Distinguishing Adaptive from Non-Adaptive Evolution Using Ashby's Law of Requisite Variety.

Anil K Seth 1,2

¹The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, CA 92121, USA ²CCNR/COGS, University of Sussex, Brighton, BN1 9QH, UK

seth@nsi.edu

Abstract— The introduction of noise into an evolutionary Iterated Prisoner's Dilemma model can promote the evolution of strategy memory. In this paper, an analysis in terms of Ashby's Law of Requisite Variety enables a distinction to be drawn between the adaptive evolution of memory and non-adaptive evolution, or drift, a distinction which in such situations is remarkably insensitive to fitness statistics. As part of this, it is demonstrated how the influence of noise can depend on its locus in the evolving system. Additional evidence is presented for a second influence of noise in facilitating drift, an influence which can be interpreted in terms of noise-induced genotype-phenotype degeneracy.

I. INTRODUCTION

The Iterated Prisoner's Dilemma (IPD) provides a convenient platform for the exploration of phenomena in evolutionary dynamics. Previous work has indicated that the introduction of noise into an IPD system can encourage the evolution of strategies with increased memory [5]-[7], however it was not clear whether the evolved memory was an adaptation to the presence of noise, or a serendipitous sideeffect, an instance of (non-adaptive) genetic drift somehow facilitated by noise. This is an example of the general problem of distinguishing adaptation from non-adaptive drift in evolving systems. Here we examine this question in the context of the IPD with an analysis inspired by W. Ross Ashby's Law of Requisite Variety (LRV) [1], which suggests how the influence of noise should critically depend on its locus in the system. On this basis, evidence is found that, in some cases, evolved memory can indeed be understood in terms of adaptation to noise. Interestingly, fitness metrics alone were not sufficiently sensitive to reveal this effect.

In other cases, however, it is argued that noise and memory are *not* adaptively linked. In these cases evidence is presented for a role for noise in facilitating genetic drift, an effect interpreted in terms of noise-induced degeneracy in genotype-phenotype mappings, where degeneracy refers to the ability of structurally distinct genotypes to yield the same phenotype [3].

A. The Iterated Prisoner's Dilemma

The IPD is an iterated non-zero-sum game, for two or more players, in which each player chooses either to *coop*erate or defect on any given iteration. Each player is ignorant of the present move (though not necessarily of the history of moves) of its opponent, with payoffs distributed according to table I. Although IPD models are renowned for providing insight into how cooperation can arise in a



Fig. 1. IPD model genotype encoding scheme. Each time a previous move in the game history (between two particular individuals) is considered, half of the genotype is (temporarily) discarded (the non-shaded areas) - one half if the move had been cooperative, the other for a defection. Thus a genotype of length 16 (as above) can encode a strategy of memory of 4 moves. For example, the black square indicates which allele would be accessed for a [c,d,c,d] history. In practice the genotype must actually be longer in order to specify the initial moves up until this memory limit is reached (in this case an extra 15 alleles). This scheme is based on Lindgren [5].

	$B \ coop$	B defect
A coop	A 3 pts - B 3 pts	A 0 pts - B 5 pts
A defect	A 5 pts - B 0 pts	<i>A</i> 1 pt - <i>B</i> 1 pt

TABLE I Standard IPD payoff table for 2 players.

population of selfish individuals [2], the present focus is on the relations between strategy memory and noise (defined later). Following previous work [6]-[7], IPD ecologies here consist of populations of variable length genotypes, each encoding a strategy for playing the IPD (fig. 1). A simple GA (population size 30) applies crossover (p = 0.95), point mutation, and 'splitting' and 'doubling' mutations (all p = 0.005) to the genotypes; doubling copies the latter half of the genotype twice again onto its own end, thereby increasing the strategy memory by 1, splitting reverses the effect. Genotype fitness is simply the accumulated payoff after a series of IPD interactions.

Again following [6]-[7], two distinct IPD models are analysed in what follows: a 'compulsory' model, in which each individual engages in the IPD with every other individual in the population (for 60 iterations per contest), and an 'IPD/CR' model in which individuals choose and refuse partners with reference to continuously updated expected payoffs that each individual maintains for every other. The IPD/CR algorithm is based on Stanley et al. [8], and is fully described in Appendix A. For present purposes, as we shall see, the significant difference between the two models is that stable cooperation evolves with much greater reliability in the latter. But first let us consider the LRV.

