Leviathan, natural selection, and ethics

Renee M. Borges

Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012, India

This paper presents a review of a conceptual framework for approaching questions on the origin, evolution and maintenance of ethical systems in animal societies. It also traces links in selected ideas in moral philosophy leading to Darwin and the implications of the Darwinian paradigm of natural selection for morality in humans and other animals.

'Mr Castle,' said Frazier very earnestly, 'let me ask you a question. I warn you, it will be the most terrifying question of your life. *What would you do if you found yourself in possession of an effective science of behavior?* Suppose you suddenly found it possible to control the behavior of men as you wished. What would you do?'...

'What would I do?' said Castle thoughtfully. 'I think I would dump your science of behavior in the ocean.'

'And deny men all the help you could otherwise give them?'

'And give them the freedom they would otherwise lose forever!'

'How could you give them freedom?'

'By refusing to control them!'

'But you would only be leaving the control in other hands.'

'Whose?'

'The charlatan, the demagogue, the salesman, the ward heeler, the bully, the cheat, the educator, the priest – all who are now in possession of the techniques of behavioral engineering.'

Walden Two (B. F. Skinner 1948)

THE debate between morality and immorality has occupied humans from time immemorial. The idea of moral relativism was disturbing but recognized early on. Herodotus recounted how Darius, King of Persia, asked some Greeks what monetary incentives would induce them to eat their ancestor's corpses. The shocked Greeks declared that no sum of money could make them perform such a heinous act. Darius then summoned people of another nation whose custom was to consume their ancestors, and asked them how much money could make them burn their ancestor's bodies. The natives recoiled in horror. Similarly, for the Yanomami of South America, enslaving the women of captured villages after slaying their men presents no ethical dilemmas.

Man is an animal with a long evolutionary history. However, it was only in the eighteenth and nineteenth centuries that the animal nature of man was grudgingly acknowledged. This paper traces the course of selected secular ideas in the field of ethics leading to Darwin and the implications of the Darwinian paradigm for morality. It also presents some modern perspectives on ethical man. This paper is not intended as an overview of ethics and also does not make any moral prescriptions. It is unabashedly a biologist's view of how ethical questions may be approached.

e-mail: renee@ces.iisc.ernet.in

Morality and the myth of Gyges

The origin of received ethical norms using a cost-benefit analysis has been investigated by only a few philosophers who were able to derive a secular view of morality. In Plato's *Republic* (c. 370 BC), Plato's brother Glaucon challenged Socrates on the question of whether moral individuals were at a disadvantage compared to immoral ones. '... once people have experienced both committing wrong and being at the receiving end of it, they see that the disadvantages are avoidable and the benefits are unattainable; so they decide that the most profitable course is for them *to enter into a contract* [my italics] with one another, guaranteeing that no wrong will be committed or received.' Glaucon recounted the myth of the ring of Gyges in which a Lydian shepherd having found a ring with the power to render the wearer invisible, used it to seduce the king's wife, killed the king and ascended to the throne. Glaucon's conclusion was that an immoral life was better only if one could escape retribution. This immunity from punishment is a fundamental issue in the evolution and maintenance of moral systems.

Thomas Hobbes and Leviathan

Thomas Hobbes (1588–1679) was one of the most cogent developers of theory regarding rules governing human society. He declared that the primitive state of man was '*bellum omnium contra omnes*' (a war of all against all). Man was essentially self-seeking. However, as he began to live in larger groups, man would find that his self-interest could be better served by peace than by war (Hobbes' fundamental law) without which man's life would be 'nasty, brutish, and short'[1]. His second law was that to achieve peace, men should mutually transfer their rights to each other in the form of contracts. The third law defined injustice as a violation of contracts. Hobbes distinguished strict reciprocity wherein debts are paid in the same currency from notional benefits or gains where the repayment is in a currency different from the first. 'When the transferring of Right is not mutuall; but one of the parties transferreth, in hope to gain thereby friendship, or service from another, or from his friend; or in hope to gain the reputation of Charity, or Magnanimity; or to deliver his mind from the pain of compassion; or in hope of reward in heaven; This is not Contract, but GIFT, FREE-GIFT, GRACE: which words signifie one and the same thing[1].'

Hobbes believed that it would be more pragmatic for each man to give up his rights for safekeeping to an assembly of men. He called this regulating body a common-wealth or a leviathan. The leviathan is a mythical invincible beast described to Job by God in the Bible as proof of God's power over creation and supremacy over man (Job 41:1-34). Leviathan, for Hobbes, was analogous to a God or moral authority on earth.

Rousseau's Social Contract

A century later, Jean-Jacques Rousseau (1712–1778) stated that the fundamental problem he wished to solve was 'how to find a form of association which will defend the person and good of each member with the collective force of all, and under which each individual, while uniting himself with the others, *obeys no one but himself* [my italics], *and remains as free as before*'[2]. He examined the transitory nature of power (attribute of fitness in a Darwinian sense) and said, 'To yield to force is an act of necessity.... In what sense can it be a moral duty ... for once might is made to be right, cause and effect are reversed, and every force which overcomes another force inherits the right which belonged to the vanquished.... Since man has no natural authority over his fellows and since force alone bestows no right, *all legitimate authority among men must be based on covenants* [my italics]'. Rousseau clarified the nature of the contract. 'Suppose we draw up a balance sheet, so that the losses and gains may be readily compared. What man loses by the social contract is his natural liberty and the absolute right to *anything that tempts him and that he can take* [my italics]; what he gains by the social contract is civil liberty and the legal right of property in what he possesses.' Therefore, once men have entered into society, freedom is inseparable from virtue. Here, freedom is equivalent to free will or free choice.

