

Escalation of Memory Length in Finite Populations

Kyle Harrington^{a,b}, Jordan Pollack^a

^a*DEMO Lab, Department of Computer Science, Brandeis University, Waltham, MA*

^b*Current Address: Virtual Technology and Design, University of Idaho, Moscow, ID*

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 to the escalation of memory length. We present long-term coevolutionary simulations showing that larger population sizes tend to support greater escalation of complexity than smaller population sizes. Additionally, escalation is sensitive to perturbation during transitions of complexity. In whole, a long-standing counter-argument to the ubiquitous nature of coevolution is resolved, suggesting that the escalation of coevolutionary arms races can be harnessed by computational simulations.

Keywords: Coevolution, Cooperation, Escalation

1. Coevolutionary Escalation

1 The escalation of complexity and accretion of knowledge within an evolving population are
2 poorly understood ideas. Yet the study of coevolution and open-ended evolution represents
3 some of the most ambitious research agendas [1] with implications for directed evolution in
4 synthetic biology [2, 3], evolutionary robotics [4], and automatic programming [5]. Long-
5 term evolution studies have been conducted in microbiological systems [6]; however, studies
6 of the evolutionary dynamics of complex strategies in cooperative games have not achieved
7 the same degree of success [7].

8 The theory of natural selection is historically associated with phyletic gradualism, the
9 slow transformation of one species to another. However, Eldridge and Gould proposed that
10 new species emerge rapidly in punctuated equilibria [8]. These punctuated equilibria are
11 generally associated with an allopatric (geographic) mechanism of species emergence, whereby
12 relocation to a novel environment leads to a change in selection pressure and often a change
13 in population capacity. In this work we show how these innovative evolutionary phenomena
14 can arise solely from coevolutionary interactions, specifically competitive coevolution.

15 Coevolution describes the dynamics that arise from interactions between species over evo-
16 lutionary timescales. "Coevolution" was first coined by Ehrlich and Raven as an approach
17 to the study of community evolution [9]. The study of coevolution encompasses many types
18 of community interactions, be they antagonistic, neutral, or symbiotic. A cornerstone of
19

20 coevolution is reciprocal selection, where selection on one species reciprocates to other fea-
21 tures and members of the ecology. Reciprocal selection has been shown to cause evolutionary
22 arms races in Natural systems [10], where Yucca moth exhibited features indicative of re-
23 ciprocal adaptation with the Yucca plant. Yet, coevolutionary dynamics have some notable
24 pathologies that make the maintenance of such evolutionary arms races non-trivial.

25 The literature on computational co-evolution has demonstrated a range of pathologies.
26 Coevolutionary simulations have been plagued by a history of mediocre results and stable
27 states [11, 12]. In one such study the inability of evolutionary game theory to model the
28 dynamics of an evolutionary algorithm with a fitness structure defined for the classic evo-
29 lutionary hawk-dove game was presented [13]. It was shown that the failure was primarily
30 caused by an insufficient finite population size [14]. This led to the formalization of finite
31 evolutionary stable states within the field of evolutionary game theory [15]. To facilitate the
32 study of coevolutionary pathologies Watson and Pollack developed the Numbers Game [16],
33 which exhibits a range of fundamental coevolutionary pathologies: loss of gradient, focussing,
34 and relativism. Loss of gradient occurs when the fitness with respect to a sample population
35 does not reflect the absolute objective fitness. Focussing occurs when selective pressures
36 focus on a subset of traits, such that the value of other traits can be forgotten. Relativism
37 occurs when selection pressures favor traits of similar quality, relaxing pressures on more
38 advanced traits. The increased rigor in the computational study of coevolutionary dynamics
39 led to the adoption of the game theoretic tool, solution concepts, by the coevolutionary com-
40 munity [17]. Bucci and Pollack then introduced the mathematical framework of maximally
41 informative individuals [18], which resolves a number of coevolutionary pathologies by using
42 a mechanism for ordered sets reminiscent of principle component analysis.

43 A significant pathology of evolutionary histories is what has become known as the Red
44 Queen effect [19]; a species must adapt as fast as it can just to survive the typical changes
45 of the system. Specifically, after analysis of the fossil record van Valen discovered that the
46 probability of a species' extinction is generally independent of the age of the species [19].
47 While the notion of a constant extinction rate has been subject to serious review and is
48 no longer in favor [20], the majority of studies assume a positive non-zero probability of
49 extinction. In the face of a continuous pressure for extinction, how can a population evolve
50 towards higher levels of complexity?

