The implications of learning across perceptually and strategically distinct situations

Daniel Crownden, Kimmo Eriksson and Pontus Strimling

Linköping University Post Print

N.B.: When citing this work, cite the original article.

The original publication is available at www.springerlink.com:

Daniel Crownden, Kimmo Eriksson and Pontus Strimling, The implications of learning across perceptually and strategically distinct situations, 2016, Synthese.

<http://dx.doi.org/10.1007/s11229-014-0641-9>

Copyright: Springer Verlag (Germany)

<http://www.springerlink.com/?MUD=MP>

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-125995>

The implications of learning across perceptually and strategically distinct situations.

Received: date / Accepted: date

Abstract Game theory is a formal approach to behavior that focuses on the strategic aspect of situations. The game theoretic approach originates in economics but has been embraced by scholars across disciplines, including many philosophers and biologists. This approach has an important weakness: the strategic aspect of a situation, which is its defining quality in game theory, is often not its most salient quality in human (or animal) cognition. Evidence from a wide range of experiments highlights this shortcoming. Previous theoretical and empirical work has sought to address this weakness by considering learning across an ensemble of multiple games simultaneously. Here we extend this framework, incorporating artificial neural networks, to allow for an investigation of the interaction between the perceptual and functional similarity of the games composing the larger ensemble. Using this framework, we conduct a theoretical investigation of a population that encounters both stag hunts and prisoner's dilemmas, two situations that are strategically different but which may or may not be perceptually similar.

Keywords Game Theory · Learning · Multiple Games · Bounded Rationality · Framing Effects · Artificial Neural Networks

1 Introduction

The traditional game theoretic or rational actor approach to understanding how people act in strategic situations relies solely on the underlying payoff structure of situations to derive optimal or equilibria behavior, i.e. the behavior a rational agent would employ. Thus, traditional game theory provides a suitable model for decisions under the assumptions that agents are fully aware of payoff structure, that agents are readily able to differentiate between payoff structures, and that agents put sufficient cognitive effort into their decisions to

Address(es) of author(s) should be given

arrive at rational conclusions (Von Neumann and Morgenstern, 2007). Some of these assumptions are relaxed by the introduction of evolutionary game theory.¹ In evolutionary game theory, equilibrium behavior emerges not from rational considerations but simply from the replication of strategies. This replication can take the form of agents modifying their strategies based on comparisons between their own payoff and the payoff of others over the course of repeated play (Weibull, 1997), or take the form of evolutionary shifts in the composition of the population through the selective reproduction of strategies based on game payoffs (Taylor and Jonker, 1978; Schuster and Sigmund, 1983; Hofbauer and Sigmund, 2003). Related learning dynamics studied in Fudenberg and Levine (1998), show that repeated play and attention solely to one's own (not others') payoff in relation to one's actions can also lead to equilibrium behavior. Thus, evolutionary game theory allows for a relaxation of the original assumptions of game theory, requiring only payoff driven replication to explain the emergence of equilibrium behavior. Consequently the evolutionary game theory approach has proved very powerful in providing ultimate explanations of evolved behaviors (Grafen, 1984). It has also allowed for the formalization and partial explanation of important notions like social contract, communication, convention and morality (Skyrms, 2002, 2004; Binmore, 2005).

These powerful evolutionary game theory models take as their starting point a strategic situation defined by its payoffs and strategies, and examine the dynamics of the possible strategies in this situation using appropriately chosen replication or learning dynamics. This is often an excellent starting point. In this paper though we would like to take a step back, note a critical and implicit assumption in this approach and examine some of the consequences of relaxing this implicit assumption. The assumption is subtle and often overlooked; we will try and illuminate it via a coarse caricature. Consider the question "What should I do?" This question is often answered by decomposing it into "What kind of situation am I in?" and "What should I do in a given situation?" The first question is primarily a problem of perception, and the second primarily a problem of strategy or optimization. The evolutionary game theoretic approach bypasses this problem of perception, and jumps straight to a strategic analysis of the second question, though there are exceptions, which we discuss later (Harsanyi, 1967, 1968a,b; Enquist et al., 2002; Jäger, 2007; Bednar and Page, 2007; Zollman, 2008; Mengel, 2012; O'Connor, 2013a). At its simplest, evolutionary game theory requires only that replicators reliably produce a given behavior in a given situation. In such a case analyzing a single game in isolation is justified. However, for replicators that encounter more than one game the problem of contingent behavior immediately emerges. In the context of evolutionary game theory, the problem of perception *is* the problem of how a replicator reliably produces one behavior when faced with one game and reliably produces another behavior when faced with a different game.

¹ These assumptions are also relaxed by the consideration of games of incomplete information, which we will discuss later.

While studying a single strategic situation in isolation is often justified, empirical evidence suggests that for human decision makers strategic choice in one situation is often affected by perceptions of similarity between situations. Experiments show that the framing of a situation can have a large impact on subject behavior (Brewer and Kramer, 1986). Framing can be as subtle as using loaded words like “risk” (Tversky and Kahneman, 1986; Levin et al., 1998) or as conspicuous as connecting the laboratory situation to a situation that subjects are familiar with (Cronk, 2007). The inherent framing provided by a subject’s broad cultural context is also significant (Henrich et al., 2001). In all of these cases subject behavior varies significantly with framing even though the strategic fundamentals of the decision are held constant. The converse effect of frames has also been demonstrated by presenting different strategic situations using the same frame and finding that subjects behave similarly in both situations despite different strategic fundamentals (Eriksson and Strimling, 2010). Strong framing effects have been found even in the case of very transparent variation in strategic structure, suggesting that the impact of frames on behavior is at least as great if not greater than that of strategic structure. Experiments also show that in readily distinguishable games, individual differences in personality, potentially stemming from differing personal histories, explain significant variation in behavior across different strategic games (Yamagishi et al., 2013).