B. The Law of Requisite Variety

For an agent to maintain relative stability in certain (internal) essential variables (for example heart rate, body core temperature), it must prevent the transmission of environmental variability through to these essential variables. In the same way that a good thermostat prevents the transmission of environmental variations in temperature through to a particular object (for example, the interior of a refrigerator should remain at a constant cool temperature despite the fluctuating temperature of a kitchen on a midsummer day), a well adapted agent prevents the transmission of certain environmental variables (for example the prevalence or scarcity of food, the proximity, or otherwise, of predators) through to such essential internal variables as blood sugar level or heart rate.¹ With this in mind, the LRV can be easily formulated: Consider a set of possible environmental disturbances D, a set of possible responses on the part of the agent, R, and a set of possible outcomes, O. Consider also that for each D_i , there is distinct outcome O_i , and a particular response R_i , with these mappings mediated by the 'system' S. Stability in the essential variables requires minimising the variation in O, and this then requires that the variety in D is matched by the variety in R. Ashby himself provides a concise summary: "If R's move is unvarying, then the variety in outcomes will be as large as the variety in D's moves; only variety in R's moves can force down variety in the outcomes" ([1], p.206).

Ashby also reminds us that not all environmental variability need threaten the survival of the agent, arguing for two fundamental forms: "There is that which threatens the survival of the gene-pattern. This part must be blocked at all costs. And there is that which, while it may threaten the gene-pattern, can be transformed (or re-coded) through the regulator R and used to block the effect of the remainder" (ibid, p.212). According to Ashby, environmental variability can be either potentially beneficial or downright dangerous.

Now consider the IPD in terms of the components of the LRV: D, R, S, and O. D would be the set of all moves made over all iterations of the game, by all the opponents (of a given agent). R would be the strategy of the agent, S would be the IPD payoff table, and O would be the set of scores awarded. In a stable, cooperating population, the overall fitness of each agent (over many generations) will be maximised with minimum variation in O away from repeated cooperation payoffs (since mutual cooperation, in the IPD, maximises overall payoff in the population). Any deviation from stable cooperation can then be expected to have a deleterious effect on fitness, unless it is countered by

a strategy that can effectively cope with such deviations. For example, an occasional 'accidental' defection will throw a population of 'tit-for-tat' players into continual mutual defection but a 'tit-for-two-tats' population will 'absorb' such a defection, permitting general cooperative behaviour to persist throughout the population. In terms of the LRV, any variety in D can only be prevented from affecting O if it is countered through the action of R on S; that is, if the variety in R (given S) matches that of D. Given the assumption that environmental variability can be associated with noise ('dangerous' environmental disturbance, which need not be structured or predictable), this construal of the IPD allows us to predict that noise (variety) on D may lead to the evolution of longer memories (supporting variety in R) to counter this disturbance, but that noise on O will not. Noise on the outcome, after all, is precisely what the strategies would be expected to prevent, and if the noise is applied directly to O, then, by definition, no strategy can provide an effective response, because payoffs have already been determined.

Two types of noise must therefore be distinguished: (1) M-noise; on each iteration, for each player, there is a certain probability that the move specified by the genotype is flipped, and only then are the payoff scores calculated, and (2) O-noise; on each iteration, for each player, there is a certain probability that the payoff awarded is selected randomly from the 4 possible payoff values. By the above argument, *if* evolved memory is adaptively significant, Mnoise should encourage the evolution of longer memories, but O-noise should not.

Importantly, both types of noise can be thought of as inducing genotype-phenotype degeneracy, but in different ways. M-noise makes the phenotype expressed by any given genotype slightly less predictable, and thus may tend to obscure the differences between phenotypes (in the limit, 100% M-noise would entail total genotype-phenotype degeneracy, all phenotypes would be identical). O-noise tends to homogenise fitnesses of genotypes/phenotypes, thus obscuring phenotypic differences from the point of view of selection (in the limit, 100% O-noise would entail all phenotypes had the same fitness).

