Escalation of Memory Length in Finite Populations

Kyle Harrington*
University of Idaho
  Virtual Technology and Design
  kharrington@uidaho.edu

Jordan Pollack
Brandeis University
  Department of Computer Science
  pollack@brandeis.edu

Abstract   The escalation of complexity is a commonly cited benefit of coevolutionary systems, but computational simulations generally fail to demonstrate this capacity to a satisfactory degree. We draw on a macroevolutionary theory of escalation to develop a set of criteria for coevolutionary systems to exhibit escalation of strategic complexity. By expanding on a previously developed model of the evolution of memory length for cooperative strategies by Kristian Lindgren, we resolve previously observed limitations on the escalation of memory length by extending operators of evolutionary variation. We present long-term coevolutionary simulations showing that larger population sizes tend to support greater escalation of complexity than smaller ones do. Additionally, we investigate the sensitivity of escalation during transitions of complexity. The Lindgren model has often been used to argue that the escalation of competitive coevolution has intrinsic limitations. Our simulations show that coevolutionary arms races can continue to escalate in computational simulations given sufficient population sizes.

1   Introduction

The escalation of complexity and accretion of knowledge within an evolving population are poorly understood ideas. Yet the study of coevolution and open-ended evolution represents some of the most ambitious research agendas [39], with implications for directed evolution in synthetic biology [3, 45], evolutionary robotics [25], and automatic programming [21]. Long-term evolution studies have been conducted in microbiological systems [22]; however, studies of the evolutionary dynamics of complex strategies in cooperative games have not achieved the same degree of success [24].

The theory of natural selection is historically associated with phyletic gradualism, the slow transformation of one species to another. However, Gould and Eldridge proposed that new species emerge rapidly in punctuated equilibria [15]. These punctuated equilibria are generally associated with an allopatric (geographic) mechanism of species emergence, whereby relocation to a novel environment leads to a change in selection pressure and often a change in population capacity. In this work we show how these innovative evolutionary phenomena can arise solely from coevolutionary interactions, specifically competitive coevolution.

Coevolution describes the dynamics that arise from interactions between species over evolutionary time scales. "Coevolution" was first coined by Ehrlich and Raven as an approach to the

---

* Corresponding author.

study of community evolution [9]. The study of coevolution encompasses many types of community interactions, be they antagonistic, neutral, or symbiotic. A cornerstone of coevolution is reciprocal selection, where selection on one species reciprocates to other features and members of the ecology. Reciprocal selection has been shown to cause evolutionary arms races in natural systems [35]: The yucca moth exhibited features indicative of reciprocal adaptation with the yucca plant. Yet, coevolutionary dynamics have some notable pathologies that make the maintenance of such evolutionary arms races nontrivial.

The literature on computational coevolution has demonstrated a range of pathologies. Coevolutionary simulations have been plagued by a history of mediocre results and stable states [7, 12]. In one such study the inability of evolutionary game theory to model the dynamics of an evolutionary algorithm with a fitness structure defined for the classic evolutionary hawk-dove game was presented [14]. It was shown that the failure was primarily caused by an insufficient finite population size [10]. This led to the formalization of finite evolutionary stable states within the field of evolutionary game theory [32]. To facilitate the study of coevolutionary pathologies, Watson and Pollack developed the *numbers game* [44], which exhibits a range of fundamental coevolutionary pathologies: loss of gradient, focusing, and relativism. Loss of gradient occurs when the fitness with respect to a sample population does not reflect the absolute objective fitness. Focusing occurs when selective pressures focus on a subset of traits, such that the value of other traits can be forgotten. Relativism occurs when selection pressures favor traits of similar quality, relaxing pressures on more advanced traits. The increased rigor in the computational study of coevolutionary dynamics led to the adoption of the game theoretic tool *solution concepts* by the coevolutionary community [11]. Bucci and Pollack then introduced the mathematical framework of maximally informative individuals [6], which resolves a number of coevolutionary pathologies by using a mechanism for ordered sets reminiscent of principal-component analysis.

A significant pathology of evolutionary histories is what has become known as the Red Queen effect [40]: A species must adapt as fast as it can just to survive the typical changes of the system. Specifically, after analysis of the fossil record van Valen discovered that the probability of a species' extinction is generally independent of the age of the species [40]. While the notion of a constant extinction rate has been subject to serious review and is no longer in favor [30], the majority of studies assume a positive nonzero probability of extinction. In the face of a continuous pressure for extinction, how can a population evolve towards higher levels of complexity?

