

Is it in our self-interest to being good?*

#20387090

July 27, 2015

It is in our self-interest to be good. I will argue this thesis through two parts. First, being good reaps more long term rewards. Second, being good benefits more as a whole, rather than individually.

The main arguments are drawn from two texts: the philosophy textbook, *The Fundamentals of Ethics* ¹, and the textbook used in CS 486 Introduction to Artificial Intelligence, called *Artificial Intelligence: A Modern Approach* ² (referred as *AIMA*). There are similar ideas from two different fields, different approaches, but same conclusions.

As in the case with mathematical proofs, I also start by giving definitions to the thesis.

Definition of *our* The term *our* refers to a *rational agent*. A rational agent is one that always acts to achieve the best outcome or, when there

*word count: 1834

¹Shafer-Landau, R. (2014). *The Fundamentals of Ethics*. Oxford University Press

²Russell, S.J. & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Prentice Hall

is uncertainty, the best expected outcome³. This simplifying assumption is made because emotions often results in irrational decisions.

The philosophy text defines rationality (according to utility theory) slightly differently. Rational means to act consistently on one's preference orderings.

Both are similar, but I highlight that the *AIMA* definition includes acting under uncertainty. This will be useful later.

Definition of *good* Being *good* means having insight and looking out for the long term. Consequently, as I will explore later, it means not being selfish (which is always trying to maximizing the reward for yourself).

Part I: Long term rewards

The key insight into gaining more long-term rewards is that choosing the best action (being selfish) doesn't necessarily mean the best rewards. Some actions do not seem to be the most beneficial action at the moment, but it creates future opportunities.

Philosophy Textbook The first argument is from the course textbook; noting the best action isn't to be greedy, but to be self-interested.

The previous sections talk about different social theories, and this quotation below concludes about the underlying commonalities from those different social theories⁴. I won't go in detail about what exactly the different social

³Definition from page 4 of *AIMA*

⁴page 215 of *The Fundamentals of Ethics*

theories are, but merely look at the textbook author's conclusion on their shared commonalities:

The key to understanding [the social negotiation theories] lies in the idea that contractors are, above all, *rational and self-interested*.

Being self-interested is not the same thing as being selfish. Being self-interested is having a strong concern for how you well you are faring in life. Being selfish is placing far too much importance on your own well-being relative to the interests of others.

While being selfish maximize your own rewards, different social theories have similar commonalities in that the best decision isn't being selfish but rather being self-interested. This is saying the best action isn't the same as the selfish action.

This supports the idea that there are future benefits of being good, at a cost of gaining less benefit at-the-moment. Being selfish maximizes the present-moment reward but at a cost to the future rewards.

Artificial Intelligence Textbook The second argument is from the AIMA textbook; on *active reinforcement learning*, supporting the idea that the best action isn't always best for yourself.

I define some terms for readers in this paragraph. A passive learning agent has a fixed policy that determines its behaviour, whereas an active agent must decide what actions to take. An utility function (also known as

reward function), is a measure of the agent's performance. A state encloses the result of an action. When actions and the search space are well-defined, an utility function can be learned by value iteration.

The experiment setup is to finding an optimal path through a grid. But the grid is non-deterministic. Each action the agent take may or may not result in the desired grid. (E.g. agent choses to move left, but the move happens probabilistically.)

Take it without proof that in this setup, an optimal action can be extracted by one-step look-ahead to maximize the expected utility. After the 276th value iteration and learning the optimal policy at each step, the policy converges; and the agent sticks to using that policy, never learning the utilities of the other states. We call this agent the **greedy agent**. The action it chooses is optimal for each state (plus one look-ahead). But interestingly, the agent *does not* learn the true utilities or the true optimal policy.

Repeated experiments show that the greedy agent very seldom converges to the optimal policy⁵!

How can it be that choosing the optimal action leads to suboptimal results? The answer is that the learned model is not the same as the true environment. An agent must make a tradeoff between exploitation (to maximize its reward) and exploration (to maximize its long-term well-being). Pure exploitation risks getting stuck in a rut. Pure exploration to improve one's knowledge is of no use if one never puts that knowledge in practice.

⁵pg. 839 AIMA 3rd ed.

The greedy agent's selfish strategy to always to choose the most rewarding action doesn't imply the best rewards, because you miss out on other opportunities. The optimal action is to also account of benefits from the future, through exploration.

Part I: Conclusion The two textbooks shares the conclusion that the best action isn't necessarily one that maximizes your current rewards, but one that factors in the future long term rewards.

Part II: Collective benefits

The main idea in this part is it is in our interest to be good because being good benefits the collective more as a whole. Rewards to a group of rational agents increase with cooperation.

I explore the readings from *Collective Action Problem* and cast the problem as a prisoner's dilemma problem, a well understood case of Nash equilibrium. Then applying the results from what we learned about the prisoner's dilemma.