The above studies do not imply that human decision makers are unable to differentiate between various situations. The participants in the cross cultural studies of Henrich et al. (2001) have not failed to distinguish between the experimental games and other more familiar situations; Rather, the perceptual features of the experimental games have triggered associations with other social contexts, i.e. games, encountered in the player’s culture. When situations are readily distinguishable, experimental evidence suggests that the selective linking of two distinct situation on the basis of shared perceptual features plays an important role in human decision making, particularly in the case of relating a novel decision to a familiar situation (Gick and Holyoak, 1980). Empirical work has also directly investigated the effects of playing multiple, strategically distinct games simultaneously. These experiments reveal that in multi game situations, behavior is often best understood in terms of spillover effects between games (Huck et al., 2011; Cason et al., 2012; Grimm and Mengel, 2012).

Taken together this empirical work points to the importance of considering a multi game context and the perceptual similarity between games within this context. Observations along these lines have also been made by philosophers and economists considering the emergence of normative behavior (Gale et al., 1995; Skyrms and Zollman, 2010), i.e. each has suggested that to understand the emergence of norms in any one particular situation, the ensemble of all situations encountered may need to be considered.

Our goal in this paper is to develop a technical framework for predicting behavior in a multi game environment. Such an environment is characterized by agents who are initially unable to discriminate between games, that must

simultaneously learn to discriminate between situations based on the perceptual features of the game, and learn to choose the correct actions in those situations. We view this work as both in the spirit of and complimentary to the earlier works investigating behavior in multi game environments (Harsanyi, 1967, 1968a,b; Bednar and Page, 2007; Zollman, 2008; Mengel, 2012). Before introducing our own framework, we briefly discuss these previous contributions.

Harsanyi (1967, 1968a,b) used a multiple game environment to investigate games of incomplete information. A game of incomplete information is a situation where players are ignorant of some, potentially all, aspects of the game they are playing. Invoking what Harsanyi terms the *Bayesian Hypothesis*, players are assumed to form subjective belief distributions over the multiplicity of possible games they might be playing. One way of interpreting this situation is that a given player engages with an ensemble of games against a population of opponents. Harsanyi shows how these belief distributions can allow for the transformation of games with incomplete information to essentially equivalent games of complete but imperfect information. These results provide powerful analytical tools for approaching situations of incomplete information. These tools though, provide no guidelines as to *how* agents form their subjective beliefs over the distribution of possible games. In other words, Harsanyi’s results allow for the analysis of a situation where an agent has a given level of ignorance and a given set of beliefs about their ignorance, but his results simply do not speak to the issue of how an agent might form this belief distribution in the first place.

Bednar and Page (2007) have the ambitious goal of developing game(s) theory as means of explaining five observed cultural phenomena:

1. Intra-individual consistency - As an individual moves from task to task he or she responds similarly.
2. Inter-agent consistency Individuals within the same community, encountering the same problems, will act like one another.
3. Contextual effects Individuals from different communities may react differently to the same problem or phenomenon.
4. Behavioral stickiness Individuals may not immediately alter their behavior despite changes to their incentives.
5. Suboptimal behavior The strategy employed by individuals within a community may be suboptimal, where individuals could benefit by acting in a different way. Formally, the behaviors are not equilibrium strategies in the repeated game or if they are equilibrium strategies, the resulting equilibrium does not belong to the set of Pareto efficient equilibria.

Bednar and Page point out that many of these phenomena simply cannot exist within the context of a single game and so propose a situation where “agents play ensembles of games, not just single games as is traditionally the case in evolutionary game theory models.” Using this model Bednar and Page are able to account for all of these phenomena, and even make predictions

about how the relative frequencies at which games are encountered might alter the course the social evolution.

With the explicit analysis of multiple games, the questions of how agents distinguish the different games from each other, and how strategies in different games are related to each other immediately arise. Bednar and Page address these questions by modeling their agents as finite state automata (FSA). They rationalize this choice as follows: “Automata are among the most basic classes of models used to represent human behavior (Kalai and Stanford 1988; Miller 1996; Rubinstein 1986), and unlike more elaborate representations, such as neural networks or genetic programming, they are mathematically tractable and computationally transparent.” Agents differentiate between games by applying different initial automata states to different games. Strategies for different games are related to each other because cognitive restrictions, i.e. the limited number of states in the FSA, forces strategies for different games to share cognitive subroutines. The simplicity and tractability of automata are put to good use in Bednar and Page’s work, allowing them to prove various claims concerning the possibilities and long run dynamics of their model, and simultaneously linking these analytic results to the outcomes of agent based simulations.

Zollman (2008) considers a simpler multi game situation where agents evolve a single strategy in response to a mixture of Nash bargaining games and ultimatum games. Zollman finds that the mixture of games favors the evolution of fairness far more frequently than the same evolutionary processes do when acting on agents playing either of the two bargaining games in isolation. This finding highlights the importance of discriminating or failing to discriminate between strategic situations. As in the FSA approach discussed above, the issue of perception is avoided; The strategy utilized in both games is assumed to be identical, i.e. no discrimination is possible. The mechanism of perception underpinning this failure to discriminate is left unspecified.