1.1 Hypothesis of Escalation

The hypothesis of escalation describes how competition between adversaries leads to an increase in complexity and/or investment [41, 43]. The dynamic can be summarized with the following example. Consider an environment with two snails, one with a thicker shell than the other, and one hungry crab. The crab attempts to consume both snails, but can only break the snail with the weaker shell. The harder-shelled snail survives and thus has future chances at reproduction. Unless other selective pressures are applied to the snail (which would be the case in a natural environment), we expect that such an encounter between snails and crabs of successive generations would bias snail morphology toward a harder shell. A similar scenario can be described for two crabs of different strengths and a hard-shelled snail. The escalation of the antagonistic traits between these species (shell thickness and crab strength) is familiar from the evolutionary arms race of Dawkins and Krebs [8]. We explore a *reduced* hypothesis of escalation that does not take account of geographic distribution, and thus does not permit allopatric speciation. Although this removes one of the primary hypothesized mechanisms of producing punctuated equilibria, genetic variation will still remain a property of our model. We will show that the key observations associated with punctuated equilibria and escalation persist.

The original hypothesis of escalation is a naturalist perspective [43], and details many features of nature that are suggested as requirements for a coevolutionary system to support the maintenance of evolutionary arms races. The original list of criteria for escalation is concisely recapitulated in

[16, 41]. We consider a reduced version of the hypothesis of escalation, where geographic features and extrinsic events are disregarded, and populations are unstructured with complete mixing. The criteria for the reduced hypothesis of escalation and their corresponding realizations within this work are:

1. *There must be competition*: Each strategy competes against many other strategies.

2. *Competition applies selective pressure*: There is limited population capacity.

3. *Strategies must be evolvable*: There is always a probability of mutation creating a new individual.

We show how adherence to these criteria allows strategies in a cooperative game to escalate in complexity, exhibiting a coevolutionary arms race.

The reduced hypothesis of escalation that we consider is indeed vastly simplified in comparison with Vermeij's original hypothesis. We do not claim that the reduced hypothesis exhibits the same rates of escalation as the original hypothesis, because, as Vermeij suggests [41], positive feedback can arise as a result of escalation across a geographic distribution of environments. The hypothesis of escalation has recently been challenged with additional statistical analysis of the fossil record [28]. These analyses have been largely invalidated on the basis of sample selection and the fossilization properties of the studied organisms [20, 42]. There still remains a debate regarding how much of evolutionary history is driven by microevolutionary antagonistic interactions, such as in the case of escalation, and macroevolutionary trends, such as punctuated equilibria. We do not attempt to resolve this question, but offer support to the microevolutionary perspective of Vermeij's hypothesis of escalation. This brings us to our computational model of escalation in a game called the *iterated prisoner's dilemma with noise*, based upon [24].

1.2 Evolution of Cooperation

The prisoner's dilemma has become the predominant model of the evolution of cooperation. In this game, two players are faced with the choice of deciding to cooperate with or defect against the opponent, but their payoff is dependent upon both players' decisions. Specifically, the best situation for a single player is to defect against a cooperative opponent; the second best situation for a single player (but best for both players combined) is for both players to cooperate. If players have no memory, then the safest assumption is that the other player is rational and will attempt to maximize payoff. Thus, a rational player with no memory will always defect. When the game is extended to multiple rounds of play, the game is called the iterated prisoner's dilemma (IPD), and that is the focus of this model. In the IPD, a player may decide to cooperate or defect based upon memory of recent encounters with their opponent. For the model presented in this article, every strategy of a given memory length encodes the response (cooperate or defect) for all possible histories.

In the early 1980s, Axelrod and Hamilton conducted a computer tournament of human-designed IPD strategies [2]. The winner of the tournament was Anatol Rappaport's tit-for-tat strategy. Since these initial tournaments, a number of researchers have embarked on the quest to find the champion evolutionarily stable strategy (ESS) for the IPD. A sequence of findings have shaped the current belief about optimal strategies in the IPD. It was shown that tit-for-tat plays a transitory role in the evolution of IPD strategies [33], and subsequent analysis led to the demonstration of the strength of the win-stay, lose-shift strategy [34]. In the case of the stochastic IPD, where the decision to cooperate or defect is determined by the flip of a genetically biased coin, a recent proof demonstrates the existence of *zero-determinant* (ZD) strategies, where a player can unilaterally specify the payoff received by its opponent [36]. This proof marks a significant discovery in the structure of the IPD; however, further research on ZD strategies has revealed that they are not evolutionarily stable (ES) [1]. It has been proven that in alternative formulations of the IPD there are no ES strategies [5, 26, 27]. However, these proofs involve features such as the discounting of future moves, which are not present in the classic IPD. Recent theoretical work has shown that longer strategies improve the average performance

of IPD strategies [23] and that longer memory lengths should evolve over time [18]; however, there have been no empirical studies that show evolutionary trajectories that satisfy this claim.