51 *1.1. Hypothesis of Escalation*

52 The hypothesis of escalation describes how competition between enemies leads to an
53 increase in complexity and/or investment [21, 22]. The dynamic can be summarized with the
54 following example. Consider an environment with 2 snails, one with a thicker shell than the
55 other, and 1 hungry crab. The crab attempts to consume both snails, but can only break the
56 snail with the weaker shell. The harder shelled snail survives and thus has future chances at
57 reproduction. Unless other selective pressures are applied to the snail (which would be the
58 case in a natural environment), we expect that such an encounter between snails and crabs of
59 successive generations would bias snail morphology toward a harder shell. A similar scenario
60 can be described for 2 crabs of varying strengths and a hard-shell snail. The escalation of
61 the antagonistic traits between these species (shell thickness and crab strength) is familiar
62 from the evolutionary arms race analogy of Dawkins and Krebs [23]. We explore a *reduced*
63 hypothesis of escalation which does not account for geographic distribution, and thus does

64 not permit allopatric speciation. Although this removes one of the primary hypothesized
65 mechanisms of producing punctuated equilibria, genetic variation will still remain a property
66 of our model. We will show that the key observations associated with punctuated equilibria
67 and escalation persist.

68 The original hypothesis of escalation is a naturalist perspective [22], and details many
69 features of Nature which are suggested as requirements for a coevolutionary system to sup-
70 port the maintenance of evolutionary arms races. The original list of criteria for escalation
71 is concisely recapitulated in [21, 24]. We consider a reduced version of the hypothesis of
72 escalation, where geographic features and [extrinsic events] are disregarded, and populations
73 are unstructured with complete mixing. The criteria for the reduced hypothesis of escalation
74 and their corresponding realizations within this work are:

- 75 1. **There must be competition**; each strategy competes against many other strategies.
- 76 2. **Competition applies selective pressure**; limited population capacity.
- 77 3. **Strategies must be evolvable**; there is always a probability of mutation creating a
78 new individual.

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

81 The reduced hypothesis of escalation that we consider is indeed vastly simplified beyond
82 Vermeij’s original hypothesis. We do not claim that the reduced hypothesis exhibits the
83 same rates of escalation as the original hypothesis, because as Vermeij suggests [21], positive
84 feedback can arise as a result of escalation across a geographic distribution of environments.
85 The hypothesis of escalation has recently be challenged with additional statistical analysis
86 of the fossil record [25]. These analyses have been largely invalidated on the basis of sample
87 selection and the fossilization properties of the studied organisms [26, 27]. There still remains
88 a debate regarding how much of evolutionary history is driven by microevolutionary antag-
89 onistic interactions, such as in the case of escalation, and macroevolutionary trends such as
90 punctuated equilibria. We do not attempt to resolve this question, but offer support to the
91 microevolutionary perspective of Vermeij’s hypothesis of escalation. This brings us to our
92 computational model of escalation in a game called the Iterated Prisoner’s Dilemma with
93 noise based upon [7].

94 2. Evolution of Cooperation

95 The Prisoner’s Dilemma has become the predominant model of the evolution of cooper-
96 ation. In this game, two players are faced with the choice of deciding to cooperate or defect
97 against their opponent, but their payoff is dependent upon both players’ decisions. Specifi-
98 cally, the best situation for a single player is to defect against a cooperative opponent; the
99 second best situation for a single player (but best for both players combined) is for both
100 players to cooperate. If players have no memory, then the safest assumption is that the
101 other player is rational and will attempt to maximize payoff. Thus, a rational player with no
102 memory will always defect. When the game is extended to multiple rounds of play the game
103 is called the Iterated Prisoner’s Dilemma (IPD), which is the focus of this model. In the IPD
104 a player may decide to cooperate or defect based upon memory of recent encounters with

105 their opponent. For the model presented in this paper, every strategy of a given memory
106 length encodes the response (cooperate or defect) for all possible histories.

107 In the early 1980's, Axelrod and Hamilton conducted a computer tournament of human-
108 designed IPD strategies [28]. The winner of the tournament was Anatol Rappaport's tit-for-
109 tat strategy. Since these initial tournaments a number of researchers have embarked on the
110 quest to find the champion evolutionarily stable strategy IPD strategy. A sequence of findings
111 have shaped the current belief about optimal strategies in the IPD. It was shown that tit-
112 for-tat plays a transitory role in the evolution of IPD strategies [29], and subsequent analysis
113 led to the demonstration of the strength of the win-stay, lose-shift strategy [30]. In the case
114 of the stochastic IPD, where the decision to cooperate or defect is determined by the flip of
115 a genetically biased coin, a recent proof demonstrates the existence of "zero-determinant"
116 (ZD) strategies, where a player can unilaterally specify the payoff received by one's opponent
117 [31]. This proof marks a significant discovery in the structure of the IPD; however, further
118 research on ZD strategies has revealed that they are not ESS [32]. It has been proven that
119 in alternative formulations of the IPD there are no ESSs [33, 34, 35]. However, these proofs
120 involve features such as discounting of future moves, which are not present in the classic IPD.
121 Recent theoretical work has shown that longer strategies improve the average performance
122 of IPD strategies [36] and that longer memory lengths should evolve over time [37]; however,
123 there have been no empirical studies that show evolutionary trajectories that satisfy this
124 claim.

