On Collective Rational Action

Jan van Eijck

A sociologist is visiting, on a mission to discuss the problem of collective rational action. The other participants are our familiar protagonists: logician, computer scientist, philosopher. Two other project visitors, an economist and a game theorist, have joined the discussion out of curiosity. The computer scientist has brought his laptop, with wireless internet connection.

Sociologist: Nice project you've got going here guys. A pity your project description fails to mention sociology as a relevant discipline. After all, the problem of collective rationality is a key issue in my field.

Computer Scientist: When the project description talks about "the social sciences", it is also meant to encompass sociology, of course. We are very glad you are visiting us, and you're most welcome to join our discussion.

Philosopher: The problem of collective rationality has been a key issue in philosophy for more than two millennia. Aristotle discusses it at length, in the *Politics*.

Computer Scientist: (Looking at his laptop.) Wait, let me google for a quote. Ah, here it is, from Book II of the *Politics*. *(Points at the screen.)*

> For that which is common to the greatest number has the least care bestowed upon it. Every one thinks chiefly of his own, hardly at all of the common interest; and only when he is himself concerned as an individual. For besides other considerations, everybody is more inclined to neglect the duty which he expects another to fulfill; as in families many attendants are often less useful than a few [1, paragraph 403, Book II].

Sociologist: Why, that is the earliest reference to the bystander effect that I've

ever heard of! This effect has been empirically studied in the late sixties [3]: solitary people usually intervene in case of an emergency, whereas a large group of bystanders may fail to intervene - everyone thinks that someone else is bound to have called the emergency hotline already, or that someone else is bound to be more qualified to give medical help.

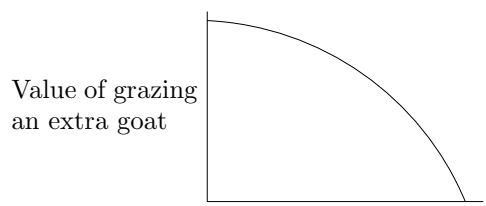
Logician: The most dramatic instance I've read about was in an article by Pacuit, Parikh and Cogan, which provided a logical analysis of how one's obligations depend on one's knowledge and vice-versa. They applied this idea to the intriguing case of Kitty Genovese, who was stabbed to death in 1964 while 38 neighbors watched from their windows but did nothing [32].

Sociologist: Actually, three social psychologists have recently shown that this Kitty Genovese story is not supported by fact, but it is more like a parable [22]. Nonetheless, I think it remains an important warning.

Game Theorist: Whatever the case may be, the more general problem that Aristotle mentions has made it to the game theory textbooks as "The Tragedy of the Commons", after an essay by Garrett Hardin [15]. Hardin, who died in 2003, was a microbiologist and ecologist, and "The Tragedy of the Commons" is his most well-known essay. Still well worth reading, by the way.

Sociologist: Yes, I know the story. If I'm not mistaken, it goes like this. Imagine a village with village greens open to all. Each farmer will try to keep as many cattle — let's say goats — on the common meadows as possible. As long as the numbers of farmers and goats stay low in relation to the carrying capacity of the land, this arrangement works fine. But there will come a time of prosperity: farmers and goats start to multiply. At some point each extra goat will lead to a marked deterioration of the greens. Still the mechanism of individual rationality will act as an encouragement for farmers to keep adding goats.

Game Theorist: That's right. Here's the tragedy in a picture [13]. *(Draws on the whiteboard.)*



Total number of goats.

Philosopher: How does that explain what goes on?

Game Theorist: On the far left you see the situation where there is still plenty, on the far right the situation where the meadows are completely destroyed. The picture shows that as we move to the right, the detrimental effect of adding extra goats keeps increasing. The value of each individual goat gets less and less, until it reaches zero.

Computer Scientist: If you want a modern version of the problem, look at the emission of carbon dioxide into the atmosphere. The Fourth IPCC Assessment report is available as a draft on the internet. Listen to what they say:

> The climate system tends to be overused (excessive GHG concentrations) because of its natural availability as a resource whose access is open to all free of charge. In contrast, climate protection tends to be underprovided. In general, the benefits of avoided climate change are spatially indivisible, freely available to all (non-excludability), irrespective of whether one is contributing to the regime costs or not. As regime benefits by one individual (nation) do not diminish their availability to others (non-rivalry), it is difficult to enforce binding commitments on the use of the climate system [17; 16]. This may result in "free riding", a situation in which mitigation costs are borne by some individuals (nations) while others (the "free riders") succeed in evading them but still enjoy the benefits ofd the mitigation commitments of the former [34, page 102].

Philosopher: What do they mean by GHG concentrations?

Computer Scientist: Green house gasses: carbon dioxide and methane. Mainly carbon dioxide, the bubbles in your Perrier water. We've started pumping it into the atmosphere in large quantities since the beginning of the industrial age, when we started burning fossil fuels in earnest. And IPCC is the International Panel on Climate Change that is trying to assess the damage.

Philosopher: But it's not really clear, yet, how dangerous this is, is it?

