IT IS A COMMON PATERN in nature for simple things to give rise to more complicated things. Great complexity arises in a sequence of incremental steps, and the resulting system acquires features that could not have existed in the beginning. One such feature is altruism. In this essay, we show how altruism, which is not sustainable in simple systems, becomes a necessary component of more sophisticated scenarios, where it actually keeps the system together. This holds true for the evolutionary dynamics of co-existing organisms. It also holds true for the spiritual evolution of individuals. As we strive to find God, truth, and love throughout our lives, we can rise to higher and higher levels of understanding and spiritual sophistication. And one of the great fruits found on this path is selflessness, the willingness to give oneself for the good of others. An outsider in the early stages of life, altruism is a familiar presence in fairy tales, games, and songs and becomes vitally important as we mature, becoming one of the pillars of the spiritual temple built in the course of one's life.

This common motivating force in the universe is the subject of this essay. We explain how cooperation could have arisen naturally, despite the always-present temptation to defect. We first present the famous Axelrod et al.'s "Prisoner's dilemma" and then talk about more complicated versions of the game, which mimic different aspects of human interactions. We see that in a complex society with a sophisticated infrastructure, altruistic behavior is necessary for survival and prosperity. It is also a necessary prerequisite for the existence of language, with all of its consequences, including our ability to share thoughts, pass on moral values, and pray. Altruism is there in the beginning as a prerequisite for what it means to be human, and it is there at the end as one of the most cherished spiritual values. Following is an attempt to make sense of this.

The "Prisoner's Dilemma"

Cooperation and mutual help are integral parts of human society. So are deceit and selfish behavior. Family members, neighbors, and colleagues often find themselves in situations where personal interests are in conflict with the interests of the larger group. The temptation is to cheat and maximize one's own benefit without concern for others. The alternative choice is to cooperate, which means accepting a (small) cost in order to help somebody else. All world religions call for an attitude that promotes such altruistic behavior, and so does evolutionary game theory.
The problem of cooperation and defection is described mathematically by the famous "Prisoner’s Dilemma" (Trivers 1971, Axelrod and Hamilton 1981). This game has two players (two prisoners that are being accused of having committed a crime and are questioned by the police in separate rooms). Each player has two choices: to cooperate (to be silent) or to defect (to betray the other). The idea is that each player gains when both cooperate (the police will not have proof, so the punishment will be reduced); but if only one of them cooperates, the one that defects will gain more (i.e., will be freed for offering evidence against the other). If both defect, both lose, but not as much as the "cheated" cooperator whose cooperation is not returned (if each of them reports on the other, the punishment for both is reduced for helping the police). The essence of the game is that the temptation to defect (five points) exceeds the reward for mutual cooperation (three points), which exceeds the payoff for mutual defection (one point), which exceeds the payoff for exploited cooperation (zero points); see Table 1.

**Table 1. Payoffs in the "Prisoner’s Dilemma"**

<table>
<thead>
<tr>
<th>You Cooperate</th>
<th>You Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Cooperate</td>
<td>I receive 3, you receive 3</td>
</tr>
<tr>
<td>I Defect</td>
<td>I receive 5, you receive 0</td>
</tr>
</tbody>
</table>

What would you do in this situation? If you cooperate without knowing what the other one is up to, you would face the risk of being severely punished in case the other betrays you. And even if the other prisoner cooperates, you are still better off defecting. So, the decision is easy: No matter what the other player does, you should defect. Thus, both players will defect—and as a result will receive only one point each. On the other hand, if they had both cooperated, they would have received three points each! This is the dilemma of cooperation.

Both in *On the Origin of Species* (1859) and in *The Descent of Man* (1871), Darwin mentions that cooperation is not easily explained by natural selection. In terms of modern evolutionary game theory, imagine a population of individuals consisting of cooperators and defectors. Individuals play the Prisoner’s Dilemma in random pairwise encounters. The payoffs are added up, and individuals reproduce proportional to their payoff. It is straightforward to see that defectors will always do better than cooperators. After some time, cooperators will become extinct. Natural selection chooses defection. Yet, as noted by Darwin, cooperation is abundant in nature. How can we explain this?

**The “Altruism” of “Selfish” Genes**

The first answer to Darwin’s question was given by William D. Hamilton (1964). Cooperation among relatives can be explained by kin selection. Genes that induce altruistic behavior are shared among relatives. Hence, genes, selfishly, promote their own survival. Hamilton’s equation demands that the coefficient of relatedness between the donor and the recipient of an altruistic act has to exceed the cost-benefit
ratio of this act. As the famous biologist J. B. S. Haldane once said, “I would jump
into the river to save two brothers or eight cousins.”

Kin selection is one of the most successful theories in evolutionary biology. But
the question remains: How do we foster cooperation among nonrelatives?