An innovative study was presented by Lindgren [24] where the set of active IPD strategies changes over time, as opposed to most studies of the IPD, where only the frequencies of a fixed set of strategies change over time. In Lindgren's study, strategies evolve by flipping between cooperation and defection based upon a history of interactions with a memory length measured by the number of actions. For example, a memory length of 4 means the strategy is dependent on two actions by each player. However, Lindgren found that the model was not able to escape an evolutionarily stable state containing strategies of memory length 4. In other words, the system did not appear to escalate beyond memory length 4. Our model alleviates this problem by using an alternative variation mechanism. In Lindgren's model, memory lengths increase by doubling and halving, which only allows for the introduction of mutant strategies that vary by the action of one player. Interestingly, these doubling and halving operations, which introduce neutral strategies, do not lead to a neutral drift toward larger population sizes. We introduce mutants with a normal distribution of memory length variations, permitting mutant strategies of any length. The limitation of changing memory length by only one interaction at a time is an inductive bias that expects that a successful mutant exists within a factor of 1 of the current distribution of memory lengths in the population. This intuition can be reinforced by the fact that a memory length extension of 1 only affects a player's behavior with respect to one role in the game. For example, tit-for-tat is a strategy of memory length 1, where it cooperates if its opponent cooperated, and defects if its opponent defected. An extension of memory length 1 allows tit-for-tat to remember not only its opponent's previous move, but also its own previous move. However, we argue that there are situations where a strategy can only be invaded by a mutant whose strategy has changed by more than 1 memory length. While Lindgren's operators do not guarantee that the third criterion of the reduced hypothesis of escalation (strategies are improvable by variation) is satisfied; our genetic operators do.

A similar observation on variation was made by Ikegami, who uses tree representations of IPD strategies [19]; his populations exhibit escalation of memory length and diversity. However, Ikegami's model is obscured by the use of a module-based evolutionary operator. This operator, akin to symbiogenesis, provides a similar variable-memory-length extension to our normally distributed extension and contraction operators. However, genetic recombination is an evolutionary transition that is expected to have emerged long after populations began to escalate [37].

It is well known that the size of a finite evolving population can have a significant effect on the fate of the population [10, 13, 32]. However, the relationship between the size of a population and its ability to support the escalation of strategy memory length remains unexplored. This is a particularly significant direction for investigation when considering the IPD with noise, which ensures that every element of a strategy has a fitness consequence. We present long-term simulation data that demonstrate a positive correlation between greater population size and the evolution of longer memory lengths, suggesting that increased population size can lead to enhanced evolution of strategic complexity.

2   Results and Discussion

Our model is an extension of Lindgren's innovative model of the IPD [24], where the use of alternative genetic operators alleviates the mediocre stable states previously observed. We both suggest that this model satisfies the criteria of the reduced hypothesis of escalation and empirically demonstrate the escalation of complexity in the model.

2.1   Model
The interactions between evolving strategies are specified by the replicator dynamics and the game. We use a standard formulation of the prisoner's dilemma,

$$P = \begin{pmatrix} (3,3) & (0,5) \\ (5,0) & (1,1) \end{pmatrix}$$

where the notation (p_1, p_2) indicates the scores of players 1 and 2, respectively. However, this payoff matrix only specifies the score of a single round of the prisoner's dilemma. In the iterated prisoner's dilemma (IPD) multiple rounds are allowed for. The standard way of accomplishing this is by iterating for a finite number of rounds and accumulating the total score for each player during each round. The IPD becomes interesting when strategies have some memory and may change their behavior depending upon the outcomes of previous rounds. This is generally accomplished by encoding lookup tables within strategies. However, the accumulated score will be sensitive to the number of iterations performed. This will be particularly true as strategies rely on memories of more encounters.

The stochastic prisoner's dilemma admits an alternative method of iteration to the previously mentioned finite iteration technique [31]. Stochastic prisoner's dilemma strategies include a noise term, whereby with a certain probability strategies will take the opposite action. This means the game is a Markov chain. We can describe the game as follows

$$\vec{H} = M\vec{H}$$

where $\vec{H}$ is the vector of probabilities of each history and M is the transfer matrix. For two strategies, s_1 and s_2, $\vec{H}$ is always of length $2^{\max(|s_1|, |s_2|)}$. The transfer matrix describes the probabilities of transitioning between histories given s_1 competing against s_2 with noise. H is called the stationary distribution of M, and represents the distribution of histories in the limit of an infinite number of rounds. We can recover the distribution of round outcomes (CC, CD, DC, and DD) by weighting all histories that end in each outcome by the corresponding payoff values. This distribution of rounds allows us to compute the scores of s_1 and s_2.

2.2 Genetic Variation

We employ the same genetic encoding as Lindgren. Strategies are represented as binary strings that encode the action to perform given a specific history. This is easily accomplished by using the observed history as an index into the genome, where the binary value stored at that position specifies the strategy's response. In the IPD with finite rounds, strategies generally also encode a sequence that specifies the *initial history*, because this historical lookup mechanism only works when the genome encodes the responses for all historical sequences. A study of the effect of memory size on the finite-round IPD was presented in [17].

