Editor Remarks to Author:

Both reviewers find much to recommend in the paper, but both have important suggestions for a major revision, which must be carried out before another review.

Reviewer Comments:

Reviewer #1:

Suitable Quality?:

Yes

Sufficient General Interest?:

No

Conclusions Justified?:

No

Clearly Written?:

Yes

Procedures Described?:

Yes

Supplemental Material Warranted?:

Yes

Comments:

This paper introduces new formal results for reactive with longer memories based on a novel algorithm to identify partner strategies in the repeated prisoners dilemma. Reactive strategies are a less complex subset of automata strategies that only condition on the co-player's last moves rather than conditioning both players' moves (e.g. Tit-for-Tat is a reactive strategy but Win-Stay-Lose-Shift is not). Prior work has mostly focused on strategies with short memory, but it is possible in practice that this is insufficient to cover the kinds of strategies used by people. This paper aims to close this gap by creating an algorithm that lets them tractably analyze reactive strategies with longer memories.

The author's key insight is that for any opponent's reactive strategy, there is an associated self-reactive strategy that yields the same payoffs. They build on this insight to find tractable conditions for testing whether a reactive-n strategy is a partner strategy (cooperative equilibrium strategy). The number of conditions seems to grow exponentially with the size of the memory, but the authors claim these conditions boil down to reducing the cooperation rate proportional to the number of defections in the memory. Finally, the formalism is applied to reactive-n counting strategies, which are a subset of reactive-n strategies, and the formalism is tested empirically.

Overall, I think this is interesting work that may move the field forward. However, I have some significant concerns and reservations about the implications of these analyses and the conclusions that can be drawn, which currently discount the the impact this work might have:

First, the authors make some strong claims about the value of memory when comparing reactive-n and reactive-n counting strategies. However, this comparison seems to be made by naively mapping the variable n as a notion of memory. The reactive-n counting strategy only need to represent a single number and increment / decrement that number based on whether they experience a defection or cooperation. To be more concrete, a reactive-256 strategy will require 256 bits of memory, while a reactive-256 counting strategy only requires 8 bits of memory. This is one of the key results in the paper, taking about 1/4 the length of the abstract - either this notion must be refined, or the paper does not support one of its key claims.

We somehow need to address this in the Model/Discussion.   
Also, do we need to tone down the Abstract/Significance statement?

Second, a key motivation for longer memory strategies (according to the authors) is the need to study longer memory strategies. They write, "Longer memory seems particularly relevant for noisy games, where people occasionally defect because of unintended errors". However, there is no discussion of whether this framework applies to the repeated prisoner's dilemma when there are errors. Do reactive-n strategies that have longer memory actually perform better under higher error rates? This should be investigated analytically and empirically since this is a key motivation for the relevance of this class of strategies.

Reply:

1. Empirical work with errors is only justification why we study longer memory. In addition, there are good theoretical reasons.
2. Analytically: Explain why results with errors are harder (mention that even for memory-1 the exact Nash are not known).   
   Results for approximate Nash in the limit of rare errors; Results on equalizers

Empirically: Figure S8: For reactive-n having a larger n is beneficial when there are errors in the parameter regions where it makes sense. For reactive-n counting strategies, higher n does not help.

Third, another conclusion from this work is that knowing the sequence of moves matters, not just the count. But it seems from looking at the conditions, to be a partner strategy, the order does not matter in a straightforward way (equation 1). For reactive-2 strategies, all that matters is the sum of p\_dc and p\_cd with equal weight. Likewise, for many of the conditions in reactive-3 strategies in equation 2. Finally, looking closely at the conditions in equation 2, it does not seem to generalize reactive-2 strategies where the cooperation rate needed to decrease by c/2b - is there some other pattern that applies, or is that intuitive explanation only valid for reactive-2?

Reply: Check what Eq. [2] looks like in the special case that we use a reactive-2 strategy.   
In addition, explain that there is some kind of pattern: proportion of D’s on the left hand side explains by how much you need to reduce cooperation probability on the right hand side.

Fourth, when looking at the simulated results, the abundance of reactive-2 strategies seems to differ substantially between p\_cd and p\_dc, but this is not predicted by theory (and neither is p\_cd < p\_dd). What explains this result? I was also hoping to learn more about what it is about the timing of cooperation that makes it important. Tit-for-two-tats and other known versions make some intuitive sense, but can these intuitions be mapped onto the conditions and the framework presented here?

Reply: This is indeed not explained. And in fact, for perfect partner strategies these two versions are equivalent. However, in simulations, players never adopt perfect partners, and in that case we numerically verified that there’s a difference between (a,b,c,d) and (a,c,b,d).

1. We confirmed the effect is robust (Figure S4)
2. The effect is due to a difference in the self-payoffs (Figure S4)
3. This difference is due to the way how quickly these strategies return to mutual cooperation. (Figure S5-S7).

Reviewer #2:

Suitable Quality?:

Yes

Sufficient General Interest?:

Yes

Conclusions Justified?:

Yes

Clearly Written?:

Yes

Procedures Described?:

Yes

Supplemental Material Warranted?:

Yes

Comments on Significance Statement:

Results are limited to reactive strategies, but conclusions are stated far more broadly, especially in the significance statement.

Significance statement is actually fine, but still slightly reword it.

Comments:

The authors characterize the Nash equilibria that sustain cooperation among the space of reactive memory-n strategies, for a two-player 2x2 iterated game. The analysis results in explicit conditions for n=2 and n=3, and also for arborary n when restricted to "counting" strategies that merely count how many times an opponent has cooperated. The authors are also interested in strategy evolution in a population, and they compare their analytical results on cooperative memory-n strategies to monte-carlo simulations in the limit of weak mutation.

