can (a) either co-exist in equilibrium in the population, or (b) one can dominate the other in the sense of always eliminating it from any mixture, or (c) they can mutually resist invasion by the other. In case (c), that of bistability, both strategies would be ESSs, but not unbeatable; in case (b), the dominating strategy is unbeatable. (In a further case (d) of selective neutrality, both strategies have equal success against each other and against themselves. In many contexts, this case is unlikely. But overlooking it causes confusion. For instance, both Hamilton (Axelrod and Hamilton, 1981) and Maynard Smith (1982) published proofs of the purported 'evolutionary stability' of the Tit For Tat strategy which, in fact, only showed that it was a Nash equilibrium).

The notion of ESS is particularly appropriate for large, well-mixed populations, for in that case any invading mutant is obviously rare. In highly structured populations, this is not necessarily the case, and it could well be that Hamilton conceived the stronger condition of unbeatability because he was used to think of viscous populations. But the main reason why the concept of evolutionary stability spread faster than that of unbeatability was probably caused, less by its greater generality, than by the fact that it was introduced in a much more explicit fashion (Maynard Smith, 1974) and in a context (pairwise conflicts) which was closer to usual examples in game theory. In Hamilton's 1967 paper, the 'unbeatability' concept was almost hidden behind too many other breakthrough ideas on intragenomic conflicts, levels-of-selection and the cost of sex, all piling on each other.

In later papers (Hamilton and May, 1977, and Comins, Hamilton, and May, 1980), Hamilton used the concept of an ESS and praised its 'combination of simplicity and generality' in helping to define 'final' strategies. But it should be stressed that there exists, by now, a confusing variety of versions of population stability which had all too often been illegitimately labelled as ESS (for discussions, see Lessard, 1990, and Eshel, 1996). Furthermore, the corresponding equilibria need not always be the ultimate outcome of evolution (Eshel and Motro, 1981, and Nowak, 1990).

Hamilton was clearly less interested in disentangling strategic and genetic viewpoints of great generality than in studying concrete examples of applications of game theory to social behaviour. It was, as he wrote (Hamilton 1996), `a complete surprise to find an exact evolutionary analogue to one of game theory's most famous puzzles coming to light right in the midst of my own work on population genetical dynamics', namely the Prisoner's Dilemma. He first discussed it in the proceedings volume of a symposium which took place in Washington DC in 1969 (Hamilton, 1971).

In this paper, Hamilton still had to struggle with the fact that game theory 'presupposes beings to think, and potentially, to communicate', which implies that 'in non-human biology, at least as regards preconsidered strategies, its relevance must be limited'. The rationality axiom of classical game theory was looming as an obstacle. But the Prisoner's Dilemma problem becomes actually easier within the evolutionary context — though Hamilton adds: 'I am doubtful whether the findings from natural selection throw any light on the problem of how it is rational to act'.

Hamilton later referred to this proceedings paper as 'my first intellectual wild oat', and warned the reader of its shortcomings — including, in his own words, 'lack of focus and lack of anything new to say' (Hamilton, 1996). Lack of focus there may be; indeed, the paper reads, from its opening Nietzsche quotation onwards, like an unbridled steeple-chase through Bill Hamilton's brain. But it darts over a vast amount of new things to say, with utterly original ideas taken up in full gallop and tossed off at the next turn, including in particular, anticipations of seminal papers by John Maynard Smith, E.O. Wilson and Robert Trivers that were soon to burst on the scene. In particular, game theory is almost casually applied to interactions in pairs, payoff-matrices are introduced in biological context and modified to take account of relatedness, ritualised fighting is addressed in the context of the group selection debate, and gregarious behaviour as well as non-random pairing are shown to modify the Prisoner's Dilemma.

Hamilton did not explicitly mention reciprocation in that paper, and neither did he consider repeated Prisoner's Dilemma games. But the 'personal communication by W.D. Hamilton' which Trivers (1971) quotes as 'offering a concise reformultation' of his own treatment of reciprocal altruism did take place at about that time. It stated that if the number of repeated interactions is sufficiently high, a very small cluster of altruists ready to retaliate could invade a population of non-altruists, and take over. This is arguably the most significant contribution of evolutionary game theory towards explaining the emergence of cooperation.

A much more disciplined return to some of the topics covered in Hamilton's 'wild oat' occured in Hamilton's next paper for conference proceedings (Hamilton 1975). It dealt squarely with a topic that had been central to Hamilton's thinking, ever since he had tried (in vain) to get accepted as a graduate student in a department of social anthropology: namely the innate social aptitudes of man — a term which, at that time was tantamount to a declaration of war on the prevailing intellectual climate. It may

have been in anticipation of the expected backlash that this paper was structured with particular care.

By 1975, game theory had established a safe bridgehead in the theory of evolution, and the notion of reciprocal altruism had been resoundingly put forward by Trivers. Moreover, as a result of his association with Price, Hamilton could deal much more concisely with the level-of-selection issue in this paper. A game-theoretic aspect of far-reaching consequence was addressed towards the end of the article. The more players participate in common good games of the Prisoner's Dilemma type, the more opportunities there usually are for defection. Hamilton pointed out that social enforcement can overcome this problem. This opened an approach which has proved highly successful since (see e.g. Boyd and Richerson, 1992, and Fehr and Gächter, 1998). In particular, Hamilton describes in concise form a 'three-person Prisoner's Dilemma': whatever the other two players are doing, a player is always best off defecting; but if two cooperators join forces to punish the cheater, it no longer pays to defect unilaterally.

