Hormonal Systems for Prisoners Dilemma Agents

Daniel Ashlock, *Senior Member, IEEE* Christopher Kuusela Nicholas Rogers

***Abstract*—A large number of studies have evolved agents to play the iterated prisoners dilemma. This study models a novel type of interaction, called the *Shopkeeper* model of interaction, in which a state conditioned agent interacts with a series of other agents without resetting its internal state. This is intended to simulate the situation in which a shopkeeper interacts with a series of customers. This study is the second to use the shopkeeper model and uses a more complex agent representation for the shopkeepers. In addition to a finite state machine for play, the shopkeepers are equipped with an artificial hormonal system that retains information about recent play. This models the situation in which an agent's current behaviour is affected by its previous encounters. We train hormonal shopkeeper prisoners dilemma agents against a variety of distributions of possible customers. The shopkeepers are found to specialize their behavior to their customers, but often fail to discover maximally exploitative behaviors against more complex customer types. This study introduces a technique called *agent-case embeddings* to demonstrate that shopkeepers adapt to their customers and that hormonal and non-hormonal agents are different. Agent-case embeddings are representation-independent maps from agents behaviors into Euclidean space that permit analysis across different representations. Unlike an earlier tool, fingerprinting, they can be applied easily to very complex agent representations. We find that the hormonal shopkeepers exhibit a substantially different distribution of evolved strategies than the non-hormonal ones. Additionally, the evolved hormonal systems exhibit a variety of hormone level patterns demonstrating that multiple types of hormonal systems evolve.**

I. Introduction

The prisoners dilemma [21], [20] is a classic in game theory. It is a two-person game developed in the 1950s by Merrill Flood and Melvin Dresher while working for the RAND corporation. Two agents each decide, without communication, whether to cooperate (C) or defect (D). The agents receive individual payoffs depending on the actions taken. The payoffs used in this study are shown in Figure 1. The payoff for mutual cooperation C is the *cooperation* payoff. The payoff for mutual defection D is the *defection* payoff. The two asymmetric action payoffs S and T, are the *sucker* and *temptation* payoffs, respectively. In order for a two-player simultaneous game to be considered prisoners dilemma, it must obey the following pair of inequalities:

$$S \leq D \leq C \leq T \quad (1)$$

and

$$2C \leq (S+T). \quad (2)$$

D. Ashlock and C. Kuusela, are at the Department of Mathematics and Statistics at the University of Guelph, Guelph On Canada, N1G 2W1. Email: $\{dashlock|ckuusela\}@uoguelph.ca$

N. Rogers is at the Department of Mathematics at McMaster University, 1280 Main Street West Hamilton, Ontario Canada L8S 4K1. Email: $rogern4@math.mcmaster.ca$

The authors thank the National Science and Engineering Research Council of Canada and the Department of Mathematics and Statistics for their support of this research.

In the *iterated prisoners dilemma* (IPD) the agents play many rounds of the prisoners dilemma. IPD is widely used to model emergent cooperative behaviors in populations of selfishly acting agents and is often used to model systems in biology [34], sociology [26], psychology [32], and economics [25].

$$\begin{array}{cc|cc} & & \mathcal{S} & \\ & & C & D \\ \hline \mathcal{P} & C & 3 & 5 \\ & D & 0 & 1 \end{array} \qquad \begin{array}{cc|cc} & & \mathcal{S} & \\ & & C & D \\ \hline \mathcal{P} & C & C & T \\ & D & S & D \end{array}$$

(1) (2)

Fig. 1. (1) The payoff matrix for prisoners dilemma used in this study – scores are earned by player $\mathcal{S}$ based on its actions and those of its opponent $\mathcal{P}$ (2) A payoff matrix of the general two player game – C, T, S, and D are the scores awarded.

This study is the second to evolve agents to play the iterated prisoners dilemma using a novel variation called the *shopkeeper* model of interaction. Many variations on the theme of evolving agents to play IPD have been published, [31], [24], [28], [30], [9], [3], [23], [2], [27], [1], [19], [17], [18], [4], [12], [5], but agents are typically "reset" (return all of their internal state variables to their initial state) in between interactions with different agents. The shopkeeper model is asymmetric between shopkeeper and customer agents. The shopkeepers undergo interaction with a series of customers without resetting their internal state variables. Customers are reset between encounters with shopkeepers, representing a relative indifference of the customer to the shopkeeper. This choice could be reversed and, in fact, there is an entire space of possible scenarios for determining when to reset an agent's internal variables.

The earlier study [8] used shopkeepers implemented with sixteen-state Mealy machines as finite state controllers. In this study a hormonal system in the form of a *side effect machine* [16], [14], [10] is added to the original finite state controller. The finite state controller can act as before, responding only to the opponent's last action, or the levels of hormones in the hormonal system can drive the transitions of the finite state controller. Details of the representation appear in Section I-A. Shopkeeper agents are trained against a variety of different distributions of potential customers. The resulting populations of shopkeeper agents are compared in a number of ways. This study tests the hypothesis that

978-1-4577-0011-8/11/$26.00 ©2011 IEEE

the hormonal shopkeepers will specialize to the customers used to train them and that the distribution of shopkeeper strategies will differ substantially from both that of agents evolved in the usual way (evaluating fitness with a round robin tournament [30]) and shopkeeper agents without a hormonal system.

