Generalized Symmetry in Stochastic Games

Martin A. Zinkevich
Computer Science Dept.
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

This paper discusses symmetries in stochastic games. In the past, people have discussed symmetric bimatrix games. However, we extend this concept so that players can have internal symmetries between their own states, as well as have states which are symmetric with the states of other players in the game. We prove that if a stochastic game possesses a symmetry, then the best response function is symmetric and there exists a symmetric Nash equilibrium.

1 Introduction and Background

Zinkevich and Balch (2001) establish that if a Markov Decision Process(MDP) is symmetric, then it possesses a symmetric optimal policy. This means that a learning agent can exploit symmetries when using Watkin's Q-learning (Watkins, 1989), or value iteration (Samuel, 1959). However, what if there are other agents in the system? Can the agents still exploit symmetry? The answer is yes, if the other agents are playing symmetrically as well.

First, let us describe what type of symmetry we are discussing. Symmetry between agents has been discussed before (Rappaport, 1973; Schelling, 1960; Thie, 1979). Here we unify this type of symmetry with symmetries within agents: actions with symmetric outcomes, states with symmetric actions, and so forth. The actual definition of symmetry becomes quite complex, but can be summarized simply. If in state s player p is in the "same" situation as player p' in state s', then these two situations are symmetric. If all the players playing the joint action a in state s is the same for p as all the players players playing the joint action a' in state s' is for p', then the two joint actions are symmetric. Finally, if player p plays an action a in state s, then if regardless of the actions of the other players, player p' playing an action a' in state s' has the same effect. This is a very intricate definition, and very hard to actually apply. Therefore, it is more recommended that it be utilized at an intuitive level, in order to guide programming decisions or the development of learning models. However, the formalization is important in that it allows one to recognize situations where the theorem does not apply.

Suppose that there are two agents playing a guessing game. Player A chooses a number between 1 and 10, and player B must guess the number A chose. If player B guesses correctly, then B wins a dollar from A[1], and if B guesses incorrectly, B loses a dollar. If A chooses a number uniformly at random, then B can do no better than guess a number uniformly at random. In fact, this is a Nash equilibrium, because A can also not improve over choosing at random if B is guessing at random.

However, suppose A always chooses 4. Then B can do far better by always choosing 4. Hence, if any agent plays according to an asymmetric strategy, some or all of the other agents may have no symmetric best responses. But there are times when we can guarantee that our opponent is going to use a symmetric strategy. Suppose there are two teams of robots. In team A, each robot is identical in its appearance as well as its abilities and goals. The robots in team B cannot discriminate between these robots. Therefore, team B must use the same strategy regardless of the permutation of the robots on team A. In such circumstances, where a player knows the other players will act symmetrically, the player can act symmetrically as well. Also, for any symmetric stochastic game, there is a guarantee that there exists some symmetric Nash equilibrium.

[1] Throughout the paper, we assume that each player prefers a situation where its expected monetary gain is higher over a situation where its expected monetary gain is lower

2 Related Work

The study of the existence of Nash equilibria is fundamental to game theory and economics. A large body of work exists establishing the conditions of the existence of Nash equilibria under various circumstances. Nash (1950) proved that there was a Nash equilibrium in N-person games in strategic form. Fink (1964) established for any stochastic game the existence of a Nash equilibria. Filar & Vrieze (1997) gave another proof of the existence of Nash equilbria in an arbitrary stochastic game, and the structure of the proof of the main theorem of this paper follows the structure of their proof.

Unfortunately, there has been little work done in attempting to establish the structure of the Nash equilibria in stochastic games Filar & Vrieze (1997). It is well known that in games in strategic form symmetry between two players results in symmetric Nash equilibria. This type of symmetry is included in the results here, but our definition of symmetry is far more general.

3 Stochastic Games

Here we present the traditional definitions relating to stochastic games, in order to clarify the notation that will be needed for the definition of symmetry introduced in section 4.

Definition 1 *A two-player* ***stochastic game*** *is a tuple* $(n, \mathcal{S}, \mathcal{A}, T, R)$, $\mathcal{S}$ *is the finite set of states,*$\mathcal{A}(s) = \mathcal{A}_1(s) \times \mathcal{A}_2(s) \times \ldots \times \mathcal{A}_n(s)$ *is the finite set of joint actions for a state* s, $\mathcal{A}_i(s)$ *is the set of actions for player* i *in a state* s, T *is a transition function from* $\bigcup_{s\in\mathcal{S}}(\{s\} \times \mathcal{A}(s)) \times \mathcal{S}$ *to* $[0,1]$ *such that for all* $s \in \mathcal{S}$, *for all* $a \in \mathcal{A}(s)$, $\sum_{s'\in\mathcal{S}} T(s,a,s') = 1$, *and* R *is a vector of functions such that for all* $i \in \{1 \ldots n\}$, $R_i : \bigcup_{s\in\mathcal{S}}(\{s\} \times \mathcal{A}(s)) \to \mathbb{R}$ *is a reward function for player* i.