The genetic operators first used by Lindgren [24] implement gene doubling, gene halving, and point mutation. Point mutation is familiar from genetic algorithms, where a single bit is flipped with some mutation probability. In gene doubling, the entire bit string is extended by a factor of 2 during duplication; because the index is based on the historical observations, a doubling event on its own does not change the meaning of a genome. Gene halving is accomplished by randomly truncating the first or second half of the genome.

We use variants of each of these genetic operators. Instead of point mutation, we use uniform mutation, where multiple bits may be flipped during a single reproductive event. To accomplish extension and contraction we draw a random number from a Gaussian distribution, and the absolute value of its integer component is taken as the number of extensions or contractions to perform. Both extension and contraction are accomplished in the same way as in Lindgren's model, but extension (contraction) may be more (less) than a factor of 2. When performing genetic operations, first the mutant genome may or may not be extended or contracted; then it subsequently may or may not be subject to uniform mutation. Thus in a given reproductive event a mutant may have been extended or contracted as well as varied with uniform mutation. We now revisit a requirement of the hypothesis of escalation: *strategies are evolvable*.

Our genetic operators ensure that it is possible to reach a large number of strategies from any population distribution, while Lindgren's operators appear to only reach a limited range of genotypes. Specifically, no mutant strategy will ever be larger by more than 1 memory length than the biggest genome in the population, or smaller by more than 1 memory length than the smallest

genome in the population. However, we hypothesize that it may be necessary to invade with a strategy outside of that range, and our results suggest this is correct. There has been some work on the invasion by pairs of strategies [5, 26, 27], but there is still no known champion IPD strategy.

While Lindgren utilized the continuous-time replicator dynamics and introduced mutants while time-stepping, we instead use the Moran process with mutation. The Moran process models evolutionary dynamics by iteratively replacing one individual at a time [29]. In the evolutionary computation literature, the Moran process is sometimes called "steady-state" evolution [38]. The Moran process offers an intuitive way of introducing mutant strategies into the population. On the other hand, the best method of introduction of mutant strategies into a mean-field model is not immediately apparent. In Lindgren's model, each strategy has a probability of introducing a single mutated variant proportional to the frequency of the parent strategy.

2.3   Dependence of Average Diversity on Population Size Over Time

In this study we simulate the previously described model with the Moran process with population sizes 5,000, 10,000, and 20,000. For each population size, 25 replicates were used. For all simulations the following parameters are used: $p_{\text{extend}} = 0.000001$, $p_{\text{contract}} = 0.000001$, $p_{\text{uniform}} = 0.001$, and $T_{\text{max}} = 6,000,000,000$ births. This number of births corresponds to 1,200,000 generations for population size 5,000, to 600,000 generations for population size 10,000, and to 300,000 generations for population size 20,000. We have also restricted the maximum length of strategies to 12; however, we never observe this limit being reached. The cost of simulating infinite games increases exponentially with the maximum memory size of the competing strategies, which is a strong motivation for prohibiting excessively long strategies.

Just as in Lindgren's study [24], we observe similarities between all simulations (and Lindgren's), especially during the initial generations as the system passes through metastable states. For example, compare Figure 1 with Lindgren's Figure 1 [24], both of which exhibit the same patterns in the initial phase of their evolutionary trajectory. While much of Lindgren's discussion regarding the evolutionary timeline remains intact, our model provides an epilogue to Lindgren's reference to open-ended evolution.

The inclusion of noise in the IPD model admits ES strategies [4]. Both Lindgren's and our model do reach ES states under some conditions, and in Lindgren's model it is unclear whether all paths will lead to such an evolutionarily stable state. While Lindgren found some evolutionary trajectories that did not get stuck in the same memory-4 ES strategy that plagued many of his simulations, he did not demonstrate evolutionary trajectories that exceeded memory length 4. Here we present simulation results for evolutionary trajectories that escalate beyond memory length 4.

