

Further results on the societal iterated prisoner's dilemma

Toby Ord

Introduction

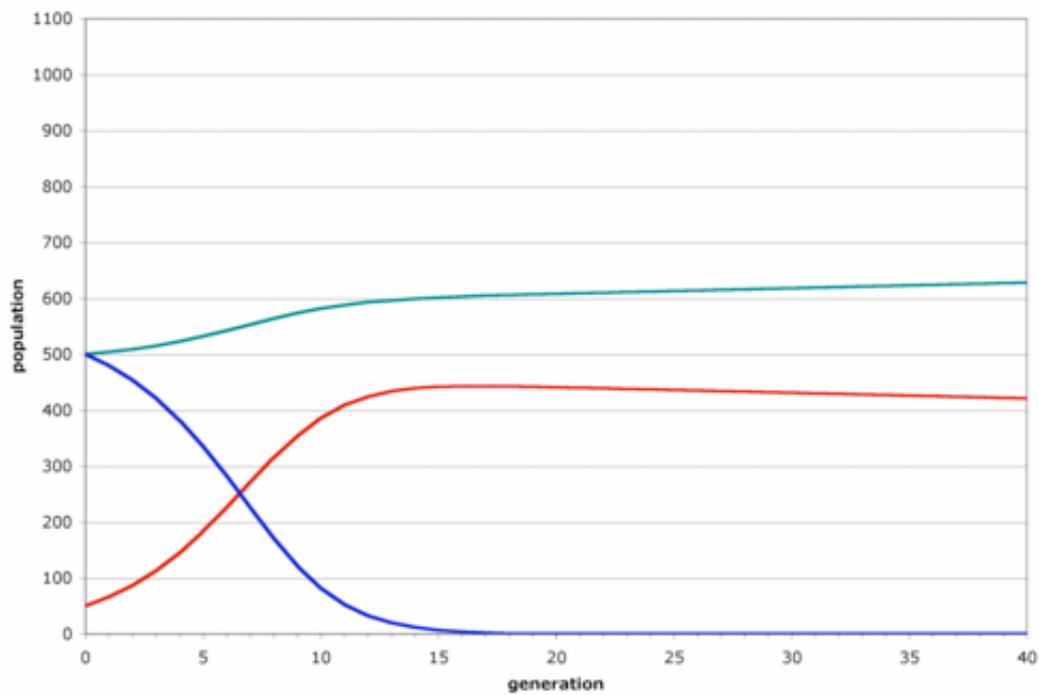
Several years ago, I did some research on a new multiple player variant of the iterated prisoner's dilemma (IPD). In this variant, which I named *Societal Iterated Prisoner's Dilemma* (SIPD), the game consists of a number of rounds in which each of the n players simultaneously choose to cooperate or to defect on each of the other players. At the end of the round, the players get a score for each of their interactions according to the standard rules for the prisoner's dilemma and then continue to the next round. This would be equivalent to the standard *Round Robin Iterated Prisoner's Dilemma* (RR-IPD) if not for one important difference: all the interactions from the past rounds are common knowledge.

This difference opens up a great many strategies which are impossible in the two standard variants of multiple player prisoner's dilemma. For example, an exploitative player could defect on all those who have failed to retaliate when defected on by others. Or, alternatively, a player could help enforce cooperation by defecting on those who defect on others without provocation. Thus, while all interactions are still pair-wise, there are strategies which exhibit interesting social behaviours.