Mengel (2012) considers an ensemble of games, repeatedly played by two players. Reinforcement learning is used to determine both how players discriminate between situations, with finer grained discriminations having an increasing cognitive cost, and which strategy to play given the discriminations made. The problems of learning to form categories of games, and learning which behavior to employ based on those learned categorizations are thus tackled simultaneously. Reinforcement learning fits well with Mengel’s desire to avoid “an exogenous measure of similarity.” However, an exogenous perceptual similarity structure were desired this learning model provides no immediate avenue for imposing such a measure.

These previous investigations of learning in multi game environments are well suited to their various purposes, but all are unable to directly address the interplay between strategic and perceptual similarity. Here, we will incorporate artificial neural networks into a multi game environment similar to those used in Bednar and Page (2007); Mengel (2012). This will allow us to address the entwined nature of perceptual and strategic similarity.

The entanglement of perceptual and strategic problems has been, at least partially, addressed within the theories of signaling, manipulation, the origins of communication and conventions, and the evolution of perceptual categories. Four analyses, each considering a generalization of a simple signaling game, stand out as closely related to our own. Enquist et al. (2002) incorporate the problem of perception by enlarging the strategy space to include rich signaling stimuli for a sending player and a complex stimulus response function for a receiving player. They find that this enrichment of the strategy space radically alters the standard outcome predicted by evolutionary game theory. Jäger (2007) considers the case where the set of possible meanings to be communicated is infinite and structured, i.e. equivalent to some n -dimensional Euclidean space. The set of signals sent and received remains finite, discrete, and inherently distinguishable. Sender strategies in such a game are mappings from the structured n -dimensional Euclidean meaning space to some finite, discrete, and distinguishable set of “signals”, and receiver strategies are mappings from this set of signals back to the meaning space. Jäger (2007) identifies properties of the evolutionary stable communications strategies of such a game under various replicator dynamics, but does not give explicit consideration to how such a strategy is implemented or modified via evolution or learning. O’Connor (2013a) builds on the work of Jäger (2007), by explicitly considering the learning dynamics that might underlie the formation and modification of sender and receiver strategies. The learning considered in O’Connor (2013a) requires a simpler set of possible meanings to be communicated than originally considered by Jäger (2007), i.e. a meaning space equivalent to some n -element subset of the integers. O’Connor (2013b) observes that the signaling games studied in O’Connor (2013a) and Jäger (2007) can be interpreted as occurring within the mind of a single agent, and that this provides some insight into the evolution of perceptual categories. O’Connor’s model formalizes arguments for the case that perceptual categories emerge in part from functional, i.e. payoff based, similarity.

Now that we have roughly outline the context of multi game analyses, and entangled strategic and perceptual problems with which our own model belongs, we are in a position to describe our own approach. Our basic model can be summarized as follows.

1. There is a population of agents, all initially naive about their environment.
2. The environment consists of a variety games.
3. Each time step agents are randomly paired and assigned a game which are also chosen at random from the environment of possible games.
4. Agents “experience” a game as the set of perceptual features associated with that game, and an awareness of the possible behaviors in that situation. Agents do not receive any further information regarding the strategic structure of the game.
5. Agents decide which of the available behaviors to employ based on the set of perceptual features associated with the game.

6. Agents receive a payoff determined by their own choice of behavior and the choices of other agent they have been paired with via the strategic structure of the game.
7. Agents re-evaluate their chosen behavior in light of received payoff.

Thus, broadly speaking, our approach is very similar to that of Bednar and Page (2007) and Mengel (2012). The important differences appear when considering points 4 and 5. There are no perceptual features in either of the previous models, and so in those previous models it is impossible for a decision to be based upon perceptual features. If agents do not experience games as a set of perceptual features, how do they experience them? In the case of Bednar and Page (2007) each game is perceived as a pointer to an initial state within an FSA. The position of the pointer is learned in conjunction with the FSA structure, and these together determine the behavior of an agent. In the case of Mengel (2012) agents do not experience the game as a set of features, but rather each game is perceived as belonging to a category. Again, this categorization is learned in conjunction with a category contingent behavioral rule, and these together determine the behavior of an agent.

Thus in our model and the models of Bednar and Page (2007) and Mengel (2012), agents consist of a behavioral function which maps perceptions of game situations to strategies, and a rule for modifying this function based on experience. To model issues of perception explicitly though, we require a behavioral function which acts on a set of perceptual features. Additionally we would like this behavioral function to respond to novel stimuli in a sensible way. We also require a learning rule that will construct this behavioral function from previous experiences. We use the set of stimuli shown in figure 1 to construct an example illustrating what constitutes sensible generalization.

Fig. 1 A set of possible stimuli.

Suppose that an agent is repeatedly exposed to stimuli a , c , d and e , and has learned to employ behavior A when experiencing stimuli a or e and to respond with behavior B to stimuli c or d . For the moment we ignore how this

learning occurs. Suppose that subsequently this agent is exposed to a novel stimuli b . In differentiating a and c from c and d , neither shading nor line placement are salient features, but only whether the central shape is a circle or a triangle. Thus an agent seeking a parsimonious decision rule is likely to have internalized the rule $\bullet \rightarrow A$, $\blacktriangle \rightarrow B$. While there is no guarantee that this rule will hold for novel stimuli, under the assumption that the world contains meaningful structure, responding with behavior A to b constitutes the sensible generalization.