Logician: What do you mean, "not clear"? There's no doubt that the planet is warming. And climatologists on the IPCC agree that it is highly likely that the increased concentration of carbon dioxide causes global warming [34]. If you want to read up on the issue you should visit their website, or if you are short of time, read Marc Maslin's *Global Warming, A Very Short Introduction*

[23]. Maslin also addresses what the skeptics say. Believe me, more scientific agreement than this you are not going to get.

Sociologist: Yes, expecting every single scientist to agree would be quite unreasonable. Compare this to the question of whether smoking causes lung cancer. This is commonly accepted as scientific fact. Insurance companies use it to adjust their fees for smokers. But if I search the internet it's not at all difficult to find mavericks who deny the connection.

Philosopher: But surely on the issue of global warming there are believers and dissidents. Al Gore's film *An Inconvenient Truth* gives a voice to the believers. A much less well-known documentary, *The Great Global Warming Swindle*, was shown on British television by Channel 4.

Logician: Please, be reasonable. Scientific opinion has gravitated to the conclusion that global warming is real, that it's man-made, and that it's dangerous. Obviously, the public does not want to hear this. So there's a huge demand for denial. And there will always be journalists, scientists and documentary makers that cater for this demand. And, to add insult to injury, right-wing political groups such as the Republican Party in the US actively promote biased media coverage of the scientific discussion about global warming [26]. In the Netherlands we have science journalist Simon Rozendaal [18]. In the US there are Fred Singer [35] and Richard Lindzen [19]. Then there is the Danish skeptical environmentalist, Bjørn Lomborg [20]. The public laps it up, of course.

Computer Scientist: Hang on! Singer is not a scientist but a lobbyist.

Sociologist: And Lomborg is not denying the reality of global warming, or that it is man-made. He is skeptical about the proposed solutions. But otherwise I agree. It is only natural to prefer a comforting lie to an inconvenient truth.

Logician: It may be natural, but it ain't rational. Not long ago, in a public lecture, I mentioned global warming as an example of a phenomenon calling for collective rational action. A distinguished professor from the University of Amsterdam urged me not to worry. The earth had seen higher temperatures and greater atmospheric concentrations of carbon dioxide during the Cretaceous period, she said.

Game Theorist: What, a hundred million years ago? Well, the whole scene looked a lot different then. Dinosaurs may have liked it hot, but there were few mammals. If you take such a broad perspective then, indeed, we need

not worry. The earth is a tough old lady, she will surely get over what we did to her, in a few million years. And given that there are so many of us it is unlikely that no humans will survive, so the human species is also not in immediate danger. But whether civilization as we know it will survive, that is a different matter [21].

Logician: Global warming is what Edward Tenner calls a revenge effect of industrialization, an unintended consequence [40]. Revenge effects are everywhere. Building new roads solves a short-term problem but generates more traffic congestion in the long run as it causes suburbia to spread. Large-scale use of antibiotics causes emergence by natural selection of antibiotic-resistant bacteria. RSI is a revenge effect of office automation. And so on. Only this time the consequences might be more serious than usual.

Computer Scientist: (Listening to the conversation while crawling the internet) Yes, I found it! The Royal Society, probably the most respected and certainly the oldest learned organization in the world, issued a press release to respond to *The Great Global Warming Swindle*. This is what Martin Rees, their president, said:

> Global temperature is increasing. This warming threatens the future health and wellbeing of many millions of people throughout the world. This is especially true of those in the developing countries who are the least able to adapt and who are likely to be the worst affected. Many factors play a part in global warming but there is significant scientific evidence that greenhouse gas emissions, particularly CO_2, are responsible for most of the temperature rise. If present trends continue the projected climate change will be far greater than that already experienced. Greenhouse gas emissions are something that we can and must take action on.
>
> `http://www.royalsoc.ac.uk/news.asp?id=6089`

Logician: So there you have it. Do we want to side with the scientists, or with the mavericks? That seems to be the question.

Philosopher: I beg your pardon. Professor Hendrik Tennekes, former head of the Royal Netherlands Meteorological Institute (KNMI), is not a maverick. He is deeply worried about the arrogance of climate scientists, who erroneously think they can predict the climate [29].

Computer Scientist: Tennekes has turned against KNMI. He calls his former

colleagues "civil servants" who are telling their politician-masters what they want to hear.

Philosopher: Well, he is in a position to speak his mind, isn't he? I mean, he has retired, there is nothing at stake for him. His former KNMI colleagues have to keep their institute running. They know that the policy makers that provide their funding want climate forecasts, so it would not be prudent for them to admit that their computer models are flaky or fake. On the other hand, the fact that we are breaking weather record after weather record should worry Tennekes too. The winter of 2007 was the mildest one in the last three centuries [42].

Logician: I am not sure what Tennekes is playing at. He may have an axe to grind. He claims he was forced to retire because of his unfashionable views on the topic of climate change [43].

Computer Scientist: There have been attempts to discredit the scientific findings, by lobby groups indirectly linked to Exxon-Mobile [24; 25]. Not a new trick. Philip Morris tried the same before, to discredit a report on the dangers of passive smoking [25].

Logician: Martin Rees has urged scientists to get more involved in public debate, to speak out against minority "maverick" views [27]. Only those who understand how science works—and I suppose that includes all of us—can appreciate the difference between peer-reviewed papers in top-ranking scientific journals and mere pamphlets on the internet. We are skilled in distinguishing false from true in scientific matters, and I believe that this skill comes with responsibilities. We can see that there is a consensus on climate change. The scientific consensus is that there is global warming, that it is to a large extent anthropogenic, and that it is dangerous.