A. Agents with Hormones

Hormonal agents are agents where a standard representation, finite state machine in this study, are augmented with a continuous state-conditioned system. This second system mimics biological hormones which are triggered or increased by an environmental event, e.g. outcome of of game play, but then decay if no other stimuli occur. An example of a hormonal agent is shown in Figure 2. The agents have a 16-state finite state controller that encodes their nominal play strategy. The transitions are all driven by a Boolean test; this test is either "did the opponent defect last time?" or a comparison of the value of a hormonal register to a constant. The hormonal system is a side effect machine [13], [10], a finite state machine with a numerical register associated with each state. These registers are initialized to zero and the transitions of the hormonal side effect machine (henceforth: hormonal system) are driven by the payoff, S, D, C, or T, that the agent received on the previous time-step. Each time the hormonal system makes a transition to a state, that state's numerical value is increased by 1.0. After every play, the hormonal registers are all multiplied by a decay constant $\omega = 0.95$. This combination of arithmetic increase and exponential decay gives the hormonal system a time horizon beyond which events have no effect. The distance of the time horizon is set by the choice of ω with larger values representing longer time horizons. Examples of the variation of hormone levels over time while playing a single opponent appear in Figure 3.

The constants used to permit the hormonal system to control agent behavior are chosen with an unusual distribution. If an agent always used a single register, then the numerical value would be an arithmetic sequence with a value less than:

$$\frac{T}{1-\omega} = 100 \quad (3)$$

This is a weak upper bound, computed by summing the ideal geometric series of hormone levels, and is unlikely to occur in practice. The constants are therefore generated according to the distribution

$$\frac{X}{2^Y} \quad (4)$$

where X is uniformly distributed in the range $[0, 100]$ and Y has a binomial distribution with $n = 6$ and $p = 0.5$. In effect a uniform random variable in the range between zero and the upper bound is generated and then six coins are flipped. For each coin that flips heads, the original value is divided in half. This distribution takes into account the potential for the hormonal system to use between one and four states and so have a wide variety of possible numerical ranges for the hormonal registers.

```
Start: D->10
     If(F) If(T)      Test
  0) D-> 0 C->15 (R[0]>1.23131)
  1) D-> 4 D->14 (R[2]<36.4598)
  2) D->14 D-> 9 (R[2]<0.0418594)
  3) C-> 8 C->14 (Opponent Def.)
  4) D-> 4 D-> 9 (R[1]>9.28591)
  5) D-> 5 D-> 3 (R[2]>11.141)
  6) C->12 C->15 (R[1]>1.11704)
  7) D->13 D-> 0 (Opponent Def.)
  8) D-> 7 C-> 6 (R[1]>12.7847)
  9) D-> 5 C-> 5 (R[3]>2.60719)
 10) C-> 8 D-> 1 (R[0]<43.9427)
 11) C->12 D-> 9 (R[0]<0.472249)
 12) C->11 C-> 6 (R[3]<1.78318)
 13) C->14 D->15 (R[2]>0.530241)
 14) D->14 C-> 5 (R[0]>21.139)
 15) D-> 3 D-> 8 (R[2]>8.2214)

Hormonal System:

           (Score Recieved)
Register C      S      T      D
  R[0]   0      1      3      2
  R[1]   0      0      3      1
  R[2]   0      2      1      3
  R[3]   3      0      2      0
```

Fig. 2. An example of an evolved hormonal side effect machine. The finite state play controller initially defects and begins in state 10. The hormonal system's side effect machine always starts in state 0 and its transitions are driven by the score the agent receives.

Earlier versions of agents with augmentations similar to hormonal systems have appeared in [12], [13]. In the first of these studies the hormonal system was characterized as emotions, in the second as long-term memories. The hormonal system in this study is more complex and versatile than those used in [12] which, in effect, used a single hormonal register. These are simpler and easier to interpret than those used in [13] because play is driven by single registers rather than comparison of pairs of registers. In the second of these studies [13] it was found that possessing the long term memories (analogous to hormonal registers) gave agents a substantial competitive advantage compared with agents without such memories or registers. The name of the system (emotion, long-term memory, hormonal system) depends on the interpretation of the underlying modeling goal. The term hormonal system is preferred to emotion or long-term memory because it better captures the summary nature of the information in the hormonal registers.

B. Agent-Case Embeddings

Fingerprinting [4], [5] is a technique used to associate a representation-independent functional signature with agents that play the iterated prisoners dilemma. The fingerprint depends only on the way that the agent plays, not the way it is encoded, and so is intended to enable cross-representational comparisons. A great frustration in the studies [12], [13], which contained precursors to the hormonal agents in this study, was that the mathematical derivation of fingerprints for these more complex agent representations was not implementable. *Agent-case embeddings* (ACES) are intended as an alternative to fingerprinting.

The fingerprint function captures the behavior of an agent playing for infinite time against an infinite test suite of opponents. In order to obtain a substitute for the functional signature of a fingerprint, an agent-case embedding records the performance of an agent playing for finite time against a finite set of opponents. These vectors of scores form a numerical-vector signature or set of features. The first example of ACEs (pre-dating the name) appears in [15]. In this study, virtual robots for the *Tartarus* task and their fitness on various instances of the Tartarus task were used to place a metric-space structure on agents. The vector of scores an agent receives is an embedding of the agent into Euclidean space with the same dimension as the number of cases. Distance is computed in the usual fashion.

In this study a "case" to be solved by a prisoners dilemma playing agent is another prisoners dilemma playing agent. Three finite sets suites were tried. In all cases the average payoff for playing 150 rounds of iterated prisoners dilemma was used as the score the agent received for a given opponent. The first set of opponents (case set 1) consisted of the strategies Always Cooperate, Always Defect, Tit-for-tat, Tit-for-two-tats, Two-tits-for-tat, Pavlov, Fortress-3, and Psycho (case set 1). These definition of these strategies is given in Section I-C. The second set of opponents consisted of the 100 most-fit agents drawn from 100 replicates of a round-robin fitness evolutionary algorithm training hormonal agents (case set 2). The third consisted of a set of 25 strategies drawn from the fingerprinting research (case set 3). We now define these IPD strategies.