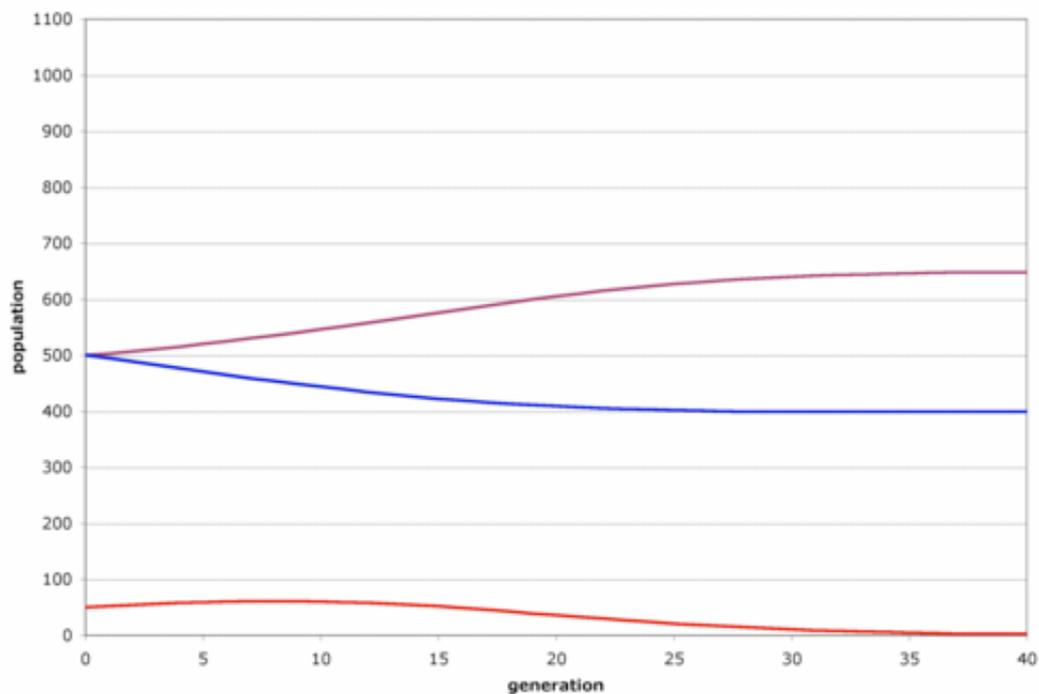
Most of the results of my research can be found in ([Ord and Blair 2002](#)), and I strongly encourage you to read that paper before continuing with this note. After writing that paper, I continued researching SIPD for a brief period and found a few more interesting results, which I feel are worth sharing here.

As we saw, *peacekeeping* strategies have considerable promise for fostering the rise of cooperation within a group by jointly punishing those who exploit others. Moreover, similar behaviours are seen in real life, such as in trade sanctions and police action. However, as noted, a problem arises when we consider the motivation for such peacekeeping. It is better (for the society) to have peacekeepers around, but better for an individual to not be a peacekeeper and to instead stay out of other players' conflicts like TFT does. As I mentioned in the paper, this is a kind of meta level prisoner's dilemma. To see how this works in practice, I shall again use some graphs of evolutionary tournaments (where there are populations of each strategy and these vary over time in accordance with their performance).

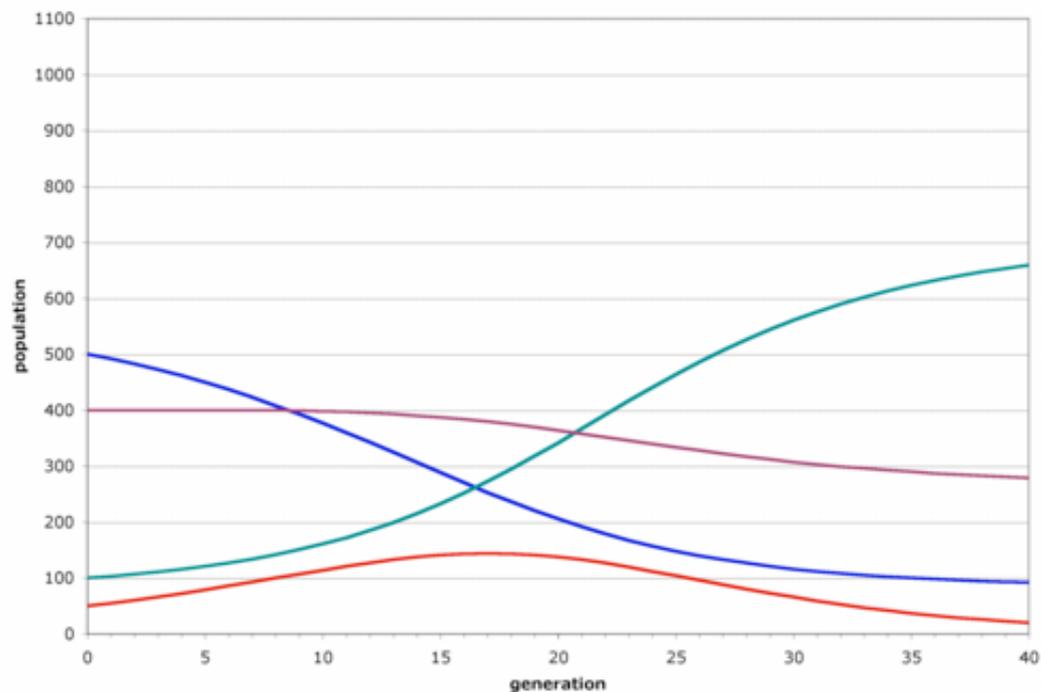
First recall the following two graphs from the paper:



The above tournament begins with large populations of TFFT (teal) and All-C (blue), as well as a small population of Spiteful-Bullies (red). The Spiteful-Bullies score very well initially at the expense of the All-Cs. This leads to the Spiteful-Bullies getting a strong hold in the population while eliminating the All-Cs. TFFT gains population due to out-performing the All-Cs, but the gain is much smaller than that of the Spiteful-Bullies.



When the TFFTs are replaced by Police (purple), things are quite different. The Police serve a peacemaking role, retaliating on the Spiteful-Bullies when they exploit the All-Cs. This leads to the extinction of the Spiteful-Bullies and an overall small gain for the Police. However, while the Police gain about as much in population as the TFFTs did, they were actually having much worse scores during this period, since they were having long periods of mutual defection with the Spiteful-Bullies. The reason that they gain is just that every other strategy is also doing quite poorly in this situation. This becomes very clear if we now replace 100 of these Police with TFFTs:



Here the 400 Police still manage to keep the Spiteful-Bullies from gaining control (although it is a close thing). However, their peacekeeping took its toll on their scores, letting the TFTs gain a great deal of population. The TFTs get the benefits of having the Police in the population and pay none of the price.

As remarked in the paper, one way to keep the Police present in the population is to eliminate this meta-level prisoner's dilemma, through *meta-peacekeeping*: having strategies that defect on a strategy like TFT which does not perform a peacekeeping role. We can indeed come up with strategies which are both peacekeepers and meta-peacekeepers, but even this just shifts the problem up another level. For in a population with meta-peacekeepers, peacekeepers, TFTs, the meta-peacekeeper population will decay into regular peacekeepers (to avoid having mutual defection with TFT) and then these will decay into TFTs. This is not merely a problem within the evolutionary framework that I am using here (though it does show up very clearly in this way). It is a general problem in organising people to implement sanctions against exploitative parties.

Absolutist

What we really need is a strategy which is a peacekeeper, a meta-peacekeeper, a meta-meta-peacekeeper and so on. Rather surprisingly, there exists such a strategy which is expressible in the strategy specification language that I have been using. Moreover, it is very simply expressible. Continuing with the tradition of rather anthropomorphised strategy names, I call this strategy *Absolutist*.

Absolutist:

Defect on c iff c has ever cooperated with someone when you defected, or *vice versa*.