A stochastic game is a very general type of system with discrete time steps. If the system is in a state $s \in \mathcal{S}$, and for all $i \in \{1 \ldots n\}$ the player i executes action $a_i \in \mathcal{A}_i$, $T(s, (a_1, \ldots, a_n), s')$ is the probability that the system will be in state $s' \in \mathcal{S}$ at the next time step. If the system is in a state $s \in \mathcal{S}$, and for all $i \in \{1 \ldots n\}$ the player i executes action $a_i \in \mathcal{A}_i$, for all $p \in \{1 \ldots n\}$, player p receives a reward of $R_p(s, (a_1, \ldots, a_n))$.

Definition 2 *For a stochastic game* $(\mathcal{S}, \mathcal{A}, T, R)$, *a* ***strategy*** *for a player* $i \in \{1 \ldots n\}$, *is a function* $\sigma_i : \mathcal{S} \times \mathcal{A}_i \to [0,1]$, *where for all* $s \in \mathcal{S}$, $\sum_{a\in\mathcal{A}_i} \sigma_i(s,a) = 1$. *The set of all strategies for player* i *is* Σ_i. *A strategy profile is a tuple* $\sigma \in \Sigma_1 \times \Sigma_2 \times \ldots \times \Sigma_n$. *For all* $s \in \mathcal{S}$, *for all* $a \in \mathcal{A}$, *define* $\sigma(s,a) = \prod_{i=1}^n \sigma(s, a_i)$.

A strategy profile defines how the agents will act in a game. For all $p \in P$, for all $s \in \mathcal{S}$, $a \in \mathcal{A}_i$, $\sigma_i(s,a)$ is the probability that agent i will execute action a from state s. Given the system is in a state $s \in \mathcal{S}$, for all $a \in \mathcal{A}$, $\sigma(s,a)$ is the probability that the players will execute the joint action a. Therefore, if the agents are using a strategy tuple σ, given the system is in state $s \in \mathcal{S}$, the probability that the system will be in state s' at the next time step is $P_{s,s'} = T(s,\sigma,s') = \sum_{a\in\mathcal{A}} \sigma(s,a)T(s,a,s')$, and the expected reward for agent i is $R_i(\sigma, a) = \sum_{a\in\mathcal{A}} \sigma(s,a)R_i(s,a)$. The matrix formulation is convenient for the purposes of easily representing multiple transitions as an exponentiated matrix.

Definition 3 *Given a stochastic game* $(n, \mathcal{S}, \mathcal{A}, T, \mathbb{R})$, *and a strategy tuple* σ, *the* ***expected value*** *of a state* $s \in \mathcal{S}$, *for a player* $i \in \{1 \ldots n\}$ *is:*

$$V_{\sigma,i}(s) = \sum_{s'\in\mathcal{S}} \gamma^t (P^t)_{s's} \sum_{a\in\mathcal{A}} \sigma(s,a) R_i(s,a)$$

Observe that

$$V_{\sigma,i}(s) = R_i(s,\sigma) + \gamma \sum_{s'\in\mathcal{S}} T(s,\sigma,s')V(s')$$

Suppose that for every player plays according to a strategy profile σ but player p. We can represent this situation by $\sigma|_p\sigma'_p$, which means "replace the strategy of p in σ with σ'_p". Also, it is possible to write $\sigma|_p(\sigma_p|_s a)$, which indicates that all one should change the strategy of p so that it plays action a in state s with probability 1. We can simplify this by writing $\sigma|_{p,s}a$.

Definition 4 *Given a stochastic game* $(n, \mathcal{S}, \mathcal{A}, T, \mathbb{R})$, *and strategies* $\sigma = \sigma_1, \ldots, \sigma_n$, *a best response of player* j *is a strategy* σ'_j *such that for all* $\sigma''_j \in \Sigma_j$, *for all* $s \in \mathcal{S}$,

$$V_{(\sigma|_j\sigma'_j),j}(s) \geq V_{(\sigma|_j\sigma''_j),j}(s)$$

. *Define* $BR_j(\sigma)$ *to be the set of best responses in* Σ_j. *Define* $BR(\sigma) = BR_1(\sigma) \times \ldots \times BR_n(\sigma)$. *A* ***Nash equilibrium*** *is a strategy tuple* σ *such that* $\sigma \in BR(\sigma)$.

The existence of Nash equilibria is one of the principal studies in Game Theory. It is known that for any stochastic game there exists a Nash equilibrium.

4 Symmetry

What is symmetry in a stochastic game? In order for it to be useful, it should have certain properties.

For instance, symmetric players should be able to use the same strategies in some sense. Also, symmetric states should have the same Equilibrium values. Zinkevich and Balch (2001) studied symmetry in MDPs. In that paper, we considered sets of symmetric states and symmetric actions in symmetric states. Now, because there are multiple players, we will consider symmetric states for players, and symmetric joint actions in symmetric states for players. We argued that two actions in MDPs should be considered symmetric if the results were themselves symmetric. This is still the foundation of our discussion for stochastic games.

It is also possible to consider symmetric players. Two players A and B are symmetric if for every state of player A, there exists a symmetric state of player B, and for every state of player B, there exists a symmetric state of player A. If A knows a good strategy for B, and he is some state s, then A can look at the symmetric state s' of B and then translate the action of B in state s' back into an action for A in state s.