II. ANALYSIS OF THE MODEL

Both compulsory and IPD/CR models were analysed in each of 3 conditions; 2% M-noise, 2% O-noise, and zero noise. In each condition (in each model) 12 evolutionary runs of 10,000 generations each were performed. All populations were initialised with a universal memory of 1. For the compulsory model, Fig. 2 shows that with zero noise, long memories never really evolve: evolved memory tends to stay either at the initial level of 1, or drop to zero. With 2% M-noise, long memories *do* evolve; not always, and not always to the maximum, but it does happen. However 2% O-noise has the *same* effect as 2% M-noise, suggesting that the evolved memory is *not* adaptively significant. These observations are statistically significant (by t-test, p < 0.05between M-noise and zero noise, p < 0.05 between O-noise and zero noise, and p > 0.5 between M-noise and O-noise).

¹Of course, these variables can and do vary within certain limits, but trespass beyond these limits is severely maladaptive.



Fig. 2. Evolution of memory in compulsory (a,c,e) and IPD/CR (b,d,f) models. Each figure superimposes results from 12 evolutionary runs. Average population memory is shown. See text for details.

A very different pattern emerges from the IPD/CR model (also fig. 2), for which there is significantly more evolved memory in the 2% M-noise condition than with 2% O-noise or zero noise (p < 0.01 in each case), but no significant difference between the zero noise and the 2% O-noise conditions (p > 0.1). These results suggest that the evolution of memory in the IPD/CR model is indeed adaptively significant. (It should also be recorded that in this model, initial observations indicated that the population rapidly reached maximum memory in all conditions, and so, in order to differentiate the conditions, a small fitness penalty on evolved memory was applied in all conditions: 0.0025% of total fitness per memory unit. The evolution of memory under *M*-noise despite this cost further attests to its adaptive significance. To control for this modification,

the same fitness cost was also introduced to compulsory models under 2% M-noise and O-noise: the evolution of memory was completely abolished in both cases: fig. 3.)

A second set of experiments corroborated these results. Modified compulsory and IPD/CR models were analysed in which, although heritable memory was allowed to vary normally, a maximum *functional* memory of 1 was imposed. In these models, then, there is no possible functional difference, with regard to the IPD, between strategies of memory 1 and of memory 6. Therefore, if memory is evolving as an adaptation to the presence of M-noise, it may be expected to be abolished by this modification. As fig. 4 clearly illustrates, whilst this is indeed the case for the modified IPD/CR model (with 0.0025% fitness cost per memory unit), it is not the case for the modified compul-



Fig. 3. Compulsory model with 0.0025% fitness cost per memory unit. Average population memory is shown. Each figure superimposes results from 12 evolutionary runs.



Fig. 4. Compulsory and IPD/CR models with a functional memory limit of 1, 2% M-noise. Average population memory is shown. Results from 12 evolutionary runs (10,000 generations each) are superimposed.



Fig. 5. Fitness and memory in compulsory and IPD/CR models under 2% M-noise. See text for details.



Fig. 6. Average and standard deviation cooperation in compulsory and IPD/CR models (statistics generated from 12 evolutionary runs, 10,000 generations each).

sory model.

We can begin to understand these differences by inspecting how the two kinds of model perform in general. Average cooperation in the IPD/CR model is high, mandating application of the LRV framework (fig. 6b; the slightly lower - but still high - average cooperation under M-noise may be understood to result from the evolved strategies failing to cope perfectly with noise.) However in the compulsory model average cooperation is low (fig. 6a; for present purposes we may overlook the differences between O-noise and the other conditions, all are low when compared to 6b). And if there is no stable cooperating population to begin with, there is no reason (with respect to the LRV) to expect long memories to evolve to cope with noise. The question therefore arises: why do long memories *ever* evolve in the compulsory model in either 2% O-noise or 2% M-noise conditions?

A. The insufficiency of fitness metrics

Before addressing this question it is worth forestalling a possible objection. Why, one might ask, bother with the LRV at all? Surely adaptive and non-adaptive evolution can be distinguished simply from fitness statistics?

