

# Give a Hard Problem to a Diverse Team: Exploring Large Action Spaces

Leandro Soriano Marcolino<sup>1</sup>, Haifeng Xu<sup>1</sup>, Albert Xin Jiang<sup>1</sup>, Milind Tambe<sup>1</sup>, Emma Bowring<sup>2</sup>

<sup>1</sup>University of Southern California, Los Angeles, CA, 90089, USA

{sorianom, haifengx, jiangx, tambe}@usc.edu

<sup>2</sup>University of the Pacific, Stockton, CA, 95211, USA

ebowring@pacific.edu

## Abstract

Recent work has shown that diverse teams can outperform a uniform team made of copies of the best agent. However, there are fundamental questions that were not asked before. When should we use diverse or uniform teams? How does the performance change as the action space or the teams get larger? Hence, we present a new model of diversity for teams, that is more general than previous models. We prove that the performance of a diverse team improves as the size of the action space gets larger. Concerning the size of the diverse team, we show that the performance converges exponentially fast to the optimal one as we increase the number of agents. We present synthetic experiments that allow us to gain further insights: even though a diverse team outperforms a uniform team when the size of the action space increases, the uniform team will eventually again play better than the diverse team for a large enough action space. We verify our predictions in a system of Go playing agents, where we show a diverse team that improves in performance as the board size increases, and eventually overcomes a uniform team.

## Introduction

Team formation is crucial when deploying a multi-agent system (Nair and Tambe 2005; Liemhetcharat and Veloso 2012). Many researchers emphasize the importance of diversity when forming teams (LiCalzi and Surucu 2012; Lamberson and Page 2012; Hong and Page 2004; Marcolino, Jiang, and Tambe 2013). However, there are many important questions about diversity that were not asked before, and are not explored in such models.

LiCalzi and Surucu (2012) and Hong and Page (2004) propose models where the agents know the utility of the solutions, and the team converges to the best solution found by one of its members. Clearly in complex problems the utility of solutions would not be available, and agents would have to resort to other methods, such as voting, to take a common decision. Lamberson and Page (2012) study diversity in the context of forecasts, where the solutions are represented by real numbers and the team takes the average of the opinion of its members. Domains where the possible solutions are discrete, however, are not captured by such a model.

Copyright © 2014, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Marcolino, Jiang, and Tambe (2013) study teams of agents that vote in discrete solution spaces. They show that a diverse team of weaker agents can overcome a uniform team made of copies of the best agent. However, this phenomenon does not always occur, and they do not present ways to know when we should use diverse teams. They also lack a formal study of how the performance of diverse teams change as the number of agents and/or actions increase.

In this paper we shed a new light into this problem, by presenting a new, more general model of diversity for teams of voting agents. Our model captures better than the previous ones the notion of a diverse team as a team of agents that tend not to agree on the same actions, and allows us to make new predictions. Our main insight is based on the notion of *spreading tail (ST)* and *non-spreading tail (NST)* agents. As we will show, a team of *ST* agents has a diverse behavior, i.e., they tend to not agree in the same actions. Hence, we can model a diverse team as a team of *ST* agents, and show that the performance improves as the size of the action space gets larger. We also prove upper and lower bounds on how fast different teams converge. The improvement can be large enough to overcome a uniform team of *NST* agents, even if individually the *ST* agents are weaker. As it is generally hard to find good solutions for problems with a large number of actions, it is important to know which teams to use in order to tackle such problems. Moreover, we show that the performance of a diverse team converges to the optimal one exponentially fast as the size of the team grows. We show synthetic experiments that provide even further insights about our model. Even though the diverse team overcomes the uniform team in a large action space, the uniform team eventually will again play better than the diverse team as the action space keeps increasing if the best agent does not behave exactly like an *NST* agent.