Philosopher: Maybe this whole discussion about climate change misses the point. Nobody can deny that we are putting great strain on our living environment. Or that we should limit our use of fossil fuels. Even if CO_2 in the air is not a problem, the supply of fossil fuels is limited, and our economy is hooked on it. And this addiction creates huge political problems, because the main natural gas and oil reserves happen to be in politically unstable regions.

Sociologist: Or maybe regions where enormous fuel reserves are detected have a tendency to become politically unstable because dominant powers move in from elsewhere to get a share of the pie.

Philosopher: The Arab peninsula was a backward region well before oil was detected there. And when huge natural gas reserves where found in the Netherlands and Norway, this did not cause any destabilization, or did it?

Sociologist: Never mind. Jared Diamond suggests that the collapse of a civilisation always has something to do with mismanagement of natural resources [5]. The environment is our common meadow, and it is about to be depleted. We are at the limits of growth, and the public doesn't want to hear it.

Economist: (With a wry smile) Only economists and fools believe in continued exponential growth in a finite world, right?

Computer Scientist: That's completely right, yes. *(With a side glance at the sociologist)* I believe the socialization procedures for social scientists are different from those for computer scientists. For sure, every first year economy course has an explanation of exponential growth at some point, and a story about what the prisoner's dilemma predicts for cases where individual interests are at odds with the common interest.

Economist: The students quickly understand which side their bread is buttered on: Future employers are not a bit interested in limits to the growth of their market share.

Logician: In the fields of logic and computer science the situation is dramatically different. We also explain to our students what an exponential function is, but with us the message is that they should never ever forget. If an algorithm is exponential in the size of the input, then this is a fundamental limitation that no investment in hardware will cure. Different socialization, indeed.

Computer Scientist: Formal sciences have an additional benefit. They prepare you for the view that there are insights that are "for all eternity", so to speak. Expressive formal languages are undecidable: You can use them to state questions that no computer can answer. No computer scientist in his right mind would ever dream of wanting to refute Turing and Church. Sociologists seem to think that because their subject matter is empirical, any insights in social reality have to be refutable by further evidence.

Sociologist: Social phenomena have both empirical and formal aspects. I do not deny that a mathematical look at social reality might reveal eternal truths. In fact, one of my favorite books gives examples of this. *(Shows and opens a copy of a book by Mancur Olson, [31]).* Let me quote:

> But it is not in fact true that the idea that groups will act in their self-interest follows logically from the premise of rational and self-interested behavior. It does not follow, because all of the individuals in a group would gain if they achieved their group objective, that they would act to achieve that objective, even if they were all rational and self-interested. Indeed unless the number of individuals in a group is quite small, or unless there is coercion or some other special device to make individuals act in their common interest, rational, self-interested individuals will not act to achieve their common or group interests. [31, p. 2]

Computer Scientist: Ahem—less optimistic about the emergence of an optimal outcome than Adam Smith's famous invisible hand, the hidden mechanism that fuses actions motivated by individual interests into a self-regulating social mechanism. Another well-known quote that should be easy to find on the internet. *Smith benevolence butcher* should be enough for Google. Yes, there it is:

> It is not from the benevolence of the butcher, the brewer or the baker that we expect our dinner, but from their regard to their own interest. We address ourselves not to their humanity but to their self-love, and never talk to them of our necessities but of their advantages. Nobody but a beggar chooses to depend chiefly upon the benevolence of their fellow-citizens.
>
> Adam Smith, [36, Book 1, Chapter II]

Economist: Smith never wanted to defend self-interested behavior; he just wanted to argue that it is not necessarily bad. You may care to know that Smith also wrote a *Theory of Moral Sentiments* [37], where he extolls the virtue of selflessness. Also, when he died, most of his wealth went to charity. Hardly the act of a selfish man.

Game Theorist: In game theory, as in economics I suppose, maximization of individual interest is used as a means of abstraction. It turns out that one can explain quite a lot about human behavior by assuming that every individual is pursuing his own interest. What that says about human nature is beside the point for now. But what is interesting, is that the theory of individual rationality sometimes yields funny results. Most game theory books have a chapter on the prisoner's dilemma. See for example [13] or [38]. The dilemma illustrates that perfect individual rationality may lead to a non-optimal outcome. Should one keep silent or betray the other, that is the

question. *(Draws on the whiteboard.)*

	B Keeps Silent	B Betrays
A Keeps Silent	six months in jail for each	10 years in jail for A B goes free
A Betrays	A goes free 10 years in jail for B	2 years in jail for each

Sociologist: That's consistent with what Olson has to say.

Computer Scientist: Yes, we all know the example. Betrayal pays off, whatever happens. Suppose I am prisoner A. If B keeps silent I get six months if I also keep silent, and I am free if I betray. So it is in my interest to betray. And if B starts talking, I get ten years if I keep silent, but only two years if I also talk. So again it is in my interest to betray. But what is the big deal? Why are you guys going on and on about this?

Logician: I suppose we should first establish why this is a dilemma.