By a rational process, both Hobbes and Rousseau arrived at the paradoxical conclusion that cooperation could actually serve the selfish interests of an individual. They were, therefore, able to conceptualize systems of order in society without resort to any divine presence. Leviathan was a God on earth.

The sympathy of David Hume

At the same time as Rousseau, David Hume (1711–1776) proposed the derivation of a moral sense from an innate 'sympathy' that humans felt for each other. By sympathy Hume meant a capacity to be affected by the happiness or suffering of others. For Hume, morality was maintained by sentiment and virtue was defined as '*whatever mental action or quality gives to a spectator the pleasing sentiment of approbation*; and vice the contrary'[3]. Hume's contemporary, David Hartley, proposed a materialistic theory according to which a nervous association of experiences provided a physiological basis for the origin of moral behaviour[4]. In modern terms, Hume and Hartley seem to have been considering proximal mechanisms for the development of ethical behaviour, both hinging their views on positive or negative feedback being experienced by an individual as a result of his behaviour. Here, ethical behaviour is equivalent to socially sanctioned behaviour.

Morality and Adam Smith

David Hume exerted a powerful influence. Even Adam Smith wrote *The Theory of Moral Sentiments* (1759), in which he borrowed heavily from Hume and Hartley. Smith declared that 'a man of rank and fortune ... is obliged to a very strict observation of that species of morals, *whether liberal or austere* [my italics], which the general consent of this society prescribes to persons of his rank and fortune. But as soon as he [a man of low condition] comes into a great city, he is sunk in obscurity and darkness. *His conduct is observed and attended to by nobody* [my italics], and he is therefore very likely to abandon himself to every sort of low profligacy and vice'[5].

The Darwinian paradigm

A century after Hume and Rousseau and nearly two centuries after Hobbes, Charles Darwin gave the world the paradigm of natural selection which results in the survival of the fittest or the best adapted. This simple rule could be used to investigate any attribute subject to evolution provided that there is variation in the attribute, a mechanism to transmit the variation, and an open competition between the variants.

Darwin consistently cited Hume's sympathy and disapprobation as guiding factors in the development of a sensitivity to fellow beings[6]. Darwin was, however, perplexed by the fact that sympathy was evoked to a greater degree by 'a beloved, than by an indifferent person'. The idea of kin selection was yet to be born and would appear only a century later. 'With mankind, selfishness, experience and imitation, probably add ... to the power of sympathy; for we are led by the hope of receiving good in return to perform acts of sympathetic kindness to others; and sympathy is much strengthened by habit[6].' By adding the processes of social facilitation, learning and reinforcement to innate selfish inclinations, Darwin was taking an all-inclusive view of the various components of a modern theory of behavioural evolution. It would still be a century before the theory of reciprocal altruism would be formalized.

Huxley's dilemma

It soon became evident that Darwin's paradigm had serious implications for human behaviour, for if humans are also subject to natural selection, each individual is necessarily selfish and could be driven to commit selfish acts to ensure his own survival and successful reproduction at the cost of other members of society. Darwin's devoted champion, Thomas Henry Huxley, bemoaned the 'dark, cosmic forces' and realized that man would always have to struggle against the beast within to function 'morally' within society[7].

From the Renaissance onwards, following the scientific revolutions of Kepler, Galileo and Newton, it was becoming popular to believe that there was order and harmony in nature and that nature was either in a perfect state or was progressing towards perfection. In the late eighteenth century there were protagonists like the Marquis de Condorcet and William Godwin who believed in the organic perfectability of man. For them, the vices of man were not invincible and man was capable of perpetual improvement and benevolence. This was very different from the pessimistic views of Thomas Malthus and Darwin regarding the perpetual struggle between men for a limited amount of resources and the potential such circumstances could have for a decline in benevolent fellow-feeling.

Like Rousseau, Huxley realized that 'fitness' in the Darwinian sense did not have any moral authority since the current conditions of existence determined the fittest. Therefore, 'fitness' and 'morality' were necessarily uncoupled and within such a framework there could be no absolute moral norms. Huxley attempted to find a way out of this impasse by arguing that since morality and environment were linked, perhaps if man could change his environment so that resources were sufficient to prevent an internecine struggle, then an absolute code of morality could exist and man could confidently and successfully battle the internal forces of selfish intent[7]. Following the industrial revolution, Huxley had great faith in the ability of man to create artefacts that could help him win the Malthusian battle. By implying that man could escape from natural selection, Huxley was in effect attempting to create a gap between man and animals because the implications of natural selection for morality were too terrible for him[4]. Huxley stated his view very clearly on this issue. 'There are two very different questions which people fail to discriminate. One is whether evolution accounts for morality, the other whether the principle of evolution in general can be adapted as an ethical principle. The first, of course, I advocate.... The second I deny, and reject all so-called evolutional ethics based upon it[8].' He discussed the Golden Rule 'Do as you would be done by' and opined that by refusing to punish law-breakers, the Golden Rule actually subverted the interests of society[7] (see the Prisoner's Dilemma later). He stated that 'the followers of the "golden rule" may indulge in hopes of heaven, but they must reckon with the certainty that other people will be masters of the earth[7]'!