Finally, we test our predictions by studying a system of voting agents, in the Computer Go domain. We show that a uniform team made of copies of the best agent plays better in smaller board sizes, but is overcome by a diverse team as the board gets larger. Moreover, we analyze the agents and verify that weak agents have a behavior closer to *ST* agents, while the best agent is closer to an *NST* agent. Therefore, we show that the predictions are verified in a real system, and can effectively be used while forming a multi-agent team.

## Related Work

We have already discussed models of diversity in team formation in the previous section, so here we start with other models of team formation in the multi-agent literature.

In classical team formation research, the team with maximum expected utility is chosen, based on a model of the capabilities of each agent (Nair and Tambe 2005; Guttman 2008). However, in many domains we do not have such a model. The study of “ad-hoc” teamwork deals with multi-agent teams with absence of information (Barrett et al. 2013; Agmon and Stone 2012). They focus, however, on how a new agent must decide its behavior in order to cooperate with agents of unknown type, not on picking the best team.

Concerning voting, this paper is related to the view of voting as a way to discover an optimal choice (or ranking). Classical models study this view of voting only for teams of identical agents (List and Goodin 2001; Conitzer and Sandholm 2005). However, more recent works are also considering agents with different probability distribution functions. Caragiannis, Procaccia, and Shah (2013) study which voting rules converge to a true ranking as the number of agents (not necessarily identical) goes to infinity. In Soufiani, Parkes, and Xia (2012) the problem of inferring the true ranking is studied, assuming agents with different pdfs, but drawn from the same family. Even though recent works on voting are not assuming identical agents, they still do not provide a way to find the best teams of voting agents.

More related works can be found in the study of ensemble systems. These are the construction of a strong classifier by combining multiple weak classifiers, for example by voting (Polikar 2012). An important problem is how to form the ensemble system, i.e., how to pick the classifiers that lead to the best predictions (Fu et al. 2012). Our model, based on the notion of *spreading tail* and *non-spreading tail* agents allow us to make many predictions about teams as the action space and/or number of agents change, and also compare the rate of change of the performance of different teams. To the best of our knowledge, there is no model similar to ours in the ensemble system literature.

### Model for Analysis of Diversity in Teams

Consider a problem defined by choosing an action  $a$  from a set of possible actions  $\mathbf{A}$ . Each  $a$  has an utility  $U(a)$ , and our goal is to maximize the utility. We always list the actions in order from best to worst, therefore  $U(a_j) \geq U(a_{j+1}) \forall j$  ( $a_0$  is the best action). This does not mean that newly added actions have lower utility, as we can always reorder the actions. In some tasks (like in Section Experimental Analysis), a series of actions are chosen across different states, but here we focus on the decision process in a given state.

Consider a set of agents, voting to decide over actions. The agents do not know the utility of the actions, and vote for the action they believe to be the best according to their own decision procedure, characterized by a probability distribution (pdf). We write as  $p_{i,j}$  the probability of agent  $i$  voting for action  $a_j$ . We denote by  $p_{i,j}(m)$ , when we explicitly talk about  $p_{i,j}$  for an action space of size  $m$ . If the pdf of one agent is identical to the pdf of another agent,

they will be referred to as copies of the same agent. The action that wins by plurality voting is taken by the team. Ties are broken randomly, except when we explicitly talk about a tie breaking rule. Let  $\mathbf{D}_m$  be the set of suboptimal actions ( $a_j, j \neq 0$ ) assigned with a nonzero probability in the pdf of an agent  $i$ , and  $d_m = |\mathbf{D}_m|$ . We assume that there is a bound in the ratio of the suboptimal action with highest probability and the one with lowest nonzero probability, i.e., let  $p_{i,min} = \min_{j \in \mathbf{D}_m} p_{i,j}$  and  $p_{i,max} = \max_{j \in \mathbf{D}_m} p_{i,j}$ ; there is a constant  $\alpha$  such that  $p_{i,max} \leq \alpha p_{i,min} \forall$  agents  $i$ .