However, it is interesting to consider that there may be meaningful symmetries where two players are not wholly symmetric but have meaningful symmetric states. For instance, consider a game where player A has a choice between playing rock, paper, scissors with player B, or playing matching pennies with player C. Obviously, A is not initially in a symmetric state with B or C. However, if A decided to play rock, paper, scissors with player B, A is now symmetric with player B. If A decided to play matching pennies with player C, then A is now in a state symmetric with the state of player C. This is a result of the Markov Property, that future events are only dependent on the current state of the game.

Also, consider another game where player A decides first to play game I or II. If A decides to play game I, then A chooses an integer between 1 and 10, and then B attempts to guess the number A chose. If B guesses the number correctly, then A gives B a dollar, and the game is over. If B guesses incorrectly, then B gives A a dollar, and the game is over. If A decides to play game II, A can now choose an integer between 1 and 20.

Now for every joint action that A and B play in game I, there exists a joint action in game II with the same consequences. However, these games are not symmetric in the sense that A should play game II if A wants a greater chance of winning a dollar. If A chooses uniformly at random over all integers between 1 and 20, then A will receive an expected reward of $(+1)(19/20)+(-1)(1/20)=9/10$. If A plays game I and chooses uniformly at random over all integers between 1 and 10, then A will receive an expected reward of 8/10. This violates the intuition that symmetric states should have symmetric values.

Therefore, it is important that not only symmetric states have symmetric joint actions, but that these symmetric joint actions be configured in a similar fashion.

Definition 5 *A* ***symmetry*** *is an ordered tuple* $(E_{\mathcal{S}}, E_{\mathcal{A}}, E_{\mathcal{A}^*})$*,* $E_{\mathcal{S}}$ *is an equivalence relation over* $\mathcal{S}\times\{1\ldots n\}$*,* $E_{\mathcal{A}}$ *is an equivalence relation over* $\bigcup_{s\in\mathcal{S}}(\{s\}\times\{1\ldots n\}\times\mathcal{A}(s))$*,* $E_{\mathcal{A}^*}$ *is an equivalence relation over* $\bigcup_{s\in\mathcal{S}}(\{s\}\times\bigcup_{p\in\{1\ldots n\}}(\{p\}\times\mathcal{A}_p(s)))$ *and the following properties hold:*

1. *If* $((s,p),(s',p'))\in E_{\mathcal{S}}$*, then there exists a bijection* $g:\mathcal{A}(s)\to\mathcal{A}(s')$ *such that for all* $a\in\mathcal{A}(s)$*,* $((s,p,a),(s',p',g(a)))\in E_{\mathcal{A}}$
2. *If* $((s,p),(s',p'))\in E_{\mathcal{S}}$*, then there exists a bijection* $g:\mathcal{A}_p(s)\to\mathcal{A}_{p'}(s')$ *such that for all* $a''\in\mathcal{A}_p(s)$*,*$((s,p,a''),(s',p',g(a'')))\in E_{\mathcal{A}^*}$ *.*
3. *If* $((s,p,a),(s',p',a'))\in E_{\mathcal{A}}$*, then there exists a bijection* $h:\{1\ldots n\}\to\{1\ldots n\}$ *such that for all* $p''\in\{1\ldots n\}$*,* $((s,p'',a_{p''}),(s',h(p''),a'_{h(p'')}))\in E_{\mathcal{A}^*}$*.*
4. *If* $((s,p,a),(s',p',a'))\in E_{\mathcal{A}^*}$ *and* $((s,p,a''),(s',p',a'''))\in E_{\mathcal{A}}$*, then* $((s,p,a''|_p a),(s',p',a'''|_{p'}a'))\in E_{\mathcal{A}}$*.*
5. *If* $((s,p,a),(s',p',a'))\in E_{\mathcal{A}^*}$*, then* $((s,p),(s',p'))\in E_{\mathcal{S}}$
6. *If* $((s,p,a),(s',p',a'))\in E_{\mathcal{A}}$*, then* $((s,p),(s',p'))\in E_{\mathcal{S}}$*.*

If $((s,p),(s',p'))\in E_{\mathcal{S}}$, then player p in state s is in situation equivalent to player p' in state s'. It is crucial to understand that situations as opposed to states in games are symmetric. For instance, being in jail is one thing for a criminal and quite another for a jailer.

If $((s,p,a),(s',p',a'))\in E_{\mathcal{A}}$, then player p in state s when the joint action a is played is in a symmetric situation to p' in state s' when the joint action a' is played. Note that the transitions and rewards of the system are based on thes joint actions. In order for two states to be symmetrical, symmetric joint actions must be present.

However, no individual player can select the joint action on its own, it only has control over its own action. Therefore, for two players to be in the same situation, they must have the same actions in those situations. If $((s,p,a),(s',p',a'))\in E_{A^*}$, then player p in state s playing action a is the same as player p' in state s' playing action a'. Here is the truly confusing aspect of the symmetry: what does it mean for two players' actions to be symmetrical? The outcome of a single player's action is not well defined: one must know the actions of all the players in order to determine the outcome. Suppose that player p in state s plays action a, and player p' in state s' plays action a'. Suppose that everyone else in state s plays according to joint action a'', and everyone in state s' plays joint action a'''. Then the outcomes should be the same: if two players play symmetric actions in situations where everyone is playing symmetrically, then the outcomes should be symmetric.