**Direct Reciprocity**

In a single round of a nonrepeated game of the Prisoner’s Dilemma, there is no
incentive to cooperate. If the game is repeated, however, cooperation can become a
viable option. The repeated game of the Prisoner’s Dilemma admits an infinite
number of possible strategies. A strategy has to specify whether to cooperate or to
defect given any history of the game. The simplest and least cooperative strategy is
Always Defect (AD). Interestingly, more cooperative strategies such as Tit-for-Tat
(TFT) can take over AD. In TFT, a player cooperates on the first move and then
does whatever the other player did in the previous round. If the other player cooper-
ated, the second player cooperates. If the other player defected, the second player
defects. Hence, these moves somehow embody the harsh advice referred to in the
New Testament: “an eye for an eye, a tooth for a tooth” (Matt 5:38). TFT can take
over AD if the initial abundance of TFT players exceeds a certain threshold (the
“invasion barrier”) or if the TFT players form clusters.

TFT does well against many other strategies, but it has a weakness. If two TFT
players interact and one makes a mistake and defects, then both players will be
locked in a series of alternating defection and cooperation. Another mistake can
bring the sequence to all-out defection. In the long run, in a world such as ours
where mistakes are possible, two TFT players perform as poorly as players who
choose cooperation or defection randomly with a 50 percent chance on each move.
Hence, TFT’s unforgiving retaliation is its Achilles’ heel.

The population of TFT players can be invaded by more forgiving strategies such as
Generous TFT (GTFT), which always answers cooperation with cooperation and
sometimes answers defection with cooperation (Nowak, May, and Sigmund 1995).
The optimum level of forgiveness for the payoff values in the Table is that one in
three defections of the other player are followed by cooperation. The rule of GTFT is
“never forget a good move, but sometimes forgive a bad move.”

GTFT in turn can be undermined by Always Cooperate (AC). In mixed popula-
tions of AC and GTFT, there is (almost) only cooperation, and the payoff is the
same for every player. Random drift can lead to populations dominated by AC. This
is reminiscent of a peaceful society that loses any mechanism to retaliate or punish
defection. The outcome is clear: After some time, AD will invade again. A cycle of
war and peace is closed, as shown in Figure 1. The evolution of altruism displays
cycles of cooperation and defection.

A simple strategy such as Win-Stay, Lose-Shift (WSLS) can break this basic cycle.
In WSLS, a player will stick with his or her move from the previous round if this was
successful (e.g., earning five or three points) and change if it wasn’t (e.g., earning
only one or zero points). WSLS can correct errors and hence is as forgiving as GTFT.
In addition, WSLS can exploit AC; hence, WSLS populations cannot be undermined
by AC.
INDIRECT RECIPROCITY

The model of direct reciprocity assumes repeated interactions between two players. This is a good assumption in many situations, but it does not hold for all social interactions. People often are willing to cooperate with those whom they are unlikely to meet again. The motivation for such behavior can be explained by models of indirect reciprocity. The basic idea is that cooperation leads to a reputation that elicits help from others (Nowak and Sigmund 1998). In this setting, cooperation pays because it confers the image of a valuable community member on the cooperating individual.

EFFECT OF NEIGHBORHOODS

Finally, we note that cooperation is easier to maintain in a sedentary population (Nowak, May, and Sigmund 1995). Defectors can thrive in an anonymous crowd, but mutual aid is frequent among neighbors. It is interesting that territorially structured interactions promote cooperation even if no follow-up encounter is expected. This result favors cooperation even for the seemingly hopeless single round of the Prisoner’s Dilemma. Let us suppose that each member of the population is constrained to a node of a square grid and interacts only with his or her eight closest neighbors. We further assume that each player is either a pure cooperator or a pure defector. After each round of the game, each player is replaced by a copy of the winner. It is easy to see that a lone cooperator will be exploited by the surrounding defectors and succumb. Four cooperators in a block, however, may hold their own, because each of them interacts with three cooperators, and a defector from the outside can reach and exploit only two. If the bonus for cheating is not too high, clusters of cooperators will grow! On the other hand, lone defectors will do very well at first because they can exploit all their neighbors; but as soon as they spread, defectors will surround themselves with their like, and so diminish their returns.
Cooperation, Defection, and Communication

We have seen that game-theoretic models, crude as they are, illustrate how cooperation and altruism might arise and be maintained in a population of individuals. Sophisticated creatures such as humans may be drawn to follow strategies that encourage cooperation because of repeated interactions among individuals that can recognize and remember one another or by virtue of self-organizing structures generated by interactions with neighbors. Once cooperation is in place, it can give rise to a new level of highly complex interactions among individuals. One of the most amazing examples of that is human language.