Game Theorist: Usually, in strategic situations, it is important to predict what others will do. Not so here: Whatever B does, it is always in A's interest to betray. This shows that betrayal is what is called a dominant strategy. You are *always* better off by betraying. The other guy reasons as you do. He will also betray you.

Logician: So here is the dilemma: By both acting rationally, namely by defecting, the two prisoners are worse off than if they had both stayed silent.

Game Theorist: That's correct. The prisoner's dilemma is a non zero-sum game where defection yields a Nash equilibrium that is not an optimal solution. It is not a Pareto optimum.

Computer Scientist: Pareto optimum? Nash equilibrium? Can you remind us of the definitions, please?

Game Theorist: We discussed all these notions extensively in our discourses on Social choice theory and on Game theory, logic, and rational choice. But as a reminder, a *Pareto optimum* is an outcome that cannot be improved upon without hurting at least one player. A *Nash equilibrium* is a set of strategies (one for each player) such that no player has an incentive to unilaterally change her action.

Computer Scientist: I see. The point is that to get at a better outcome, both

prisoners have to change strategies at the same time, and there is no other way.

Game Theorist: OK. Now let's change this to an $N+1$ prisoner game. I am in prison. There are N other prisoners. They all, like me, have the choice between keeping silent or talking, to incriminate the others in order to get a lighter sentence. I do not know what they are going to do. What do I do? *(Draws on the whiteboard again.)*

	Number of Others Keeping Silent										
	0	1	2	3	4	5	6	7	8	9	10
I keep silent	-2	-1	0	1	2	3	4	5	6	7	8
I betray	-1	0	1	2	3	4	5	6	7	8	9

Look at the picture, which gives the case for N=10. Whatever the others do, it is *always* in my interest to betray. The more of the others keep silent, the better for me. I suppose the question is this: If *you* are in this situation, would you keep silent or talk?

Computer Scientist: Can we assume there are no repercussions?

Game Theorist: Assume that the other prisoners will never know that it is because of *you* that they are serving these long jail sentences.

Computer Scientist: Then I would betray them.

Philosopher: Shame on you.

Computer Scientist: Why make this a moral issue? "Es ist eine Krankheit, das schlechte Gewissen, das unterliegt keinem Zweifel, aber eine Krankheit, wie die Schwangerschaft eine Krankheit ist."

Philosopher: You are quoting Nietzsche, aren't you? "Bad conscience is an illness, there is no doubt about it, but an illness in the same way that pregnancy is an illness." That is from the *Genealogy of Morals* [28, p. 19]. What Nietzsche is trying to say is that it is others who instill a conscience in us.

Computer Scientist: Like a pregnancy, a conscience is at least partially caused from outside. And it is always a good idea to try and find out in whose interest the moral behavior is that others have instilled in me. In the $N+1$ prisoner's dilemma it is in my interest to instill a conscience in the other prisoners, so they all stay silent, and I get maximum benefit from betraying them all.

Philosopher: There is also an evolutionary argument against moral behavior. Those without conscience have an advantage if they convince others to act on their conscience. It increases their likelihood of survival. And the wimps that let themselves be talked into following their conscience are at a disadvantage with respect to those without one.

Sociologist: This would predict that moral behavior has to die out, in the long run. But does it? Not what we see, is it?

Philosopher: How can anyone tell? Maybe all we see is an appearance of morality. Of course, in the struggle for survival it is an advantage to *appear* moral and altruistic. As Machiavelli aptly remarked, many people live by appearances anyway, and if one wants to have one's way with people it is wise to take this fact of life into account.

Game Theorist: Anyway, in Garrett Hardin's essay [15] there is a nice analysis of the moral appeal as a double bind. A double bind is a contradiction between an overt and a hidden message. In the case of the tragedy of the commons, the overt message is: You are bad, bad, bad if you thrive at the expense of the community. But the hidden message is quite different.

Computer Scientist: You are silly, silly, silly, if you let yourself be talked into carrying the burden of others.

Philosopher: The upshot of this seems to be this. When hearing a moral appeal, find out who is talking. In particular, find out whether they have an interest at stake.

Computer Scientist: Never believe anyone who has something to sell.

Logician: That rules out a lot of people, and maybe too many. But there is a Dutch saying that may help: "Als de vos de passie preekt, boer pas op je kippen." When the fox takes up preaching, farmer watch your chickens.

Economist: How does one recognize foxes?

Logician: Simple. Their moral appeals are always aimed at preventing the institution of a binding regulation. Binding regulations would go against their interests.

Philosopher: That's right. Foxes are always in favor of giving both foxes and chickens free range.

Sociologist: But now we are faced with an empirical problem. According to the

game-theoretic analysis we have seen so far, behavior that is in the collective interest will not easily emerge. In many cases that is what we actually see. In the early Middle Ages, the farmers in West European coastal villages were quite ineffective at fighting off the Nordic invaders. Once the Erics and Olavs had disembarked, they should have been at a disadvantage against a well-organized collective of determined farmers. But for a long, long time these farmers did not get their act together.

Game Theorist: Well, that is only in accordance with the findings of game theory, isn't it?