So, let us consider what it means for a stochastic game to have a symmetry. First of all, if an agent turns eight degrees right on the right side of the soccer field, then it should have the same effect if the agent turns eight degrees left on the left side of the soccer field. This does not mean that the agent ends in exactly the same position in both cases, only that the new positions where it arrives are symmetrical to each other.

Definition 6 *Given a stochastic game $(n, \mathcal{S}, \mathcal{A}, T, R)$ and a symmetry $(E_{\mathcal{S}}, E_{\mathcal{A}}, E_{\mathcal{A}^*})$:*

1. *A reward function is symmetric if for all $((p, s, a), (p', s', a')) \in E_{\mathcal{A}}$, $R_p(s, a) = R_{p'}(s', a')$.*
2. *A transition function is symmetric if for all $((p, s, a), (p', s', a')) \in E_{\mathcal{A}}$, for all $s'' \in \mathcal{S}$:*
$$\sum_{(s''',p)\in E_{\mathcal{S}}(s'',p)} T(s,a,s''') = \sum_{(s''',p')\in E_{\mathcal{S}}(s'',p)} T(s',a',s''')$$
3. *A stochastic game is symmetric if the reward function and the transition function are symmetric.*
4. *A strategy profile is symmetric if for all $((p, s, a), (p', s', a')) \in E_{\mathcal{A}^*}$, $\sigma_p(s, a) = \sigma_{p'}(s', a')$.*
5. *A best response function is symmetric if for all symmetric strategy profiles σ there exists a symmetric strategy profile σ' such that $\sigma' \in BR(\sigma)$.*

Theorem 1 *Given a stochastic game $(n, \mathcal{S}, \mathcal{A}, T, R)$ which is symmetric with respect to a symmetry (p, s, a) then:*

1. *The best response function is symmetric.*
2. *There exists a symmetric Nash equilibrium.*

In order to prove the first part, we first discuss how the problem of discovering a best response function can be reduced to discovering the set of optimal actions in an MDP. Then, we show how when we begin with a symmetric stochastic game and a symmetric strategy profile is being used, the resulting MDP is symmetric. Then we refer to the fact that a symmetric MDP has a set of symmetric optimal actions (Zinkevich & Balch, 2001). Finally, we establish that a symmetric set of optimal actions implies the existence of at least one symmetric best response strategy profile.

In order to prove the second part, we use Kakutani's Fixed-Point Theorem and some properties of symmetric correspondences. The proof will be in Sections 6-8. Before we give this, we will establish some examples of symmetry.

5 Examples of Symmetry

Consider the following situation from (Zinkevich & Balch, 2001): Wacko Foods has two stores in Fruitytown. Each day, each store can sell either apples or oranges, but not both. If one store sells apples and the other sells oranges, they sell them all and Wacko Foods makes ten dollars. If both stores sell apples or both stores sell oranges, they sell half and Wacko Foods makes five dollars. If we defines this as a stochastic game with one state, the players are symmetric. Therefore, there exists a symmetric Nash equilibrium in which the two players do the same thing: specifically, they each sell apples and oranges with equal probability. However, there are Nash equilibria which are significantly better.

Another more drastic example is this: suppose that one has a set of symmetric robots that are playing soccer. Then it is simple to find the symmetric equilibrium - they forfeit! If all but one player decides not to play, then that player gains nothing by playing. However, suppose we consider the players as being controlled by a single agent. Then a forfeit is no longer part of a Nash equilibrium, because the agent can order that the players all take the field. Thus, we are guaranteed a symmetric Nash equilibrium in this case. However, if we consider the players to be all controlled by the same agent, then they must have some shared source of randomness. Another concern is that they must be able to read from this shared source. For instance, if they are all in the same state, and the random action chosen requires them to do different things, how are the actions delegated? However, if they are always in distinct states, then they can easily know which action applies to them.

Thus, we state as a corollary to the above theorem:

Corollary 1 *Given a homogeneous set of agents in distinct states with a common source of randomness, there exists a homogeneous Nash equilibrium for their controller.*

I state this without formalism or proof, because I believe that the intuitive meaning is clear, and the proof follows directly from the proof of symmetry in stochastic games. For a formalism of homogeneity, see (Zinkevich & Balch, 2001).

6 Symmetric Joint Strategies

Understanding the nature of symmetric joint strategies is crucial to the proof of the main theorem.

Lemma 1 *Suppose that $\sigma \in \Sigma$ is symmetric with respect to $(E_{\mathcal{S}}, E_{\mathcal{A}}, E_{\mathcal{A}^*})$, then for all $((s,p,a),(s',p',a')) \in E_{\mathcal{A}}$, $\sigma(s,a) = \sigma(s',a')$.*

The proof follows directly from the definition of symmetry.