In 1978 Hamilton, who was disappointed by the lack of recognition at London's Imperial College, moved to the University of Michigan. In retrospect, it looks almost inevitable that he should have joined forces with a young professor of political science working there, Robert Axelrod, who was at the time running a spectacular series of computer tournaments on the repeated Prisoner's Dilemma, soon to be published in the Journal of Conflict Resolution (Axelrod, 1980). Hamilton quickly grasped the importance of what Axelrod termed 'the shadow of the future'(the probability of a further round of the game), and of the huge diversity of mechanisms — discrimination, territoriality etc — for extending this shadow. In particular, he suggested that in symbiotic associations, enfeeblement of the host (caused, for instant, by ageing or illness) can act as a signal that the association will end, and result in an increase of virulence: symbionts turn parasites. Other applications were found within the field of intra-genomic conflict, the earliest of Hamilton's scientific obsessions.

The outcome of the collaboration between Axelrod and Hamilton was the now-classical paper on 'The Evolution of Cooperation' (Axelrod and Hamilton, 1981) which won the Newcomb-Cleveland Prize of the American Association for the Advancement of Science. It contained relatively little that had not already been derived by one or the other of the two authors. But simply putting it together made a spectacularly convincing case for the importance of reciprocal altruism, and spawned what Richard Dawkins, in a foreword to a later edition of Axelrod's book, termed 'a whole new research industry' with many hundreds of papers (see e.g. Dugatkin, 1997). Since Hamilton was, almost by constitution, inclined to choose the opposite of the majority strategy, this success may have been a reason for him not to return to evolutionary game theory in his later years.

References

Axelrod, R. (1980) Effective choice in the Prisoner's Dilemma, Journal of Conflict Resolution 24, 3-25.

Axelrod, R. (1984) The Evolution of Cooperation, Basic Books, New York (also published 1990, Penguin, Harmondsworth).

Axelrod, R. and W.D. Hamilton (1981) The evolution of cooperation, Science 211, 1390-6.

Boyd, R. and R. Richerson (1992) Punishment allows the evolution of cooperation (and anything else) in sizable groups. Ethology and Sociobiology 13, 171-95.

Comins, H.N., R.M. May and W.D. Hamilton (1980) Evolutionarily stable dispersal strategies, Journal of Theoretical Biology 82, 205-30.

Dugatkin, L.A. (1997) Cooperation among Animals: an Evolutionary Perspective, Oxford University Press.

Eshel, I. (1996) On the changing concept of evolutionary population stability as a reflection of a changing point of view in the quantitative theory of evolution, Journal of Mathematical Biology 34, 485-510.

Eshel, I. and A. Motro (1981) Kin selection and strong evolutionary stability of mutual help, Theor. Pop. Biol. 19, 420-33.

Fehr, E. and S. Gächter (1998) Reciprocity and Economics: the economic implications of homo reciprocans. European Economic Review 42, 845-59.

Fisher, R.A. (1930) The Genetical Theory of Natural Selection, Oxford University Press (revised edition 1958, Dover, New York).

Fisher, R.A. (1934) Randomisation and an old enigma of card play, Mathematical Gazette 18, 294-7.

Hamilton, W.D. (1967) Extraordinary sex ratios, Science 156, 477-88.

Hamilton, W.D. (1971), Selection of selfish and altruistic behavior in some extreme models, in J.F. Eisenberg and W.S. Dillon (eds) Man and Beast: Comparative Social Behavior, Smithsonian Press, Washington, 57-91.

Hamilton, W.D. (1975), Innate social aptitudes of man: an approach from evolutionary genetics, in R. Fox (ed.), Biosocial Anthropology, Malaby Press, London, 133-53.

Hamilton, W.D. (1996) Narrow Roads to Gene Land, vol I, Freeman, New York.

Hamilton, W.D. and R.M. May (1977) Dispersal in stable habitats, Nature 269, 578-81.

Lessard, S. (1990) Evolutionary stability: one concept, several meanings. Theoretical population biology 37, 159-70

Lewontin, R.C. (1960) Evolution and the theory of games, Journal of Theoretical Biology 1, 382-403.

Luce, R.D. and H. Raiffa (1957) Games and Decisions, Wiley, New York.

MacArthur, R.H. (1965) Ecological consequences of natural selection, in T. Waterman and H. Horowitz (eds.) Theoretical and Mathematical Biology, Blaisdell, New York, 388-97.

Maynard Smith, J. and G.R. Price (1973), The Logic of Animal Conflict, Nature 246, 15-8.

Maynard Smith, J. (1974) The Theory of Games and the Evolution of Animal Conflicts, Journal of Theoretical Biology 47, 209-21.

Maynard Smith, J. (1982), Evolution and the Theory of Games, Cambridge University Press.

von Neumann, J. and O. Morgenstern (1944) The theory of Games and Economic Behavior, Princeton University Press.

Nowak, M.A. (1990) An evolutionarily stable strategy may be inaccessible, Journal of Theoretical Biology 142, 237-41.

Trivers, R.L. (1971), The Evolution of Reciprocal Altruism, Quarterly Review of Biology, 46, 35-57.

Verner, J. (1965) Natural selection for the sex-ratio, American Naturalist 99, 419-21.