125 An innovative study was presented by Lindgren [7] where the set of active IPD strategies
126 change over time, as opposed to most studies of the IPD where only the frequencies of a
127 fixed set of strategies change over time. In Lindgren's study, strategies evolve by flipping
128 between cooperation or defection as based upon a history of interactions with a memory
129 length measures the number of actions. For example, a memory length of 4 means the
130 strategy is dependent on 2 interactions between both players. However, Lindgren found that
131 the model was not able to escape an ESS containing strategies of memory length 4. In other
132 words, the system did not appear to escalate beyond memory length 4. Our model alleviates
133 this problem by using an alternative variation mechanism. In his model, memory lengths
134 increase by doubling and halving, which only allows for the introduction of mutant strategies
135 that vary by the action of one player from the current population. Instead we introduce
136 mutants with a normal distribution of memory length variations, permitting the mutant
137 strategies of any length. The limitation of changing memory length by only 1 interaction
138 at a time is an inductive bias which expects that a successful mutant exists within a factor
139 of 1 of the current distribution of memory lengths in the population. This intuition can be
140 reinforced by the fact that a memory length extension of 1 only affects a player's behavior
141 with respect to one role in the game. For example, tit-for-tat is a strategy of memory length
142 1, where TFT cooperates if its opponent cooperated, and defects if its opponent defected.
143 An extension of memory length 1 allows TFT to remember not only its opponent's previous
144 move, but also its own previous move. However, we argue that there are situations where
145 a strategy can only be invaded by a mutant who's strategy has changed by more than 1
146 memory length. While Lindgren's operators do not guarantee that the third criteria of the
147 reduced hypothesis of escalation (strategies are improvable by variation) are satisfied, our
148 genetic operators do.

149 A similar observation on variation was made by Ikegami who uses tree-representations

150 of IPD strategies [38], where his populations exhibit escalation of memory length and di-
151 versity. However, Ikegami’s model is obscured by the use of a module-based evolutionary
152 operator. This module-based operator, akin to symbiogenesis, provides a similar variable
153 memory length extension to our normally-distributed extension/contraction operators. How-
154 ever, genetic recombination is an evolutionary transition that is expected to have emerged
155 long after populations began to escalate [39].

156 It is well known that the size of a finite evolving population can have a significant impact
157 on the fate of the population [40, 14, 15]. However, the relationship between the size of
158 a population and its ability to support the escalation of strategy memory length remains
159 unexplored. This is a particularly significant direction when considering the IPD with noise,
160 which ensures that every element of a strategy has an fitness consequence. We present long-
161 term simulation data that demonstrate a positive correlation between greater population size
162 and the evolution of longer memory lengths, suggesting that increased population size can
163 lead to enhanced evolution of strategic complexity.

164 3. Model

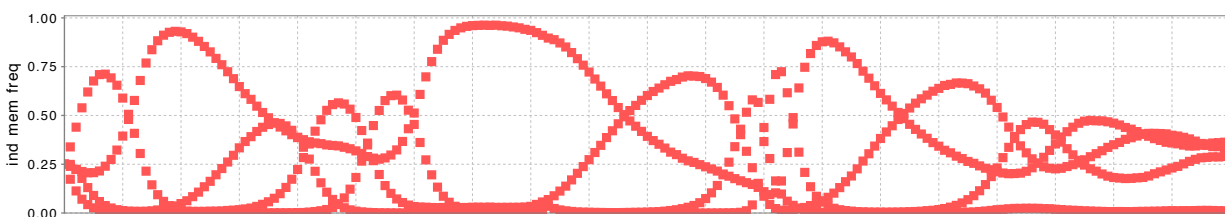


Figure 1: Example of initial evolution.

165 Our model is an extension of Lindgren’s innovative model of the IPD [7], where the use of
166 alternative genetic operators alleviates the mediocre stable-states previously observed. We
167 both suggest that this model satisfies the criteria of the reduced hypothesis of escalation and
168 empirically demonstrate the escalation of complexity in the model.

169 3.1. Prisoner’s Dilemma

170 The interactions between evolving strategies are specified by the replicator dynamics and
171 the game. We use a standard formulation of the Prisoner’s Dilemma

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

172 where the notation (p_1, p_2) indicates the scores of players 1 and 2, respectively. However,
173 this payoff matrix only specifies the score of a single round of the Prisoner’s Dilemma. In
174 the Iterated Prisoner’s Dilemma (IPD) multiple rounds are accounted for. The standard
175 way of accomplishing this is by iterating for a finite number of rounds and accumulating the
176 total score for each player during each round. The IPD becomes interesting when strategies
177 have some memory and may change their behavior depending upon the outcomes of previous
178 rounds. This is generally accomplished by encoding lookup-tables within strategies. However,

179 the accumulated score will be sensitive to the number of iterations performed. This will be
180 particularly true as strategies rely on memories of more encounters.