To see that the very existence of language is impossible without cooperation, let us perform the following thought experiment. Assume that individuals in the population can exchange information. At each interaction, one person acts as a “speaker” and the other as a “listener.” The speaker has a choice of either giving away some piece of useful information (telling the truth) or lying. Telling the truth may be associated with some cost to the speaker (for instance, telling others about a food source may harm the donor in the future), and it confers a positive payoff to the listener. On the other hand, telling a lie may be beneficial to the speaker and harmful to the listener. This sets up a game-theoretic scenario in which different strategies may have more or less success in the evolutionary dynamics. Let us suppose for a moment that the AD strategy (always tell a lie) wins over the population. It is immediately clear that a language cannot evolve under such circumstances because there is no point for listeners to learn to decode the messages of the speakers! In a population of liars, those who can pass on messages and understand the messages of others have no advantage compared with individuals who do not have the ability to communicate, which shows that evolving a signaling system requires some level of cooperation.

Interesting examples can be found in biology. Bees have a highly sophisticated signaling system in which they can “tell” other members of the hive about the location, and even the quality, of food sources. This is hardly surprising considering the high level of cooperation that exists in bee colonies. The biological explanation of cooperation is kin selection (bees in a hive are closely genetically related, and thus by helping others they help spread their own genes). In its turn, cooperation leads to the development of language. Similarly, one could argue that the existence of human language relies on a high level of cooperation and altruism in human society; the difference is, of course, that the reason for altruism is not kin selection, but other, more sophisticated interactions, some of which we have examined.

If cooperation exists in a population of individuals, then we can argue that a coordinated signaling system can arise. Humans use words as a basic unit of communication. If the cognitive abilities of the individuals are high enough, then the population will follow self-organizing dynamics until all individuals have a common lexicon (Komarova and Nowak 2001). The next step is to clump words together and construct sentences following a common system of generative rules, which eventually leads to the emergence of syntactic communication. We argue that this is pos-
possible only in a highly cooperative society in which it pays to have the ability to communicate (Nowak, Komarova, and Niyogi 2002). Once a communication system is established, it leads to a wealth of sophisticated social behaviors and promotes the emergence of complex social structures, morals, and a common system of values—altruism usually being one of them.

Conclusion

We have seen that evolution as we know it would be impossible without altruism; altruistic/cooperative behavior arises naturally (and by necessity) and gives rise to a great variety of phenomena we observe in nature and in society. At the same time, altruism appears to be one of the highest and most desirable moral values of modern religions. A person on his or her spiritual quest will not advance far without embracing an altruistic attitude. This common pattern, which has arisen in a rather unexpected way, suggests that there are "favorite" themes in the universe. Our efforts to find them and internalize them should come from both the theological and scientific perspective. Future interdisciplinary research should include mathematical modeling in ecology and sociology, performed in close contact with experts in human spirituality, cultural history, and religion. This will shed more light on the emergence of altruism and on broader questions of human spiritual evolution.

Martin A. Nowak, Ph.D., is Professor of Mathematics and of Biology at Harvard University since 2003, Director of the newly founded Center for Theoretical Biology, and the Faculty Director of the Program for Evolutionary Dynamics. He received his Ph.D. from the University of Vienna in 1989, where he studied biochemistry and mathematics. His diploma thesis was on quasi-species theory, and his Ph.D. thesis was on the evolution of cooperation. After graduating, he went to the University of Oxford as the Erwin Schrödinger scholar to work with Robert May. In 1992, he became a Wellcome Trust Senior Research Fellow and in 1997 Professor of Mathematical Biology. In 1998, Professor Nowak moved to Princeton to establish the first program in Theoretical Biology at the Institute for Advanced Study. He is interested in all aspects of mathematical biology, particularly the dynamics of infectious diseases, cancer genetics, the evolution of cooperation, and human language. In an effort to describe the evolution of human language, he designed a mathematical approach bringing together formal linguistics, learning theory, and evolutionary dynamics. Professor Nowak has published more than two hundred papers and is on the editorial board of Philosophical Transactions of the Royal Society London, Journal of Theoretical Biology, and Journal of Theoretical Medicine. His first book, Virus Dynamics (together with Robert May), was published by Oxford University Press in 2000. Professor Nowak is a corresponding member of the Austrian Academy of Sciences and has won numerous awards.
Natalia L. Komarova, Ph.D., is Assistant Professor, Department of Mathematics, University of California, Irvine, and a member of the Institute for Advanced Study, Princeton. She studied theoretical physics at Moscow State University, where her master’s thesis was on chaos control in 2-D maps. She received her doctorate in Applied Mathematics in 1998 from the University of Arizona, where her thesis title was “Essays on Nonlinear Waves.” In 1998–1999, Dr. Komarova was a Research Fellow at the Mathematics Institute, the University of Warwick, United Kingdom, where she continued her work on natural pattern formation and competition between nonlinearity and randomness. She went to IAS in 1999 as a member of the School of Mathematics, where she became interested in problems of mathematical biology. In 2000, Dr. Komarova joined the Program in Theoretical Biology at IAS, and in 2003 she became part of the faculty in the Department of Mathematics at Rutgers University. She is interested in applying mathematical tools to describe natural phenomena. She received the 2002 Prize for Promise from the Student Achievement & Advocacy Services Corporation.

References