In this case not, or at least not easily. First, consider a single run of the compulsory model under 2% M-noise. Fig. 5(a) scatter-plots fitness and memory of the entire population every 200 generations for 10,000 generations (as defined in section I-A). At this gross level, no relation is discernible. Fig. 5(b) represents the same data in a different way, showing only those generations for which there is a non-zero diversity in memory across the population, normalising the recorded range in such cases to [0.0, 1.0], and plotting this data against ranked and normalised fitness values; once again there is no discernable relation. So far so good - we already have reason to believe that evolved memory in the compulsory model is not adaptive. However, repeating this analysis for the IPD/CR model (under 2% M-noise) reveals exactly the same pattern (figs. 5c and d). Therefore there is no evidence from these fitness measures that longer memories, in the IPD/CR model with M- noise, provide any reliable selective advantage over shorter memories. One may consider the interesting implication that adaptive evolution cannot always be distinguished from non-adaptive evolution with confidence by relying on simple fitness metrics. (The possibility must nevertheless be left open that other statistical fitness measures and/or visualisation techniques than those deployed here may enable some discrimination to be made. We must be content to say that fitness measures do not *readily* illuminate such discriminations.)

None of this is to say that longer memories *do not* provide fitness advantages in the IPD/CR model with M-noise. At some level of analysis they must, if it is true that they evolve as adaptations. However, because IPD fitness landscapes are extremely rugged and prone to rapid change, these advantages may show themselves only rarely and unpredictably. For example, in a population undergoing a rapid change from high fitness to low fitness, it may not be appropriate to draw any conclusions from observing the memory of the 'fittest' individual. In other words, the large-scale population dynamics of IPD ecologies may well disguise any direct selective advantages of longer memories, without implying that these selective advantages do not exist.

B. The facilitation of genetic mobility

To return - finally - to the puzzling and apparently nonadaptive evolution of memory in the compulsory model, here we briefly evaluate the hypothesis that this is an example of enhanced genetic drift promoted by noise. A compulsory IPD model is considered in which strategy memory is genetically limited to being either 0 or 1 (the average population memory can range continuously from 0.0 to 1.0), with the motivation of investigating differences between the three noise conditions that are relatively independent of the evolution of memory.² In other words, to try to distinguish between the causes of the evolved memory, and a mixture of the causes and consequences. If strategy memory is given free reign to evolve, the causes and consequences of its evolution cannot be rigorously distinguished. The following experiments allow such a distinction to be drawn - if not rigorously, then at least in outline. Twelve evolutionary runs were performed in each of the 3 conditions; zero noise, 2% O-noise, and 2% M-noise. Fig. 7 illustrates that in both O-noise and M-noise conditions, average evolved memory is near to the expected average (0.5). But with zero noise, in each case the average evolved memory (usually) either stays near the initial level of 1.0, or is very close to 0.0. This indicates that both O-noise and M-noise are permitting evolution to wander easily within the prescribed area of genotype space (determined by the maximum memory limit of 1), but with zero noise, the evolving population often becomes 'stuck' on either memory 1 or memory 0. Furthermore, the means of these average memories are not significantly different between any of the conditions. The average (of the average memory) in the M-noise condition is 0.26; in the O-noise, 0.34; and with zero noise, 0.41, and pairwise t-tests between all sets of means all give p > 0.05, indicating that neither type of noise is actually encouraging (or discouraging) the evolution of memory per se in this

 $^{^{2}{\}rm Of}$ course, this genetic limitation entails a corresponding functional limitation.



Fig. 7. Compulsory model; memory genetically limited to the range [0,1]. A - zero noise, B - 2% M-noise, C - 2% O-noise.

model. All this strongly suggests that noise *does* facilitate genetic drift, at least in the dimension reflected by strategy memory in the present model.

What of the origins of this effect? We have seen that another way to consider noise in an evolving system is as inducing genotype-phenotype degeneracy, and this may be enough. Consider an analogy with the Baldwin effect, which describes how lifetime learning can 'smooth out' fitness landscapes [4]. The principle of this effect is that if individuals vary genetically in their ability to learn some trait, those most able to do so will leave the most descendants, and over time the trait itself will become genetically assimilated. Hinton & Nowlan provide a simple simulation model of this effect [4], in which learning is instantiated as random search, which can be informally construed as noise, or genotype-phenotype degeneracy, through its effect that any given genotype will be awarded a greater diversity of fitnesses than would otherwise have been the case. Thus it is possible to imagine that noise, through inducing genotype-phenotype degeneracy and smoothing out sharp peaks in fitness landscapes, could enhance genetic mobility. Further work is clearly required to fully explore this possibility.