181 The stochastic Prisoner’s Dilemma admits an alternative method of iteration to the pre-
182 viously mentioned finite iteration technique [41]. Stochastic Prisoner’s Dilemma strategies
183 include a noise term, whereby with a certain probability strategies will take the opposite
184 action. This means the game is a Markov chain. We can describe the game as follows

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

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

192 3.2. Genetic Variation

193 We employ the same genetic encoding as Lindgren. Strategies are represented as binary
194 strings that encode the action to perform given a specific history. This is easily accomplished
195 by using the observed history as an index into the genome, where the binary value stored
196 at that position specifies the strategy’s response. In the IPD with finite rounds, strategies
197 generally also encode a sequence that specifies the “initial history” because this historical
198 lookup mechanism only works when the genome encodes the responses for all historical
199 sequences. A study of the effect of memory size on the finite-round IPD was presented in
200 [42].

201 The genetic operators first used by Lindgren [7] implement gene-doubling, gene-halving,
202 and point mutation. Point mutation is familiar from genetic algorithms, where a single bit
203 is flipped with some mutation probability. In gene-doubling, the entire bitstring is extended
204 by a factor of 2 during duplication, because the index is based on the historical observations
205 a doubling event on its own does not change the meaning of a genome. Gene halving is
206 accomplished by randomly truncating the first or second half of the genome.

207 We use variants of each of these genetic operators. Instead of point mutation, we use
208 uniform mutation, where multiple bits may be flipped during a single reproductive event.
209 To accomplish extension and contraction we draw a random number from a Gaussian dis-
210 tribution, and the absolute value of the integer component is taken as the number of ex-
211 tensions/contractions to perform. Both extension and contraction are accomplished in the
212 same way as Lindgren’s model, but extension/contraction may be more/less than a factor
213 of 2. When performing genetic operations, first the mutant genome may or may not be
214 extended/contracted, then it subsequently may or may not be subject to uniform mutation.
215 Thus in a given reproductive event a mutant may have been extended/contracted as well as
216 varied with uniform mutaiton. We now revisit a requirement of the hypothesis of escalation:
217 “strategies are evolvable.”

218 Our genetic operators ensure that it is possible to reach a large number of strategies from
219 any population distribution, while Lindgren’s operators appear to only reach a limited range

220 of genotypes. Specifically, no mutant strategy will ever be larger than 1 memory length longer
221 than the biggest genome in the population, or smaller by more than 1 memory length than
222 the smallest genome in the population. However, we hypothesize that it may be necessary
223 to invade with a strategy outside of that range, and our results suggest this is correct. There
224 has been some work on the invasion by pairs of strategies [33, 34, 35], but there is still no
225 known champion IPD strategy.

226 While Lindgren utilized the continuous-time replicator dynamics and introduced mutants
227 while time-stepping, we instead use the Moran process with mutation. The Moran process
228 models evolutionary dynamics by iteratively replacing one individual at a time [43]. In the
229 evolutionary computation literature the Moran process is sometimes called “steady-state”
230 evolution [44]. The Moran process offers an intuitive way of introducing mutant strategies
231 into the population. On the other hand, the best method of introduction of mutant strategies
232 into a mean-field model is not immediately apparent. In Lindgren’s model, each strategy has
233 a probability of introducing a single mutated variant proportional to the frequency of the
234 parent strategy.

235 4. Results

236 *Average diversity per population*
237 *size over time.* In this study we
238 simulate the previously described
239 model with the Moran process using
240 a range of population sizes.
241 For all simulations the following
242 parameters are used: $p_{extend} =$
243 0.000001 , $p_{contract} = 0.000001$,
244 $p_{uniform} = 0.001$, and $T_{max} =$
245 $100,000$ generations. We have also
246 restricted the maximum length of
247 strategies to 12; however, we never
248 observe this limit being reached.
249 The cost of simulating infinite
250 games increases exponentially with
251 the maximum memory size of the
252 competing strategies, which is a
253 strong motivation for prohibiting excessively long strategies.

254 Just as in Lindgren’s study [7], we observe similarities between all simulations (and Lind-
255 gren’s), especially during the initial generations as the system passes through meta-stable
256 states. For example compare Figure 1 to Lindgren’s Figure 1 [7], both of which exhibit the
257 same patterns in the initial phase of their evolutionary trajectory. While much of Lindgren’s
258 discussion regarding the evolutionary timeline remains intact, our model provides an epilogue
259 to Lindgren’s allusion to open-ended evolution.