Philosopher: I don't agree that the Vikings were at a disadvantage. Just imagine: they arrived in ships that needed at least twenty men to handle, in heavy weather. Cooperation was of the essence, and "One for all and all for one" must have been natural to them. Not so for the farmers, who had perhaps only learnt to be surly and mind their own business.

Sociologist: Game theory predicts that *no* collective ever gets their act together. But this is contradicted by what we see in actual life. What we see is that suddenly the collective structures emerge that allow successful coping with emergencies. How can we explain that this happens?

Logician: What is needed is emergence of common knowledge and common intention. The issue of common knowledge we have discussed at length already (see the Chapter starting on page **??**). The theme of collective action is quite hot in theories of multi-agent interaction these days. In [8] there is a nice overview.

Sociologist: But let me ask again. How do common knowledge and collective intention arise in the first place?

Logician: As I said, we have talked about common knowledge before. Basically, there are just two ways: public announcement and common experience. Mind you, public announcement is also a kind of common experience. All of us hear the same proclamation, and we are aware of the fact that we are all hearing it.

Sociologist: And collective intentions?

Logician: Collective intentions have a clear motivational component, so a public announcement or a common experience does not suffice. After all, a collective intention among a team means that its members not only intend to

do their best to achieve the goal as part of the team, but they also need a mutual intention, which means that they all intend that they all intend that they all intend ... and so on, ad infinitum, to achieve the goal together. Often a team leader will need to persuade each individual potential team member, in order to create such a heavy-duty mutual intention. After that, the team leader can indeed use a public announcement to create a common knowledge or common belief about this mutual intention. Finally, after both these kinds of communication, the collective intention is in place and the team can start to work [6].

Sociologist: Communication and ostentation certainly play a role in the emergence of collective social structures. The way in which the care of the poor got organized in medieval Europe is a nice example. The clergy played a role in this, and they may have liked to believe that people started to contribute to collective charities out of a moral sense instilled by the Church. But in fact what the clergy did was much more effective. They created a communication structure where everyone knew how much everyone else was contributing to poor relief. Making a contribution to the collective soon became a matter of honor. One might say that collectives come into being as a result of common knowledge and common action [39].

Logician: In modern logical theories of the effects of communication, one can study the difference in effects between private acts of communication and public events. These differences turn out to be vast. And in the logic of action, there is an agreement that effective collective action can never be the sum of individual actions. At least three ingredients seem needed: (i) common knowledge of the moral stature of those influencing the group, (ii) common knowledge of what is the interest of the group as a whole, (iii) common knowledge of the collective willingness to take action. I suppose social structures for this are all structures that foster the sense of community.

Philosopher: Some societies were much more effective at this than others. Picture the life in a Greek city state, on a festival day. The whole city would gather for a day at the theatre, for three tragedies and a comedy, in a place where everyone can see that everyone is there, to be part of the same overwhelming experience. A very powerful way to shape a community and keep it together.

Logician: You are a romantic. But you are right, going to the theatre is much more civilized than watching ferocious animals slaughter innocent Christians,

as the Romans were fond of doing.

Philosopher: Which in turn is only slightly worse than watching hooligans disturb football matches, I suppose.

Sociologist: It is clear that you are not great admirers of popular culture. Perhaps we should return to our main question. How does collective rational action emerge? That is not an easy one to explain. Charles Darwin himself believed he had a ready explanation [4]. He cheerfully applied the principle of natural selection to groups, assuming that groups compete just like individuals and that the features that make some groups more successful than others — altruism, courage, selfless acting in the interest of the community — were perfected through natural selection of the most successful groups. But this explanation simply cannot work. For within such communities, the free riders who make use of the spirit of self-sacrifice of others are at a huge evolutionary advantage.

Game Theorist: That's completely right. Modern evolutionary game theory has analyzed some interesting solutions though, which I will explain in a while [30].

Computer Scientist: (Who is running Google queries all the time) Wait, wait, I have found an online version of Darwin's *Descent of Man*. Chapter 5 discusses precisely this issue. Listen, here it is:

> But it may be asked, how within the limits of the same tribe did a large number of members first become endowed with these social and moral qualities, and how was the standard of excellence raised? It is extremely doubtful whether the offspring of the more sympathetic and benevolent parents, or of those who were the most faithful to their comrades, would be reared in greater numbers than the children of selfish and treacherous parents belonging to the same tribe. He who was ready to sacrifice his life, as many a savage has been, rather than betray his comrades, would often leave no offspring to inherit his noble nature. The bravest men, who were always willing to come to the front in war, and who freely risked their lives for others, would on an average perish in larger numbers than other men. Therefore, it hardly seems probable that the number of men gifted with such virtues, or that the standard of their excellence, could be increased through natural selection, that is, by the survival of the fittest; for we are not here speaking of one tribe being victorious over another [4, Chapter 5].

Sociologist: (Dryly) Darwin is quite eloquent in expressing the refutation of his own explanation. How does the chapter go on to refute the refutation?

Computer Scientist: Well, he talks about men developing foresight, and learning that giving aid commonly meant receiving aid in return. And more importantly, about successful tribes learning to use praise and blame to regulate the behavior of the tribe members.