We define strength as the expected utility of an agent and/or a team. The probability of a team playing the best action will be called  $p_{best}$ . We first consider a setting where  $U(a_0) \gg U(a_j) \forall j \neq 0$ , hence we can use  $p_{best}$  as our measure of performance. We will later consider more general settings, where the first  $r$  actions have a high utility.

We define team formation as selecting from the space of all agents a limited number of agents that has the maximum strength by voting together to decide on actions. We study the effect of increasing the size  $m$  of the set of possible actions on the team formation problem. Intuitively, the change in team performance as  $m$  increases will be affected by how the pdf of the individual agents  $i$  change when  $m$  gets higher. As we increase  $m$ ,  $d_m$  can increase or not change. Hence, we classify the agents as *spreading tail (ST)* agents or *non-spreading tail agents (NST)*.

We define *ST* agents as agents whose  $d_m$  is non-decreasing on  $m$  and  $d_m \rightarrow \infty$  as  $m \rightarrow \infty$ . We consider that there is a constant  $\epsilon > 0$ , such that for all *ST* agents  $i$ ,  $\forall m, p_{i,0} \geq \epsilon$ . We assume that  $p_{i,0}$  does not change with  $m$ , although later we discuss what happens when  $p_{i,0}$  changes.

We define *NST* agents as agents whose pdf does not change as the number of actions  $m$  increases. Hence, let  $m_{i0}$  be the minimum number of actions necessary to define the pdf of an *NST* agent  $i$ . We have that  $\forall m, m' \geq m_{i0}$ ,  $\forall j \leq m_{i0} p_{i,j}(m) = p_{i,j}(m')$ ,  $\forall j > m_{i0} p_{i,j}(m) = 0$ .

We first give an intuitive description of the concept of diversity, then define formally diverse teams. By diversity, we mean agents that tend to disagree. In Marcolino, Jiang, and Tambe (2013), a diverse team is defined as a set of agents with different pdfs. Hence, they disagree because of having different probabilities of playing certain actions. In this paper, we generalize their definition to capture cases where agents disagree on actions, regardless of whether their pdfs are the same or not. Formally, we define a diverse team to be one consisting of a set of *ST* agents (either different *ST* agents or copies of the same *ST* agent). In our theoretical development we will show that this definition captures the notion of diversity: a team of *ST* agents will tend to not agree on the same suboptimal actions. We call uniform team as the team composed by copies of an *NST* agent. This is an idealization to perform our initial analysis. We will later discuss more complex domains, where the agents of the uniform team also behave like *ST* agents.

### A Hard Problem to a Diverse Team

We start with an example, to give an intuition about our model. Consider the agents in Table 1(a), where we show the pdf of the agents, and  $p_{best}$  of the uniform team (three

(a) With 2 actions, the uniform team plays better than the diverse team

Agents	Action 1	Action 2
Agent 1	0.6	0.4
Agent 2	0.55	0.45
Agent 3	0.55	0.45
Uniform $p_{best}$ :	0.648	
Diverse $p_{best}$ :	0.599	

(b) When we add one more action, the diverse team plays better than the uniform team

Agents	Action 1	Action 2	Action 3
Agent 1	0.6	0.4	0
Agent 2	0.55	0.25	0.2
Agent 3	0.55	0.15	0.3
Uniform $p_{best}$ :	0.648		
Diverse $p_{best}$ :	0.657		

Table 1: The performance of a diverse team increases when we increase the number of available actions.

copies of agent 1) and the diverse team (one copy of each agent). We assume agent 1 is an *NST* agent, while agent 2 and 3 are *ST* agents. In this situation the uniform team plays better than the diverse team. Now let's add one more action to the problem. Because agent 2 and 3 are *ST* agents, the probability mass on action 2 scatters to the newly added action (Table 1(b)). Hence, while before the *ST* agents would always agree on the same suboptimal action if they both did not vote for the optimal action, now they might vote for different suboptimal actions, creating a tie between each suboptimal action and the optimal one. Because ties are broken randomly, when this happens there will be a 1/3 chance that the tie will be broken in favor of the optimal action. Hence,  $p_{best}$  increases when the probability of the *ST* agents agreeing on the same suboptimal actions decreases, and the diverse team now plays better than the uniform team, even though individually agents 2 and 3 are weaker than agent 1.

