

# Evolution Reinforces Cooperation with the Emergence of Self-Recognition Mechanisms: an Empirical Study of the Moran process for the Iterated Prisoner's Dilemma - Response to reviews

January 11, 2018

## 1 Comments from editor

We have a problem with reviewers declining to re review the paper, which forced us to invite first time reviewers for the revision. The new reviewers have a number of problems with the manuscript. In particular you need to work much harder to justify the relevance of your work to understanding iterated games.

We open our second response to the reviews by highlighting that in the original review no comments were raised about the relevance of the work. **Specifically all previous reviewers highlighted that the work was noteworthy and worth publishing.**

Regarding the relevance of the work, comments from each reviewer were:

- The paper is well written, the results seem solid and overall very interesting. The main issues with this work, in my opinion, lie in the presentation:
- The analysis in the paper is well designed and clearly explained. It does an excellent job of leveraging the work of many years and many different researchers to create new and significant results. And, the paper serves as a good model of how to make all code and data available, making the results easily reproducible and expandable to other researchers curious about the behaviour of additional strategies.
- The reported experiments fill an important hole in the understanding of IPD evolution and are sorely needed. The authors also take pains to ensure the data are publicly available, which is highly appreciated.

Most requests were on the subject of presentation of the results as well as requests for specific pieces of further analysis. All of these were addressed.

This latest set of comments from the reviewers are mainly of two types:

1. Further minor requests for presentation modifications (specifically the remaining reviewer from the first round only made requests of this type).
2. Questions about the suitability/impact of the work.

As we detail below we have addressed all requests of the first type. For the requests of the second type, from the two new reviewers who are perhaps not experts in the IPD, we have made some modifications to the manuscript but our main responses are to answer the questions posed by the reviewers.

Specifically we describe how the Moran process is a well-studied model of evolutionary processes and that this paper is a major contribution to the literature. It can be thought of as a tournament on Moran processes equivalent to the seminal work by Robert Axelrod in the 1980s.

Robert Axelrod's work on the Prisoner's Dilemma tournaments (described in the literature review) which has been cited more than 34,000 times considered pair-wise matches between strategies. Our work here is much more computationally-intensive and accesses deeper questions of evolutionary stability.

Furthermore, Moran processes for the IPD have been frequently studied from a theoretical point of view and our results extend that to strategies that are difficult (or impossible) to study analytically.

## 2 Comments from first referee

Figures are much improved! It was a good idea to move some of them to the Supplementary Material. I still think the title is not representative of the work -- you should reconsider at least changing the subtitle. You are not studying the Moran process; you are studying Prisoner's Dilemma strategies. I

Thank you for agreeing to review the paper again.

We have changed the subtitle to be: 'Empirical Study of strategies in the Moran process for the Iterated Prisoner's

also still think the description on page 3 of how you found the TF1, TF2, and TF3 strategies is confusing. It has a frustrating level of detail. You should either have less detail or more detail. This is not a crucial point, however.

The description of how we obtain the strategies is quite minimalistic and all training code is documented and available online (archived according to best practice). The rest of the description in this section is specifically describing each strategy (accompanied by the corresponding diagram). This description is quite important as it highlights the handshake mechanism which is one of the main findings of the work. We have considered changing this further but as this is not a crucial point we would prefer to leave it as it is.

## 3 Comments from second referee

I couldnt tell how each round is played. Does each individual play against only one other (randomly chosen) individual or against all  $N - 1$  other individuals?

Yes they play against all  $N - 1$  other individuals. This is a standard Moran process of this type and is described in Algorithm 2. The mathematical equations (2) and (3) also ensures this to be the case. We have added a sentence clarifying this:

"Individuals are **selected** according to a given fitness landscape, in this work it is their total utility against all other individuals in the population."

In the introduction the authors state "The results and insights contained in this paper would be difficult to derive analytically. But in Section II the authors present several formulas (Eq. (2)(7)) and analytically compare a couple of strategies. Why cant these formulas be used to analytically compare any pair of strategies listed in the paper?