Definition 1: For a strategy S for playing the prisoners dilemma the corresponding *Joss-Anne strategy* with parameters α and β is denoted

$$JA(S, \alpha, \beta)$$

This strategy cooperates with probability α, defects with probability β and otherwise plays as the strategy S would.
Case Set 3 consists of $JA(TFT, \alpha, \beta)$ and $JA(Psycho, \alpha, \beta)$ for all $\alpha + \beta < 1$ that are positive multiples of $\frac{1}{5}$ together with those random strategies with a probability of cooperation that a positive multiple of $\frac{1}{5}$ less than one. A random strategy is characterized by it probability of cooperation; if it does not cooperate, it defects.

To evaluate the quality of these sets of problem cases, the 25-value fingerprint [4] was extracted for a set of 100 most-fit non-hormonal agents. These 25 values were used to embed the agents into the Euclidean space $\mathbb{R}^{25}$ with the Euclidean metric. The cases were used to generate score vectors for each of these agents as well. In all cases, score vectors were viewed as embedding agents into Euclidean space. Distance matrices were derived from each of the different embeddings of agents into Euclidean space and the Pearson's correlation (Equation 5) of the entries of the distance matrices of the agent-case embeddings with the fingerprint-based distance were computed. All three correlations, shown in Table II, were highly significant ($p < 10^{-6}$) but the correlation with Case Set 3 was significantly better than the others. This agent-case set was used to generate embeddings of agents into $\mathbb{R}^{25}$ for all agents throughout the remainder of the study.

TABLE I
PEARSON CORRELATION OF AGENT-CASE DISTANCE WITH FINGERPRINT BASED DISTANCE FOR THE THREE AGENT-CASE SETS.

Case Set	Pearson Correlation
1	0.502533
2	0.576927
3	0.914981

TABLE II
GIVEN ARE THE COMPOSITION OF CUSTOMER SETS USED FOR THE SHOPKEEPER EXPERIMENTS IN THIS STUDY.

Code	Composition
AllC	All Always cooperate
Mix 1	50% TFT, 50% TF2T
Mix 2	33% TFT, 33% TF2T, 34% AllC
Mix 3	25% TFT, 25% TF2T, 25% AllC, 25% AllD
Evo	Evolved hormonal agents, round robin fitness.
EvoNE	Evolved non-hormonal agents, round robin fitness.
For3	All Fortress 3
Pavlov	All Pavlov
Rand	100 distinct randomly generated emotional agents.
TF2T	Tit-for-two-tats.
TFT	Tit-for-tat.

C. Definitions of Strategies

Following are the definitions of the strategies used to build the customer sets. Always cooperate and always defect are defined well by their names. Tit-for-tat cooperates initially and repeats its opponent's action thereafter. Tit-for-two-tats (TF2T) defects only if its opponent has defected on the last two actions. Pavlov cooperates initially and cooperates thereafter if it and its opponent made the same move. Fortress-3 defects twice and, if its opponent did the same, will cooperate thereafter if its opponent does; any deviation from this sequence restarts the sequence of two defects. The random agents are hormonal agents of the same sort generated for initial populations, by filling in all values at random. The evolved agents, both hormonal and non-hormonal, are the best-of-run agents from the 100 replicates in the first two experiments.

$$correlation = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{(n-1)s_x s_y} \quad (5)$$

(Note that for $z \in \{x, y\}$, $\overline{z}$ denotes the sample mean and s_z denotes the sample variance.)

D. Nonlinear Projection

Nonlinear projection (NLP) [33] is a visualization technique used on high dimensional data, such as the ACE vectors in this study. It is an evolutionary form of non-dimensional multi-metric scaling [29] also called *multidimensional scaling* [22]. The goal of NLP is to provide a projection of points from a high-dimensional space into a two-dimensional space that distorts the inter-point distances as little as possible. The projection forms a visualization of the higher dimensional data set. The original form of NLP used evolutionary computation to minimize the squared error between the distance matrix of the original data and the Euclidean distance matrix of the projection of the data into two dimensions. The representation consists of a simple list of point coordinates. The optimization problem of locating a good projection is treated as a standard real-valued evolutionary optimization.

In this study we minimize the *Pearson correlation*, given in Equation 5, of the 25-dimensional distance matrices for the sampled fingerprints with their two dimensional projections. Since this correlation is scale-free it reduces the complexity of the evolutionary search. An issue that arises when minimizing squared error between original and projected distances is that the optimizer must get the inter-point distances correct. In order to do this it must estimate the scale of the data. This is possible, but is an added, unnecessary, complication. Pearson correlation is invariant under translation and scaling of either of the data sets and so permits the evolutionary algorithm to solve the problem of relative distance without regard to scale. The resulting projection does not, itself, have a scale, rather it depicts only relative distances between points.

The evolutionary algorithm used to perform NLP uses a population of ten tentative projections stored as lists of points (x, y). The points are initially generated to lie within the unit square with corners $(0, 0)$ and $(1, 1)$. The model of evolution is tournament selection of size seven. Variation operators are two point crossover of the lists of points (points are treated as atomic objects that cannot be split by crossover) and two mutation operators, each used 50% of the time. The first mutation operator randomly replaces a point with a new point selected uniformly at random within the unit square. The second adds a Gaussian random variable with a standard deviation of 0.1 to both coordinates of a point. From 1-3 mutations, with the number selected uniformly at random, are performed on any new structure. The algorithm is run for 40,000 mating events, a number chosen by looking at when the fitness for a small initial set of runs leveled off. In this study, non-linear projection is a tool used for exploratory analysis and so we make no attempt to explore the parameter space of its algorithm. Rather, we use parameters that have worked well in the past.