Given our requirements for generalization to novel stimuli, a brain inspired *connectionist* or *artificial neural network* behavioral function is an appealing candidate. There is a vast literature debating the relative merits and limitations of artificial neural networks as models of learning and behavior. A comprehensive review of this literature is well beyond the scope of this paper, but we direct the interested reader to Enquist and Ghirlanda's recent treatment of the issue (Enquist and Ghirlanda, 2005). For us, the primary appeal of an artificial neural network approach is that, once trained, artificial networks generalize to novel stimulus in precisely the manner discussed above. This should be contrasted with the learning approaches used in Bednar and Page (2007) and Mengel (2012), where such generalization is simply not possible. In addition to generalizing to novel stimuli in a sensible manner, the types of discrimination problems which human or primate learners require relatively many trials to master, typically non-linear XOR type discriminations (Smith et al., 2011), are precisely those discriminations for which artificial neural networks also require relatively more learning trials to master.

The manner in which humans and other animals learn to use the perceptual features of a situation to engage in the appropriate behavior for that situation has been the focus of intense study over the past century and beyond (Ghirlanda and Enquist, 2003). Similarly individual learning in single specific game situations, and to a lesser extent in multi game situations, has also been well studied. Here we take some first steps at integrating the problems of perceptual learning and learning in strategic situations.

We begin by describing the technical implementation of our agents and the environment. Once this technical framework is established, we will illustrate some of its possible applications using a two simple simulation studies. We will conclude with a brief discussion of the empirical work suggested by this study and other possible implications of this framework.

2 Technical Framework

2.1 Environment

The environment consists of a set of games and a probability distribution over this set, determining which games are encountered most frequently by agents. In addition to their strategic structure, each game is also associated with a vector of perceptual features. A game's strategic structure is opaque to the

Fig. 2 Network Structure

agents (barring their implicit knowledge of the behavioral options available), and so agents must use the associated set of perceptual features to choose among the possible behaviors. How precisely agents do this is described in the following subsection.

2.2 Agents

Our agents consist of a feed forward neural network and a set of learning parameters which govern how this network is modified by experience.

An agent's network can be thought of as a function $\mathbf{f}(\cdot)$ which maps a vector, $\mathbf{x}$, of real valued perceptual features, to a vector of action probabilities, $\mathbf{y}$, via an intermediate layer of "hidden" neurons, $\mathbf{h}$, and an estimate of the value of each action in that situation $\mathbf{a}$. The general structure of this network function is illustrated in figure 2. This network structure is more formally defined below. Note that all vectors, $\mathbf{x}$, $\mathbf{h}$, $\mathbf{y}$, etc. are column vectors, and $\cdot$ denotes standard matrix multiplication.

The activity of the hidden layer, $\mathbf{h}$, is computed by feeding the perceptual vector, $\mathbf{x}$, through a matrix of connections, W_0 .

$$\mathbf{h} = \sigma(W_0 \cdot \mathbf{x} + \mathbf{b}_0) \quad (1)$$

Here, W_0^{ji} denotes the connection strength between perceptual feature x^i and hidden unit h^j , $\mathbf{b}_0$ are the biases of the hidden units, and σ denotes the logistic sigmoid function applied element wise to a vector.

$$\sigma(x) = \frac{1}{1 + \exp(-x)}$$

The use of a sigmoidal function is inspired by neural spiking activity, which is a sigmoidal function of the sum of incoming excitatory and inhibitory activity.

The logistic sigmoid is often used in artificial neural networks because of its relatively simple derivative.

The agent's anticipated payoff from each action, $\mathbf{a}$, is then computed by feeding the activities of the hidden layer, $\mathbf{h}$, through a second matrix, W_1 .

$$\mathbf{a} = W_1 \cdot \mathbf{h} + \mathbf{b}_1 \quad (2)$$

$$(3)$$

Finally these expected payoffs are converted into action probabilities using a softmax activation function.

$$y^k = \frac{\exp(\mathbf{a}^k/\tau)}{\sum_i \exp(\mathbf{a}^i/\tau)}. \quad (4)$$

The softmax activation function is a convenient way of transforming real valued payoff estimates into action probabilities. The parameter τ , often referred to as the temperature parameter, determines the extent to which high valued actions will be chosen over low valued actions. As $\tau \rightarrow \infty$ all actions are chosen with equal probability and as $\tau \rightarrow 0$ the highest valued action is chosen with certainty. Thus τ can be adjusted to modulate an agent's willingness to experiment with potentially dubious behaviors versus sticking to behaviors that it believes are high valued. Typically agents will be relatively exploratory at the beginning of a simulation, i.e. τ will be relatively high valued, and over the course of the simulation τ will decrease so that agents focus less on exploration and more on exploiting those behaviors which are most likely to return a high payoff.

Now that the network aspect of an agent is laid out clearly we are in a position to describe the way in which this network is modified by the experiences of the agent. When an agent performs an action they receive a payoff p . While an agent can never know whether it chose the best of all possible actions, it can seek to improve its estimate of the payoff received for the particular action taken. Suppose that an agent engages in behavior k . Prior to engaging in this action the agent estimated the payoff from engaging in k to be $\mathbf{a}^k$. The error e in the agent's estimate is then $e = p - \mathbf{a}^k$. What the agent requires is a way of modifying the parameters of the feed forward neural function, i.e. the elements of W_0 , W_1 , $\mathbf{b}_0$ and $\mathbf{b}_1$, so that over time the network function will produce better estimates of the reward from a given action in a given situation.