> A man who was not impelled by any deep, instinctive feeling, to sacrifice his life for the good of others, yet was roused to such actions by a sense of glory, would by his example excite the same wish for glory in other men, and would strengthen by exercise the noble feeling of admiration. He might thus do far more good to his tribe than by begetting offspring with a tendency to inherit his own high character [4, Chapter 5]..

Philosopher: To me, this sounds more like cultural selection than natural selection. Surely, if both these mechanisms are at work, they have to be quite different. Darwin presents cultural selection as a kind of minor variation on natural selection. Very strange.

Sociologist: What we are looking for is a mechanism that explains how the tragedy of the commons can be avoided. It seems to me that the Darwin quote does not provide that.

Game Theorist: Let me give you the common understanding in game theory. Most game theorists do not buy Darwin's account of group selection. Rather, to get out of the tragedy, they devised mechanisms that make altruism pay. Two mechanisms for that are kin altruism and reciprocal altruism. The first was eloquently defended in a famous book by George Williams [44]. It tries to explain the emergence of altruistic behavior by assuming that "altruistic genes" (genetic treats that cause cooperative behavior) favor families that have them over families that do not, on the assumption that the altruistic behavior is limited to next-of-kin. The theory of reciprocal altruism, first proposed by Robert Trivers [41], explains the emergence of collaboration because it pays off, for the individual. A game-theoretic version of this argument was later given by Robert Axelrod [2].

Philosopher: Aha, I see. You guys simply explain altruism away as a rather special kind of egoism. Immanuel Kant would have been shocked: to him, altruistic behavior emerges from application of the golden rule or categorical imperative — "treat others as you would like to be treated yourself" — a

commandment of reason. For Kant, doing things because you *like* doing them, or even worse, because they serve your interest, could never be moral behavior.

Game Theorist: Axelrod studied a version of the prisoner's dilemma where the same two players repeatedly interact, the so-called iterated prisoner's dilemma. He even organized iterated prisoner's dilemma tournaments, where colleagues were asked to propose strategies for this game. It then turned out that a very simple strategy, tit for tat, devised by Anatol Rapoport, was the most successful [33].

Sociologist: I suppose what makes the iterated prisoner's dilemma so different from the simple prisoner's dilemma is the fact that there are other consequences besides the immediate payoff. Other players will remember how you treated them, and can retaliate in later installments of the game. The tit for tat rules are very simple. Always collaborate, unless provoked. When provoked, retaliate immediately. Be quick to forgive.

Computer Scientist: That's interesting, for it neatly sums up how I handle my colleagues. I don't mind if they treat me the same, but I doubt whether that is an application of Kant's categorical imperative.

Sociologist: If you are sure that this is your last encounter with a particular player you can get away with being selfish. People with long experience in business are understandably wary when dealing with colleagues on the brink of retirement. These guys may be tempted to play you a departure trick, knowing that you cannot get back at them.

Game Theorist: In modern times, binary reciprocity is just the simplest kind of helping one another. Societies thrive best if there is *generalized reciprocity*, or paying it forward: If I scratch your back, you don't need to scratch mine, as long as you scratch someone else's. Recently, Nowak and Sigmund have published interesting research on such generalized reciprocity in evolutionary game theory. It turns out that two things help generalized reciprocity to emerge: shared information and a reputation mechanism by which an agent's social score depends on whether they are free riders or are paying it forward. Agents who are known to be free riders, are not helped anymore [30].

Sociologist: In the Netherlands, Egas and Riedl conducted interesting experiments about such "altruistic punishments" via a public goods game on the Internet, in which almost a thousand people participated. It turned out that

participants dealt out many more corrective punishments when it was cheap for themselves to do so and had high impact on the free-riders, than when it was expensive and had low impact [11].

Philosopher: (To the sociologist) Surely, there are alternative sociological views on the emergence of group behavior. Wasn't Emile Durkheim, the founding father of your discipline, the one who said that the idea of society is the soul of religion?

Sociologist: (Quoting from memory) "Car on sait aujourd'hui qu'une religion n'implique pas nécessairement des symboles et des rites proprement dits, des temples et des prêtres [...]. Essentiellement, elle n'est autre chose qu'un ensemble de croyances et de pratiques collectives d'une particulière authorité" [10, p. 270].

Computer Scientist: Translation, if you please.

Sociologist: (Smiling) "For we know today that a religion does not necessarily imply symbols and rites in the narrow sense of those terms, or temples and priests. Essentially, it is nothing other than a body of collective beliefs and practices endowed with a certain authority." What Durkheim intends to say is that religion is what bonds people together in a community. Mind you, in the same essay he goes on to say that a "religion of individualism" is the only viable form of religion in modern society.

Game Theorist: David Sloan Wilson, in [45], tries to revive Darwin's argument for group selection, and identifies religion as the determining factor. Wilson is not a defender of religion. He is not blind to its dark side: the tendency to blend in-group morality with hostility towards outsiders. In fact, his analysis of religious practices as a set of group-forming operations that serve to enhance the survival value of the group squares well with this. But his book gives a minority report, although Robin Dunbar expresses some similar ideas about the importance of group cohesion and the role of religion in his *Grooming, Gossip and the Evolution of Language* [7]. Most game-theorists and evolutionary biologists still prefer kin altruism and reciprocal altruism as explanations of emerging group behavior.