$$\sigma(s,a) = \prod_{i=1}^{n} \sigma_i(s, a_i)$$

Since $((s,p,a),(s',p',a')) \in E_{\mathcal{A}}$, there exists a bijection $h : \{1 \dots n\} \to \{1 \dots n\}$ such that for all $p'' \in \{1 \dots n\}$, $((s,p'',a_{p''}),(s',h(p''),a'_{h(p'')})) \in E_{\mathcal{A}^*}$. Since σ is symmetric, for all $p'' \in \{1 \dots n\}$, $\sigma_{p''}(s, a_{p''}) = \sigma_{h(p'')}(s', a'_{h(p'')})$. Therefore:

$$\sigma(s,a) = \prod_{i=1}^{n} \sigma_{h(i)}(s', a'_{h(i)})$$

Since h is a bijection:

$$\sigma(s,a) = \prod_{j=1}^{n} \sigma_j(s', a'_j) = \sigma(s', a')$$

∎

7 A Related Markov Decision Process

For this section let us fix a stochastic game $(n, \mathcal{S}, \mathcal{A}, T, R)$ which is symmetric with respect to $(E_{\mathcal{S}}, E_{\mathcal{A}}, E_{\mathcal{A}^*})$. If one specifies the strategies for all but one player, a stochastic game reduces to an MDP[2]. Suppose that for every player plays according to a strategy σ but player p. First, the set of states for the MDP is $\mathcal{S}$ and the set of actions is $\mathcal{A}_p$. The transition function is $T'(s,a,s') = T(s, \sigma|_{p,s} a, s')$. The reward function is $R'(s,a) = R_p(s, \sigma|_{p,s} a)$. Observe that for the above equations, it is only really important that we are taking the action a at state s.

[2]Observe that these MDPs differ from those in Zinkevich & Balch (2001), in that the action is dependent on the state. However, it is trivial to extend the definitions and proof in that paper to such games

The optimal value function of a MDP is the expected discounted reward if the agent begins in that state and acts optimally. So, the set of all optimal actions for a state s is those that maximize $R'(s,a) + \gamma T'(s,a,s')V^*(s')$. Observe that this equals $R_i(s, \sigma|_p a) + \gamma T(s, \sigma|_p a, s')V^*(s')$. In other words, a strategy which places a probability of 1 on the set of optimal actions in all states is a best response in the stochastic game. Observe that this set of strategies is convex: i.e., the mixture of two best response strategies results in a best response strategy.

We can make such a MDP for every player. We can also combine these into a single MDP. Let the set of states in the MDP be $\mathcal{S} \times \{1 \dots n\}$, and the set of actions for state (s,p) be $\mathcal{A}_p(s)$. The transition function is $T'((s,p),a,(s',p')) = T(s, \sigma|_{p,s} a, s')$ when $p = p'$ and zero otherwise. The reward function is $R'((s,p),a) = R_p(s, \sigma|_{p,s} a)$. So, if we discover the optimal actions in this MDP, we can find the best responses for all of the players. We will now correlate the symmetry in the stochastic game to a symmetry in the MDP.

Lemma 2 *Consider the MDP symmetry $(E'_{\mathcal{S}}, E'_{\mathcal{A}})$ such that $E'_{\mathcal{S}} = E_{\mathcal{S}}$ and for all $p,p' \in \{1 \dots n\}$, for all $s,s' \in \mathcal{S}$, for all $a \in \mathcal{A}_p(s)$, for all $a' \in \mathcal{A}_{p'}(s')$, $(((s,p),a),((s',p'),a')) \in E'_{\mathcal{A}}$ if and only if $((s,p,a),(s',p',a')) \in E_{\mathcal{A}^*}$. R' is symmetric with respect to $(E'_{\mathcal{S}}, E'_{\mathcal{A}})$ if σ is symmetric with respect to $(E_{\mathcal{S}}, E_{\mathcal{A}}, E_{\mathcal{A}^*})$.*

For all $(((s,p),a),((s',p'),a')) \in E'_{\mathcal{A}}$:

$$\begin{aligned} R'((s,p),a) &= R_p(s, \sigma|_{p,s} a) \\ &= \sum_{a'' \in \mathcal{A}(s)} (\sigma|_{p,s} a)(s, a'') R_p(s, a'') \\ &= \sum_{a'' \in \mathcal{A}_p(s)} \sigma(s, a'') R_p(s, a''|_p a) \end{aligned}$$

Now, $((s,p,a),(s',p',a')) \in E_{\mathcal{A}^*}$. Therefore, $((s,p),(s',p')) \in E_{\mathcal{S}}$. There exists a mapping $g : \mathcal{A}(s) \to \mathcal{A}(s')$ such that for all $a'' \in \mathcal{A}(s)$, $((s,p,a''),(s',p',g(a''))) \in E_{\mathcal{A}}$. Therefore:

$$R'((s,p),a) = \sum_{a'' \in \mathcal{A}(s)} \sigma(s', g(a'')) R_p(s, a''|_p a)$$

Also, because $((s,p,a''|_p a),(s',p',g(a'')|_{p'} a')) \in E_{\mathcal{A}^*}$:

$$R'((s,p),a) = \sum_{a'' \in \mathcal{A}(s)} \sigma(s', g(a'')) R_p(s', g(a'')|_{p'} a')$$

Because g is a bijection:

$$\begin{aligned} R'((s,p),a) &= \textstyle\sum_{a'''\in\mathcal{A}(s')} \sigma(s',a''')R_p(s',a'''|_{p'}a') \\ &= R'((s',p'),a') \end{aligned}$$