The naturalistic fallacy

Just as Huxley had realized that man must constantly struggle against his 'baser' instincts, other moral philosophers such as John Stuart Mill were also voicing their concern over what has come to be called the naturalistic fallacy[9]. This is the false inference of what man should do based on what nature does or the incorrect inference of 'ought' from 'is'. Mill railed against this view in an essay entitled *Nature* and said that if nature and man were both the works of a perfect God, that God intended nature as a scheme to be amended, not imitated by man[10].

Despite this dark despair, others believed that a struggle for existence need not always lead to confrontation. Petr Kropotkin wrote *Mutual Aid: A Factor in Evolution* in 1902 in reply to Huxley's *The Struggle for Existence in Human Society* (1888) in which Huxley had cited overpopulation and the Malthusian struggle as the prime drivers of human behaviour[4]. Kropotkin felt that there were two types of struggles, one in which individuals compete directly with each other for scarce resources and the other in which individuals come together in cooperation to combat the physical environment. Kropotkin saw mutual aid as being beneficial not only to entire species or populations but also to individuals in the classical Darwinian sense. He admitted that he felt mutual aid was more important than competition because in the vast expanses of Russia he found very little evidence for the fearful Malthusian struggle[11]. Darwin's ideas of a Malthusian competition were born of a setting of overpopulation and resource shortages in the small island of England, a situation very different from the vast, but harsh and underpopulated Russian landscapes where cooperation to survive the rigours of the physical environment was probably more relevant.

First principles and game theory

So far, personages like Hobbes, Rousseau, Hume and Darwin have used verbal arguments to account for cooperation. Is it possible to start from first principles and to examine mathematically whether a cost-benefit approach can really influence behaviour between two interactants resulting in a stable form of reciprocity leading to cooperation? Game theory, which is essentially an analysis of conflict, has been found to be invaluable for this purpose. The pioneers of game theory were John von Neumann and Oskar Morgenstern whose treatise on the use of game theory in economics[12] was very important in operations research during World War II. The theory of games involves strategies for each player of the game and payoffs associated with each strategy. The application of game theory to evolutionary problems was largely formalized by John Maynard Smith[13] and in evolutionary terms, a payoff is the increase in Darwinian fitness achieved by the adoption of a particular strategy.

In a two-person zero-sum game, one player's gain is exactly equal to the other player's loss. In such games, a minimax pair of strategies is obtained at the equilibrium point of the game such that a player who deviates unilaterally from his equilibrium strategy in this pair will receive a worse expected payoff than if he did not deviate[14]. The following example will illustrate the point. In a mock battle between two groups of men, raiding parties from clan A and clan B are converging towards each other in one of two possible routes from their respective villages – either around the hill or around the lake. Clan A wants to reach and capture clan B's village; clan B has the opposite intention. The success of the capture by either party will depend on weakening the raiding party by engaging it in combat. This will depend on how long the two parties are in contact. Based on the distances involved in either route relative to the positions of the villages, it is known to both clans in advance what rewards in terms of contact time each clan would get by using either path depending on the route decision of the other clan. A matrix of payoffs in this game in terms of number of contact days could look like the one below.

		Clan B		
		Around hill	Around lake	Row minimum
Clan A	Around hill	3	3	3
	Around lake	1	4	1
	Column maximum	3	4	

(Modified from Casti[14])

The numbers are the rewards to clan A and indicate the number of days clan A can score hits against clan B by going either around the hill or the lake. Each player in turn can be considered to be a maximizing player; the other is then considered a minimizing player, and vice versa. If clan A is first considered as the maximizing player, then clan A wants to maximize the minimum number of days it encounters clan B and clan B simultaneously wants to minimize the maximum number of days it is encountered by clan A. The maximum minimum for clan A = 3 (maximum of row minimum), and the minimum maximum for clan B = 3 (minimum of column maximum). In this case the maxmin and the minmax strategy coincide with the decision of going around the hill for either clan. If both clans pick the option of going around the hill, then they would both be indulging in rational and risk-averse play as they are ensuring for themselves a certain minimal payoff irrespective of what the opponent decides to do. Since this game is a zero-sum game, cooperation between the two players is inconceivable.

Cooperative non-zero sum games

Another important type of game is a non-zero sum game which is a game in which one player's loss is not necessarily the other player's gain. Such games are in the realm of cooperation because two players could actually cooperate to get higher payoffs. [According to Wright[10], Mill was unknowingly advocating a non-zero sum game in his 'greatest happiness principle' by assuming that maximization of general happiness was the goal and by declaring that this could be achieved by self-sacrifice or by obeying the Golden Rule!]