The smallest population size that we consider is 5,000 (Figure 2a). In this case a number of simulations are unable to escape the initial metastable states, and populations remain at low memory lengths. However, some simulations do reach populations consisting primarily of memory length 7. We will revisit this observation for a larger population size. The number of species grows for the first 70,000 generations, then plateaus just below 300 species. However, even during this plateau of species diversity, some escalation can still be observed, as strategies of memory length 7 are still on the rise at the end of the simulation.
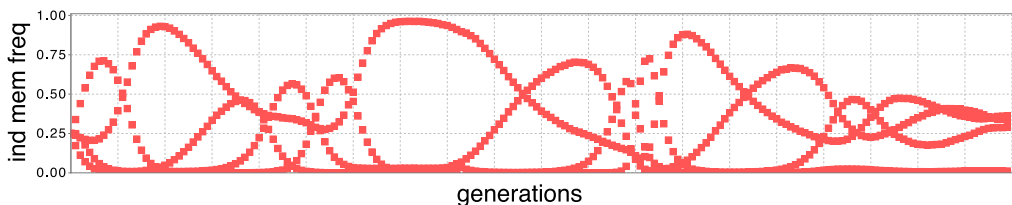

Figure 1. Example of initial evolutionary trajectory. Patterns of initial evolutionary dynamics are nearly identical to those shown in [24]. As evolutionary time progresses, evolutionary trajectories of the finite-population models diverge from the continuous-time model in [24].
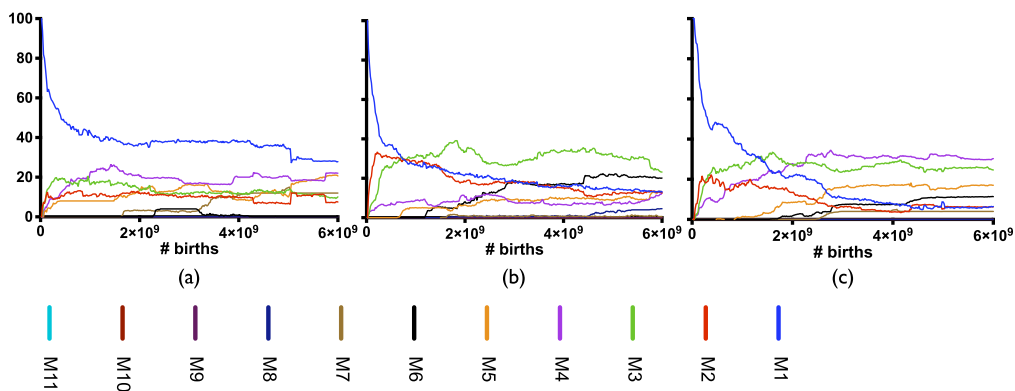
Figure 2. Density of strategy memory lengths as a function of population size. Larger population sizes tend to include longer-memory-length strategies. Panels correspond to different population sizes: a = 5,000; b = 10,000; c = 20,000.

Our results for population sizes 10,000 and 20,000 show the most escalation in memory length (Figure 2b, c). The population size of 20,000 shows runs that are beginning to be dominated by memory length 8. These runs are the most escalated of all the experiments we conducted. Across the panels of Figure 2 we show the average percentage of the population that is dominated by each memory length. On average, population size 5,000 is dominated by memory length 1, and to a lesser extent by memory length 4; population size 10,000 is dominated by memory length 3, seconded by memory length 6; population size 20,000 is dominated by memory length 4. From this it is clear that larger population sizes tend to support longer memory lengths. It is also important to note that the stochastic nature of the Moran process contributes to the variability of times when different memory lengths emerge.

By extending Lindgren's model with alternative genetic operators we have cleared the path to open-ended evolution in the IPD model. We explore the model using finite-population evolutionary dynamics, as opposed to Lindgren's use of continuous-time replicator dynamics. The model continues to exhibit similar evolutionary trajectories to those presented by Lindgren, which suggests that it is not our use of the Moran process that leads to the escape from the memory-4 metastable states that appeared to limit Lindgren's original model. While we see that larger population sizes are capable of supporting a larger number of species, larger population size does not eliminate the possibility of getting stuck in an evolutionary equilibrium. This leads to the suggestion that achieving greater escalation is not simply a matter of using a larger population size.

Now let us consider a specific example trajectory from a population-size-20,000 run. In Figure 3 we have a timeline showing the evolutionary history after 2×10^9 birth events. Over the course of
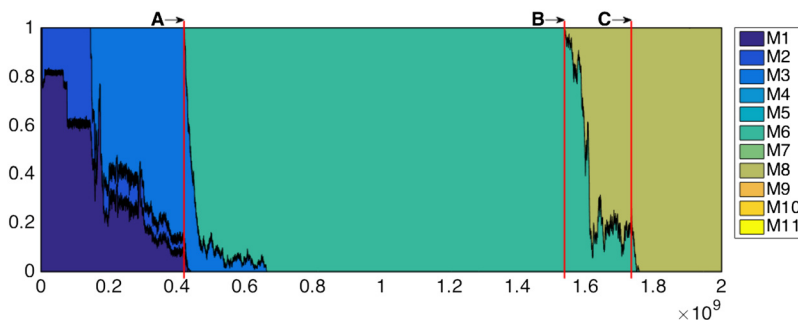

