A Cybernetic Perspective on the Role of Noise in the Iterated Prisoner's Dilemma.

CSRP 506

Anil K Seth

Centre for Computational Neuroscience and Robotics
and School of Cognitive and Computing Sciences
University of Sussex
Brighton, BN1 9QH, UK
anils@cogs.susx.ac.uk

January 21, 1999

Abstract

An interpretation of the evolution of complexity in the Iterated Prisoner's Dilemma (IPD) is developed, based on Ashby's "law of requisite variety". It is demonstrated that the influence of *noise* on the evolutionary dynamics of this system is critically dependent on the locus of this noise. It is also argued that noise in such an evolving system is not merely, (or necessarily) a source of variation that must be adapted to, but, in certain circumstances, can facilitate the evolutionary exploration of increased areas of genotype space.

This convergence between artificial life/game theory and cybernetics holds implications for how we understand the general relationship

1   Introduction

The general principle that there is organism complexity by virtue of environmental complexity has been well explored philosophically, [6],[13],[1],[4], and there have been several recent contributions, within artificial life, to an empirical foundation for this position, [5],[8],[9],[11],[10].

This report extends work presented in [9] demonstrating that the introduction of *noise*

that pertains to this research; Ashby's "Law of Requisite Variety" (LRV), [1].

Ashby also reminds us that not all environmental variability need threaten the survival of the agent, and argues that variety comes in 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, then, environmental variability can be either poten-

	player 2 cooperates	*player 2 defects*
player 1 cooperates	1:R=3 2:R=3	1:S=0 2:T=5
player 1 defects	1:T=5 2:S=0	1:P=1 2:P=1

Table 1: IPD payoff matrix; note that the actual scores are unimportant as long as $T > R > P > S$ and $2R > T + S$.

evolution of *complex* strategies. The ecology consisted of a population of variable length genotypes, with each genotype coding for a particular strategy for playing the IPD, and with the length of each genotype determining the number of previous iterations it could take account of when delivering its next move. This strategy "memory" can be taken to be a metric of complexity[3]. Noise was incorporated by introducing a certain probability that the move-to-be-made on any given iteration (as specified by the genotype) was "flipped" before the payoffs were calculated.

Two models were developed (for details, see [9]); a *compulsory* model, in which every member was forced to interact with every other member, and a *choice and refusal* model (the IPD/CR), in which each member maintained a set of expected payoff values for every other member, and could choose and refuse who to interact with. This latter model is drawn from [14] and is redescribed in [9]. A tournament-style genetic algorithm was employed in both models, with each model

away from stable cooperation can then be expected to have a deleterious effect on fitness, unless it is countered by strategy that can effectively cope with such variation. 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 Ashby's 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. This construal of the IPD allows us to predict that noise (variety) on D may lead to the evolution of more complex strategies (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.

We can therefore distinguish two types of noise:

- *M-noise:* on each iteration, and for each player, there is a certain probability, ($p = 0.01$ in these experiments), that the move specified by the genotype is flipped, and only then are the payoff scores calculated with reference to the payoff table.

- *O-noise:* on each iteration, and for each player, there is a certain probability, ($p = 0.01$ in these experiments), that the payoff awarded is altered. Each time this occurs, the actual payoff awarded is selected randomly from the four possible payoff values.

We can now predict that M-noise may lead to the evolution of more complex strategies, but O-noise will not. This hypothesis will henceforth be referred to as the *LRV hypothesis*. Furthermore, we may expect that this hypothesis will only be confirmed in stable cooperating IPD populations, and *not* in unstable populations.

3 The LRV Hypothesis

The first set of experiments were conducted with a compulsory IPD model, the parameters of which are given in Appendix 1. Ten evolutionary runs,

of 10,000 generations each, were performed in each of 3 conditions; M-noise, O-noise, and zero noise.[4]

Figure 1 shows that without any noise, complex strategies never really evolve; the evolved memory tends to stay either at the initial level of 1, or drop to zero. Figure 2 illustrates that with M-noise, complex strategies *do* evolve; not always, and not always to the maximum, but it does happen. However, contrary to the LRV hypothesis, figure 3 illustrates that O-noise has the *same* effect as M-noise. Applying t-tests to the sets of means from each condition, these observations are statistically significant. There is significantly more evolved memory in the M-noise and O-noise conditions than in the zero noise conditions ($t = 2.52, df = 18, p < 0.05$ and $t = 2.51, df = 18, p < 0.05$ respectively). But there is *no* significant difference between the M-noise and the O-noise conditions ($t = 0.22, df = 18, p > 0.5$).