The Prisoner's Dilemma (PD) is the most famous of the non-zero sum games. In this game, two partners in crime are being held separately for interrogation. Each is told that if he defects or tells on his partner, he will get a reward of 5 units (T = temptation to defect), while if he cooperates with his partner in concealing information from the police, he will get a payoff of 3 units (R = reward for mutual cooperation). If both he and his partner defect, the reward will be only 1 unit (P = punishment for mutual defection), and finally if he conceals information while his partner defects, the payoff will be 0 units (S = sucker's payoff). The game is defined by the following inequalities: $T > R > P > S$. The matrix of payoffs could look like the following.

		Player B	
		Cooperation	Defection
Player A	Cooperation	(3,3) R,R	(0,5) S,T
	Defection	(5,0) T,S	(1,1) P,P

The minimax pair of strategies consists of mutual defection (*Always D* or *All D*) at which point both players have their maximum minimal payoffs. Therefore neither player will want to deviate from this strategy unilaterally. *All D* is rational play from each individual's point of view. Paradoxically, if both players cooperate, they can get higher payoffs than that obtained from individual rationality. Mutual cooperation is a collectively rational decision. Yet *All D* is the equilibrium strategy of the game and is therefore equivalent to an evolutionary stable strategy (ESS) because if all players of the PD adopt it, it cannot be invaded by any other mutant strategy. If this is the case how can cooperation ever get started?

Iterated non-zero sum games and evolutionary ethics

Apparently, there is a world of difference between a one-shot or once-off PD game, which is a game played only once between two opponents who do not know each other, and an iterated PD game (IPD) in which the same opponents play the game over and over again. It turns out that *All D* is the best strategy to use in the one-shot PD. However, as the same partners meet repeatedly in the iterated version, there may be strategies other than *All D* that could prevail especially if the number of iterations is not fixed in advance and if the interactants get to learn the predilections of their opponents. In 1979, to see if any strategy other than *All D* was successful in iterated games, Robert Axelrod, a political scientist, had the novel idea of conducting a computer tournament. He invited scientists to submit strategies that could be set against each other in the IPD wherein the same two interactants would meet again with a probability value 'w'. The strategies could be such as 'defect in every fifth round, otherwise cooperate' or 'keep cooperating unless the other has defected twice, then defect'. Fifteen strategies were sent in by game theorists in economics, mathematics, sociology and political science. Axelrod also introduced a strategy of randomly made choices into the fray. The tournament was designed to assess the following: a) a strategy's robustness, i.e. what type of strategy can compete successfully in an environment where individuals are already using a variety of strategies, b) its stability, i.e. once a strategy has gone to fixation in an environment, what is its resistance to invasion by a mutant strategy (is the strategy an ESS?), c) how could any cooperative strategy get started in an environment where all individuals are uncooperative and are playing *All D*[15]?

Of the entries submitted, the highest average score was obtained by a three-line program submitted by Anatol Rapaport. The winning strategy was TIT for TAT (TFT) which consists of cooperating on the first move and then doing whatever the other player did on the previous move. TFT was found to be robust because it was never the first to defect; it punished defectors, and was forgiving as it retaliated only once and would cooperate if the defector cooperated. Completely unforgiving strategies fared the worst in the tournament.

Shortly after the first tournament, Axelrod solicited entries for a second round-robin tournament. This time 62 entries came in from six countries, and Axelrod once again introduced the strategy of random play. TFT won again, after an analysis of 3 million possible choices. Once TFT had gone to fixation it was thought to be an ESS if and only if the number of iterations between the same individuals was sufficiently large, i.e. if the value of 'w' was sufficiently great. This then led to the crux of the issue – in a world of all uncooperatives, how could TFT get established in the first place? There are two possible theories[15]. In the genetic relatedness or kin selection theory, relatives or kin could be expected to cooperate with each other with greater likelihood. Therefore, even in a society of selfish individuals, groups of related individuals could benefit from cooperative acts and increase the frequency of the players adopting TFT. In the cluster theory, if individuals with a predilection for cooperation happen to come together or interact with each other in a cluster, then there is a greater frequency of interaction between the 'mutant' cooperating individuals than among the non-cooperatives. Such an increased interaction frequency could again give cooperation a foothold.

Despite TFT's popularity, it is subject to certain restrictions. For example, unconditional cooperators can jeopardize the cooperative state regulated by TFT because they could be exploited by defectors or cheats (recall Huxley's predictions!), thus allowing the population of defectors to grow. In this sense, TFT is not evolutionarily stable. TFT is also jeopardized by occasional mistakes. If an interactant mistakenly defects, TFT will also defect and this could set in a long series of mutual retributions. Once again Huxley[7] seems to have thought it all out before when he said, '... civilisation could not advance far without the establishment of a capital distinction between the case of involuntary and of wilful misdeed; between a merely wrong action and a guilty one.'