III. SUMMARY

In this paper, analysis of both the compulsory IPD model and the IPD/CR model indicated that noise encourages the evolution of strategy memory. However, only in the IPD/CR model can this memory be appropriately interpreted as an adaptive response, an interpretation justified by (a) the abolition of such evolution if M-noise is replaced by O-noise, or under the application of a functional memory limit, and (b) the evolution of memory despite a fitness penalty. The importance of the LRV framework in making this distinction - interestingly, fitness metrics alone were not sufficient - suggests that similar analyses may help contrast adaptive from non-adaptive evolution in other difficult cases, and help explore the influence of noise in evolving systems in general.

Further analysis of the compulsory model provided evidence for a second influence of noise: the facilitation of genetic drift. This is not anticipated by the LRV framework, which views noise as either potentially useful or potentially harmful (section I-B). One interpretation is that noise can be thought of as inducing genotype-phenotype degeneracy which may 'smooth out' fitness landscapes allowing greater mobility throughout genotype space.

A final comment concerns the term 'redundancy', which is often employed in discussions of genotype-phenotype mappings where I have used 'degeneracy'. As Edelman & Gally point out, however, 'redundancy' refers to the same function (or outcome) realised by multiple *identical* structures. They argue that this term is often misused in place of 'degeneracy', which refers to "the ability of elements that are structurally different to perform the same function or yield the same output" [3] (p.13763). It is clear, at least in the present case and perhaps more generally throughout the relevant literature, that degeneracy is the more appropriate term.

Acknowledgements

I am grateful to the CCNR and to the Neuroscience Research Foundation for financial support, and to my colleagues in the CCNR and the Neurosciences Institute, also my anonymous reviewers, for constructive advice.

Appendix A: the IPD/CR algorithm

Each individual maintains a set of 'expectation values' E_i for every other individual (indexed by i). These values reflect the expected outcome of a prisoner's dilemma interaction, and each E_i is initially set to 3.0. For each subsequent iteration, each individual signals its willingness to play to n of its most preferable partners (determined by the values of E_i for all *i*; *n* is a constant); the recipients thus form lists of those individuals who have proposed to them. Every individual now has a list of potential partners, and engages in IPD interactions with those partners for whom $E_i \geq \omega$, where ω is a fixed threshold value common to all individuals; both participants are awarded payoffs according to table I. If, however, $E_i < \omega$, the refusee is awarded a payoff of value 1.0 (mutual defection payoff). After each interaction or refusal, values of E_i are updated according to $E_i[t+1] =$ $\gamma E_i[t] + (1 - \gamma)p[t]$, where γ is a (fixed) learning rate, and p[t] is the payoff awarded to that individual during the IPD interaction or refusal event at time t (note that only the refuse is updated after a refusal). After all accepted interactions have occurred, individuals who have not interacted at all are awarded a 'wallflower' payoff ω . Here, $\gamma = 0.7, n = 2, \omega = 1.6$.

References

- W.R. Ashby. An introduction to cybernetics. Chapman Hall, London, 1956.
- [2] R. Axelrod. The evolution of cooperation. Basic Books, New York, 1984.
- [3] G.M. Edelman and J. Gally. Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences*, 98(24):13763–13768, 2001.
- [4] G.E. Hinton and S.J. Nowlan. How learning can guide evolution. Complex Systems, 1:495–502, 1987.
- [5] K. Lindgren. Evolutionary phenomena in simple dynamics. In C.G. Langton, J.D. Farmer, S. Rasmussen, and C. Taylor, editors, *Artificial life II*, volume XVI of *Santa Fe Institute studies in the sciences of complexity*, pages 295–312. Addison-Wesley, Reading, MA, 1991.
- [6] A.K. Seth. Interaction, uncertainty, and the evolution of complexity. In P. Husbands and I. Harvey, editors, *Proceedings of the Fourth European Conference on Artificial Life*, pages 521–530, Cambridge, MA, 1997. MIT Press.
- [7] A.K. Seth. The evolution of complexity and the value of variability. In C. Adami, R.K Belew, H. Kitano, and C.E. Taylor, editors, Artificial Life VI: Proceedings of the Sixth International Conference on the Simulation and Synthesis of Living Systems, pages 209–221, Cambridge, MA, 1998. MIT Press.
- [8] E.A. Stanley, D. Ashlock, and L. Tesfatsion. Iterated prisoner's dilemma with choice and refusal of partners. In C.G. Langton, editor, Artificial life III, volume XVII of Santa Fe Institute studies in the sciences of complexity, pages 131–175. Addison-Wesley, Reading, MA, 1994.