We now present our theoretical work. First we show that the performance of a diverse team converges when  $m \rightarrow \infty$ , to a value that is higher than the performance for any other  $m$ .

**Theorem 1.**  $p_{best}(m)$  of a diverse team of  $n$  agents converges to a certain value  $\tilde{p}_{best}$  as  $m \rightarrow \infty$ . Furthermore,  $\tilde{p}_{best} \geq p_{best}(m)$ ,  $\forall m$ .

*Proof.* Let  $p_{i,min} = \min_{j \in \mathbf{D}_m} p_{i,j}$ ,  $p_{i,max} = \max_{j \in \mathbf{D}_m} p_{i,j}$  and  $\mathbf{T}$  be the set of agents in the team. By our assumptions, there is a constant  $\alpha$  such that  $p_{i,max} \leq \alpha p_{i,min}$  for all agents  $i$ . Then, we have that  $1 \geq 1 - p_{i,0} = \sum_{j \in \mathbf{D}_m} p_{i,j} \geq d_m p_{i,min}$ . Therefore,  $p_{i,min} \leq \frac{1}{d_m} \rightarrow 0$  as  $d_m$  tends to  $\infty$  with  $m$ . Similarly,  $\alpha p_{i,min} \rightarrow 0$  as  $d_m \rightarrow \infty$ . As  $p_{i,j} \leq \alpha p_{i,min}$  we have that  $\forall j p_{i,j} \rightarrow 0$  as  $d_m \rightarrow \infty$ . We show that this implies that when  $m \rightarrow \infty$ , weak agents never agree on the same suboptimal action. Let  $i_1$  and  $i_2$  be two arbitrary agents. Without loss of generality, assume  $i_2$ 's  $d_m$  ( $d_m^{(i_2)}$ )

is greater than or equal  $i_1$ 's  $d_m$  ( $d_m^{(i_1)}$ ). The probability ( $\sigma_{i_1,i_2}$ ) of  $i_1$  and  $i_2$  agreeing on the same suboptimal action is upper bounded by  $\sigma_{i_1,i_2} = \sum_{a_j \in \mathbf{A} \setminus a_0} p_{i_1,j} p_{i_2,j} \leq d_m^{(i_2)} p_{i_1,max} p_{i_2,max} \leq d_m^{(i_2)} \alpha p_{i_2,min} p_{i_1,max} \leq \alpha p_{i_1,max}$  (as  $d_m^{(i_2)} p_{i_2,min} \leq 1$ ). We have that  $\alpha p_{i_1,max} \rightarrow 0$  as  $p_{i_1,max} \rightarrow 0$ , because  $\alpha$  is a constant. Hence the probability of any two agents agreeing on a suboptimal action is  $\frac{\sum_{i_1 \in \mathbf{T}} \sum_{i_2 \in \mathbf{T}, i_2 \neq i_1} \sigma_{i_1,i_2}}{2} \leq \frac{n(n-1)}{2} \max_{i_1,i_2} \sigma_{i_1,i_2} \rightarrow 0$ , as  $n$  is a constant.

Hence, when  $m \rightarrow \infty$ , the diverse team only chooses a suboptimal action if all agents vote for a different suboptimal action or in a tie between the optimal action and suboptimal actions (because ties are broken randomly). Therefore,  $p_{best}$  converges to:

$$\tilde{p}_{best} = 1 - \prod_{i=1}^n (1 - p_{i,0}) - \sum_{i=1}^n (p_{i,0} \prod_{j=1, j \neq i}^n (1 - p_{j,0})) \frac{n-1}{n}, \quad (1)$$

that is, the total probability minus the cases where the best action is not chosen: the second term covers the case where all agents vote for a suboptimal action and the third term covers the case where one agent votes for the optimal action and all other agents vote for suboptimal actions.