The remainder of the study is structured as follows. The details of experiments performed are given in Section II as are details of the analysis techniques used. Results and discussion are presented in Section III. Conclusions and possible next steps for both the improvement of fingerprinting techniques and applications are given in Section IV.

II. Experimental Design

Two sets of experiments were performed to evaluate the new hormonal agent representation. These replicated the Miller-type evolutionary experiment [30]. The first, using *non-hormonal agents*, whose actions are determined by the opponent's previous action. The second used the full hormonal agent representation. A population of 36 agents, a number chosen for compatibility with previous studies [11], [6], [7] were evolved. Agent fitness is assessed by a round-robin tournament in which each pair of players engage in 150 rounds of the iterated prisoners dilemma. Evolution continues for 200 generations. Reproduction is elitist with an elite of the 24 highest scoring strategies, another choice that maintains consistency with past studies. When constructing the elite, ties are broken uniformly at random. Twelve pairs of parents are picked by fitness-proportional selection with replacement on the elite. Parents are copied, and the copies are subjected to crossover and mutation. The hormonal agent crossover is either two-point crossover of the list of states of the main finite state machine or of the list of states of the hormonal system. Each new strategy is subjected to a single mutation with a 10% chance of generating either a new initial action or initial state in the main finite state machine, a 30% chance of mutating a transition which changes the destination of a transition two-thirds of the time, the type of the Boolean comparison one-sixth of the time, or the value of the constant for comparison one-sixth of the time, a 30% chance of mutating an action (C becomes D, D becomes C), and a 30% chance of changing the destination of a transition in the hormonal system. For both these experiments, one hundred replicates are performed to supply evolved hormonal and non-hormonal agents to be used as customers.

A set of eleven experiments were performed, similar to the first two, except that a shopkeeper fitness function was used. These experiments each consisted of 30 replicates. A population of 100 customers, in the form of IPD-playing agents, is chosen. The shopkeeper agents have their internal state information reset and then undergo a series of 40 encounters with customers. Each encounter lasts for 2-6 iterations of prisoners dilemma with the duration selected uniformly at random and kept the same for all shopkeepers in a given encounter. Before each encounter the customers are shuffled, their internal state information is reset, and one customer is assigned to each shopkeeper. They then play for the selected number of iterations. The shopkeeper agents fitness is the average of the scores obtained from customers over all the encounters in a given fitness evaluation. The eleven different experiments were differentiated by the customer sets used. These customer sets are given in Table II.

III. Results and Discussion

The first two experiments are used to compare the hormonal and non-hormonal agents evolved with the more usual round-robin fitness function. Fitness over the course of evolution for all one hundred replicates of both experiments is shown in Figure 4. The color scheme is chosen for

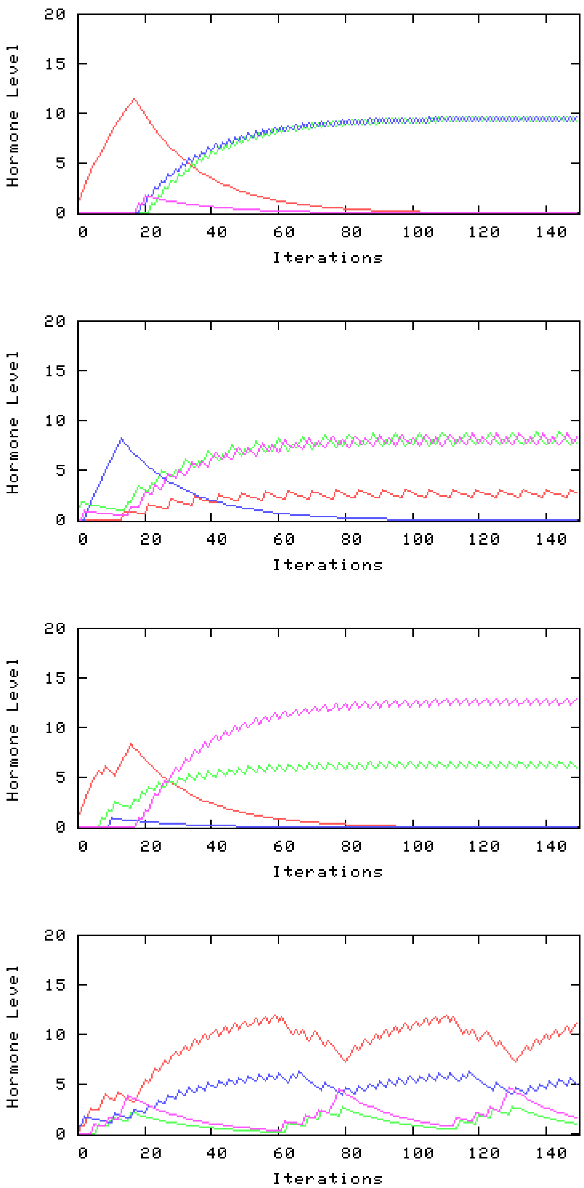

Fig. 3. Shown are examples of the level of hormonal registers of agents evolved with a round-robin fitness function over 150 iterations of prisoners dilemma. The opponent is playing the strategy Pavlov.

consistency with earlier studies. In agreement with results from [12], [13] the hormonal agents are, on average, less cooperative than non-hormonal agents and many have scores near the always-defect Nash equilibria for iterated prisoners dilemma. Roughly 40% of the hormonal agents avoided the always-defect part of the score space. The hormonal agents also exhibited much more complex fitness trajectories than the non-hormonal agents suggesting a richer set of behaviors.