260 The inclusion of noise in the IPD model admits evolutionarily stable strategies [45]. Both
261 Lindgren’s and our model do reach evolutionarily stable states under some conditions, and
262 in Lindgren’s model it is unclear whether all paths will lead to such an ESS. While Lindgren

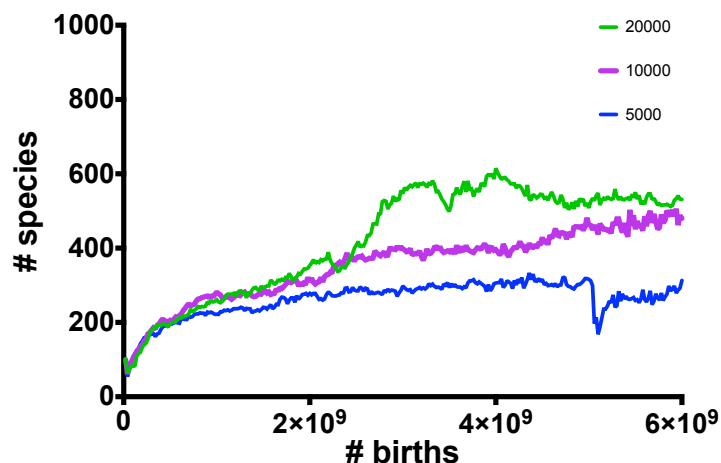


Figure 2: Species diversity for population sizes 5,000, 10,000, and 20,000 averaged over 25 simulations.

263 found some evolutionary trajectories that did not get stuck in the same memory 4 ESS
 264 that plagued many of his simulations, he did not demonstrate evolutionary trajectories that
 265 exceeded memory length 4. Here we present simulation results for evolutionary trajectories
 266 that escalate beyond memory length 4.

267 We conducted experiments using the Moran process with population sizes: 5,000 and
 268 10,000. 25 replicates were used for each population size. Although we were not able to sim-
 269 ulate all population sizes for the full 100,000 generations, we present results where a number
 270 of evolutionary trajectories pass the memory 4 meta-stable state. When comparing results
 271 each timestep represents a generation, which is N breeding events, where N is the popu-
 272 lation size. Therefore, when 2 simulations are compared with different population sizes, at
 273 any given timestep each population will have experienced a different number of reproductive
 274 events (i.e. $N * t$).

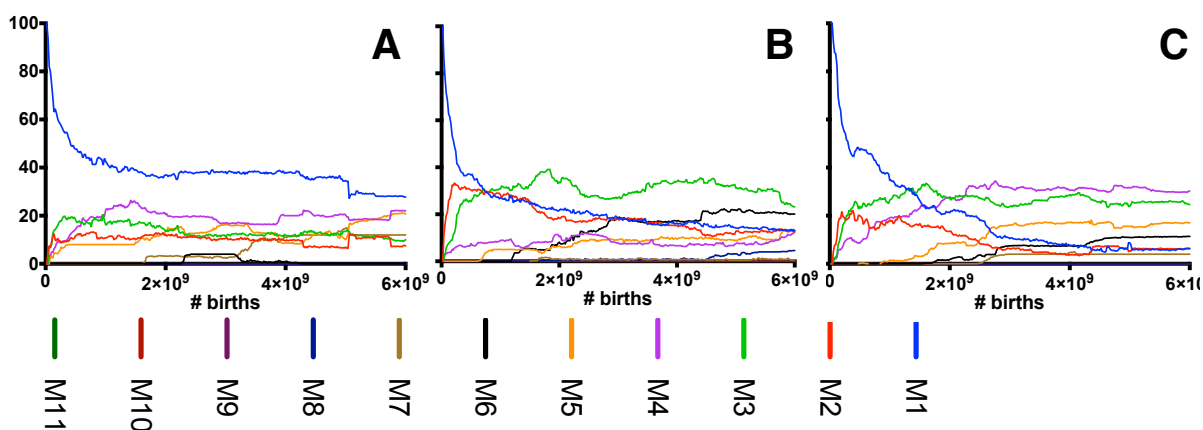


Figure 3: Fraction of population composed of each memory length for Population size 5,000, 10,000, 20,000. Runs are averaged over 25 simulations.

275 The smallest population size that we consider is 5,000 (Figure 3A). In this case a number
 276 of simulations are unable to escape the initial meta-stable states, and populations remain at
 277 low memory lengths. However, some simulations do reach populations consisting of primarily
 278 memory length 7. We will revisit this observation for some of a larger population size. The
 279 number of species grows for the first 70,000 generations, then plateaus just below 300 species.
 280 However, even during this plateau of species diversity, some escalation can still be observed
 281 as strategies of memory length 7 are still on the rise at generation 100,000.