When  $m$  is finite, the agents might choose a suboptimal action by agreeing over that suboptimal action. Therefore, we have that  $p_{best}(m) \leq \tilde{p}_{best} \forall m$ .  $\square$

Let  $p_{best}^{uniform}(m)$  be  $p_{best}$  of the uniform team, with  $m$  actions. A uniform team is not affected by increasing  $m$ , as the pdf of an *NST* agent will not change. Hence,  $p_{best}^{uniform}(m)$  is the same,  $\forall m$ . If  $\tilde{p}_{best}$  is high enough so that  $\tilde{p}_{best} \geq p_{best}^{uniform}(m)$ , the diverse team will overcome the uniform team, when  $m \rightarrow \infty$ . Therefore, the diverse team will be better than the uniform team when  $m$  is large enough.

In practice, a uniform team made of copies of the best agent (the one with highest expected utility) might not behave exactly like a team of *NST* agents, as the best agent could also increase its  $d_m$  as  $m$  gets larger. We discuss this situation in Section Experimental Analysis. In order to perform that study, we derive in the following corollary how fast  $p_{best}$  converges to  $\tilde{p}_{best}$ , as a function of  $d_m$ .

**Corollary 1.**  $p_{best}(m)$  of a diverse team increases to  $\tilde{p}_{best}$  in the order of  $O(\frac{1}{d_m^{min}})$  and  $\Omega(\frac{1}{d_m^{max}})$ , where  $d_m^{max}$  is the highest and  $d_m^{min}$  the lowest  $d_m$  of the team.

*Proof.* We assume here the notation that was used in the previous proof. First we show a lowerbound on  $p_{best}(m)$ . We have that  $p_{best}(m) = 1 - \psi_1$ , where  $\psi_1$  is the probability of the team picking a suboptimal action.  $\psi_1 = \psi_2 + \psi_3$ , where  $\psi_2$  is the probability of no agent agreeing and the team picks a suboptimal action and  $\psi_3$  is the probability of at least two agents agreeing and the team picks a suboptimal action. Hence,  $p_{best}(m) = 1 - \psi_2 - \psi_3 = \tilde{p}_{best} - \psi_3 \geq \tilde{p}_{best} - \psi_4$ , where  $\psi_4$  is the probability of at least two agents agreeing. Let  $\sigma^{max} = \max_{i_1,i_2} \sigma_{i_1,i_2}$ , and  $i_1^*$  and  $i_2^*$  are the agents whose  $\sigma_{i_1^*,i_2^*} = \sigma^{max}$ . We have that  $p_{best}(m) \geq \tilde{p}_{best} - \frac{n(n-1)}{2} \sigma^{max} \geq \tilde{p}_{best} - \frac{n(n-1)}{2} d_m^{(i_2^*)} p_{i_1^*,max} p_{i_2^*,max} \geq$

$$\begin{aligned} \tilde{p}_{best} - \frac{n(n-1)}{2} d_m^{(i_2^*)} \alpha p_{i_1^*, min} \alpha p_{i_2^*, min} &\geq \tilde{p}_{best} - \\ \frac{n(n-1)}{2} \alpha^2 \frac{1}{d_m^{(i_1^*)}} &\text{ (as } p_{i, min} \leq \frac{1}{d_m} \text{). Hence, } p_{best}(m) \geq \\ \tilde{p}_{best} - \frac{n(n-1)}{2} \alpha^2 \frac{1}{d_m^{min}} &\rightsquigarrow \tilde{p}_{best} - p_{best}(m) \leq O\left(\frac{1}{d_m^{min}}\right). \end{aligned}$$