For a given perceptual stimulus $\mathbf{x}$, a chosen action k and a received payoff p , the square of the error is,

$$e^2 = (p - \mathbf{a}^k)^2 \quad (5)$$

$$= \left(p - \sum_j W_1^{kj} \cdot \mathbf{h}^j \right)^2 \quad (6)$$

$$= \left(p - \sum_j W_1^{kj} \cdot \sigma \left(\sum_i W_0^{ji} \cdot \mathbf{x}^i \right) \right)^2 \quad (7)$$

An agent that minimizes this squared error over all perceptual stimuli $\mathbf{x}$, and action choices k , will ultimately be able to consistently choose rewarding actions. The most straightforward, though potentially fraught, way to minimize this squared error is to use the chain rule to compute the gradient of e^2 with respect to the parameters of the network. Once the gradient has been computed the network parameters can be shifted in the direction of the gradient, scaled by a factor referred to as learning rate. This updating method, known as the back-propagation method (Rumelhart et al., 1986), is guaranteed to eventually converge on a locally error minimizing parameter configuration. However, this method cannot guarantee the global optimality of the parameter configuration achieved. Indeed, there is always a risk of agents becoming stuck in parameter configurations which, while locally optimal, are quite poor from a global perspective. Given that in our simulations there will be many agents all simultaneously learning in the same environment, if a small fraction of agents become stuck in poor local optima during the learning processes this will have only a marginal effect on the population level phenomena we are interested in.

In our simulations we will also incorporate a technique, known as *momentum*, to accelerate the learning process. With the momentum technique, the weight changes prescribed by a given error are spread out over future updates in an exponentially decaying fashion. Effectively, this means that only weight changes that are persistently prescribed for many different types of errors are actually implemented, leading to more efficient weight changes. For a more thorough treatment of momentum and its use see Rumelhart et al. (1986).

Thus an agent is completely described by its network function, the schedule according to which τ is initialized and lowered over time (modulating how agents transition from exploratory to conservative over time), and the learning parameters of the back-propagation method used to modify the network function.

3 Illustrative Studies

Our goal with these studies is to illustrate some of the possibilities of our general technical framework within a simple and transparent setting. To this end we exclusively consider two particular instances of the *stag hunt* and *prisoner's*

	C	D		C	D
C	3	1	C	3	4
	3	0		3	0
D	1	1	D	4	1
	1	1		4	1
	Stag Hunt			Prisoner's Dilemma	

Fig. 3 Payoff Matrices

dilemma games. These games are chosen for their central role in both the development of biological and evolutionary theories of altruism and cooperation (Axelrod and Hamilton, 1981; Hamilton, 1963) and in the formalization of political and philosophical ideas of social structure, social contracts, morality, communication, and ethics (Skyrms, 2002, 2004; Binmore, 2005). The payoff structures for these games are shown in figure 3. Notice that for these particular payoff structures, when players coordinate the payoffs received do not differ between games. It is only when players fail to coordinate that the payoff differences between these games occur, and these difference are only for the defecting player. Thus the payoff structures shown in figure 3 create a worst case scenario for discrimination between games. Restricting our attention to these two games gives us an environment consisting entirely of symmetric, two-player, two-strategy games, with identical strategy sets. These restrictions are chosen for simplicity. The technical framework described in section 2 does allow for environments consisting of games which are any combination of asymmetric, sequential, many player, and many strategy.

3.1 Study: Variation in cooperation depends on both the ratio between stag hunts and prisoner's dilemmas in the environment and the degree of perceptual similarity between stag hunt and prisoner's dilemma games.

Our basic hypothesis is that strategic situations sometimes share perceptual features, and that on the basis of these shared perceptual features behavioral tendencies learned in one situation may spill over to another. Here we use our general framework to investigate the implications of this hypothesis with regard to variations in perceptual similarity. To do this we measure the frequency of cooperative behavior as we vary both the perceptual similarity of the games and the relative frequencies at which stag hunts and prisoner's dilemma games are encountered within the environment.

3.1.1 Methods

We consider an environment consisting of ten games, where each game is assigned a 10-dimensional vector of binary perceptual features.² We vary the

² The choice of 10 features is a balance between a desire for gradations in perceptual similarity, requiring more features, and a desire for agents with smaller neural networks and hence less computational cost for simulations. Although binary and real valued features are

degree of perceptual similarity between games by assigning perceptual features to games using five different methods.

1. Each of the ten games in the environment is assigned the same vector of perceptual features. In this case agents lack the perceptual ability to distinguish between games and so must learn a strategy which works well for the particular mixture of games they encounter. (As in Zollman (2008) contingent behavior is impossible.)
2. Each game is assigned a random vector of perceptual features, drawn from the same distribution, regardless of the strategic structure of the game, i.e. whether it is a stag hunt or a prisoner's dilemma. In this case there is no correspondence between the functional structure of a game and its perceptual features, and so while discrimination is technically possible, it may be very difficult to learn. We use a distribution where each feature is present with a probability of 0.5, independent of which other features are or are not present.
3. Stag hunt games are assigned random perceptual features from one distribution, and prisoner's dilemma games are assigned perceptual features from a different distribution. Specifically the first five features in the label of a stag hunt game have a 0.6 probability of being present, and the later five features have a 0.4 probability of being present. The pattern is reversed for prisoner's dilemma games where the first five features have a 0.4 probability of being present and the last five features have a 0.6 probability of being present.
4. As in the previous case stag hunt games are assigned random perceptual features from one distribution, and prisoner's dilemma games are assigned perceptual features from another. In this case we decrease the similarity of these two distributions. Specifically the first five features in the label of a stag hunt game have a 0.9 probability of being present, and the later five features have a 0.1 probability of being present. The pattern is reversed for prisoner's dilemma games where the first five features have a 0.1 probability of being present and the last five features have a 0.9 probability of being present.
5. Stag hunt games are all assigned the same vector of perceptual features consisting of five 1's followed by five zeros. Similarly prisoner's dilemma games are all assigned the same vector of perceptual features consisting of five zeros followed by five ones. As the two games have diametrically opposed perceptual features this creates one of the easiest possible cases for discrimination.