In the first order language that I used in the paper to formally define strategies, we can define absolutist as follows:

$$\exists t \exists j$$

$$(\begin{aligned} & D(r, j, t) \wedge \neg D(c, j, t) \\ & \vee \\ & \neg D(r, j, t) \wedge D(c, j, t) \end{aligned})$$

Or, if we let \oplus denote 'exclusive or', then the strategy is simply:

$$\exists t \exists j D(r, j, t) \oplus D(c, j, t)$$

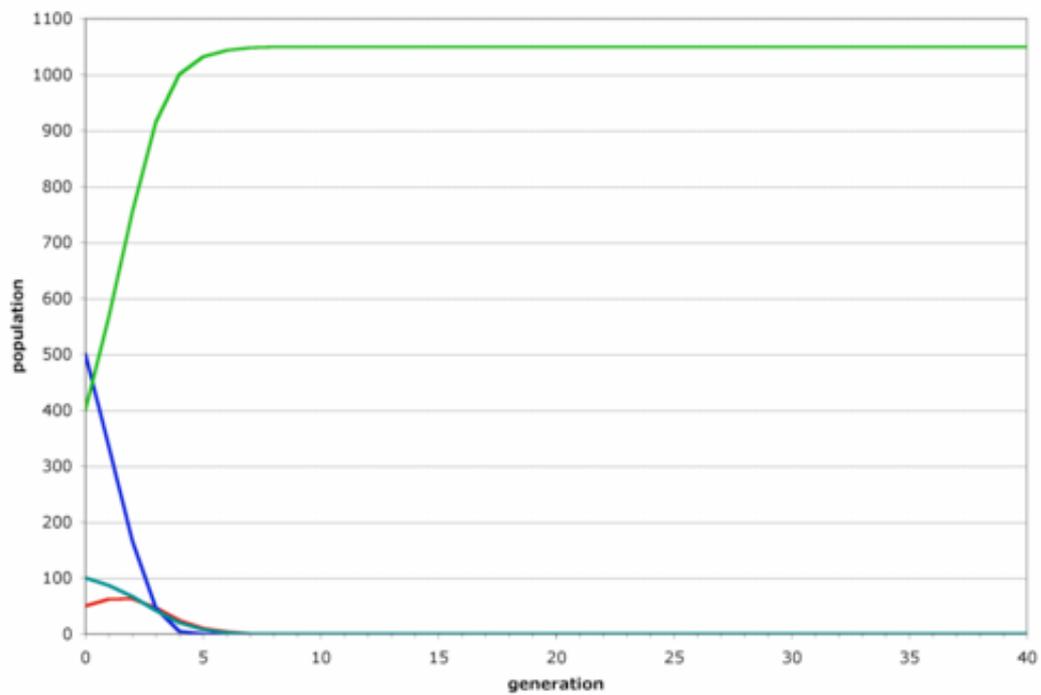
Since nothing has happened before the first round, Absolutists will initially cooperate with everyone. On the second round, they cooperate with everyone who cooperated with all and defect on those who defected on anyone. In this they are performing a peacekeeping role. On the third round, they again defect on anyone who didn't act like they have done so far: thus they defects on those who did not perform this kind of peacekeeping and are thus meta-peacekeepers...

Now note that Absolutist is a *loyal* strategy for it can never defect on another player who uses the same strategy. Interestingly, it is also the minimal *loyal* strategy: it is impossible for any strategy to defect more often than Absolutist and yet still be loyal. We can see this because Absolutist defects on someone as soon as they do anything differently and never stops defecting on them after that.