Figure 3 shows the change in the hormonal registers over 150 iterations for four evolved hormonal agents playing against a Pavlov opponent. The patterns of change in hormone level are distinct for the four agents and support the notion that the hormonal agents can exhibit a greater diversity of behaviors. The initial spikes in the first two hormone traces, followed by decay, represent transient states in the hormonal system. The saw toothed patterns in the first three traces represent tight cycles though sets of states. The complex pattern of traces in the last example represent a more complex pattern of cycling among all four states in the hormonal system.

Figure 5 is a non-linear projection of ACE vectors for the hormonal and non-hormonal agents. It shows a central cluster for each agent type, but the behavioral regions, as measured with the ACEs, do overlap to some extent. Figures 4 and 5 support the hypothesis that the hormonal agents will evolve differently from non-hormonal agents.

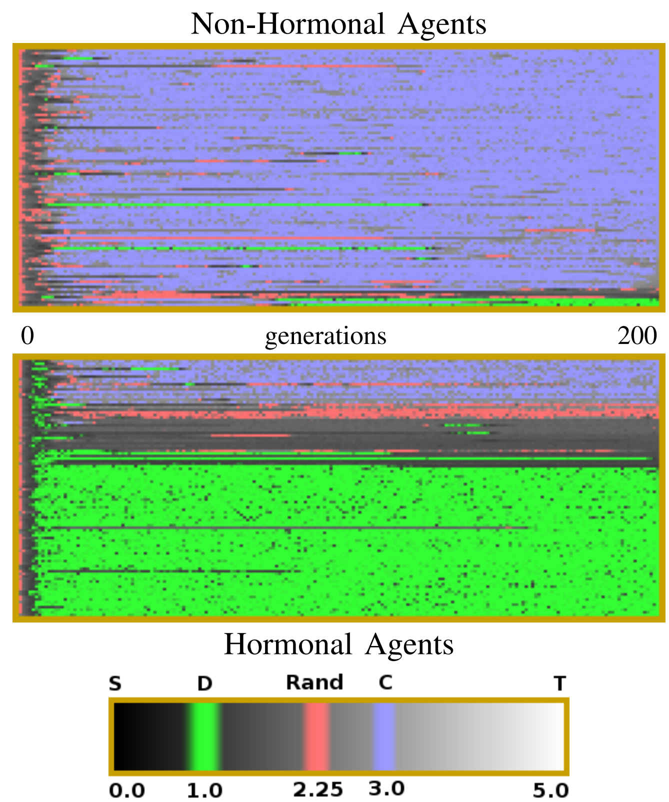

Fig. 4. Shown are the mean population scores for 100 populations of agents evolved for 200 generations using a round robin fitness function. The upper panel shows non-hormonal agents while the lower panel shows agents with active hormonal systems. Color coding for scores is given by the bar at the bottom of the figure; Rand (shown in red) is the score a random player gets against itself. The populations are sorted by their mean score in the final generation with the highest score at the top.

Figure 6 shows a nonlinear projection of agent-case embeddings for the eleven shopkeeper experiments. This figure supports the hypothesis that the shopkeepers will adapt to the customers they are trained against. Notice that shopkeepers evolved against different sets of customers appear in different parts of the projected space in many cases. The shopkeepers trained against random and both sorts of evolved agents are tightly clustered on the left side of the figure. These shopkeepers mostly defect against their customers. These three sets of customers are the most complex and the least uniform. Given that the shopkeepers have from 2-6 iterations to interact with the customers, it is not surprising that they would adopt a strategy similar to always defect against a diverse set of complex opponents. The solid violet squares representing the shopkeepers trained against Fortress-3 customers are, mostly, in the same large cluster of defectors. Three of them are outside the cluster - these represent shopkeeper populations that learned to elicit some cooperation from Fortress-3 customers in spite of the short interaction time. Many of the shopkeepers trained against customers that always cooperate are also in this cluster always defect as this is optimal behavior against always cooperate.

A second large cluster appears in Figure 6, near the middle. It consists of shopkeepers that have learned to exploit the TF2T and AllC players in the three sets of mixed, simple customers. The agents trained against the mixes, while common in this cluster, have a wide distribution. The shopkeepers that are trained against tit-for-tat customers are the most widely scattered, but include a modest cluster on the right side of Figure 6. Looking at the distribution of shopkeepers trained against evolved customers, none (or perhaps one) of those trained against non-hormonal evolved agents are outside of the cluster of defecting shopkeepers; three of the shopkeepers trained against hormonal evolved agents are outside. This suggests that is it easier to learn to elicit cooperation from evolved hormonal agents than their non-hormonal counterparts during encounters of short duration.

Figure 7 is a nonlinear projection of shopkeepers trained against four of the customer types. This projection is far less cluttered than Figure 6 and does a better job of showing the distribution of shopkeepers trained against tit-for-tat agents, the interpenetration of the Mix1 and Mix3 trained shopkeepers, and the relatively low scatter of the shopkeepers trained against evolved hormonal agents.

In the course of the research on fingerprints it was found that the space in which the fingerprints are distributed is 6-dimensional but with the space having very low diameter in four of its dimensions, yielding a space that projects very well onto two-dimensional space. The Pearson correlations across thirty different evolutionary runs used to generate visualizations in this study had a mean Pearson correlation in excess of 0.99; this suggests that the agent-case embeddings chosen also occur in a subspace in which most variation is in two directions.