Although TFT in small clusters can invade populations of defectors, another strategy called Generous TIT for TAT (GTFT) actually does better in invasion[16]. GTFT always cooperates after an interactant cooperates but defects only with a certain probability after the interactant defects. GTFT could, therefore, be more 'forgiving' especially of mistaken defections. Also, since it is impossible to predict exactly when GTFT will cooperate after a defection, it is difficult to exploit its generosity[17]. A newly formulated strategy called PAVLOV (name derived from a type of behavioural conditioning) is now a serious contender for the best performer in the PD game[16,18]. PAVLOV uses the rule: if win, then stay; if lose, then shift. In other words, PAVLOV cooperates if and only if both players, the PAVLOV player and its interactant, made the same decision on the previous move (whether a defection or a cooperation). Therefore, a PAVLOV player repeats its former move if rewarded by T or R points (same meanings for T and R as in the PD game), but switches strategies if it is punished by S or P points. Therefore, if a PAVLOV player defects and experiences defection in return, it will switch to cooperation unlike TFT which could have got locked into endless rounds of defections.

Although PAVLOV outperforms TFT in the IPD game, it cannot invade a population of defectors precisely because it switches to cooperation after experiencing mutual defection. Such a population would have to be invaded by a strategy like TFT which essentially would act as a catalyst for the achievement of a cooperative state. However, once established, PAVLOV has distinct advantages over TFT. It is tolerant of mistaken defections. Also, PAVLOV will exploit unconditional cooperators, and therefore, cannot be invaded by unconditional cooperators which would lead to the rise of exploiters or cheats. In other words, if a sucker (unconditional cooperator) is accidentally responded to with defection (by a player playing PAVLOV), and if PAVLOV does not experience defection in the next move from the sucker, it will continue the exploitation, thus ultimately leading to the decline of the suckers. Therefore unlike TFT, PAVLOV is unforgiving towards suckers, and can prevent them from subverting a cooperative state based on reward and punishment. Interestingly, it has been shown that exploitation is a successful strategy in either large or patchy populations because itinerant exploiters (called ROVERs) can move from one group or subgroup to another to escape retribution once they have been discovered to be cheats[19].

Despite the recent excitement about game theory, two-person games may be too simplistic to model cooperation within groups because payoffs to the two participants may also depend partly on payoffs to other members of the group[20]. Moreover, the currencies of the cooperative acts between two or more players may be different and difficult to quantify[21].

Evolutionary ethics: kin selection and reciprocal altruism

Hamilton's theory of kin selection referred to earlier provided a powerful explanatory perspective for seemingly altruistic acts in humans and other animals. This theory emphasized a gene view of selection and stated that altruism towards relatives could be expected because it would facilitate the survival of genes shared between the altruist and its relative, and that the probability of altruism should be directly proportional to the relatedness (proportion of shared genes) between the altruist and the individual benefitting from the altruistic act[22]. Kin selection would explain Darwin's bewilderment as to why the emotion of 'sympathy' was greater towards one's own kin than towards someone distant!

In 1971, inspired by Darwin's speculations and those of George C. Williams[23] on altruism, Robert Trivers published a seminal paper on reciprocal altruism in which he tried to explain seemingly altruistic acts performed towards non-relatives in various animal societies and attempted to take the 'altruism out of altruism'[24]. Trivers used the framework of the PD game and a cost-benefit approach and determined that an individual should be altruistic only when the benefits to the recipient of the altruistic act were much greater than the costs to the altruist of performing the act. Trivers provided a conceptual framework for analysing behavioural conflicts. For example, Trivers also speculated that once moralistic aggression had been selected for to protect against cheating, selection could favour sham moralistic aggression as a new form of cheating. This in turn could lead to selection for the ability to distinguish between the true and false versions and to develop protective measures against sham aggressors and could result in an unending escalation of the complexities within the games humans play. The theory of reciprocal altruism inspired Axelrod, Hamilton and many others to explore the origin and evolution of cooperative behaviours in greater depth. Today, besides reciprocity and kin selection, mutualisms of various kinds are also considered to be important factors promoting group living in animal societies[18,21].

Culture and the Darwinian paradigm

It appears then that the Lockian blank slate is a mythical construct. However, the Darwinian paradigm was rejected by many anthropologists and social scientists such as Franz Boas and Emile Durkheim of the early twentieth century. According to their view, human behaviour was not subject to natural selection and could not be analysed from an evolutionary perspective. Margaret Mead, trained in the Boasian tradition, studied adolescence and sexual behaviour in a group of Pacific islanders on Samoa and found striking contrasts with Western society. She made a strong case for cultural determinism and asserted that there could be no universal view of the inner workings of human societies[25]. It was later found that the Darwinian paradigm does form a framework for Samoan society and that Mead was unable to make cross-cultural generalizations because of her anti-evolutionary perspectives[26]. Perhaps the difference between the cultural determinists and the Darwinists is just a matter of time frame[27]. At a proximal level an individual's behaviour is determined by the culture in which he was raised, but from a teleonomic perspective, it is individuals who determine their culture in evolutionary time.

Natural selection can act on any self-replicating entity, even behavioural rules or cultural elements called memes[28]. Although the transmission of genes is well understood, theoretical analysis of the cultural transmission of memes is still at an early stage and is considerably more difficult to model because of the vertical and horizontal transmission of memes and the potential uncoupling of genetic and cultural evolution[29–30]. It seems logical, however, that the fitness of memes that have strong survival value for the individual carrying them should in general be correlated with the genetic fitness of individuals.