∎

Before we prove that the transition function is symmetric, we present a function W: this function does not have a very intuitive meaning: however, it has a structural place in the proof. Given a stochastic game and a strategy σ, define a function W : $(\bigcup_{s\in\mathcal{S},p\in\{1\ldots n\}}\{s\}\times\{p\}\times\mathcal{A}(s)\times\mathcal{S}\times\{1\ldots n\}\times\mathcal{A}_p(s))\to\mathbb{R}$:

$$\begin{aligned} &W(s,p,a,s'',p'',a'') \\ &\qquad = \sum_{(s^{iv},p)\in E_{\mathcal{S}}(s'',p'')} \sigma(s,a)T(s,a|_p a'',s^{iv}) \end{aligned}$$

Lemma 3 *If σ is symmetric, then for all $s''\in\mathcal{S}$, $W(s,p,a,s'',p'',a'')=W(s',p',a',s'',p'',a''')$.*

By definition:

$$\begin{aligned} &W(s,p,a,s'',p'',a'') \\ &\qquad = \sigma(s,a) \sum_{(s^{iv},p)\in E_{\mathcal{S}}(s'',p'')} T(s,a|_p a'',s^{iv}) \end{aligned}$$

Observe that $((s,p,a|_p a''),(s',p',a'|_{p'}a''')) \in E_{\mathcal{A}}$. Therefore:

$$\begin{aligned} &W(s,p,a,s'',p,a'') \\ &\qquad = \sigma(s,a) \sum_{(s^{iv},p')\in E_{\mathcal{S}}(s'',p'')} T(s',a'|_{p'}a''',s^{iv}) \end{aligned}$$

Because $((s,p,a),(s',p',a'))\in E_{\mathcal{A}}$:

$$\begin{aligned} &W(s,p,a,s'',p,a'') \\ &\qquad = \sigma(s',a') \sum_{(s^{iv},p')\in E_{\mathcal{S}}(s'',p'')} T(s',a'|_{p'}a''',s^{iv}) \end{aligned}$$

$$W(s,p,a,s'',p,a'') = W(s',p',a',s'',p'',a''')$$

∎

Lemma 4 *Consider the MDP symmetry $(E'_{\mathcal{S}},E'_{\mathcal{A}})$ such that $E'_{\mathcal{S}}=E_{\mathcal{S}}$ and for all $p,p'\in\{1\ldots n\}$, for all $s,s'\in\mathcal{S}$, for all $a\in\mathcal{A}_p(s)$, for all $a'\in\mathcal{A}_{p'}(s')$, $(((s,p),a),((s',p'),a'))\in E'_{\mathcal{A}}$ if and only if $((s,p),a),((s',p'),a'))\in E_{\mathcal{A}}$. T' is symmetric with respect to $(E'_{\mathcal{S}},E'_{\mathcal{A}})$ if σ is symmetric with respect to $(E_{\mathcal{S}},E_{\mathcal{A}},E_{\mathcal{A}^*})$.*

For any $(((s,p),a),((s',p'),a'))\in E'_{\mathcal{A}}$, $((s,p,a),(s',p',a'))\in E_{\mathcal{A}^*}$. Observe that

$$\begin{aligned} T'((s,p),a,(s',p)) &= \sum_{a''\in\mathcal{A}(s)} (\sigma|_{p,s}a)(s,a'')T(s,a'',s') \\ &= \sum_{a''\in\mathcal{A}(s)} \sigma(s,a'')T(s,a''|_p a,s') \end{aligned}$$

It follows directly:

$$\begin{aligned} &\sum_{(s'',p)\in E_{\mathcal{S}}(s'',p'')} T'(s,a,s'') \\ &\qquad = \sum_{a''\in\mathcal{A}(s)} W(s,p,a'',s'',p'',a) \end{aligned}$$

Observe that $((s,p),(s',p'))\in E_{\mathcal{S}}$. Therefore, there exists a function $g:\mathcal{A}(s)\to\mathcal{A}(s')$ such that for all $a''\in\mathcal{A}(s)$, $((s,p,a''),(s',p',g(a'')))\in E_{\mathcal{A}}$. Thus, by the previous lemma:

$$\begin{aligned} &\sum_{(s'',p)\in E_{\mathcal{S}}(s'',p'')} T'(s,a,s'') \\ &\qquad = \sum_{a''\in\mathcal{A}(s)} W(s',p',g(a''),s'',p'',a') \end{aligned}$$

Because g is a bijection:

$$\begin{aligned} &\sum_{(s'',p)\in E_{\mathcal{S}}(s'',p'')} T'(s,a,s'') \\ &\qquad = \sum_{a''\in\mathcal{A}(s')} W(s',p',a'',s'',p'',a') \\ &\qquad = \sum_{(s'',p')\in E_{\mathcal{S}}(s'',p'')} T'(s',a',s'') \end{aligned}$$

So T' is symmetric. ∎

Lemma 5 *Consider the MDP symmetry $(E'_{\mathcal{S}},E'_{\mathcal{A}})$ such that $E'_{\mathcal{S}}=E_{\mathcal{S}}$ and for all $p,p'\in\{1\ldots n\}$, for all $s,s'\in\mathcal{S}$, for all $a\in\mathcal{A}_p(s)$, for all $a'\in\mathcal{A}_{p'}(s')$, $(((s,p),a),((s',p'),a'))\in E'_{\mathcal{A}}$ if and only if $((s,p),a),((s',p'),a'))\in E_{\mathcal{A}}$. The MDP constructed above is symmetric with respect to $(E'_{\mathcal{S}},E'_{\mathcal{A}})$ if σ is symmetric with respect to $(E_{\mathcal{S}},E_{\mathcal{A}},E_{\mathcal{A}^*})$.*

This follows directly from Lemma 2 and Lemma 4. ∎

Finally, we shall establish that the best response function is symmetric. Since we have established that the MDP is symmetric, we can use Theorem 2 from Zinkevich and Balch (2001) to establish that the sets of optimal actions are symmetric. Now we wish to mix these optimal actions in order to construct a symmetric strategy profile. The technique we will use is to choose an action from the set of optimal actions uniformly at random.