The purpose of this section (II) is to show that the simulations match the theory well in cases that are analytically accessible as a validation. This is in addition to the extensive test suite for the software which covers 100% of the Axelrod library.

As an example of a case where a theoretical result is more difficult to obtain, we note that when one or more of the strategies is stochastic then the utility function is also stochastic and would need to be analyzed statistically, assuming that the utility distribution could be analytically derived, which is unrealistic in many cases. Additionally long memory strategies are difficult to analyze theoretically, which is why they are often avoided in the literature, and why a number

of misconceptions about them have persisted (including the strength of Zero determinant strategies, a point raised by the reviewers). We approximate the utilities with a large number of simulations, a commonly used technique.

After careful examination of how this section is written we do not see how it can be improved and feel it is clear. Specifically this sentence: ‘This demonstrates that assuming a given interaction between two IPD strategies can be summarised with a set of utilities as shown in (1) is not correct.’.

The authors should avoid jargon because it doesn't help readability. For instance, in Section III(c) the authors state this is explained by the need of strategies to have a steady hand when interacting with their own kind. Acting stochastically increases the chance of friendly fire.. What is a steady hand and what is friendly fire in an IPD game?

We have replaced this comments with more standard language.

Here are some of the questions I have after reading the paper. (The first two points deal with the realism of the models; the last two with exploiting the papers findings.)

- Suppose 10 people decide to play an IPD game. The authors only allowed pairs of strategies in the population. If  $k$  individuals happen to pick the same strategy, what is the likelihood the other  $10 - k$  people will pick the same alternative strategy?
- The strategies evolved under a Moran process. Under a Moran process each round one player must change his strategy 1 . Is there an example of a real-world social dilemma or a human experiment where strategy changes resulted from a Moran process or something similar? (This issue is particularly important because the authors specifically trained some IPD strategies to excel at the Moran process.)

This is the standard Moran process and the reviewer is asking about one of the many extensions of the standard process, e.g. to multi-strategy populations or other dynamics (such as a best-reply dynamic) which may allow players to dynamically choose new strategies each round. The standard is an evolutionary game where players have a fixed strategy throughout their lives with the biological analogy of bacteria with different genetically coded traits and also cancer, the latter of which is referenced in the introduction.

Suppose the population is using either strategy A or strategy B. Mutation was not allowed in the authors simulation so no other strategies can be used. How useful then is it to know some other strategy C can successfully invade strategy B if strategy C isn't available?

In biological situations it's often assumed that mutations are relatively rare so it makes sense to run a process to fixation rather than look at many variants at once in that context. In fact, the rate of mutation is typically set to  $1 / N$ , and the population runs for  $O(N)$  iterations before converging to one of the absorbing states (and often far fewer), so there would typically be very few simultaneous existing mutations. There are other contexts where there would be more than two strategies at a time and they are interesting, but not the focus of this work.

Additionally, if there is mutation then different techniques are needed. The Markov process no longer has absorbing states the stationary distribution has to be studied instead. This is an interesting case but would be a work of similar length, and this work is the limiting case where mutation goes to zero.

How useful is it knowing that strategy A can or cannot resist invasion by strategy B when players rarely know exactly what strategy their opponent is using? In particular, before the game begins players won't know what strategy others intend to use.

This feels like a question about the entirety of the research field on the IPD.

We refer the referee to the prior literature in this field and in particular highlight that one of the main goals of this area are to understand properties of environment and strategies that enable for emergence of a given type of behaviour. In the case of this work, we show that in evolutionary environments, handshake mechanisms can emerge which reinforces the emergence of cooperation.

We have added a sentence to the paper further explaining how this work offers a population based perspective.

## 4 Comments from third reviewer

I saw no significant contribution in the field of evolutionary computation itself. The whole paper would have been a great *\*experimental section\** after some new idea or scheme.

We have justified the contribution in many ways but agree that this paper is a strong numerical experimental paper (as evidenced by the title) following the novel idea of creating tournaments with Moran processes as ways to score pairs of strategies.

