

conflict: insights from game theory

this handout with links to all research studies was posted to www.stressedtozest.com on 6.06.11

I've written a series of posts on conflict in the last several days – most recently *"Conflict: not too much, not too little - when to get real & problem solve in close relationships"*. I was away on holiday in France last month and I read Matt Ridley's slightly dated, but fine book *"The origins of virtue"*. As A. S. Byatt commented *"Matt Ridley's splendid book studies co-operation (and conflict) from the genes themselves to modern technological societies ... 'Our minds have been built by selfishness, but they have been built to be social, trustworthy and co-operative. That is the paradox this book has tried to explain.' It has done it brilliantly"*. And Richard Dawkins wrote *"If my The Selfish Gene were to have a Volume Two devoted to humans, The Origins of Virtue is pretty much what I think it ought to look like."*

Chapter three of *"The origins of virtue"* is about game theory and I found it fascinating. The helpful website *"Gametheory.net"* states *"Game theory is the study of how people interact and make decisions. This broad definition applies to most of the social sciences, but game theory applies mathematical models to this interaction under the assumption that each person's behavior impacts the well-being of all other participants in the game. These models are often quite simplified abstractions of real-world interactions but offer a tractable way of predicting likely outcomes"* and Wikipedia comes up trumps with many pages on game theory. It comments *"In mathematics, game theory models strategic situations, or games, in which an individual's success in making choices depends on the choices of others ... Today, 'game theory is a sort of umbrella or 'unified field' theory for the rational side of social science ... Game theory has been widely recognized as an important tool in many fields. Eight game-theorists have won the Nobel Memorial Prize in Economic Sciences, and John Maynard Smith was awarded the Crafoord Prize for his application of game theory to biology."*

Matt Ridley looked at game theory in *"The origins of virtue"* – particularly a classic "game" called *"The prisoner's dilemma"*. The prisoner's dilemma mathematically explores competition and cooperation. A standard example runs: *"Two suspects are arrested by the police. The police have insufficient evidence for a conviction, and, having separated the prisoners, visit each of them to offer the same deal. If one testifies for the prosecution against the other (defects) and the other remains silent (cooperates), the defector goes free and the silent accomplice receives the full one-year sentence. If both remain silent, both prisoners are sentenced to only one month in jail for a minor charge. If each betrays the other, each receives a three-month sentence. Each prisoner must choose to betray the other or to remain silent. Each one is assured that the other would not know about the betrayal before the end of the investigation. How should the prisoners act?"*

Assuming that the prisoners are rational and motivated primarily to minimise their own time in jail - they will obviously choose to "defect". And this, in a sense, is the "bad taste" left by the mathematics of the classic prisoner's dilemma. It seems as though intelligence "should" always choose selfishness. Then along came research on repeated prisoner's dilemma games - where the game is played a random number of times or indefinitely and there is an opportunity to "punish" the other player for their non-cooperative choices. Here something very interesting emerges. As Wikipedia puts it: *"Interest in the iterated prisoners' dilemma (IPD) was kindled by Robert Axelrod in his book The Evolution of Cooperation (1984). In it he reports on a tournament he organized of the N step prisoner dilemma (with N fixed) in which participants have to choose their mutual strategy again and again, and have memory of their previous encounters. Axelrod invited academic colleagues all over the world to devise computer strategies to compete in an IPD tournament. The programs that were entered varied widely in algorithmic complexity, initial hostility, capacity for forgiveness, and so forth."*

[Cont.]

Axelrod discovered that when these encounters were repeated over a long period of time with many players, each with different strategies, greedy strategies tended to do very poorly in the long run while more altruistic strategies did better, as judged purely by self-interest. He used this to show a possible mechanism for the evolution of altruistic behaviour from mechanisms that are initially purely selfish, by natural selection.

The best deterministic strategy was found to be "tit for tat", which Anatol Rapoport developed and entered into the tournament. It was the simplest of any program entered, containing only four lines of BASIC, and won the contest. The strategy is simply to cooperate on the first iteration of the game; after that, the player does what his or her opponent did on the previous move. Depending on the situation, a slightly better strategy can be "tit for tat with forgiveness." When the opponent defects, on the next move, the player sometimes cooperates anyway, with a small probability (around 1-5%). This allows for occasional recovery from getting trapped in a cycle of defections. The exact probability depends on the line-up of opponents.

By analysing the top-scoring strategies, Axelrod stated several conditions necessary for a strategy to be successful.

nice: *The most important condition is that the strategy must be "nice", that is, it will not defect before its opponent does (this is sometimes referred to as an "optimistic" algorithm). Almost all of the top-scoring strategies were nice; therefore a purely selfish strategy will not "cheat" on its opponent, for purely self-interested reasons first.*

retaliating: *However, Axelrod contended, the successful strategy must not be a blind optimist. It must sometimes retaliate. An example of a non-retaliating strategy is Always Cooperate. This is a very bad choice, as "nasty" strategies will ruthlessly exploit such players.*

forgiving: *Successful strategies must also be forgiving. Though players will retaliate, they will once again fall back to cooperating if the opponent does not continue to defect. This stops long runs of revenge and counter-revenge, maximizing points.*

non-envious: *The last quality is being non-envious, that is not striving to score more than the opponent (impossible for a 'nice' strategy, i.e., a 'nice' strategy can never score more than the opponent)."*

Fascinating! Game theory simplifies real life situations. In doing so it can lose some of the richness and possibility of our actual choices – see, for example, the handout "*Honesty, transparency & confrontation*". However game theory can provide helpful insights too. "*Tit-for-tat with forgiveness*" strategies make a good deal of sense and fit well with research findings like those described in blog posts such as 4.06.11's "*Conflict: not too much, not too little – the importance of assertiveness in close relationships*".