Figure 3. Density of strategy memory lengths for a highly escalated example with population size 20,000. Knockouts are performed at three points during a highly escalated evolutionary trajectory (indicated by vertical lines: A, B, and C).
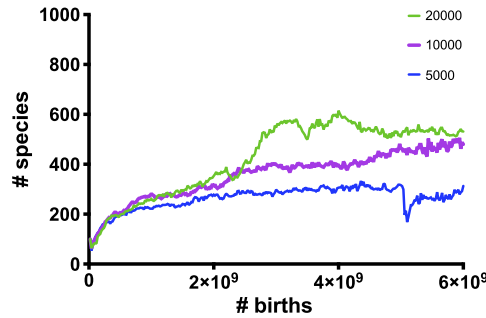
Figure 4. Species diversity for population sizes 5,000, 10,000, and 20,000 averaged over 25 simulations. Species diversity generally increases over evolutionary time, with occasional rapid changes. Larger population sizes support greater diversity.

this evolutionary trajectory the population transitions to the previously observed limit of memory length 4 to memory length 6, and on to memory length 8. As we note in Figure 4, the diversity of species increases significantly toward the end of evolution, making analysis of individual strategies challenging. When comparing the individual trajectory shown in Figure 3 with the aggregate statistics shown in Figure 2, a key difference to note is the presence of multiple memory lengths at any point in time for the aggregate statistics, while the individual trajectory shows long periods of a single dominant memory length. Due to the stochastic nature of the Moran process, evolutionary trajectories transition to different memory lengths at different points in evolutionary history. In general we observed that populations were often composed of one dominant memory length for the majority of their evolutionary history. To evaluate the sensitivity of populations to individual strategies during transitions between memory lengths, we perform a species knockout analysis at multiple points within the evolutionary trajectory.

2.4   Species Knockout Analysis
An analysis of population stability was performed via species knockout, where a given strategy is eliminated from the population. The population is rebalanced by uniformly allocating the previously occupied fraction of the population to the remaining strategies. After performing the knockout, the simulation is evaluated for 4×10^5 birth events; then the distribution of memory lengths is investigated. For each timepoint a species knockout is performed with respect to each strategy in the population. The timepoints of knockouts are indicated in Figure 3. The figure itself shows the evolutionary trajectory in the absence of knockouts.

Knockout A is performed at the transition from memory length 4 to 6. Memory length 4 is the level where Lindgren found a tendency for populations to stabilize [24]. The knockout is performed at the first generation where memory length 6 strategies appear. Of the 17 species knockouts performed, two lead to a collapse of escalation, while the remaining knockouts continue to support memory-length-6 strategies.

Knockout B is performed at the first timepoint where there are more than 50 individuals with memory-length-8 strategies. It was necessary to choose such a timepoint because mutations ephemerally introduce strategies of memory length 8 that are not capable of triggering a transition to memory length 8. Nevertheless, for all 41 knockouts the populations revert to memory length 6. This suggests that the fitnesses of strategies are highly interdependent during this particular transition to greater complexity, which motivates us to consider a knockout after the transition from memory length 6 to 8 has progressed further.

Knockout C is performed when the majority of the population is occupied by memory-length-8 strategies (approximately 30% to 70%). Here we find that all 157 knockouts maintain populations with memory-length-8 strategies. This suggests that the interdependence observed in knockout B

has stabilized, and the population has become more robust to the distribution of strategies that it contains.

These three knockout studies highlight a key point. Knockout A is performed immediately following the transition from memory length 4 to 6, and still many knockout populations are capable of escalating to greater memory lengths. Knockout B is performed close to the transition from memory length 6 to 8, and none of the knockout populations escalate to greater memory lengths. Finally, knockout C is performed much later, in the transition from memory length 6 to 8, and all knockout populations continue to escalate to greater memory length. While it is possible that longer evaluation of knockout populations may lead to observations of eventual escalation to greater memory length, in this example the point remains that that escalation of greater complexity is more vulnerable to destabilizing knockouts.

3   Conclusions

The study of coevolutionary arms races has had a challenging history plagued with premature mediocre stabilization [12] and other coevolutionary pathologies [44]. These pathologies were previously related to observations of limited escalation of complexity in simple evolutionary models of cooperative games [24]. In our study we have drawn inspiration from the macroevolutionary theory of the escalation [43] to show that previous observations of limited evolution in the iterated prisoner's dilemma with noise [24] were due to a lack of evolvability. By conducting long-term evolutionary simulations we have shown that an improved model can lead to continued escalation of strategic complexity. We have also shown that strategies escalate in complexity faster in larger populations. Coevolutionary simulation can drive the escalation of complexity, and that escalation can be amplified in larger population sizes. Furthermore, the escalation of complexity can be sensitive to species knockouts during transition periods. Thus, the stabilization of species and maintenance of large population sizes are viable mechanisms to support the escalation of strategic complexity.