Note that there are really only two games in method 1, a single stag hunt and a single prisoners dilemma, and they are perceived identically by the

both compatible with artificial neural networks, we have chosen binary perceptual features here as they are readily interpretable either at the lowest sensory level as the spiking of a single sensory neuron, or at a more complex cognitive level as the presence or absence of some learned perceptual feature. Such high level features might be, for example, whether a green light is on, whether a con-specific is present, or whether someone has mentioned the word risk in recent conversation.

agents. Similarly, there are really only two games in method 5, a single stage hunt and a single prisoners dilemma, this time with diametrically opposed perceptual features. Additionally, in method 4 there will typically be fewer than ten distinct feature vectors, as the distributions from which these feature vectors are drawn are relatively low variance. Nevertheless we persist with the ten game construct so that we can discuss all of the methods together, using the same terms.

We also vary the frequency with which prisoners dilemmas and stag hunts are encountered by the population. To do this we systematically vary the proportion of stag hunts and prisoners dilemmas in the 10 game environment. Taken together there are five methods of varying the degree of similarity, and eleven different ratio's of stag hunt to prisoner's dilemmas considered, for a total of 55 different treatments. For each of these treatments we run ten simulations. At the beginning of each simulation the ten games in the environment are assigned perceptual features using one of the methods described above. The perceptual features will remain fixed throughout the simulation, but may change from simulation to simulation. In each simulation a population of 100 agents will play 20000 rounds of games. Each round, every agent is randomly paired with a partner from the population. These pairs of agents play a game that is chosen uniformly and at random from the ten games comprising the environment. The strategies that agents employ are functions solely of the 10-element perceptual features of the games they encounter. In each simulation we record the total frequency of cooperation. The average level of cooperation over these simulations is plotted for each treatment in figure 4.

Agents learn as described in section 2.2. Each agent's network consists of 10 input units corresponding to the perceptual features of the games, 20 hidden processing units, and 2 output units corresponding to the cooperate and defect strategies in each game. Agents use a the same learning parameters in each treatment. The learning rate is 0.01 for the parameters W_1 and $\mathbf{b}_1$ and a learning rate is 0.005 for the parameters W_0 and $\mathbf{b}_0$. Agents have a momentum parameter of 0.6. The initial temperature of the agents is $\tau = 10.0$. Temperature decreases by 0.005 each round until it reaches a temperature of $\tau = 0.01$.³

3.1.2 Results

In the case where both games appear identical to the agents, it is impossible for agents to learn to discriminate between games. Consequently, the population converges to a state where cooperation is either ubiquitous or non-existent.

³ The selection of this cooling regime is critical to the results of these simulations. This particular regime was chosen as roughly the quickest cool-down that still allowed for some discrimination in the case where perceptual similarity as created by method 3 and where there an equal number of stag hunts and prisoner's dilemmas in the environment. A quicker cooling rate entails less exploration and hence less discrimination, whereas a slower cooling rate entails more exploration and hence more discrimination. The particular choice of cooling regime is unimportant for the demonstration of our point that perceptual similarity is important in determining the outcomes of learning in a multi game environments.

Fig. 4 Cooperation as a function of the number of prisoner’s dilemmas versus stag hunts in a 10 game environment

There is no middle ground.⁴ In contrast the case where the stag hunts and prisoner’s dilemmas are maximally distinguishable, i.e. have the exact opposite perceptual features, the agents consistently learn to discriminate perfectly. Consequently cooperation is a simple linear function of the proportion of games in the environment which engender cooperation and games which do not. As we vary the degree of perceptual similarity between these two extremes of perceptually identical and totally perceptually distinct we observe levels of cooperation intermediate between the two. In these cases of intermediate perceptual similarity, these intermediate levels of cooperation are the result of populations learning to discriminate less than perfectly, and sometimes not at all, between the two games in the environment.

3.1.3 Discussion

The “all or nothing” cooperation observed when the games are indistinguishable can be understood as follows. Initially agents play each strategy, in each game, with roughly equal probability. If there are few enough prisoner’s dilemmas, agents learn that on average cooperation pays. With relatively few prisoner’s dilemmas in the environment, a higher prevalence of cooperation increases the expected value of cooperation in a randomly chosen game more

⁴ The intermediate level of cooperation observed when 0.3 of the games encountered are prisoner’s dilemmas is the result of the population converging on always cooperating in some simulations, and converging on always cooperating in other simulations. This is to be expected. When 0.3 of the games encountered are prisoner’s dilemmas and initially the population plays cooperate and defect with equal probability, the expected value of playing cooperate and defect are very close, 1.5 and 1.45 respectively, making these particular simulations very sensitive to initial stochasticity.

than it increases the expected value of defection. This creates a feedback loop which increases the frequency of cooperation. In addition, because a prisoner’s dilemma and a stag hunt appear the same to players who cooperate, when cooperative agents do play prisoner’s dilemma they do not disabuse each other of the notion that cooperation is a good idea, but rather reinforce this notion.