Cognition, learning and evolutionary ethics

Game theory has provided insight into evolutionarily sucesssful behavioural strategies. However, in the real world, what sort of mental tools would humans and other animals need to be able to find and employ the most successful rule? The first prerequisite is the ability to make causal inferences, that is, to be able to relate cause and effect. The many studies done on conditioning in 'lower' animals like rats and pigeons have shown that this tool is well developed. Another prerequisite is the ability to classify information so that stimuli which are slightly different from earlier ones but have the same general characteristics can be perceived as leading to similar effects. This ability of associative learning has major survival implications in social contexts, because in the real world a potential bully may come in different shapes and sizes. Pigeons and chimpanzees can associate various levels of complexity of the same type of stimulus with the appropriate response[31,32]. The next and probably the most important prerequisite, especially in a long-lived animal, is the ability to store and retrieve information from memory. The memory of sensations felt long after a particular behaviour, the recollection of the context of the act, and the long-term retention of the identity of the interactant will ensure appropriate behavioural responses. Cheats can then either be punished[33] or ignored, altruists can be favoured and unconditional cooperators can be exploited for individual benefit. Vampire bats have been experimentally shown to exhibit reciprocal altruism involving long-term memory of the identities of interacting bats[34].

Man's closest living relatives are the great apes (orangutans, gorillas and chimpanzees) and primatologists have long known that primate societies are complex and dependent on the establishment and nurturing of many complex social relationships[35,36]. In 1976, Humphrey postulated that 'individual intelligence' in primates should be positively correlated with 'social complexity'[36]. Enhanced behavioural skills also require enhanced behavioural equipment in the form of complex neuronal circuitry for computation and storage. Consequently, the neocortex is well developed in primate brains. The neocortex has been added on to the 'old' or reptilian brain in evolution and constitutes about 60% of total brain volume in Old World monkeys, 70% in chimpanzees and 80% in humans. Andrew Whiten and Richard Byrne formulated the Machiavellian intelligence hypothesis (a contemporary version of Humphrey's social intelligence hypothesis) according to which primate intelligence and associated neocortical changes evolved in the context of the challenges posed by social circumstances[37]. Neocortical size within haplorhine primates (monkeys and apes) has been shown to be positively related to social group size after controlling for the effects of phylogeny and overall brain size[38]. Interestingly, diurnality and frugivory are also important independent correlates of neocortical size and this probably suggests a strong relation between visual perception and social cognition[38–40]. Efforts to rigorously define 'social complexity' are needed and are on-going[41].

In addition, a 'theory of mind' (ToM) could also be involved in the development of complex behaviours in great apes like the chimpanzee[42]. An individual is said to have ToM or to mindread if it acts as if it understands that other individuals have mental states and if it is able to interpret these states in the appropriate context and intentionally modify its behaviour accordingly[37]. Behaviours originating from ToM are, therefore, different from purely reflexive behaviours or those originating from simple conditioning à la Skinner. So far, only humans are unquestionably regarded to possess ToM while naturalistic observations of potential mindreading abilities in the great apes[42–43] and the bottle-nosed dolphin[37], need rigorous experimental scrutiny. One type of argument for ToM in great apes involves autistic human children. Autistic children (and normal children below the age of 3–4 years) are thought to be mindblind (i.e. to possess no ToM)[44]. Autistic children are often socially impaired while great apes exhibit considerable social abilities. Therefore, by comparison, it is argued that it should not seem unreasonable to attribute ToM to great apes based even on the naturalistic observations[43]. The great apes are even thought to be capable of tactical deceit[37,45,46]. Yet, whether monkeys and apes can indeed intentionally manipulate the minds of interactants is still to be completely validated experimentally as these behaviours could also be parsimoniously interpreted as resulting from trial-and-error learning[47–49].

Does the human animal possess unlimited cognitive abilities or does it face any cognitive constraints in the development of behaviour? In 1966, Peter Wason devised the Wason Selection Task which used alphabets and numerals to investigate the power of logical reasoning in humans. The psychologists Leda Cosmides and John

Tooby modified the Wason Selection Task and couched the same logical problems in social contexts; the alphabets were changed into people and the numerals into their ages or habits[50]. The results were striking. Human subjects fared much better in the Cosmides-Tooby modified version than in the original Wason test and better still if the test featured situations requiring the detection of cheats[51,52]. Experiments such as these are revealing that the human mind has evolved to interpret social situations rapidly and accurately, and that human cognitive abilities are designed within the framework of specific social domains rather than within abstract contexts.