Will this result in a symmetric joint strategy? Consider an element in $((s,p,a),(s',p',a')) \in E_{\mathcal{A}^*}$. Since the set of optimal actions is symmetric, if a is optimal in (s,p), then a' is optimal in (s',p'). If the cardinality of the set of optimal actions in (s,p) equals the cardinality of the set of optimal actions in (s',p'), then $\sigma_p(s,a) = \sigma_{p'}(s',a')$. Observe that $((s,p),(s',p')) \in E_{\mathcal{S}}$. Therefore, there exists a bijection $g : \mathcal{A}_p(s) \to \mathcal{A}_{p'}(s')$ such that for all $a'' \in \mathcal{A}(s)$, $((s,p,a''),(s',p',g(a''))) \in E_{\mathcal{A}^*}$. Since every optimal action in $\mathcal{A}_p(s)$ maps to an optimal action in $\mathcal{A}_{p'}(s')$, and every action which is optimal in $\mathcal{A}_{p'}(s')$ is mapped by g^{-1} to an optimal action in $\mathcal{A}_p(s)$, then the cardinality of the set of optimal actions in (s,p) is equal to the cardinality of the set of optimal actions in (s',p'). Thus, if σ is symmetric, we can construct a symmetric best response. This concludes the proof of the first portion of Theorem 1.

8 Symmetric Correspondences

The best response function is a correspondence: it maps elements to sets. A fixed point of the best response function, a strategy which is a best response to itself, is a Nash equilibrium. In this section, we will discuss what a symmetric correspondence is, and show how Kakutani's fixed point theorem can be extended to prove that symmetric correspondences have symmetric fixed points. For more information on fixed point theory and its applications to game theory, see (Border, 1985).

Definition 7 *Given $S \subseteq \mathbb{R}^m$ and $T \subseteq \mathbb{R}^n$, a correspondence $\gamma : S \rightarrow\rightarrow T$ is upper semi-continuous if for all sequences $x^k \to x$ in S and $y^k \to y$ in T, $y \in f(x)$.*

The double arrow($\rightarrow\rightarrow$) indicates a correspondence. This is similar to the concept of continuity of continuous functions. Observe that if $f : S \to T$ is a continuous function, then $\gamma(x) = \{f(x)\}$ is an upper semi-continuous correspondence.

Lemma 6 *Let $S \subseteq \mathbb{R}^m$, $T \subseteq \mathbb{R}^n$, $S' \subseteq S$, and $T' \subseteq T$. Then $\lambda : S' \rightarrow\rightarrow T'$ defined as $\lambda(x) = \gamma(x) \cap T'$ is an upper semi-continuous correspondence.*

Suppose that $\{x^k\} \to x$ is in S' and $\{y^k\} \to y$ is in T', and for all k, $y^k \in \lambda(x^k)$. Then they are in S and T respectively, so $y \in \gamma(x)$. Since $y \in T'$, then $y \in \lambda(x)$. ∎

Theorem 2 Kakutani's Fixed Point Theorem Let $K \subset \mathbb{R}^m$ be compact and convex and $\gamma : K \rightarrow\rightarrow K$ be closed or upper semi-continuous with nonempty, convex, compact values. Then γ has a fixed point.

For a proof, see (Border, 1985).

In (Filar & Vrieze, 1997), when they prove that a stochastic game has a Nash equilibrium, they establish that the best response function is upper semi-continuous and has nonempty, convex, compact values.

Definition 8 *Consider an equivalence relation E over $\{1 \dots m\}$.*

1. *a vector $\vec{v} \in R^m$ is symmetric if for all $(i,j) \in E$, $v_i = v_j$.*
2. *a set $S \subseteq R^m$ is symmetric if there exists a symmetric vector $\vec{v} \in S$.*
3. *if $S \subseteq R^m$ and $T \subseteq R^m$, a correspondence $\gamma : S \rightarrow\rightarrow T$ is symmetric if for all symmetric $\vec{x} \in S$, there exists a symmetric $\vec{y} \in \gamma(\vec{x})$.*

Later, we will show how symmetric strategies and a symmetric best response functions correlate to the definitions of symmetry presented here.

Theorem 3 *Let E be an equivalence relation over $\{1 \dots m\}$. The set of all symmetric vectors in $\mathbb{R}^m$ is closed and convex.*

See (Zinkevich & Balch, 2001) for a proof.

Theorem 4 *Let E be an equivalence relation over $\{1 \dots m\}$, $K \subset R^m$ be a symmetric, convex, compact set, and $\gamma : K \rightarrow\rightarrow K$ be a upper semi-continuous, symmetric correspondence with convex, compact values. Then γ has a symmetric fixed point.*

Proof: We will restrict γ to a version defined only on symmetric vectors, and prove that this function has a fixed point.