Sociologist: Durkheim and Wilson would have gotten along well. Durkheim was deeply aware of the fact that people need a community to belong to. In a truly groundbreaking work [9], he studied suicide rates in a great diversity of populations across Europe, and found a clear correlation between lack of social

constraints — what Durkheim called 'anomie' — and likelihood to kill oneself. Durkheim drew the conclusion that people need obligations and constraints to instill their lives with structure and meaning. People who are religious, married and with children are much less likely to kill themselves. Suicide, at first sight the most individual act one can imagine, is explained in terms of what links — or fails to link — an individual to society.

Game Theorist: Let's come back to "The Tragedy of the Commons". It may interest you what Hardin himself advocated as a solution. His recipe is what he called "mutual constraint, mutually agreed upon." He did believe in pledges, promises, laws and ... sanctions. Like Sigmund Freud, he was well aware that civilization comes at a cost; it invites or forces us to suppress part of our nature, in the interest of the community, and deep, deep down, civilized individuals do resent this [12]. But there is no other way.

Philosopher: Now apply this to problem of global warming. This project has made quite a contribution to carbon dioxide emission, with flying in colleagues from around the world: the United States, New Zealand.

Logician: Do you think we should have set an example? What would be the point? I mean, if everyone else is blazing by in petrol guzzling SUVs, why should I be the only one to ride a bicycle? My solitary plodding along on my bike is not going to save the world. I am quite willing to give up my intercontinental flights and my car. Maybe even my house in France—it is getting too hot there anyway. But only on condition that others do the same.

Sociologist: Preaching self-imposed abstinence or attempting to convert by setting an example are wastes of time.

Logician: Ahem—I admire those who are setting an example, but the trouble is there are too few of them. What we need is mutual constraint, mutually agreed upon. We individual citizens have to convince our governments that it is time to constrain us. This may sound paradoxical, but it makes sense. Not only that, but it is the only way.

Philosopher: Constrain us? How? By imposing a system of individual carbon dioxide emission rights [24]? My good man, we emit carbon dioxide every time we exhale. They can't forbid us to breathe.

Logician: First we need the insight that it is urgent to limit our ecological footprint, to live wisely on this planet. Next, we need to see that we have to be forced to live wisely, that we cannot do it without communities that support

us and keep us on track. So we need to build and rebuild our communities, for they will have to impose the mutual constraints.

Computer Scientist: For all that to happen there has to be a universal sense of urgency, like the sense of urgency that was felt in the US after Pearl Harbor. Not very likely.

Philosopher: One can see how the $N + 1$ prisoner's dilemma now resolves itself. Suppose a group knows that they are a group, and that they are under mutual constraint, mutually agreed upon. Assume there are 11 prisoners, and the individual payoffs are as before: *(Points at the whiteboard)*

	Number of Others Keeping Silent										
	0	1	2	3	4	5	6	7	8	9	10
I keep silent	-2	-1	0	1	2	3	4	5	6	7	8
I betray	-1	0	1	2	3	4	5	6	7	8	9

Then the collective payoff of everyone keeping silent is $11*8$, and the collective payoff of everyone defecting is $11 * -1$, so we get:

Everyone keeping silent	88
Everyone incriminating each other	-11

Logician: Speaking about pledges: In the year 2000 there was an interesting *Unesco* initiative. The *Unesco Manifesto 2000* invites us to make a pledge to devote our honest efforts to fostering community sense and make our contribution to community building. One year after this call to action, 9/11 happened, and initiatives like this may have seemed pointless for a while. But for building communities one has to start with small steps. Anyway, have a look at `http://www3.unesco.org/manifesto2000/`. If it appeals to you, you are in the excellent company of the Dalai Lama and other winners of the Nobel Peace Prize.

Game Theorist: Expect to hear a lot more about pledges and contracts on environmental issues in the near future [14].

Computer Scientist: (Starts googling immediately.)

References

[1] Aristotle. *The Politics of Aristotle: Translated into English with Introduction, Marginal Analysis, Essays, Notes and Indices*, volume 1. Clarendon Press, Oxford, 1885. Translated and annotated by B. Jowett.

[2] Robert Axelrod. *The Evolution of Cooperation*. Basic Books, New York, 1984.

[3] J.M. Darley and B. Letane. Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8:5377–383, 1968.

[4] Charles Darwin. *The Descent of Man, and Selection in Relation to Sex*. John Murray, London, 1871. Many online editions.

[5] Jared Diamond. *Collapse: How Societies Choose to Fail or Succeed*. Viking Books, New York, 2005.

[6] F. Dignum, B. Dunin-Kęplicz, and R. Verbrugge. Creating collective intention through dialogue. *Logic Journal of the IGPL*, 9:145–158, 2001.

[7] R. Dunbar. *Grooming, Gossip and the Evolution of Language*. Faber and Faber, London, 1996.

[8] Barbara Dunin-Kęplicz and Rineke Verbrugge. A logical view on teamwork. In Jan van Eijck and Rineke Verbrugge, editors, *Games, Actions and Social Software*. College Publications, 2008. To appear.

[9] Émile Durkheim. *Suicide*. Free Press, 1897/1951.

[10] Émile Durkheim. *La Science Sociale et l' Action*. Collection SUP. Presses Universitaires de France, Paris, 1987.