Finding or characterizing Nash equilibria for any sizeable space of strategies in an iterated game is a hard problem. The authors have made real progress on such a problem, when restricted to "reactive" strategies that condition only on your opponent's plays. This is very solid work, and it is made possible by extending the approach of Press & Dyson, Akin 2012, and also Park.

Perhaps the most intuitive and elegant result has to do with the requirements for a reactive counting strategy to be Nash and ensure full cooperation in the donation game. Such a strategy must reduce the chance of cooperating by a constant factor (c/nb), for every co-player's defection observed across the prior n-rounds. This is more than simply a characterization of Nash equilibria for reactive counting strategies, but also a very nice intuition for how cooperative Nash equilibria must respond to instances of defection in the opponent.

As the authors are surely aware, finding Nash equilibria does not always predict the outcome of evolution in population -- hence the notions of ESS and ESS\_N in finite populations. In this part of their paper, though, the authors resort to simulations to determine what reactive strategies dominate in populations undergoing payoff-biased imitation. (As it happens, at least among reactive strategies, that Nash conditions do a fairly good job a predicting the types of strategies that are evolutionary robust.)

Reply: Perhaps elaborate on the usefulness of different equilibrium concepts for repeated games (Nash vs ESS vs. RAII). Perhaps also include a paragraph in the Discussion.

My main critique has to do with how the authors treat prior literature on this topic of long-memory strategies in iterated games. Part of the problem here is that they mis-represent prior work as being purely based on simulation, when in fact much prior work is strictly analytical and in even greater generality than the reactive strategies studied here.

Reply: Our shortcoming, we now provide a literature review.

The other part of the problem is that the authors have not compared their qualitative results observed in evolutionary simulations to the analytical results in prior literature for how memory length effects the volume of cooperative evolutionary stable strategies.

Reply: Also correct, we do this now. To this end, we now also characterize all reactive-n defecting strategies (see Propositions XYZ). We use this to numerically estimate the relative volumes of partners and defecting strategies (Figure S1). We find indeed (as Stewart & Plotkin), that the relative volume of parnters increases. Interestingly, however, we also find in the case of counting strategies that these relative volumes alone do not predict evolutionary success.

I am supportive of publication after the authors give a complete account of prior work and make effort (see specific suggestions below) to contextualize their results (especially the results that seem at odds with prior studies of long-memory strategies).

Critique 1) Regarding prior work on long-memory strategies, the authors state that "previous studies considered simulations for small n (56-59), or they analyzed the properties of a few selected higher-memory 134 strategies (60-62)." This is not an accurate description of prior work.

In particular, McAvoy & Nowak (2019) [not cited in this sentence] contains very detailed analytical work on a fascinating class of learning strategies, that is strict super-set of memory-1 strategies. McAvoy shows, eg, that such learning strategies can punish a defecting opponent over multiple rounds; and that such strategies have much greater power to shape the region of feasible payoffs compared to simple memory-1 strategies (which results in convex hulls). The authors are clearly aware of this prior work on long-memory strategies, and so they should discuss it and not dismiss it as purely simulation, when it is detailed analysis.

Likewise, the results of Ueda 2021 (memory-2 strategies) are not discussed or compared to the authors' own results.

Likewise, the authors describe Stewart et al (Ref 57) as based purely on simulation, which is false. Stewart et al derives analytical conditions for the space of memory-n strategies that result in either pure cooperation or pure defection, and that resist selective invasion by any mutant strategy in a populations of size N. (The authors are clearly aware that Ref 57 is based on mathematical analysis, because they cite Stewart as having previously developed a generalization of ZD strategies for memory-n, in their supplement.)

In general, the authors should discuss prior analytical work on long-memory strategies for iterated games, especially in evolving populations, and compare those prior results to their own results. There are some striking differences (see below), which must be reconciled and discussed.

Reply: Fair, we rephrased the main text to give more credit to previous work. In addition, we provide a comprehensive literature in the SI.   
Find a constructive way to highlight why the nature of our results is different from Stewart & Plotkin.

Critique 2 (which is related): Figure 4 seems to report evolutionary simulations only for reactive memory-n strategies. But I am not sure if the main results of the figure (notably the distinction between panels C and D) will hold beyond reactive strategies. I suggest the authors repeat the exploration in Figure 4 with the full space of memory-2 or memory-3 strategies.

Reply: Fair, done. We still get the same results (result: Figure S9, explanation: Figure S10). Note, however, that within the space of memory-n strategies it is infeasible to do systematic simulations for n=3.

This is an important question, I feel, because the authors use the results of Figure 4 to conclude that the order of events must be remembered, not just the number of cooperative events, for a longer memory to benefit cooperation. But this general conclusion is not justified beyond the (restrictive) space of reactive strategies. In fact, Ref 57 has shown (analytically, and verified by simulation) that the opposite trends holds when counting strategies are not just reactive: among the space of strategies that track how many times you cooperated and your opponent cooperated, the volume of evolutionary robust strategies that induce full cooperation is made larger, relative to volume of robust defecting strategies, as memory length increases. In other words, strategies that count how many times both you and your opponent cooperated (which is a very natural thing for humans to count!) are very different than simply counting how many times your opponent cooperated. The authors should discuss this distinction, and they must certainly temper their conclusions about the need to keep track of order of events, which is limited to reactive strategies alone.

Reply: We completely agree that (i) this previous work is relevant, and (ii) that based on a relative-volume analysis, counting strategies should perform well with larger n.   
However, this advantage with respect to volumes does not necessarily translate in evolutionary success. That’s what we show with Figures 4 for reactive-n strategies, and Figure S9 for memory-n strategies.

Nikoleta: Read & write literature review.

Christian: Read literature.   
Write error-section in the SI.   
Write point-by-point reply.   
Revise main text and SI based on Nikoleta’s notes.