The second set of experiments were conducted with an IPD model incorporating choice and refusal (IPD/CR). In order to differentiate the conditions in this model, a small *cost* on complexity was applied (in all the conditions) by levying a fitness penalty on genotype length. Again, 10 runs were performed, of 5,000 generations each, in each condition. Here we clearly see that with either zero noise (figure 4) or O-noise (figure 6), complex strategies do not often arise. However, with M-noise (figure 5), the evolution of complex strategies is considerably more noticeable. Again, t-tests reveal these observations to be highly significant. There is significantly more evolved memory in the M-noise condition than with O-noise or zero noise ($t = 3.20, df = 18, p < 0.005$ and $t = 3.79, df = 18, p < 0.005$ respectively). But there is *no* significant difference between the zero noise and the O-noise conditions ($t = 1.42, df = 18, p > 0.1$).

The LRV hypothesis, in this model, is seen to hold true. The evolution of complexity only responds to variety in the environment, not to variety on the outcome.

So why is it that only the IPD/CR model produces results consistent with the LRV hypothesis? Table 2 clearly indicates that *only* the IPD/CR model

---

[4]Each run, of 10000 generations, took approximately 1 hour on a 143MHz single user Sun UltraSparc. The 5000 generations of the IPD/CR model required 3 hours of CPU time for each run.
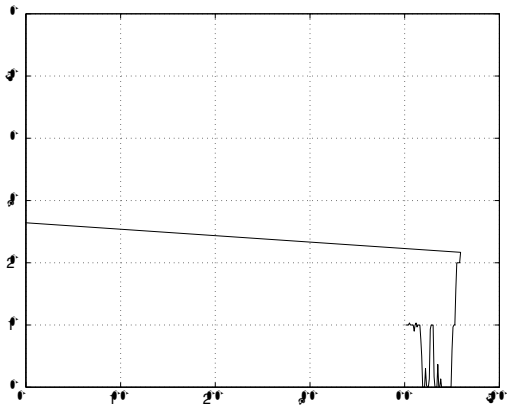
this model. We first consider evidence that the evolution of complexity that we observe with
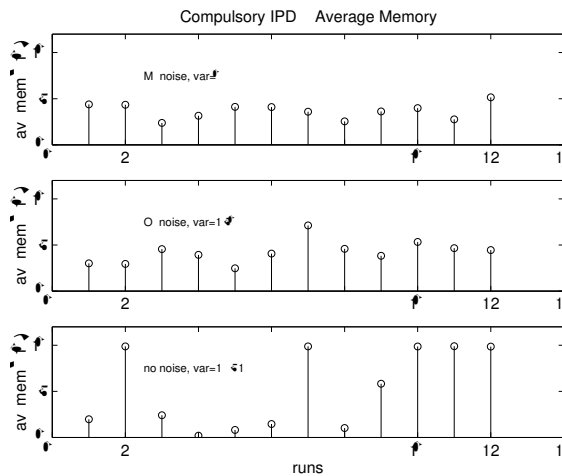
Figure 10: *Compulsory IPD; memory limited to 0-1 range. Both M-noise and O-noise permit average memory to approach the expected average; without noise, the average memory appears to be "stuck" at one extreme or the other.*

zero noise $t = -0.78, df = 22, p > 0.1$). So neither type of noise is actually encouraging (or discouraging) the evolution of complexity *per se* in this model.

4.2   Influence of Noise on Cooperation

It was also observed that both O-noise and M-noise increase the amount of cooperation in this memory-limited compulsory IPD model. Both M-noise and O-noise might be expected to have such an effect since, given a predominantly defecting population, both kinds of noise would permit a cooperative move to occasionally score well. Fig 7 indicates that typical cooperation in the zero-noise compulsory IPD model is very low, and indeed the average prevalence of cooperation over all ten runs was only 4.43% (standard deviation 2.43), and also seems to indicate that both M-noise and O-noise enhance population cooperation.