Is man then, a prisoner of his past? What of free will? What forces influence man's behaviour today? The answer may be both glaringly simple and tortuously complex. Yes, man has an 'old' mind shaped by the ancient history of his genes and a 'new' mind subject to new experiences, and influenced by superior cognitive abilities resulting in advanced forms of learning, rationalization and even self-awareness. Nature combines with nurture to give each individual a unique answer to his own particular social situation. Therefore, man is 'free' because he is confronted with an array of choices at every stage and he can select any of these choices within the constraints of 'real' or 'perceived' self-interest, and within the limits of his own biological and cultural potential. Many humans profess to perform altruistic acts not in the hope of future repayment in this life but from the perception of an expected reward in an after-life. However, if such acts are unlinked, even unconsciously, to expectations of some reward in this life, for example an elevation in social hierarchy, and are therefore not performed to alter the balance and outcome of social relationships ultimately leading to an increase in individual and inclusive Darwinian fitness, then and only then are they truly altruistic. Such acts would go against the concept of individual fitness, but would be a manifestation of man exercising his free will. Yet if such freedom involves a countering of the evolutionary canon of self-interest, can 'ethical' man ever be truly free?

There can thus only be a bewilderment in man as he watches the leviathan of natural selection in action, and as he looks both forwards into the future and backwards into his past. This is the cost of self-awareness. How manifold the ultimate benefits of the human mind remains to be seen.

1. Hobbes, T., *Leviathan*, 1651, Penguin edition, England, 1968.
2. Rousseau, J-J., *The Social Contract*, 1762, Translated by Maurice Cranston, Penguin, England, 1968.
3. Hume, D., *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, 1751, Reprinted, Oxford University Press, Oxford, 1975.
4. Paradis, J., in *Evolution and Ethics* (introductory essays to T. H. Huxley's 1894 manuscript), Princeton University Press, Princeton, 1989, pp. 3–55.
5. Smith, A., *An Inquiry into the Nature and Causes of the Wealth of Nations*, 1776, Reprint, Random House, New York, 1937.
6. Darwin, C., *The Descent of Man and Selection in Relation to Sex*, John Murray, London, 1871, revised edition 1901.
7. Huxley, T. H., *Evolution and Ethics*, 1894, Reprint, Princeton University Press, 1989.
8. Huxley, L., *Life and Letters of Thomas Henry Huxley*, Appleton, New York, 1900, vol. 2, pp. 360.
9. Moore, G. E., *Principia Ethica*, 1903, Cambridge University Press, Cambridge, paperback edition, 1980.
10. Wright, R., *The Moral Animal*, Abacus, London, 1996.
11. Gould, S. J., in *Bully for Brontosaurus*, Penguin, London, 1989, pp. 325–339.
12. von Neumann, J. and Morgenstern, O., *Theory of Games and Economic Behaviour*, Princeton University Press, Princeton, 1944.
13. Maynard Smith, J., *Evolution and The Theory of Games*, Cambridge University Press, Cambridge, 1982.
14. Casti, J. L., *Five Golden Rules. Great Theories of 20th Century Mathematics – and Why They Matter*, John Wiley, New York, 1996.
15. Axelrod, R. and Hamilton, W. D., *Science*, 1981, **211**, 1390–1396.
16. Nowak, M. and Sigmund, K., *Nature*, 1993, **364**, 56–58.
17. Miller, G. F., in *Machiavellian Intelligence II. Extensions and Evaluations* (eds Whiten, A. and Byrne, R. W.), Cambridge University Press, Cambridge, 1997, pp. 312–340.
18. Connor, R. C., *TREE*, 1995, **10**, 84–86.
19. Enquist, M. and Leimar, O., *Anim. Behav.*, 1993, **45**, 747–757.
20. Legge, S., *TREE*, 1996, **11**, 2–3.
21. Brembs, B., *Oikos*, 1996, **76**, 14–24.
22. Hamilton, W. D., *J. Theor. Biol.*, 1964, **7**, 1–52.
23. Williams, G. C., *Adaptation and Natural Selection*, Princeton University Press, Princeton, New Jersey, 1966.
24. Trivers, R., *Q. Rev. Biol.*, 1971, **46**, 35–56.
25. Mead, M., *Coming of Age in Samoa. A Psychological Study of Primitive Youth for Western Civilization*, Morrow Quill Paperbacks, 1928.
26. Freeman, D., *Margaret Mead and Samoa: The Making and Unmaking of an Anthropological Myth*, Harvard University Press, Cambridge, Massachusetts, 1983.
27. Dunbar, R. I. M., in *Human Adaptation* (ed Harrison, G. A.), Oxford University Press, Oxford, 1993, pp. 73–98.
28. Dawkins, R., *The Selfish Gene*, Oxford University Press, Oxford, 1976.
29. Cavalli-Sforza, L. L. and Feldman, M. W., *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton University Press, Princeton, 1981.
30. Boyd, R. and Richerson, P., *Culture and the Evolutionary Process*, University of Chicago Press, 1985.
31. Hernstein, R. D., *Philos. Trans. R. Soc. London*, 1985, **B308**, 129–143.
32. Matsuzawa, T., *J. Human Evol.*, 1985, **14**, 283–291.
33. Clutton-Brock, T. H. and Parker, G. A., *Nature*, 1995, **373**, 209–216.
34. Wilkinson, G. S., *Nature*, 1984, **308**, 181–184.
35. Jolly, A., *Science*, 1966, **153**, 501–506.
36. Humphrey, N. K., in *Growing Points in Ethology* (eds Bateson, P. P. G. and Hinde, R. A.), Cambridge University Press, Cambridge, 1976, pp. 303–317.
37. Byrne, R., *The Thinking Ape. Evolutionary Origins of Intelligence*, Oxford University Press, Oxford, 1995.
38. Barton, R. A., *Proc. R. Soc. London*, 1996, **B263**, 173–177.
39. Barton, R. A., Purvis, A. and Harvey, P. H., *Philos. Trans. R. Soc. London*, 1995, **B348**, 381–392.
40. Barton, R. A., and Dunbar, R. I. M., in *Machiavellian Intelligence II. Extensions and Evaluations* (eds Whiten, A. and Byrne, R. W.), Cambridge University Press, Cambridge, 1997, pp. 240–263.