Now we show an upper bound:  $p_{best}(m) = \tilde{p}_{best} - \psi_3 \leq \tilde{p}_{best} - \psi_5$ , where  $\psi_5$  is the probability of at least two agents agreeing and no agents vote for the optimal action. Let  $\sigma^{min} = \min_{i_1, i_2} \sigma_{i_1, i_2}$ ;  $i_1^*$  and  $i_2^*$  are the agents whose  $\sigma_{i_1^*, i_2^*} = \sigma^{min}$ ; and  $p_{max,0} = \max_{i \in \mathbf{T}} p_{i,0}$ . Without loss of generality, we assume that  $d_m^{(i_2^*)} \geq d_m^{(i_1^*)}$ . Hence,  $p_{best}(m) \leq \tilde{p}_{best} - \frac{n(n-1)}{2} \sigma^{min} (1 - p_{max,0})^{n-2} \leq \tilde{p}_{best} - \frac{n(n-1)}{2} d_m^{(i_1^*)} p_{i_1^*, min} p_{i_2^*, min} (1 - p_{max,0})^{n-2} \leq \tilde{p}_{best} - \frac{n(n-1)}{2} d_m^{(i_1^*)} \frac{p_{i_1^*, max} p_{i_2^*, max}}{\alpha^2} (1 - p_{max,0})^{n-2} \leq \tilde{p}_{best} - \frac{n(n-1)}{2} \alpha^{-2} \frac{1}{d_m^{i_2^*}} (1 - p_{max,0})^{n-2} \leq \tilde{p}_{best} - \frac{n(n-1)}{2} \alpha^{-2} \frac{1}{d_m^{max}} (1 - p_{max,0})^{n-2} \rightsquigarrow \tilde{p}_{best} - p_{best}(m) \geq \Omega\left(\frac{1}{d_m^{max}}\right)$ .  $\square$

Hence, agents that change their  $d_m$  faster will converge faster to  $\tilde{p}_{best}$ . This is an important result when we consider later more complex scenarios where the  $d_m$  of the agents of the uniform team also change.

Note that  $\tilde{p}_{best}$  depends on the number of agents  $n$  (Equation 1). Now we show that the diverse team tends to always play the optimal action, as  $n \rightarrow \infty$ .

**Theorem 2.**  $\tilde{p}_{best}$  converges to 1, as  $n \rightarrow \infty$ . Furthermore,  $1 - \tilde{p}_{best}$  converges exponentially to 0, that is,  $\exists$  constant  $c$ , such that  $1 - \tilde{p}_{best} \leq c(1 - \frac{\epsilon}{2})^n$ ,  $\forall n \geq \frac{2}{\epsilon}$ . However, the performance of the uniform team improves as  $n \rightarrow \infty$  only if  $p_{s,0} = \max_j p_{s,j}$ , where  $s$  is the best agent.

*Proof.* By the previous proof, we know that when  $m \rightarrow \infty$  the diverse team plays the optimal action with probability given by  $\tilde{p}_{best}$ . We show that  $1 - \tilde{p}_{best} \rightarrow 0$  exponentially as  $n \rightarrow \infty$  (this naturally induces  $\tilde{p}_{best} \rightarrow 1$ ). We first compute an upper bound for  $\sum_{i=1}^n (p_{i,0} \prod_{j=1, j \neq i}^n (1 - p_{j,0}))$ :  $\sum_{i=1}^n p_{i,0} \prod_{j=1, j \neq i}^n (1 - p_{j,0}) \leq \sum_{i=1}^n p_{i,0} (1 - p_{min,0})^{n-1} \leq n p_{max,0} (1 - p_{min,0})^{n-1} \leq n(1 - \epsilon)^{n-1}$  for  $p_{max,0} = \max_i p_{i,0}$ ,  $p_{min,0} = \min_j p_{j,0}$ .