This intuition was tested in the memory-limited model, and fig 11 provides summary data, collected from the 12 runs in each condition, indeed indicating that both O-noise and M-noise significantly increase the level of population cooperation from that present with zero noise ($t = 3.39, df = 22, p < 0.001$, and $t = 5.49, df = 22, p < 0.001$ respectively), with the
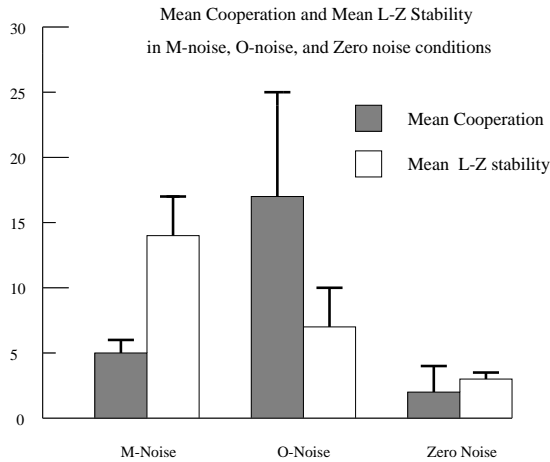
Figure 11: *Compulsory IPD; memory limited to 0-1 range; O-noise enhances population cooperativity and M-noise enhances population L-Z complexity. Standard deviations indicated by error-bars.*

effect being significantly more pronounced for the O-noise condition ($t = 4.62, df = 22, p < 0.001$). This graph also illustrates how population *stability* varies across the different conditions. Stability was measured by calculating the *Lempel-Ziv* complexity of the cooperation for each run, (see [7]). This calculation involved noticing that in the compulsory IPD model, the population, at any time, was generally either completely cooperating or completely defecting. Thus, the population cooperation over x generations can be recast as a binary string of length x. The Lempel-Ziv (L-Z) complexity measure delivers the size of the minimum program required to generate a binary string, and this is minimal for a uniform string and maximal for a completely random string. Fig 11 is therefore indicating that both M-noise and O-noise lead to significantly greater levels of population cooperation instability than the zero noise condition ($t = 10.16, df = 22, p < 0.001$, and $t = 4.19, df = 22, p < 0.001$ respectively), with the effect being significantly more pronounced in the M-noise condition ($t = 5.60, df = 22, p < 0.001$).

So both O-noise and M-noise influence the population stability dynamics, but in different ways. Both increase the overall level of cooperation, but whereas for O-noise this effect is characterised by longer periods of steady cooperation, for M-noise increased population instability is the more pronounced effect. Note that this effect is also visible in the standard (non-

14
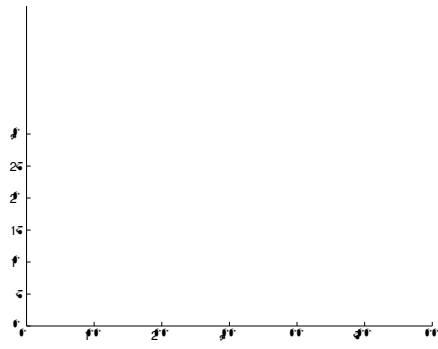
case that the variety of the response is *matching* the variety of the M-noise, in the strict sense implied by the LRV (Section 1.1). Indeed, figure 2 illustrates that the complexity response to a steady level of M-noise is generally unpredictable. It is therefore not possible to say that the complex strategies are adapting *to* the M-noise, and to that alone. What we can say is that the complex strategies are adapting to some aspect of the environment dependent on the introduction of M-noise, and which is almost certainly heavily influenced by the co-evolutionary nature of the IPD model. The exact nature of the variety of the environment presented by (and to) any given population will not be predictable (thanks to co-evolutionary dynamics), and so a stable and predictable response to the introduction of M-noise will not be observed. So, although the LRV undoubtedly

other adaptive evolving systems in use in artificial life research, *b)* a broad scope to the principles of cybernetic theory, in particular the LRV, and *c)* a distribution of noise in a broad class of evolving systems.

In other words, the models developed here serve best as a bridge between artificial life and cybernetics from which some new theoretical and empirical territory, concerning the role(s) of noise, can be surveyed. Of course, in a more immediate sense this work also constitutes a contribution

[10] 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, *Proceedings of the 6th International conference on Artificial Life (ALIFE VI)*, pages 209–221. MIT Press, 1998.

[11] A.K. Seth. Noise and the evolution of complexity; a study in evolutionary robotics. In P. Husbands and J.A. Meyer, editors, *Proceedings of the 1st European Workshop on Evolutionary Robotics (EVOROB 98)*. Springer Verlag, 1998.

[12] C.E. Shannon and W. Weaver. *The mathematical theory of communication*. University of Illinois Press, 1949.

[13] H. Simon. *The Sciences of the Artificial (3rd edition)*. MIT Press, 1988.

[14] E.A. Stanley, D. Ashlock, and M.D. Smucker. Iterated prisoner's dilemma with choice and refusal of partners: Evolutionary results. In F. Moran,