282 Our results for population sizes 10,000 and 20,000 show the most escalation in memory
 283 length (Figures 3B-C). Results the population size of 20,000 show runs that are beginning to
 284 be dominated by memory length 8. These runs are the most escalated of all experiments we
 285 conducted. The average number of species is similar across the runs to the extent that the
 286 runs are completed. However, the correlation between the completion of runs and number
 287 of species is clear. Spikes in the number of species significantly slow simulation speeds. For
 288 this reason, we cannot make clear statements regarding the number of species supported by
 289 each population size.

290 By extending Lindgren’s model with alternative genetic operators we have cleared the
291 path to open-ended evolution in the IPD model. We explore the model using finite pop-
292 ulation evolutionary dynamics, as opposed to Lindgren’s use of continuous-time replicator
293 dynamics. The model continues to exhibit similar evolutionary trajectories to those pre-
294 sented by Lindgren, which suggests that it is not our use of the Moran process that leads to
295 the escape from the memory 4 meta-stable states that appeared to limit Lindgren’s original
296 model. The computational cost of simulating large population sizes causes us to present
297 partial results. While we see that larger population sizes are capable of supporting a larger
298 number of species, larger population size does not eliminate the possibility of getting stuck
299 in an evolutionary equilibrium. This leads to the suggestion that achieving greater escalation
300 is not simply a matter of using a larger population size.

301 Now let us consider a specific example trajectory from a population size 20,000 run. In
302 Figure 4 we have a timeline showing the evolutionary history after 2×10^9 birth events. Over
303 the course of this evolutionary trajectory the population transitions to the previously observed
304 limit of memory length 4 to memory length 6 and on to memory length 8. As we noted in
305 Figure 2, the diversity of species increases significantly toward the end of population making
306 analysis of individual strategies challenging. To this end we perform a species knockout
307 analysis at multiple points within the evolutionary trajectory.

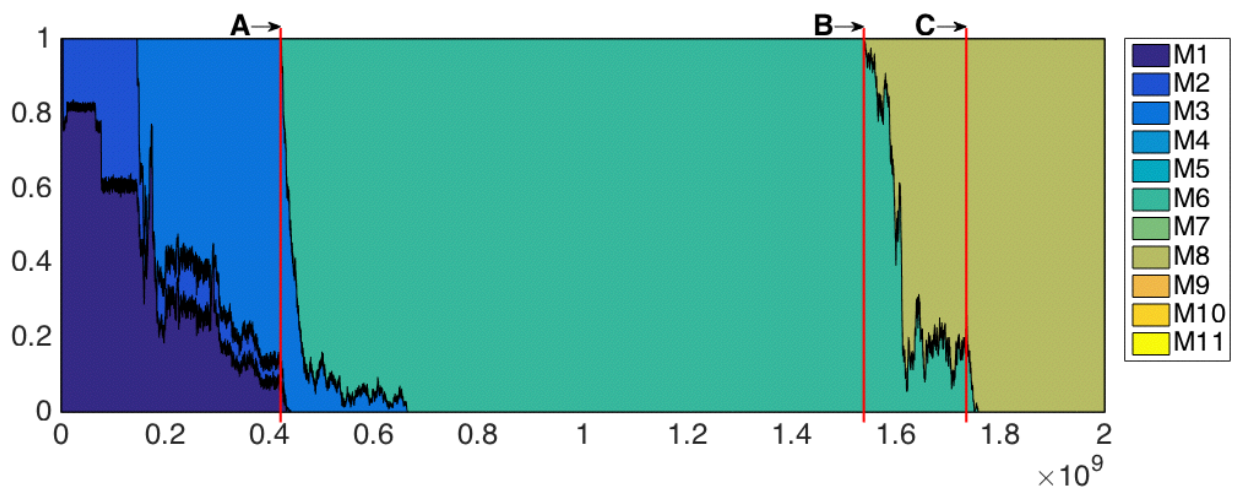


Figure 4: Fraction of population composed of each memory length for an example run with population size of 20,000. Annotated vertical lines indicate populations where specie knockout analyses are performed.

308 *Species knockout analysis.* An analysis of population stability was performed via species
309 knockout, where a given strategy is eliminated from the population. The population is
310 rebalanced by uniformly allocating the previously occupied fraction of the population to the
311 remaining strategies. After performing the knockout, the simulation is evaluated for 4×10^5
312 birth events, then the distribution of memory lengths are investigated. For each timepoint a
313 species knockout is performed with respect to each strategy in the population. The timepoints
314 of knockouts are denoted in Figure 4.

315 Knockout A is performed at the transition from memory length 4 to 6. Memory length 4
316 is the level where Lindgren found a tendency for populations to stabilize [7]. The knockout

317 is performed at the first generation where memory length 6 strategies appear. Of the 17
318 species knockouts performed, 2 knockouts lead to a collapse of escalation, while the remaining
319 knockouts continue to support memory length 6 strategies.