This extension of Lindgren's seminal work on the open-ended evolution of strategic complexity in the noisy IPD [24] has significant implications for the general agenda of open-ended evolution research. The original results from Lindgren suggested that the noisy IPD would not be an appropriate model for the study of open-ended evolution. This was a striking observation because the structure of the noisy IPD represents a fundamental evolutionary system. If a simple cooperative game cannot yield an escalating arms race, then there may be no simple dynamics that can support the evolution of complex cooperative strategies. However, by adapting Lindgren's work we have shown that our model has the capacity to escalate, and we suggest that it may serve as a model problem for open-ended evolution. We arrived at our adaptation of Lindgren's model by considering a reduced version Vermeij's hypothesis of escalation: There must be competition, competition applies selective pressure, and strategies must be evolvable. We suggest that the reduced hypothesis of escalation represent a minimum criterion for designing a system that can support coevolutionary arms races.

4   Future Work

While we have presented a model that resolves the limitations of escalation in Lindgren's model [24], there is ample possibility for future work. A detailed study of variation operators under different intensities of selection, including neutral drift, would lead to a greater understanding of the interplay between selection and variation in the escalation of strategic complexity. In this study and further experimentation, we have found that population size has a significant effect on the tendency of a population to escalate in memory length. To support future endeavors in the study of escalation, we make the code for our model freely available: https://github.com/kephale/EscalationIPD.

Acknowledgment

References

1. Adami, C., & Hintze, A. (2012). Winning isn't everything: Evolutionary stability of zero determinant strategies. *Nature Communications*, *4*(10), 1038.

2. Axelrod, R., & Hamilton, W. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390.

3. Boder, E., Midelfort, K., & Wittrup, K. (2000). Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. *Proceedings of the National Academy of Sciences of the U.S.A.*, *97*(20), 10701–10705.

4. Boyd, R. (1989). Mistakes allow evolutionary stability in the repeated prisoner's dilemma game. *Journal of Theoretical Biology*, *136*(1), 47–56.

5. Boyd, R., & Lorberbaum, J. (1987). No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature*, *327*, 58–59.

6. Bucci, A., & Pollack, J. B. (2002). A mathematical framework for the study of coevolution. In A. H. Wright, M. D. Vose, K. A. D. Jong, & L. M. Schmitt (Eds.), *Foundations of genetic algorithms*, Vol. 7 (pp. 221–235). New York: Springer.

7. Cliff, D., & Miller, G. F. (1996). Co-evolution of pursuit and evasion II: Simulation methods and results. In P. Maes, M. Mataric, J. Meyer, J. Pollack, & S. W. Wilson (Eds.), *From animals to animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior* (pp. 506–515). Cambridge, MA: MIT Press.

8. Dawkins, R., & Krebs, J. (1979). Arms races between and within species. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, *205*(1161), 489–511.

9. Ehrlich, P., & Raven, P. (1964). Butterflies and plants: A study in coevolution. *Evolution*, *18*(4), 586–608.

10. Ficici, S., & Pollack, J. (2000). Effects of finite populations on evolutionary stable strategies. In L. D. Whitley, D. E. Goldberg, E. Cantú-Paz, L. Spector, I. C. Parmee, & H. Beyer (Eds.), *Proceedings of the Genetic and Evolutionary Computation Conference* (pp. 927–934). New York: Springer.

11. Ficici, S. G. (2004). *Solution concepts in coevolutionary algorithms*. Ph.D. thesis, Brandeis University.

12. Ficici, S. G., & Pollack, J. B. (1998). Challenges in coevolutionary learning: Arms-race dynamics, open-endedness, and mediocre stable states. In C. Adami, R. K. Belew, H. Kitano, & C. E. Taylor (Eds.), *Proceedings of the Sixth International Conference on Artificial Life* (pp. 238–247). Cambridge, MA: MIT Press.

13. Fogel, D., & Fogel, G. (1995). Evolutionary stable strategies are not always stable under evolutionary dynamics. In J. R. McDonnell, R. G. Reynolds, & D. B. Fogel (Eds.), *Evolutionary programming IV* (pp. 565–577). Cambridge, MA: MIT Press.

14. Fogel, D. B., Fogel, G. B., & Andrews, P. C. (1997). On the instability of evolutionary stable strategies. *Biosystems*, *44*(2), 135–152.

15. Gould, S., & Eldredge, N. (1972). Punctuated equilibria: An alternative to phyletic gradualism. In F. J. Ayala & J. C. Avise (Eds.), *Essential readings in evolutionary biology* (pp. 82–115). Baltimore: Johns Hopkins University Press.

16. Harrington, K. I. S. (2014). *Escalation of coevolutionary arms races*. Ph.D. thesis, Brandeis University.

17. Hauert, C., & Schuster, H. G. (1997). Effects of increasing the number of players and memory size in the iterated prisoner's dilemma: A numerical approach. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *264*, 513–519.