IV. Conclusions and Next Steps

This study introduces a new representation for prisoners dilemma playing agents, the hormonal agent representation. It demonstrates that the hormonal agents evolve to discover a very different distribution of strategies than non-hormonal agents. This study also generalizes an earlier study using the shopkeeper model of fitness for iterated prisoners dilemma, using hormonal agents in place of a simpler agent representation. As with the earlier study, the hormonal shopkeepers are found to specialize to their customers.

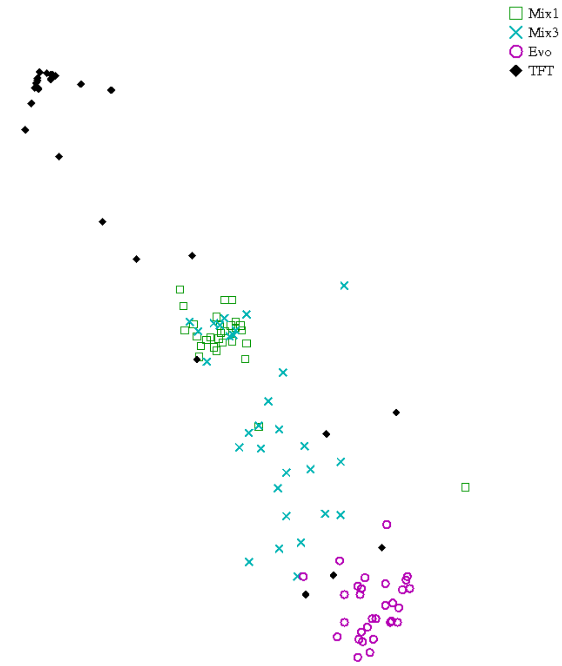

Fig. 7. Shown is a non-linear projection of the agent-case embedding vectors for the most fit shopkeeper for all thirty replicates for shopkeepers trained against customers from mix 1 and 3, tit-for-tat customers, and evolved hormonal agents.

Another novel contribution of this study is the introduction of agent-case embeddings as a substitute for fingerprinting. Agent-case embeddings are a representation-independent method of associating a point in Euclidean space with prisoners dilemma playing agent. The agent-case embedding used was chosen because the distance on the embedding into Euclidean space has a high correlation with the induced fingerprint distance. This suggests that it shares with the fingerprint distance the property that agents with nearby embeddings have similar behaviours.

The use of short encounter durations in this study, and in the earlier study on non-hormonal shopkeepers, is an arbitrary choice. Available space does not permit experimentation with different distributions of encounter duration; this remains an early priority for future research. If encounter durations are lengthened then shopkeepers will obtain more information about their customers and so will probably be able to adapt to them more efficiently. Another arbitrary choice, driven by the values used in earlier studies, is the choice to have sixteen states in the main finite state controller and four hormonal registers. These are both parameters that can be explored in future studies.

The hormonal system traces shown in Figure 3 suggest that these patterns of change may be worth mining for additional information. While the agent-case embedding provides a useful feature set for analysis of evolved agents, the hormone levels of a fixed opponent against various opponents may also be a useful feature set for understanding those opponents. The previous work on side effect machines, in which similar

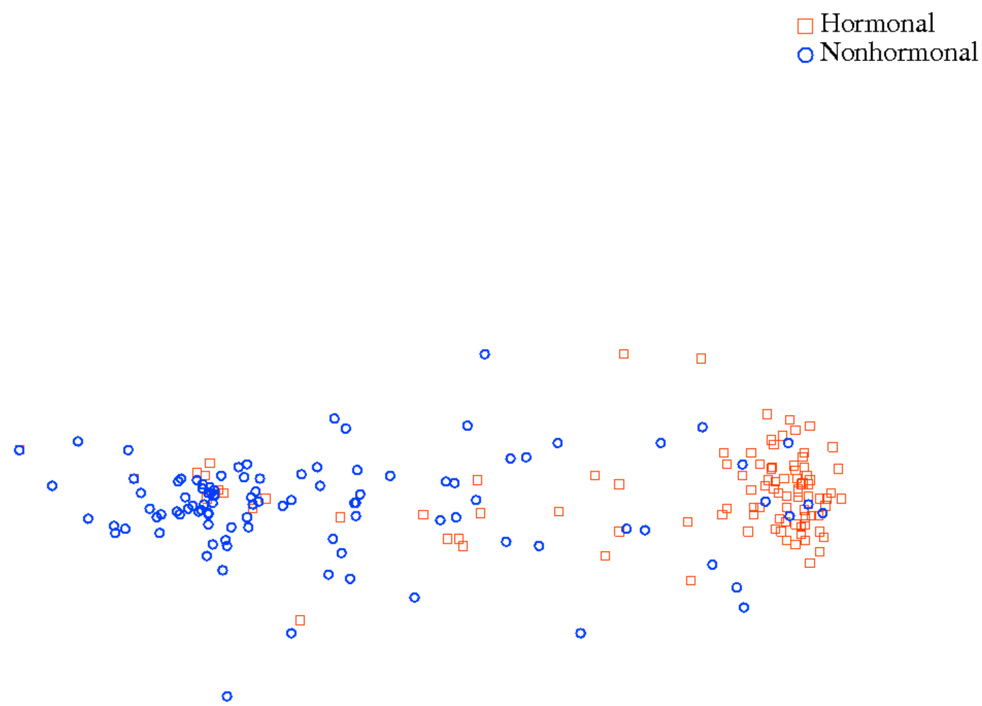

Fig. 5. Shown is a non-linear projection of the agent-case embedding vectors for the most fit shopkeeper for all one hundred replicates for hormonal and non-hormonal agents using a round-robin fitness function. The Pearson correlation for this projection is 0.995.