Collective Action Problem In the *Collective Action Problem* ⁶, people behave irrationally because it is their best action, but at a cost to others.

The reading suggests that we are lazy individuals who don't want to be the one putting in the effort to be good, unless everyone is being good. And

⁶p. 366 Constellations Volume 7, Number 3, 2000 Ideology and Irrationality. <http://homes.chass.utoronto.ca/~jheath/ideology.pdf> Accessed June 3, 2015

if everyone is being good, then certainly it is easier to not be good and reap the fruits of others' labour. A catch-22 situation.

This problem can be casted to a prisoner's dilemma problem, a well known case of Nash equilibrium, so we can apply the results learnt from prisoner's dilemma. The cast is as follows.

1. Everyone being lazy and you are lazy be interpreted as you both defect.
2. Everyone being good and you are lazy be interpreted as your friend cooperates and you defect.
3. Everyone being good and you are good be interpreted as you both cooperate.
4. Everyone being lazy and you are good be interpreted as your friend defects and you cooperate.

Now I apply results learnt from analysis on the prisoner's dilemma.

Prisoner's Dilemma A key assumption made is the prisoners have no opportunity to reward or punish their partner and that their decision will not affect their reputation in the future.

This assumption is not realistic in the real world, unless the punishment is death. Similarly, this applies to the assumption in the collective action problem that you won't be noticed from not cooperating. Realistically there would be negative impression from the collective good on your behaviour, which is a negative reward, thus it benefits you to cooperate with your peers.

But what if the assumption holds?

Since betraying a partner offers a greater reward than cooperating with him, all purely rational selfish prisoners would betray the other, and so the only possible outcome for two such prisoners is for them to betray each other. Interestingly, once the time element is added to the consideration, as in the case in iterated prisoner's dilemma, then cooperation is the only rational outcome, as well explained in a section on the iterated version of prisoner's dilemma, in the *Multi-agent Interactions* chapter ⁷.

The game of the prisoner's dilemma is played a number of times. Each play is referred to as a round. Critically, it is assumed that each agent can see what the opponent did on the previous round: player i can see whether j defected or not, and j can see whether i defected or not. Now, for the sake of argument, assume that the agents will continue to play the game forever: every round will be followed by another round. Now, under these assumptions, what is the rational thing to do?

If you know that you will be meeting the same opponent in future rounds, the incentive to defect appears to be considerably diminished, for two reasons.

Reason 1: If you defect now, your opponent can punish you by also defecting. Punishment is not possible in the one-shot pris-

⁷p. 118 Multi-agent Interactions, in An Introduction to MultiAgent Systems.

oner's dilemma.

Reason 2: If you 'test the water' by cooperating initially, and receive the sucker's payoff on the first round, then because you are playing the game indefinitely, this loss of utility (one util) can be 'amortized' over the future rounds.

When taken into the context of an infinite (or at least very long) run, then the loss of a single unit of utility will represent a small percentage of the overall utility gained. *So, if you play the prisoner's dilemma game indefinitely, then cooperation is a rational outcome.*

The summary is that selfish rationality creates the worst possible result. Maximizing individual reward leads to the least possible reward. There are more benefits, both individually and collectively, if they both cooperated. This supports my argument that being good benefits more as a whole, rather than individually.

Interestingly, humans display a systematic bias towards cooperative behavior in this, which goes against the selfish rationales⁸. It makes sense because the idea of reward/punishment comes into play, because the assumption that prisoners have no ability to reward or punish their partner is not realistic.

⁸Tversky, Amos; Shafir, Eldar (2004). Preference, belief, and similarity: selected writings. (PDF). Massachusetts Institute of Technology Press. Retrieved July 26, 2015.

Part II: Conclusion A rational agent's self-interest is to be good because it is more rewarding. Rational agents in a group making decisions forms a Nash equilibrium, so being good means cooperation and results in more rewards. In the face of punishment from other agents, making selfish decisions decreases future expected rewards, because other agents are less inclined to cooperate when facing a selfish defecting agent.

Conclusion

I conclude that it is in our self-interest to be good, as analyzed in the above two parts. Part I concluded that being good reaps more long term rewards, and Part II concluded being good benefits more as a whole.

Works Cited

- [1] Shafer-Landau, R. (2014). *The Fundamentals of Ethics*. Oxford University Press
- [2] Russell, S.J. & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Prentice Hall
- [3] p. 366 Constellations Volume 7, Number 3, 2000 *Ideology and Irrationality*. <http://homes.chass.utoronto.ca/~jheath/ideology.pdf> Accessed June 3, 2015
- [4] Wooldridge, M. (2002). *An Introduction to MultiAgent Systems*. Wiley
- [5] Tversky, Amos; Shafir, Eldar (2004). *Preference, belief, and similarity: selected writings*. (PDF). Massachusetts Institute of Technology Press. Retrieved July 26, 2015.