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

327 Knockout C is performed when the majority of the population is occupied by memory
328 length 8 strategies (approx 30% to 70%). Here we find that all 157 knockouts maintain pop-
329 ulations with memory length 8 strategies. This suggests that the interdependence observed
330 in knockout B has stabilized, and the population has become more robust to the distribution
331 of strategies that it contains.

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

341 5. Conclusion

342 The study of coevolutionary arms races has had a challenging history plagued with prema-
343 ture mediocre stabilization [12] and other coevolutionary pathologies [16]. These pathologies
344 were previously related to observations of limited escalation of complexity in simple evolu-
345 tionary models of cooperative games [7]. In our study we have drawn inspiration from the
346 macroevolutionary theory of the escalation of coevolutionary interactions [22] to show that
347 previous observations of limited evolution in the Iterated Prisoner's Dilemma with noise [7]
348 were due to a lack of evolvability. By conducting long-term evolutionary simulations we
349 have shown that an improved model can lead to continued escalation of strategic complex-
350 ity. We have also shown that strategies escalate in complexity faster in larger populations.
351 Coevolutionary simulation can drive the escalation of complexity and that escalation can be
352 amplified in larger population sizes. Furthermore, the escalation of complexity can be sen-
353 sitive to species knockouts during transition periods. Thus, the stabilization of species and
354 maintenance of large population sizes are viable mechanisms to supporting the escalation of
355 strategic complexity.

356 6. Acknowledgements

357 We thank Sevan Ficici and Anthony Bucci for insightful discussions, and Kristian Lind-
358 gren for sharing his code.

- 359 [1] T. Taylor, M. Bedau, A. Channon, D. Ackley, W. Banzhaf, Open-Ended Evolution:
360 Perspectives from the OEE1 Workshop in York, *Artificial Life* (2016).
- 361 [2] Y. Yokobayashi, Directed evolution of a genetic circuit, *Proceedings of the National
362 Academy of Sciences* 99 (2002) 16587–16591.
- 363 [3] E. Boder, K. Midelfort, K. Wittrup, Directed evolution of antibody fragments with
364 monovalent femtomolar antigen-binding affinity, *Proceedings of the National Academy
365 of Sciences* 97 (2000) 10701–10705.
- 366 [4] H. Lipson, J. B. Pollack, Automatic design and manufacture of robotic lifeforms, *Nature
367* (2000) 974–978.
- 368 [5] J. Koza, Genetic programming: on the programming of computers by means of natural
369 selection, 1992.
- 370 [6] R. Lenski, M. Rose, S. Simpson, S. Tadler, Long-term experimental evolution in *Es-
371 cherichia coli*. I. Adaptation and divergence during 2,000 generations, *American Natu-
372 ralist* (1991).
- 373 [7] K. Lindgren, Evolutionary phenomena in simple dynamics, in: *Artificial Life II*, pp.
374 295–312.
- 375 [8] S. Gould, N. Eldredge, Punctuated equilibria: an alternative to phyletic gradualism, in:
376 *Models in paleobiology*, 1972, pp. 82–115.
- 377 [9] P. Ehrlich, P. Raven, Butterflies and plants: a study in coevolution, *Evolution* 18 (1964)
378 586–608.
- 379 [10] O. Pellmyr, J. N. Thompson, J. M. Brown, R. G. Harrison, Evolution of pollination and
380 mutualism in the yucca moth lineage, *The American Naturalist* 148 (1996) 827–847.
- 381 [11] D. Cliff, G. F. Miller, Co-Evolution of Pursuit and Evasion II: Simulation Methods and
382 Results, ????
- 383 [12] S. G. Ficici, J. B. Pollack, Challenges in coevolutionary learning: Arms-race dynamics,
384 open-endedness, and mediocre stable states, in: *Proceedings of the sixth international
385 conference on Artificial life*, pp. 238–247.
- 386 [13] D. B. Fogel, G. B. Fogel, P. C. Andrews, On the instability of evolutionary stable
387 strategies, *Biosystems* 44 (1997) 135–152.
- 388 [14] S. Ficici, J. Pollack, Effects of finite populations on evolutionary stable strategies, in:
389 *Proceedings of the Genetic and Evolutionary Computation Conference*.
- 390 [15] M. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, Emergence of cooperation and evolu-
391 tionary stability in finite populations, *Nature* 428 (2004) 646–650.
- 392 [16] R. Watson, J. Pollack, Coevolutionary dynamics in a minimal substrate, in: *Proceedings
393 of the Genetic and Evolutionary Computation Conference, GECCO-01*, pp. 702–709.