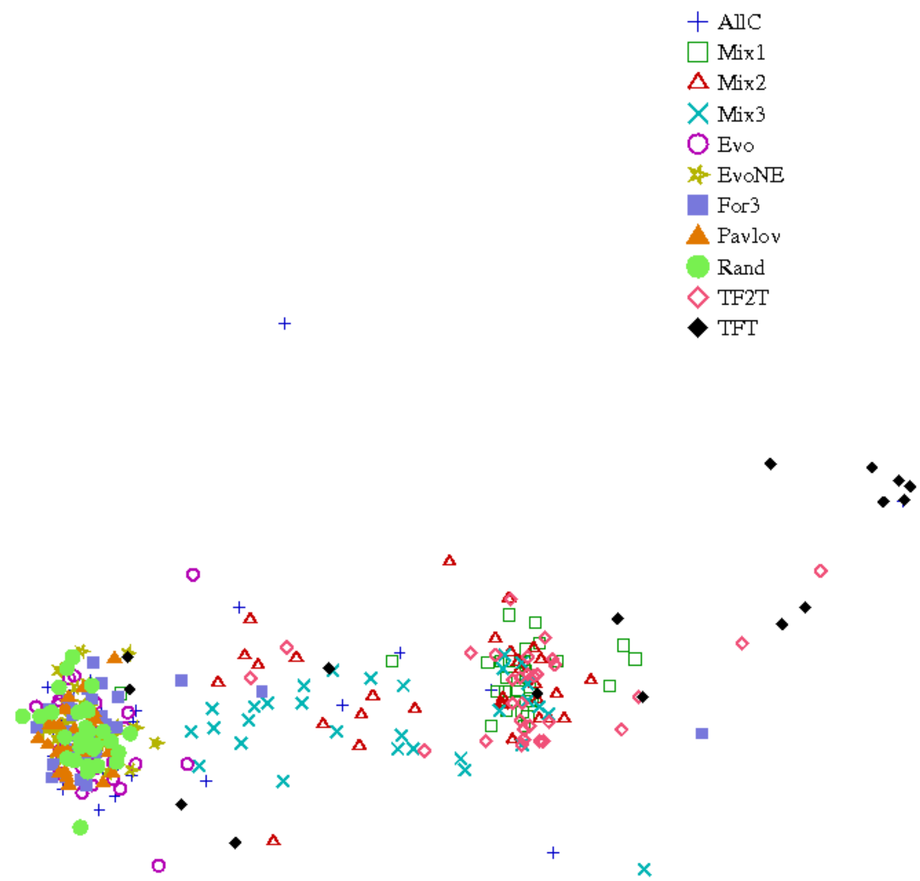

Fig. 6. Shown is a non-linear projection of the agent-case embedding vectors for the most fit shopkeeper for all thirty replicates within the eleven shopkeeper experiments. Glyphs are labeled by the customer type that the shopkeepers were trained against. The Pearson correlation for this projection is 0.994.

features were quite useful for the analysis of DNA sequences, supports this view.

Hormonal Systems in Other Types of Agents

The hormonal system incorporated into the shopkeepers in this study can be adapted to many other types of agents. Any system where an agent encounters a discrete sequence of events can use those events to drive a side effect machine like the one that is the core of the hormonal system in this study. The values of the hormone registers could then be used to drive agent behaviors. In prisoners dilemma these behaviours were limited to decisions to cooperate or defect and the updating of the internal state of the main finite state controller. In an agent engaging in social interactions, the hormonal registers could drive externally visible signals. If the agents with the hormonal system were non-player characters, the hormonal system could govern facial expressions or stance. For human or humaniform characters this would require laborious mapping of hormone registers onto various human emotions. If the non-player character were non-human then the mapping to expressions could be arbitrary and serve to give the character consistent facial features, based on recent experience, that would be consistent and learnable by a human player.

REFERENCES

[1] D. Ashlock. Training function stacks to play iterated prisoner's dilemma. In *Proceedings of the 2006 IEEE Symposium on Computational Intelligence in Games*, pages 111–118, 2006.

[2] D. Ashlock, W. Ashlock, and G. Umphry. An exploration of differential utility in iterated prisoner's dilemma. In *Proceedings of the 2005 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, pages 271–278, 2006.

[3] D. Ashlock and E. Y. Kim. The impact of cellular representation on finite state agents for prisoner's dilemma. In *Proceedings of the 2005 Genetic and Evolutionary Computation Conference*, pages 59–66, New York, 2005. ACM Press.

[4] D. Ashlock and E.-Y. Kim. Fingerprinting: Automatic analysis and visualization of prisoner's dilemma strategies. *IEEE Transaction on Evolutionary Computation*, 12:647–659, 2008.

[5] D. Ashlock, E.Y. Kim, and W. Ashlock. Fingerprint analysis of the noisy prisoner's dilemma using a finite state representation. *IEEE Transactions on Computational Intelligence and AI in Games*, 1(2):157–167, 2009.

[6] D. Ashlock, E.Y. Kim, and L. Guo. Multi-clustering: avoiding the natural shape of underlying metrics. In C. H. Dagli et al., editor, *Smart Engineering System Design: Neural Networks, Evolutionary Programming, and Artificial Life*, volume 15, pages 453–461. ASME Press, 2005.