41. Gigerenzer, G., in *Machiavellian Intelligence II. Extensions and Evaluations* (eds Whiten, A. and Byrne, R. W.), Cambridge University Press, Cambridge, 1997, pp. 264–288.
42. Premack, D. and Woodruff, G., *Behav. Brain Sci.*, 1978, **1**, 515–526.
43. Whiten, A., in *Machiavellian Intelligence II. Extensions and Evaluations* (eds Whiten, A. and Byrne, R. W.), Cambridge University Press, Cambridge, 1997, pp. 144–173.
44. Baron-Cohen, S., *Mindblindness: An Essay on Autism and Theory of Mind*, Bradford/MIT Press, Cambridge, Massachusetts, 1995.
45. de Waal, F. B. M., *Chimpanzee Politics*. Johns Hopkins University Press, Baltimore, 1982.
46. de Waal, F. B. M., *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*, Harvard University Press, Cambridge, Massachusetts, 1996.
47. Heyes, C. M., in *Animal Learning and Cognition* (ed. Mackintosh, N. J.), Academic Press, 1994, pp. 281–305.
48. Semple, S. and McComb, K., *TREE*, 1996, **11**, 434–437.
49. Hauser, M. D., in *Machiavellian Intelligence II. Extensions and Evaluations* (eds Whiten, A. and Byrne, R. W.), Cambridge University Press, Cambridge, 1997, pp. 112–143.
50. Cosmides, L. and Tooby, J., in *The Adapted Mind: Evolutionary Psychology and The Generation of Culture* (eds Barkow, J. H., Cosmides, L. and Tooby, J.), Oxford University Press, Oxford, 1993, pp. 162–228.
51. Gigerenzer, G. and Hug, K., *Cognition*, 1992, **43**, 127–171.
52. Ridley, M., *The Origins of Virtue*, Viking, London, 1996.

Received 18 October 1997; revised accepted 20 March 1998

RESEARCH ARTICLES

Chaotic dynamics of some quantum anharmonic oscillators

P. K. Chattaraj*, S. Sengupta and A. Poddar

Department of Chemistry, Indian Institute of Technology, Kharagpur 721 302, India

Quantum domain behaviour of classically chaotic systems is studied using the quantum theory of motion in the sense of classical interpretation of quantum mechanics as developed by de Broglie and Bohm. Dynamics of quantum Hénon–Heiles oscillator, Barbanis oscillator and CTW oscillator are analysed with the help of quantum Lyapunov exponent and Kolmogorov–Sinai entropy defined in terms of the distance between two initially close Bohmian trajectories. Standard diagnostics of quantum chaos like autocorrelation function and the associated power spectrum, nearest-neighbour spacing distribution, phase space volume, spectral rigidity, etc. support these results. Quantum theory of motion provides an alternative route for understanding quantum chaos. Nonlinear dynamics of integrable systems in quantum domain is also properly taken care of within this framework.

Of late, the quantum domain behaviour of classically chaotic systems has seen a great upsurge of interest[1–4]. Quantum dynamics of anharmonic oscillators like Hénon–Heiles system has been studied[5–19] extensively for this purpose. Wave packet dynamics[1–23] has been shown to be appropriate in analysing quantum manifestations of classical regular or chaotic dynamics. Various quantities which reveal the signatures of chaotic behaviour of a quantum system can be calculated once the wave packet, $\psi(r, t)$, is obtained at different time steps as a solution to the pertinent time-dependent Schrödinger's equation (TDSE), viz.

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)+V\right]\Psi=i\hbar\,\frac{\partial\Psi}{\partial t}, \tag{1}$$

where the potential V for a generalized Hénon–Heiles system takes the following form,

$$V(x,y)=\frac{1}{2}(Ax^2+By^2)+\lambda\left(Cx^2y+\frac{D}{3}\,y^3\right). \tag{2}$$

In eq. (2) λ is a parameter which measures the degree of nonlinearity and nonintegrability and may be treated as a time-like quantity[24,25]. In the conventional Hénon–Heiles potential $\lambda = A = B = C = 1$ and $D = -1$. However, it is not sacrosanct that one has to resort to the value of λ as unity only. In fact one can either take $\lambda = 1$ and vary $\hbar$, or set $\hbar = 1$ and vary λ to obtain similar results, viz. the system in higher energy levels exhibiting

*For correspondence. (e-mail: pkcj@hijli.iitkgp.ernet.in)