Since  $\prod_{i=1}^n (1 - p_{i,0}) \leq (1 - \epsilon)^n$ , thus we have that  $1 - \tilde{p}_{best} \leq (1 - \epsilon)^n + n(1 - \epsilon)^{n-1}$ . So we only need to prove that there exists a constant  $c$  such that  $(1 - \epsilon)^n + n(1 - \epsilon)^{n-1} \leq c(1 - \frac{\epsilon}{2})^n$ , as follows:  $\frac{(1 - \epsilon)^{n+1} + (n+1)(1 - \epsilon)^n}{(1 - \epsilon)^n + n(1 - \epsilon)^{n-1}} = (1 - \epsilon) \frac{1 - \epsilon + n + 1}{1 - \epsilon + n} = 1 - \epsilon + \frac{1 - \epsilon}{1 - \epsilon + n} \leq 1 - \frac{1}{2}\epsilon$ , if  $n \geq \frac{2}{\epsilon}$  (by setting  $\frac{1 - \epsilon}{1 - \epsilon + n} \leq \frac{\epsilon}{2}$ ). Hence,  $\exists c$ , such that  $(1 - \epsilon)^n + n(1 - \epsilon)^{n-1} \leq c(1 - \frac{\epsilon}{2})^n$  when  $n \geq \frac{2}{\epsilon}$ . Therefore, the performance converges exponentially.

For the uniform team, the probability of playing the action that has the highest probability in the pdf of the best agent converges to 1 as  $n \rightarrow \infty$  (List and Goodin 2001). Therefore, the performance only increases as  $n \rightarrow \infty$  if the optimal action is the one that has the highest probability.  $\square$

Now we show that we can achieve further improvement in a diverse team by breaking ties in favor of the strongest agent.

**Theorem 3.** When  $m \rightarrow \infty$ , breaking ties in favor of the strongest agent is the optimal tie-breaking rule for a diverse team.

The proof is available in the appendix (available at <http://teamcore.usc.edu/people/sorianom/aaail4-ap.pdf>). We just state here that if we break ties in favor of an agent  $s$ , the probability of voting for the optimal action is given by:

$$\tilde{p}_{best} = 1 - \prod_{i=1}^n (1 - p_{i,0}) - (1 - p_{s,0}) \left( \sum_{\substack{i=1 \\ i \neq s}}^n p_{i,0} \prod_{\substack{j=1 \\ j \neq i, j \neq s}}^n (1 - p_{j,0}) \right) \quad (2)$$

It is clear that Equation 2 is maximized by choosing agent  $s$  with the highest  $p_{s,0}$ . In our proof we show that Equation 2 is always higher than Equation 1 if  $s$  is the strongest agent. That is, we show that  $\tilde{p}_{best}$  is higher if we break ties in favor of the strongest agent than breaking ties randomly.

Next we show that with one additional assumption, not only the diverse team converges to  $\tilde{p}_{best}$ , but also  $p_{best}$  monotonically increases with  $m$ . Our additional assumption is that higher utility actions have higher probabilities, i.e., if  $U(a_j) \geq U(a_{j'})$ , then  $p_{i,j} \geq p_{i,j'}$ .

**Theorem 4.** The performance of a diverse team monotonically increases with  $m$ , if  $U(a_j) \geq U(a_{j'})$  implies that  $p_{i,j} \geq p_{i,j'}$ .

The proof of this theorem is also available in the appendix.