[11] Martijn Egas and Arno Riedl. The economics of altruistic punishment and the demise of cooperation. Tinbergen Institute Discussion Papers 05-065/1, Tinbergen Institute, June 2005.

[12] Sigmund Freud. *Das Unbehagen in der Kultur (Civilization and Its Discontents)*. Internationaler Psychoanalytischer Verlag, Wien (Vienna), 1930.

[13] Robert Gibbons. *A Primer in Game Theory*. Prentice Hall, 1992.

[14] Newt Gingrich and Terry L. Maple. *A Contract with the Earth.* Johns Hopkins University Press, 2007. Foreword by E.O. Wilson.

[15] Garrett Hardin. The tragedy of the commons. *Science*, 162:1243–48, 1968.

[16] I. Kaul, P. Conceiąo, K. Le Gouven, and R.U. Mendoz. *Providing Global Public Goods.* Oxford University Press, 2003.

[17] I. Kaul, I. Grunberg, and M.A. Stern. *Global Public Goods.* Oxford University Press, 1999.

[18] Hans Labohm, Simon Rozendaal, and Dick Thoenes. *Man-Made Global Warming: Unravelling a Dogma.* Multi-Science Publishing, 2004.

[19] Richard S. Lindzen. Some coolness concerning global warming. *Bulletin of the American Meteorological Society*, 71(3):288–299, 1990.

[20] Bjørn Lomborg. *The Skeptical Environmentalist: Measuring the Real State of the World.* Cambridge University Press, 2001.

[21] James Lovelock. *The Revenge of Gaia.* Allen Lane, 2006.

[22] R. Manning, M. Levine, and A. Collins. The Kitty Genovese murder and the social psychology of helping: The parable of the 38 witnesses. *American Psychologist*, 62:555–562, 2007.

[23] Marc Maslin. *Global Warming, a Very Short Introduction.* Oxford University Press, Oxford, 2004.

[24] George Mobiot. *Heat.* Penguin Books, 2007.

[25] Peter Montague. A new disinformation campaign. *New York Times*, April 29 1998.

[26] Chris Mooney. *The Republican War on Science.* Perseus Books, New York, 2005.

[27] BBC News. 'Maverick' risk to science debate, November 30 2006. `http://news.bbc.co.uk/2/hi/science/nature/6159371.stm`.

[28] F. Nietzsche. *On the Genealogy of Morals: A Polemic : by Way of Clarification and Supplement to My Last Book, Beyond Good and Evil.* Oxford University Press, Oxford, 1998. translated and annotated by D. Smith.

[29] NN. Klimaat: 'Kyoto is onzin' (climate: 'Kyoto is nonsense'). *Elsevier*, Wednesday, June 20, 2007.

[30] M.A. Nowak and K. Sigmund. Evolution of indirect reciprocity. *Nature*, 437:1291–1298, 2005.

[31] M. Olson. *The Logic of Collective Action; Public Goods and the Theory of Groups*. Harvard University Press, Cambridge, Mass., 1965.

[32] Eric Pacuit, Rohit Parikh, and Eva Cogan. The logic of knowledge based obligation. *Synthese*, 31:311–341, 2006.

[33] A. Rapoport. Paradoxe der entscheidungstheorie. In R. Martinsen, editor, *Das Auge der Wissenschaft. Zur Emergenz von Realität*, pages 57–73. Nomos, Baden-Baden, 1995.

[34] H.-H. Rogner, R. Zhou, R. Bradley, P. Crabbé, O. Edenhofer, B. Hare, L. Kuijpers, and M. Yamaguchi. Introduction. In B. Metz et al., editor, *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, 2007.

[35] S. Fred Singer. *The Scientific Case against the Global Climate Treaty*. The Science & Environment Policy Project, 1999. Available from `http://www.sepp.org/publications/`.

[36] Adam Smith. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Liberty Fund, Indianapolis, 1982. First published 1776.

[37] Adam Smith. *The Theory of Moral Sentiments*. Liberty Fund, Indianapolis, 1984. First published London and Edinburgh, 1759.

[38] Philip D. Straffin. *Game Theory and Strategy*. The Mathematical Association of America, New Mathematical Library, 1993. Fourth printing: 2002.

[39] Abram de Swaan. Nood en deugd: over altruïsme en collectieve actie (Distress and virtue. On altruism and collective action). *De Gids*, 147(3):139–151., 1984. In Dutch.

[40] Edward Tenner. *Why Things Bite Back; Technology and the Revenge Effect*. Fourth Estate, 1996.

[41] R.L. Trivers. The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46:35–57, 1971.

[42] G.J. van Oldenburg. The mild winter of 2007: What were the causes?, 2007. Full text available on the internet at `http://www.knmi.nl/kenniscentrum/zachte_winter_2007/index_en.html`.

[43] Thomas Vanheste. De eenzame strijd van de broeikasongelovigen (the lonely struggle of the greenhouse-infidels). *Vrij Nederland*, May 25 2005.

[44] George C. Williams. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton University Press, 1966 (reprinted 1999).

[45] David Sloan Wilson. *Darwin's Cathedral: Evolution, Religion, and the Nature of Society*. The University of Chicago Press, Chicago, 2002.