- 394 [17] S. G. Ficici, Solution concepts in coevolutionary algorithms, Ph.D. thesis, Brandeis
395 University, 2004.
- 396 [18] A. Bucci, J. B. Pollack, A mathematical framework for the study of coevolution., in:
397 FOGA, volume 7, pp. 221–235.
- 398 [19] L. van Valen, A new evolutionary law, *Evolutionary Theory* 1 (1973) 1–30.
- 399 [20] S. Nee, Birth-death models in macroevolution, *Annu. Rev. Ecol. Evol. Syst.* 37 (2006)
400 1–17.
- 401 [21] G. Vermeij, The evolutionary interaction among species: selection, escalation, and
402 coevolution, *Annual Review of Ecology and Systematics* 25 (1994) 219–236.
- 403 [22] G. J. Vermeij, *Evolution and escalation: an ecological history of life*, 1987.
- 404 [23] R. Dawkins, J. Krebs, Arms races between and within species, *Proceedings of the Royal
405 Society of London. Series B. Biological Sciences* 205 (1979) 489–511.
- 406 [24] K. I. S. Harrington, *Escalation of Coevolutionary Arms Races*, Ph.D. thesis, Brandeis
407 University, 2014.
- 408 [25] J. Madin, J. Alroy, M. Aberhan, F. Fürsich, Statistical independence of escalatory
409 ecological trends in Phanerozoic marine invertebrates, *Science* 312 (2006) 897–900.
- 410 [26] D. Jablonski, Biotic interactions and macroevolution: extensions and mismatches across
411 scales and levels, *Evolution* 62 (2008) 715–739.
- 412 [27] G. Vermeij, Escalation and its role in Jurassic biotic history, *Palaeogeography, Palaeo-
413 climatology, Palaeoecology* 263 (2008) 3–8.
- 414 [28] R. Axelrod, W. Hamilton, The evolution of cooperation, *Science* 211 (1981) 1390.
- 415 [29] M. Nowak, K. Sigmund, Tit for tat in heterogeneous populations, *Nature* 355 (1992)
416 250–253.
- 417 [30] M. Nowak, K. Sigmund, A strategy of win-stay, lose-shift that outperforms tit-for-tat
418 in the Prisoner’s Dilemma game, *Nature* 364 (1993) 56–58.
- 419 [31] W. Press, F. Dyson, Iterated Prisoner’s Dilemma contains strategies that dominate any
420 evolutionary opponent, *Proceedings of the National Academy of Sciences* 109 (2012)
421 10409–10413.
- 422 [32] C. Adami, A. Hintze, Winning isn’t everything: Evolutionary stability of Zero Deter-
423 minant strategies, *Nature Communications* 4 (2012) 1038.
- 424 [33] R. Boyd, J. Lorberbaum, No pure strategy is evolutionarily stable in the repeated
425 Prisoner’s Dilemma game, *Nature* 327 (1987) 58–59.
- 426 [34] J. Lorberbaum, No strategy is evolutionarily stable in the repeated prisoner’s dilemma,
427 *Journal of Theoretical Biology* 168 (1994) 117–130.

- 428 [35] J. Lorberbaum, D. Bohning, Are there really no evolutionarily stable strategies in the
429 iterated prisoner's dilemma?, *Journal of theoretical biology* 214 (2002) 155–169.
- 430 [36] J. Li, G. Kendall, The Effect of Memory Size on the Evolutionary Stability of Strategies
431 in Iterated Prisoner's Dilemma, *Evolutionary Computation, IEEE Transactions ...*
432 (2014).
- 433 [37] C. Hilbe, L. A. Martinez-Vaquero, K. Chatterjee, M. A. Nowak, Memory-n strategies of
434 direct reciprocity, *Proceedings of the National Academy of Sciences* (2017) 201621239.
- 435 [38] T. Ikegami, From genetic evolution to emergence of game strategies, *Physica D: Non-*
436 *linear Phenomena* 75 (1994) 310–327.
- 437 [39] J. Smith, E. Szathmary, *The major transitions in evolution*, 1997.
- 438 [40] D. Fogel, G. Fogel, Evolutionary stable strategies are not always stable under evolution-
439 ary dynamics, in: *Evolutionary Programming IV*, Cambridge, MA: MIT Press, 1995,
440 pp. 565–577.
- 441 [41] M. Nowak, Stochastic strategies in the prisoner's dilemma, *Theoretical population*
442 *biology* 38 (1990) 93–112.
- 443 [42] C. Hauert, H. G. Schuster, Effects of increasing the number of players and memory size
444 in the iterated Prisoner's Dilemma: a numerical approach, *Proc. R. Soc. Lond. B* 264
445 (1997) 513–519.
- 446 [43] P. Moran, *The statistical processes of evolutionary theory.*, 1962.
- 447 [44] G. Syswerda, A study of reproduction in generational and steady-state genetic algo-
448 rithms, in: *Foundation of Genetic Algorithms*, pp. 94–101.
- 449 [45] R. Boyd, Mistakes allow evolutionary stability in the repeated prisoner's dilemma game,
450 *Journal of Theoretical Biology* 136 (1989) 47–56.