On the other hand if there are many prisoner’s dilemmas in the environment, agents learn that on average defection pays. As the frequency of defection increases, the expected value of defection in a randomly chosen game decreases, but not as much as the expected value of cooperation decreases. This creates a feedback loop which increases the frequency of defection. Because the stag hunt is a coordination game, once the population has coordinated on the socially sub-optimal non-cooperative equilibrium in this game it is nigh impossible that they will ever learn to discriminate between stag hunts and prisoner’s dilemmas, which appear the same when played against a defector.

These non-linear all or nothing phenomena are a natural consequence of games being indistinguishable. This situation contrasts sharply with the linear mix of cooperation and defection observed when agents learn to discriminate perfectly between games.

When the stag hunts and prisoners dilemmas are perceptually distinct, but not perfectly so, the degree to which perceptual features are consistently associated with one game or the other increases the likelihood that a population will learn to discriminate between the two games. This leads to levels of cooperation intermediate between the linear levels when discrimination is perfect and the all or nothing levels when discrimination is impossible.

We hope these observations help to demonstrate the potential importance of perceptual similarity, and the perceptual mechanisms underlying learned discriminations, when considering the outcome of learning dynamics in a multi game environment.

3.2 Study: Reaction to novel stimuli.

Recall our basic hypothesis: strategic situations sometimes share perceptual features, and on the basis of these shared perceptual features behavioral tendencies learned in one situation may spill over to another. Here we use our general framework to investigate the implications of this hypothesis with regard to reactions to novel stimuli. To do this we measure the level of cooperative responses to novel stimuli from a population “trained” by playing in one of the ten game environments of the previous study.

3.2.1 Methods

We consider a single population of 100 agents that has learned through experiences of a ten game environment composed of five stag hunt games and five prisoner’s dilemma games. The games in this environment have each been assigned a ten element binary vector of perceptual features. These features

were assigned according to method 3 from the previous study, i.e. the first five features in the label of a stag hunt game have a 0.6 probability of being present, and the later five features have a 0.4 probability of being present, and the pattern is reversed for prisoner’s dilemma games. For 20000 rounds, a population of 100 agents play these games. Each round, every agent is randomly paired with a partner from the population. These pairs of agents play a game that is chosen uniformly and at random from the ten games comprising the environment. Agents learn as described in section 2.2, using the same learning parameters as in the previous study. This learning process does not always result in the agents discriminating, i.e. engaging in different behaviors, for prisoner’s dilemmas and stag hunts. However, we are interested in how discriminatory rules generalize to novel stimuli. For this reason, if the population fails to learn to discriminate we re-run the simulation until a population emerges that does, at least partially, discriminate between stag hunts and prisoner’s dilemmas. We know that this is the case if average cooperation levels fall anywhere between 10% and 90%.

After these 20000 rounds of learning we expose each agent in the population to two novel stimuli, and note their responses. The first stimulus is a perceptual feature vector consisting of five ones followed by five zeros, while the second stimulus follows the reverse pattern consisting of five zeros followed by five ones.

3.2.2 Results

In the environment these agents trained in, 1’s in the first half of the vector of perceptual features were slightly more common in stag hunt games, where as 1’s in the latter half of the vector were more common in prisoner’s dilemmas. As is expected from an artificial neural network, agents that had learned to use this regularity to discriminate between stag hunts and prisoner’s dilemmas were able to generalize from this trend, and respond consistently with cooperation to one of the stimuli and with defection to the other. To the first stimuli, five ones followed by five zeros, all of the agents respond with cooperation. To the second stimuli, five zeros followed by five ones, none of the agents respond with cooperation.

3.2.3 Discussion

The results from this study are largely unsurprising. The agents learned about a regularity between the perceptual and strategic structures of the games in the environment, and their responses to novel stimuli reflected this learning. As unsurprising as this result is, similar results simply cannot be obtained from a model which lacks explicit consideration of perceptual features.

4 Conclusion

We have extended the multi game frameworks presented in Bednar and Page (2007) and Mengel (2012), allowing for explicit consideration of perceptual similarity. We hope that the preceding studies illustrate the value this extension. The development of our framework was motivated by the empirical observation that people are sometimes more sensitive to perceptual similarities between situations than they are to strategic differences. Our technical framework helps explore the implications of this observation.

A specific application of our framework predicts that when a group of human learners repeatedly plays a mix of stag hunt games and prisoner’s dilemma games, the group is more likely to converge on all or nothing cooperation levels if the games are perceptually similar or if the mixture is dominated by one type of game, whereas discrimination is more likely to occur if stag hunts and prisoners dilemmas are perceptually dissimilar and if there is a relatively even mix of the two games. These predictions are readily testable in a standard economics game lab.

This technical framework also allows for further investigation along many possible avenues, here we briefly outline some of the most interesting.

Currently within the framework presented here, players are devoid of perceptual features, making play contingent upon the individual encountered impossible. Attaching perceptual features to players, and allowing play to be contingent both upon the perceptual features of the situation and of the other players present allows for an investigation of reputation formation across multiple strategic situations.

We investigated a case where agents choose their actions simultaneously. The framework can also be extended to sequential and stage games by attaching perceptual features to the actions chosen at previous stages in the game. Such an extension would allow for an investigation of ritualized sequences of signals across multiple strategic situations. In particular such an extension should be able to replicate the theoretical results of O’Connor (2013a) and Jäger (2007), and could be particularly useful for investigating the emergence of cross-contextual displays, as mentioned in Enquist et al. (2010) and Yabuta (2008).