[7] D. Ashlock, E.Y. Kim, and N. Leahy. Understanding representational sensitivity in the iterated prisoner's dilemma with fingerprints. *IEEE Transactions on Systems, Man, and Cybernetics–Part C: Applications and Reviews*, 36(4):464–475, 2006.

[8] D. Ashlock, C. Kuusela, and M. Cojocaru. Shopkeeper strategies in the iterated prisoner's dilemma. Accepted to the IEEE 2011 Congress on Evolutionary Computation, 2011.

[9] D. Ashlock and J. Mayfield. Acquisition of general adaptive features by evolution. In *Proceedings of the 7th Annual Conference on Evolutionary Programming*, pages 75–84, New York, 1998. Springer.

[10] D. Ashlock and A. McEachern. Nearest neighbor training of side effect machines for sequence classification. In *Proceedings of the 2010 IEEE Symposium on Bioinformatics and Computational Biology*, pages 111–118, Piscataway, NJ, 2010. IEEE Press.

[11] D. Ashlock and B. Powers. The effect of tag recognition on non-local adaptation. In *Proceedings of the 2004 Congress on Evolutionary Computation*, volume 2, pages 2045–2051, Piscataway, NJ, 2004. IEEE Press.

[12] D. Ashlock and N. Rogers. A model of emotion in the prisoner's dilemma. In *Proceedings of the 2008 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, pages 272–279, Piscataway, NJ, 2008. IEEE Press.

[13] D. Ashlock and N. Rogers. The imapact of long term memory in the iterated prisoner's dilemma. In *Intelligent Engineering Systems Through Artificial Neural Networks*, pages 245–252, 2009.

[14] D. Ashlock and E. Warner. Classifying synthetic and biological dna sequences with side effect machine. In *Proceedings of the 2008 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, pages 22–29, Piscataway, NJ, 2008. IEEE Press.

[15] D. Ashlock and E. Warner. The geometry of tartarus fitness cases. In *Proceedings of the 2008 Congress on Evolutionary Computation*, pages 1309–1316, Piscataway NJ, 2008. IEEE Press.

[16] D. Ashlock and E. Warner. Side effect machines for sequence classification. In *Proceedings of the 21st Canadian Conference on Electrical and Computer Engineering*, pages 1453–1456, 2008.

[17] D. A. Ashlock and E. Y. Kim. Fingerprint analysis of the noisy prisoner's dilemma. In *Proceedings of the 2007 Congress on Evolutionary Computation*, pages 4073–4080, 2007.

[18] W. Ashlock. Why some representations are more cooperative than others for prisoner's dilemma. In *2007 IEEE Symposium on the Foundations of Computational Intelligence*, pages 314–321, Piscataway, NJ, 2007. IEEE Press.

[19] W. Ashlock and D. Ashlock. Changes in prisoner's dilemma strategies over evolutionary time with different population sizes. In *Proceedings of the 2006 Congress On Evolutionary Computation*, pages 1001–1008, Piscataway, NJ, 2006. IEEE press.

[20] R. Axelrod. *The Evolution of Cooperation*. Basic Books, New York, 1984.

[21] R. Axelrod and W. D. Hamilton. The evolution of cooperation. *Science*, 211:1390–1396, 1981.

[22] I. Borg and P. Groenen. *Modern multidimensional scaling. Theory and applications.* Springer, New York, 1997.

[23] S. Y. Chong and X. Yao. Behavioral diversity, choices and noise in the iterated prisoner's dilemma. *IEEE Transaction on Evolutionary Computation*, 9:540–551, 2005.

[24] D.B. Fogel. Evolving behaviors in the iterated prisoners dilemma. *Evolutionary Computation*, 1(1), 1993.

[25] M. Hemesath. Cooperate or defect? Russian and American students in a prisoner's dilemma. *Comparative Economics Studies*, 176:83–93, 1994.

[26] J. M. Houston, J. Kinnie, B. Lupo, C. Terry, and S. S. Ho. Competitiveness and conflict behavior in simulation of a social dilemma. *Psychological Reports*, 86:1219–1225, 2000.

[27] Hisao Ishibuchi, Naoki Namikawa, and Ken Ohara. Effects of spatial structures of iterated prisoner's dilemma game strategies in single-dimensional and two-dimensional grids. In *Proceedings of the 2006 Congress on Evolutionary Computation*, volume 1, pages 3721–3728. IEEE Press, 2006.

[28] Lindgren K. and Nordahl M. G. Evolutionary dynamics of spatial games. In *Proceedings of the Oji international seminar on Complex systems : from complex dynamical systems to sciences of artificial reality*, pages 292–309, New York, NY, USA, 1994. Elsevier North-Holland, Inc.

[29] P. Legendre and L. Legendre. *Numerical Ecology, 2nd English Edition*. Elsevier, New York, 1998.

[30] J. H. Miller. The coevolution of automata in the repeated prisoner's dilemma. *Journal of Economic Behavior and Organization*, 29(1):87–112, January 1996.

[31] H. Mühlenbein. Darwin's continent cycle theory and its simulation by the prisoner's dilemma. *Complex Systems*, 5:459–478, 1991.

[32] D. Roy. Learning and the theory of games. *Journal of Theoretical Biology*, 204:409–414, 2000.

[33] J. Schonfeld and D. Ashlock. Filtration and depth annotation improve non-linear projection for rna motif discovery. In *Proceedings of the 2006 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, pages 352–359, Piscataway NJ, 2006. IEEE Press.

[34] K. Sigmund and M. A. Nowak. Evolutionary game theory. *Current Biology*, 9(14):R503–505, 1999.