Define S to be the set of all symmetric vectors in E. This is closed and convex according to Theorem 3. Define $S' = K \cap S$. K is closed and bounded, by the definition of compactness. Because S and K are closed, S' is closed. Because K is bounded, S' is bounded. Hence, S' is compact. S' is nonempty, because K is symmetric. Because S and K are convex, S' is convex. Now define $\lambda : S' \rightarrow\rightarrow S'$ as $\lambda(x) = \gamma(x) \cap S'$. Thus, λ is upper semi-continuous. For every $x \in S$, $\lambda(x)$ is nonempty by the definition of the symmetry of γ. $\lambda(x)$ is convex because S' and $\gamma(x)$ are convex. $\lambda(x)$ is also compact, because S' and $\gamma(x)$ are compact. Therefore, λ has a fixed point, which is a symmetric fixed point of γ. ∎

So, we can use this proof in order to prove that a symmetric stochastic game has a symmetric Nash equilibrium. Consider a strategy $\sigma \in \Sigma$ as existing in Euclidean space. First, consider the set $X = \bigcup_{s \in \mathcal{S}}(\{s\} \times \bigcup_{p \in \{1 \dots n\}}(\{p\} \times \mathcal{A}_p(s)))$. This set is finite, and assuming it is nonempty there exists a bijection $g : X \rightarrow \{1 \dots |X|\}$. We can also construct a relation E' over $\{1 \dots |X|\}$ where $(x, x') \in E'$ if and only if $(g^{-1}(x), g^{-1}(x')) \in E_{\mathcal{A}^*}$. Finally, we can construct a function $h : \Sigma \rightarrow \mathbb{R}^{|X|}$ defined as $h_i(\sigma) = \sigma_{g_2^{-1}(i)}(g_1^{-1}(i), g_3^{-1}(i))$. Therefore, symmetric strategies become symmetric vectors in $\mathbb{R}^{|X|}$, and the best response function becomes a symmetric correspondence in $R^{|X|}$.

Finally, we must show that the image of Σ in $\mathbb{R}^{|X|}$ is a symmetric set. Therefore, we need to prove that there exists a symmetric strategy. Suppose every player in every state selects an action uniformly at random. Suppose $((s,p,a),(s',p',a')) \in E_{\mathcal{A}^*}$, then $((s,p),(s',p')) \in E_{\mathcal{A}}$. Therefore, there exists a bijection $g : \mathcal{A}_p(s) \rightarrow \mathcal{A}_{p'}(s')$, so $|\mathcal{A}_p(s)| = |\mathcal{A}_{p'}(s')|$. Therefore, in these symmetric states the probability of these symmetric actions is identical, so there exists a symmetric strategy which proves that the image of Σ in $\mathbb{R}^{|X|}$ is symmetric. The image of Σ is also clearly closed and convex, as it is the cartesian product of standard closed simplexes.

This establishes the fact that there exists a symmetric Nash equilibrium, and concludes the proof of the main theorem. ∎

9 Conclusion

Symmetric stochastic games have symmetric Nash equilibria. There exists homogeneous strategies for homogeneous agents in distinct states with a common source of randomness. Symmetry can be used as a refinement to the Nash equilibrium concept. These results can guide future studies both in terms of developing programs for multiagent systems, as well as developing learning algorithms for agents in multiagent systems. The examples presented in sections 4 and 5 are in some ways more valuable than the proof itself. In them is contained the intuitive meaning of the proof, which is far more useful than the actual application of the formal mechanism.

Acknowledgements

I would like to thank Avrim Blum, Tucker Balch, Rande Shern, Pat Riley, Mike Bowling and Weng-Keen Wong for helpful discussions on this paper. This material is based upon work supported under a National Science Foundation Graduate Research Fellowship. Any opinion, findings, conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

Border, K. C. (1985) *Fixed-point theorems with applications to economics and game theory.* New York: Cambridge University Press.

Fink, A. M. (1964) Equilibrium in a Stochastic N-Person Game. *Journal of Science in Hiroshima Univ., Series A-I,* 28:89-93.

Filar, J. & Vrieze, K. (1997) *Competitive Markov decision processes.* New York: Springer Verlag.

Nash, J. F. (1950). Equilibrium Points in N-Person Games. *Proceedings, National Academy of Science, USA,* 36:48-49.

Littman, M. L. (1994). Markov Games as a framework for multi-agent reinforcement learning. *Machine Learning,* 11, (pp. 157–163).

Nash, J. F. (1950). Equilibrium Points in N-Person Games. *Proceedings, National Academy of Science, USA,* 36:48-49.

Rappaport, A. (1973). *Two-Person Game Theory.* The University of Michigan Press, Ann Arbor, Michigan.

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development,* 3, 211-229.

Schelling, T. C. (1960). *The Strategy of Conflict.* Harvard University Press. Cambridge, Massachusetts.

Thie, P. R. (1979).*An Introduction to Linear Programming and Game Theory.* John Wiley & Sons, Inc. New York.

Watkins, C. (1989). *Learning from delayed rewards.* Ph.D. dissertation, King's College, Cambridge, England.

Zinkevich, M. & Balch, T. (2001). Applications of the Theory of Homogeneous Agents To Multiagent Learning. *Eighteenth International Conference on Machine Learning.*