References

- Axelrod, R. and Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211(4489):1390–1396.
- Bednar, J. and Page, S. (2007). Can game (s) theory explain culture? the emergence of cultural behavior within multiple games. *Rationality and Society*, 19(1):65–97.
- Binmore, K. (2005). *Natural justice*. Oxford University Press, USA.

- Brewer, M. and Kramer, R. (1986). Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing. *Journal of personality and social psychology*, 50(3):543.
- Cason, T. N., Savikhin, A., and Sheremeta, R. M. (2012). Behavioural spillovers in coordination games. *European Economic Review*, 56:233–245.
- Cronk, L. (2007). The influence of cultural framing on play in the trust game: A maasai example. *Evolution and Human Behavior*, 28(5):352–358.
- Enquist, M., Arak, A., Ghirlanda, S., and Wachtmeister, C.-A. (2002). Spectacular phenomena and limits to rationality in genetic and cultural evolution. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1427):1585–1594.
- Enquist, M. and Ghirlanda, S. (2005). *Neural networks and animal behavior*. Princeton University Press.
- Enquist, M., Hurd, P., and Ghirlanda, S. (2010). Signaling. In David F. Westneat, C. W. F., editor, *Evolutionary behavioral ecology*, pages 266–284. Oxford University Press.
- Eriksson, K. and Strimling, P. (2010). The devil is in the details: Incorrect intuitions in optimal search. *Journal of Economic Behavior & Organization*, 75(2):338–347.
- Fudenberg, D. and Levine, D. K. (1998). *The theory of learning in games*. MIT press.
- Gale, J., Binmore, K. G., and Samuelson, L. (1995). Learning to be imperfect: The ultimatum game. *Games and Economic Behavior*, 8(1):56–90.
- Ghirlanda, S. and Enquist, M. (2003). A century of generalization. *Animal Behaviour*, 66(1):15–36.
- Gick, M. and Holyoak, K. (1980). Analogical problem solving. *Cognitive psychology*, 12(3):306–355.
- Grafen, A. (1984). Natural selection, kin selection and group selection. *Behavioural ecology: an evolutionary approach*, 2.
- Grimm, V. and Mengel, F. (2012). An experiment on learning in a multiple games environment. *Journal of Economic Theory*, 147(6):2220 – 2259.
- Hamilton, W. D. (1963). The evolution of altruistic behavior. *American naturalist*, pages 354–356.
- Harsanyi, J. C. (1967). Games with incomplete information played by bayesian players, i-iii part i. the basic model. *Management science*, 14(3):159–182.
- Harsanyi, J. C. (1968a). Games with incomplete information played by bayesian players part ii. bayesian equilibrium points. *Management Science*, 14(5):320–334.
- Harsanyi, J. C. (1968b). Games with incomplete information played by 'bayesian' players, part iii. the basic probability distribution of the game. *Management Science*, 14(7):486–502.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., and McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, pages 73–78.
- Hofbauer, J. and Sigmund, K. (2003). Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 40(4):479–519.

- Huck, S., Jehiel, P., and Rutter, T. (2011). Feedback spillover and analogy-based expectations: A multi-game experiment. *Games and Economic Behavior*, 71(2):351–365.
- Jäger, G. (2007). The evolution of convex categories. *Linguistics and Philosophy*, 30(5):551–564.
- Levin, I. P., Schneider, S. L., and Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational behavior and human decision processes*, 76(2):149–188.
- Mengel, F. (2012). Learning across games. *Games and Economic Behavior*, 74(2):601–619.
- O’Connor, C. (2013a). The evolution of vagueness. *Erkenntnis*, pages 1–21.
- O’Connor, C. (2013b). Evolving perceptual categories. *pre-print*.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088):533–536.
- Schuster, P. and Sigmund, K. (1983). Replicator dynamics. *Journal of theoretical biology*, 100(3):533–538.
- Skyrms, B. (2002). Signals, evolution and the explanatory power of transient information*. *Philosophy of Science*, 69(3):407–428.
- Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge University Press.
- Skyrms, B. and Zollman, K. J. (2010). Evolutionary considerations in the framing of social norms. *Politics, Philosophy & Economics*, 9(3):265–273.
- Smith, J. D., Coutinho, M. V. C., and Couchman, J. J. (2011). The learning of exclusive-or categories by monkeys (macaca mulatta) and humans (homo sapiens). *Journal of Experimental Psychology: Animal Behavior Processes*, 37(1):20–29.
- Taylor, P. D. and Jonker, L. B. (1978). Evolutionary stable strategies and game dynamics. *Mathematical biosciences*, 40(1):145–156.
- Tversky, A. and Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of business*, pages S251–S278.
- Von Neumann, J. and Morgenstern, O. (2007). *Theory of games and economic behavior (commemorative edition)*. Princeton university press.
- Weibull, J. W. (1997). *Evolutionary game theory*. MIT press.
- Yabuta, S. (2008). Evolution of cross-contextual displays: the role of risk of inappropriate attacks on nonopponents, such as partners. *Animal Behaviour*, 76(3):865–870.
- Yamagishi, T., Mifune, N., Li, Y., Shinada, M., Hashimoto, H., Horita, Y., Miura, A., Inukai, K., Tanida, S., Kiyonari, T., et al. (2013). Is behavioral pro-sociality game-specific? pro-social preference and expectations of pro-sociality. *Organizational Behavior and Human Decision Processes*, 120(2):260–271.
- Zollman, K. J. (2008). Explaining fairness in complex environments. *politics, philosophy & economics*, 7(1):81–97.