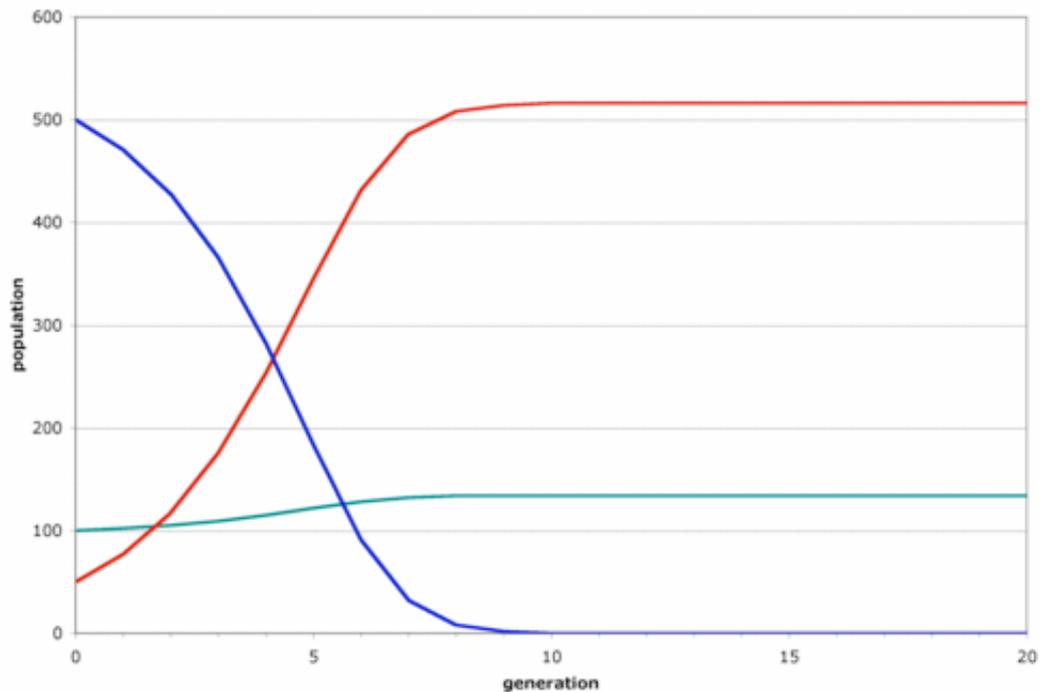
Note also that Absolutist is an *unforgiving* strategy for it never stops defecting on someone once it begins. In this it is like Grim (which happens to be the minimal individually nice strategy). In contrast, Police is a forgiving peacekeeping strategy.

Absolutist is not *individually nice* or *communally nice* for it can defect on strategies like All-C which never themselves defect.

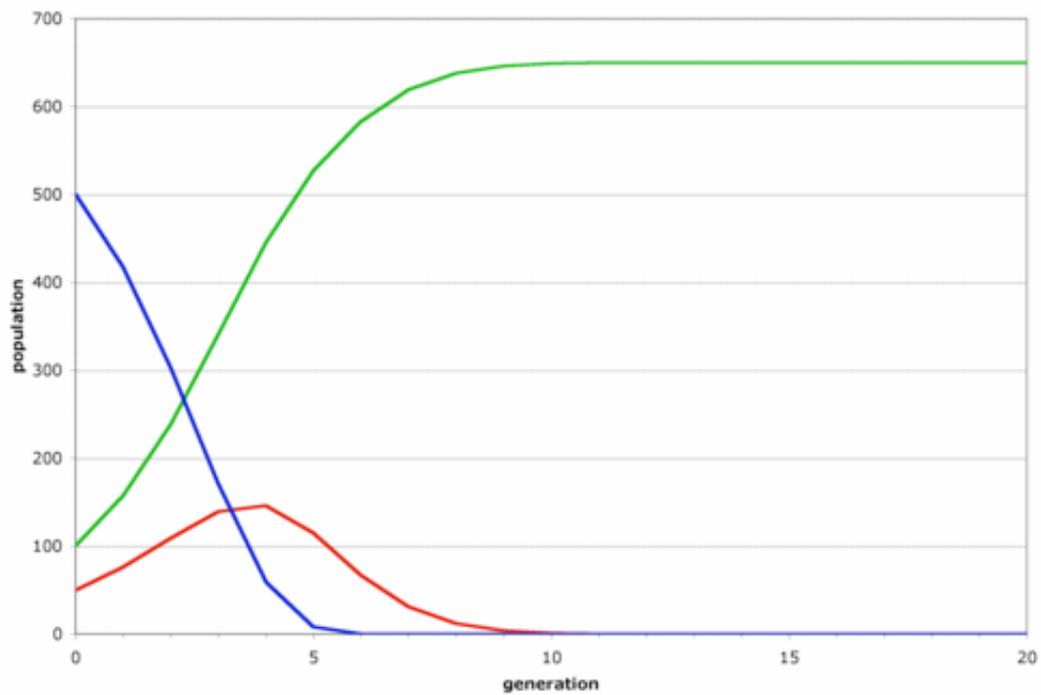
So how does Absolutist fare? Consider the previous tournament with 400 Police, 100 TFTs and 50 Spiteful-Bullies. Lets replace the 400 Police with 400 Absolutists:



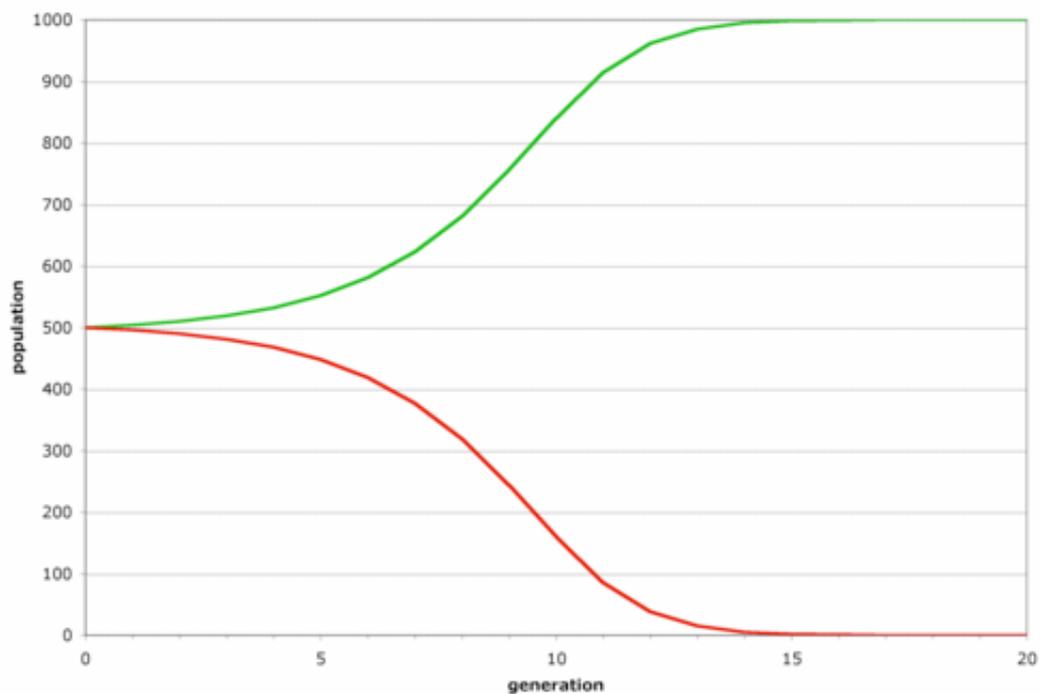
The peacekeeping of the Absolutists (green) prevents the Spiteful-Bullies from gaining a foothold, and yet there is no decay of Absolutists into TFTs. Instead, the Absolutists come to completely dominate the population within 10 generations. Moreover, this success continues even when the initial population is greatly reduced. The graph below shows the results of a mere 100 TFTs in the population with the All-Cs and Spiteful-Bullies:



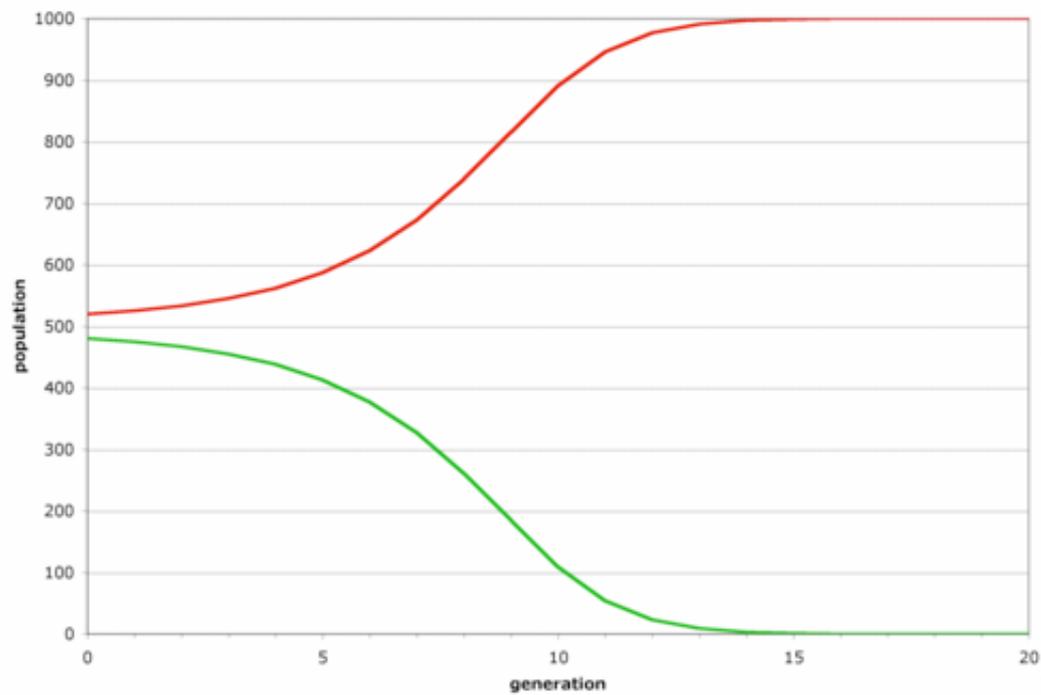
Here the TFTs survive but do not flourish, while the Spiteful-Bullies come to dominate the population.



If the TFTs are replaced by Absolutists (as above), they do remarkably well, completely eliminating all other strategies from the population. As discussed above, Absolutists defect on a strategy continually once they behave differently. In the above mix of strategies this happens after the second round of each game. Thus the Absolutists are exceptionally clannish, continually cooperating with each other and continually defecting on everyone else. In doing so, they can overcome any initially defecting strategy which starts out in lesser numbers.



This is seen clearly in the above graph, where 500 Absolutists eliminate 500 Spiteful-Bullies. It begins slowly, but once the Absolutists get enough of a lead in population, the domination of the population is very fast.



However, when the Absolutists have an initial shortfall in population (even a very minor one), the Absolutists' propensity for conflict can lead to their rapid demise.

Further observations concerning Absolutist

In the paper, I contrasted the exploitative strategies like All-D, Spiteful-Bully and Vulture with the more cooperative strategies like All-C, TFT and Police. Absolutist does not fit neatly into this dichotomy. Absolutist is loyal, but not *nice* in any of the other technical senses. It defects a lot and is in some sense intolerant of any other strategies. However, it will not initiate defection in a population and if it has enough of an initial advantage in population (as in most of the above tournaments), then it very quickly leads to a population in which there is no defection at all. Moreover, this final population is extremely evolutionarily stable.

I am thus unsure as to whether to consider Absolutist as 'one of the good guys'. It is a very powerful and interesting strategy whose presence is very dangerous: its unforgiving nature and propensity for defection means that it will lead to more defection than any other loyal strategy within a game of SIPD, but over the course of an evolutionary tournament, it may well lead to less defection overall.

Also, Absolutist is only really a tenable strategy when there is no noise in the players' information or actions. If there is a small chance of error in information or action, then Absolutists end up defecting on each other and can never break out of this. The strategy would have to be modified to take account of the probability that another strategy was trying to act as an Absolutist and this would involve considerable additional complexity to the strategy (as well as being inexpressible in my language). This is also true in cases of no noise, but where other strategies are allowed to deliberately involve randomness, or cases where the network of interaction is weakened so that not all pairs of players interact. Thus, to some extent, the power of Absolutist is an artifact of this particularly clean set-up.