18. Hilbe, C., Martinez-Vaquero, L. A., Chatterjee, K., & Nowak, M. A. (2017). Memory-n strategies of direct reciprocity. *Proceedings of the National Academy of Sciences of the U.S.A.*, *114*(18), 4715–4720.

19. Ikegami, T. (1994). From genetic evolution to emergence of game strategies. *Physica D: Nonlinear Phenomena*, *75*(1–3), 310–327.

20. Jablonski, D. (2008). Biotic interactions and macroevolution: Extensions and mismatches across scales and levels. *Evolution*, *62*(4), 715–739.

21. Koza, J. (1992). *Genetic programming: On the programming of computers by means of natural selection*. Cambridge, MA: MIT Press.

22. Lenski, R., Rose, M., Simpson, S., & Tadler, S. (1991). Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *American Naturalist*, *138*(6), 1315–1341.

23. Li, J., & Kendall, G. (2014). The effect of memory size on the evolutionary stability of strategies in iterated prisoner's dilemma. *IEEE Transactions on Evolutionary Computation*, *18*(6), 819–826.

24. Lindgren, K. (1991). Evolutionary phenomena in simple dynamics. In *Artificial life II* (pp. 295–312). Cambridge, MA: MIT Press.

25. Lipson, H., & Pollack, J. B. (2000). Automatic design and manufacture of robotic lifeforms. *Nature*, *406*, 974–978.

26. Lorberbaum, J. (1994). No strategy is evolutionarily stable in the repeated prisoner's dilemma. *Journal of Theoretical Biology*, *168*(2), 117–130.

27. Lorberbaum, J., & Bohning, D. (2002). Are there really no evolutionarily stable strategies in the iterated prisoner's dilemma? *Journal of Theoretical Biology*, *214*(2), 155–169.

28. Madin, J., Alroy, J., Aberhan, M., & Fürsich, F. (2006). Statistical independence of escalatory ecological trends in Phanerozoic marine invertebrates. *Science*, *312*(5775), 897–900.

29. Moran, P. (1962). *The statistical processes of evolutionary theory*. New York: Clarendon Press.

30. Nee, S. (2006). Birth-death models in macroevolution. *Annual Review of Ecology, Evolution, and Systematics*, *37*, 1–17.

31. Nowak, M. (1990). Stochastic strategies in the prisoner's dilemma. *Theoretical Population Biology*, *38*(1), 93–112.

32. Nowak, M., Sasaki, A., Taylor, C., & Fudenberg, D. (2004). Emergence of cooperation and evolutionary stability in finite populations. *Nature*, *428*(6983), 646–650.

33. Nowak, M., & Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature*, *355*(6357), 250–253.

34. Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, *364*(6432), 56–58.

35. Pellmyr, O., Thompson, J. N., Brown, J. M., & Harrison, R. G. (1996). Evolution of pollination and mutualism in the Yucca moth lineage. *The American Naturalist*, *148*(5), 827–847.

36. Press, W., & Dyson, F. (2012). Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences of the U.S.A.*, *109*(26), 10409–10413.

37. Smith, J., & Szathmáry, E. (1997). *The major transitions in evolution*. New York: Oxford University Press.

38. Syswerda, G. (1991). A study of reproduction in generational and steady-state genetic algorithms. In G. J. Rawlins (Ed.), *Foundation of genetic algorithms* (pp. 94–101). New York: Springer.

39. Taylor, T., Bedau, M., Channon, A., Ackley, D., Banzhaf, W., Beslon, G., Dolson, E., Froese, T., Hickinbotham, S., Ikegami, T., et al. (2016). Open-ended evolution: Perspectives from the OEE workshop in York. *Artificial Life*, *22*(3), 408–423.

40. van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory*, *1*, 1–30.

41. Vermeij, G. (1994). The evolutionary interaction among species: Selection, escalation, and coevolution. *Annual Review of Ecology and Systematics*, *25*(1), 219–236.

42. Vermeij, G. (2008). Escalation and its role in Jurassic biotic history. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *263*(1), 3–8.

43. Vermeij, G. J. (1987). *Evolution and escalation: An ecological history of life*. Princeton, NJ: Princeton University Press.

44. Watson, R., & Pollack, J. (2001). Coevolutionary dynamics in a minimal substrate. In L. Spector, E. D. Goodman, A. Wu, W. B. Langdon, & H.-M. Voigt (Eds.), *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-01* (pp. 702–709). New York: Springer.

45. Yokobayashi, Y. (2002). Directed evolution of a genetic circuit. *Proceedings of the National Academy of Sciences of the U.S.A.*, *99*(26), 16587–16591.