Indeed, results appears sensible, when not obvious. Collective strategies perform well, the usual suspects are hard to invade. However, generally speaking, the main problem is that with IPD is extremely hard to infer meaningful data from the mere statistical results, as it would be hard to draw conclusions about the "best opening" looking at students playing chess. Indeed, such an analysis makes perfect sense as soon as one restricts the domain to a specific grandmaster, or a specific top-level tournament.

A lot of the original and contemporary research in this area uses the IPD to show why/how cooperative behaviour may evolve. This has often been through large numerical analyses such as these. However, none of them (as discussed in the literature review) consider Moran processes on this scale. Furthermore, one of our main conclusions/findings is that handshake mechanisms can also appear through evolutionary processes.

This is directly analogous to "the best opening" in chess. Moreover, our numerical results call into question several common assumptions in the literature, e.g. that it's enough to only consider memory-one strategies (it's demonstrably not the case), that ZD strategies are the best yet they don't win here, etc. This work specifically identifies fruitful and valuable new directions for future work.

Finally, we show that "collective strategies" directly emerge from evolutionary processes and are reinforced by standard evolutionary mechanisms, which is a major theme of the paper and very relevant to the study of the evolution of cooperation.

The *\*Moran Processes\** should be briefly introduced, let alone mentioned in the abstract before explaining how their *\*fixation probabilities\** have been obtained.

The Moran process is briefly introduced in section A. It is also described in Fig 1, Algorithm 2 and also mathematically throughout equations (1) - (7). Furthermore numerous background references are given.

The concept of *\*fixation probabilities\** should be properly introduced, or, more simply, authors can talk about *\*the probability with which a given strategy takes over a population\** without mentioning the specific term.

We have added a description of this to the main text.

As a good journal paper could be read several years after being written, authors should avoid relative expressions like "recent" or "recently", and use absolute timings, such as "in 2012" or "in the beginning of 2010s".

We have modified this terminology.

Population sizes range from 2 to 14. While 2 is an obvious choice, being the smallest possible *\*population\**, the use of 14 is puzzling. Twice the days in a week?

There needs to be an upper bound, 14 was chosen as the highest number reasonable given the computational cost of running these models. Also, the lower values of N (greater than 2) are shown to already indicate the behaviour of higher values (see Figure 7).

\*Moreover we present a number of strategies that were created via reinforcement algorithms\* -> Which number?

16. We have added this to the text.

\*Zero determinant strategies are not particularly effective for  $N > 2$ \* -> do you mean that they \*only\* works good for  $N=2$ ? Effective as invader, defenders or both? Generally speaking, the whole text following \*While the results agree with some of the published literature, it is found that\* is almost completely unintelligible in such an early section of the paper.

We have clarified the first bullet point (both invaders and defenders). We disagree with the comment that this bullet point list is unintelligible (at present at least 4 reviewers have read this summary of the findings and no one has said that it is unintelligible). If this is a sticking point for the referee we defer to the editor to decide how and if these short sentences should be rewritten.

Use underscore, or dash, for multi word variable, eg. \*player one\* -> \*player\_one\*.

We have modified the pseudo code to follow this suggestion.

uthors use a panmictic population, but the effect of population topologies would have been interesting. Also mixed populations should be checked (eg. 50% TFT, 40% ETFT, 10% Grudger).

Both these points are already mentioned in an area for possible future work. Actually carrying out this analysis as part of this article is an unreasonable request (the paper would end up being extremely large).

Algorithm 1 has a return at line 8, nullifying the previous three for loops. Maybe \*save\* or \*record\* would be better?

We have changed this to use yield.

\*When studying evolutionary processes it is vital to consider  $N > 2$  since results for  $N = 2$  cannot be used to extrapolate performance in larger populations\* sounds more like a truism than a conclusion.

We agree. This paper presents the first empirical evidence and as discussed: in the literature this is not well understood (results are often only considered for  $N=2$ ).

In EAs, the term \*offspring\* is usually preferred to \*child\*.

We have modified this.

Do not use **bold** for emphasis.

We have removed the bold.

Don't start a section with \*Figure 1 shows a diagrammatic representation of the Moran process\*, rather, tell what a \*Moran process\* is, and then refer to figure to better illustrate the idea.

We have modified this.