We might also consider some small modifications to Absolutist. For example, could we make a forgiving version of Absolutist? We can certainly change the definition so that Absolutist doesn't look at all past rounds, but only the immediately prior one. As it happens, this doesn't change much in practice as Absolutist has such complex behaviour that very few other strategies can fall into line with it after initially doing something differently. There are other ways in which we could try to make Absolutist forgiving, but they would involve making quite a few particular choices about the strategy and I have not had time to follow it all through. Almost certainly they would require a very expressive language to implement them and quite some computational resources.

Another possibility is to emulate the so-called *Suspicious Tit-For-Tat (STFT)* which is like TFT, but begins by defecting. Thus:

Suspicious Absolutist:

The same as Absolutist, but begins by defecting.

Interestingly enough, this strategy has almost identical results to Absolutist in all the tournaments above. Indeed a group of 500 Suspicious Absolutists can beat 500 TFTs in much the same way as 500 Absolutists defeat 500 Spiteful-Bullies.

Suspicious Absolutist is not loyal, but it only defects on those of the same strategy on one round (the first). We could thus say that it is *loyal in the limit*. In sufficient numbers it eliminates pretty much any other collection of strategies, even if none of the others will initiate defection. In an evolutionary tournament, it then leads to a situation of near-perfect cooperation and even more evolutionary stability than the normal Absolutist. This strategy is more ethically dubious than Absolutist, but seems to get even more striking results. I haven't done much analysis on it, but it is certainly interesting.

In conclusion, SIPD is a model of social interaction that is very simple and very close to the standard IPD framework. However, it leads to some surprisingly sophisticated and interesting behaviour, such as that of Vulture, Police or Absolutist. My research interests are pulling in other directions at the moment and it is unlikely that I will get to investigate it further, but I do think that it would reward further inquiry.

Appendix I: a note on strategies expressed in this first order language

In my dabbling, and in my observation of evolved strategies, I have noticed that there are often many ways of achieving a certain behaviour. For example, I typically write Grim as:

$\exists t D(c, r, t)$ c has defected on r at some time

But it can be expressed just as well using the (apparently rather useless) term ' $D(r, c, p)$ ':

$D(c, r, p) \vee D(r, c, p)$ c has defected on r last round, or r has defected on c last round

This strategy will begin defecting on c the round after c first defects on r and will

continue to defect thereafter. Thus, this is just another way of writing Grim: one that uses r 's last action as its memory rather than using a quantifier.

Also note that there are many strategies that are inexpressible in this language. Anything too quantitative is inexpressible. For example, the version of *Tit for Two Tats* that waits for two defections before retaliating (no matter how far apart these come) is inexpressible. I could add predicates (such as $B(x, y)$ which is true when x is a round that is before round y), however I resisted this on grounds of the additional inelegance for such a small gain in expressibility. My language isn't perfect, but it does have some merits in the way it shows logical relationships between different strategies and easily combines multiple strategies into one.

Appendix II: my program and data

[dilemma.zip](#) is an archive containing:

- the program I used to test the strategies,
- the source code (in C),
- some help files on using the program (from the command line),
- the strategies used here (and many others),
- the tournaments run here (with their results and the graphs).

The program can also perform a form of genetic evolution of strategies. The strategies are played against each other with the best strategies going forward to the next generation, along with some new random strategies and some modified versions of the successful strategies. My results with this had been very preliminary (which is why none are included here), but TFT was very successful since it has such a good strength to complexity ratio. Populations seeded initially with many Vultures or Absolutists or Suspicious Absolutists were also very successful in avoiding invasion.

If you would like some help getting this to run, then try emailing me, though I can't promise to have all that much time to help you. I wrote this a long time ago and have not programmed much since.

References

Toby Ord and Alan Blair, 'Exploitation and peacekeeping: introducing more sophisticated interactions to the iterated prisoner's dilemma', presented at the *World Congress on Computational Intelligence*, (Honolulu, 12–17 May 2002).

This paper was originally published to the author's website on 13 Aug 2006.
<http://www.amirrorclear.net/academic/ideas/dilemma/index.html>