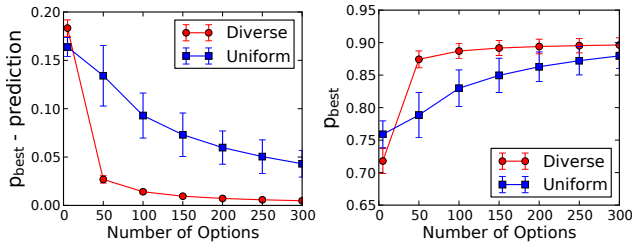
## Generalizations

In the previous theorems we focused on the probability of playing the best action, assuming that  $U(a_0) \gg U(a_j) \forall j \neq 0$ . We show now that the theorems still hold in more general domains where  $r$  actions ( $\mathbf{A}_r \subset \mathbf{A}$ ) have a significant high utility, i.e.,  $U(a_{j_1}) \gg U(a_{j_2}) \forall j_1 < r, j_2 \geq r$ . Hence, we now focus on the probability of playing any action in  $\mathbf{A}_r$ . We assume that our assumptions are also generalized, i.e.,  $p_{i,j} \geq \epsilon \forall j < r$ , and the number  $d_m$  of suboptimal actions ( $a_j, j \geq r$ ) in the  $\mathbf{D}_m$  set increases with  $m$  for  $ST$  agents.

**Theorem 5.** The previous theorems generalize to settings where  $U(a_{j_1}) \gg U(a_{j_2}) \forall j_1 < r, j_2 \geq r$ .

*Proof Sketch.* We give here a proof sketch. We just have to generate new pdfs  $p'_{i,j}$ , such that  $p'_{i0} = \sum_{j=0}^{r-1} p_{i,j}$ , and  $p'_{i,b} = p_{i,b+r-1}, \forall b \neq 0$ . We can then reapply the proofs of the previous theorems, but replacing  $p_{i,j}$  by  $p'_{i,j}$ . Note that this does not guarantee that all agents will tend to agree on the same action in  $\mathbf{A}_r$ ; but the team will still tend to pick any action in  $\mathbf{A}_r$ , since the agents are more likely to agree on actions in  $\mathbf{A}_r$  than on actions in  $\mathbf{A} \setminus \mathbf{A}_r$ .  $\square$

Now we discuss a different generalization: what happens when  $p_{i,0}$  decreases as  $m$  increases ( $\forall$  agents  $i$ ). If  $p_{i,0} \rightarrow \tilde{p}_{i,0}$  as  $m \rightarrow \infty$ , the performance in the limit for a diverse team will be  $\tilde{p}_{best}$  evaluated at  $\tilde{p}_{i,0}$ . Moreover, even if  $p_{i,0} \rightarrow 0$ , our conclusions about relative team performance are not affected as long as we are comparing two  $ST$  teams that have similar  $p_{i,0}$ : the same argument as in Corollary 1 implies that the team with faster growing  $d_m$  will perform better.



(a) Convergence of  $p_{best}$  to predicted value. (b) Actual performance value.

Figure 1: Comparing diverse and uniform when uniform also increases  $d_m$ .

## Experimental Analysis

### Synthetic Experiments

We present synthetic experiments, in order to better understand what happens in real systems. We generate agents by randomly creating pdfs and calculate the  $p_{best}$  of the generated teams. The details are available in the appendix.

As we said earlier, uniform teams composed by *NST* agents is an idealization. In more complex domains, the best agent will not behave exactly like an *NST* agent, its  $d_m$  will also increase. We perform synthetic experiments to study this situation. We consider that the best agent is still closer to an *NST* agent, therefore it increases its  $d_m$  at

a slower rate than the agents of the diverse team. We can see the average result for 200 random teams in Figure 1, where in Figure 1(a) we show the difference between the performance in the limit ( $\tilde{p}_{best}$ ) and the actual  $p_{best}(m)$  for the diverse and the uniform teams; in Figure 1(b) we show the average  $p_{best}(m)$  of the teams. As can be seen, when the best agents increase their  $d_m$  at a slower rate than the agents of the diverse team, the uniform teams converge slower to  $\tilde{p}_{best}$ . Even though they play better than the diverse teams for a small  $m$ , they are surpassed by the diverse teams as  $m$  increases. However, because  $\tilde{p}_{best}$  of the uniform teams is actually higher than the one of the diverse teams, eventually the performance of the uniform teams get closer to the performance of the diverse teams, and will be better than the one of the diverse teams again for a large enough  $m$ .

This situation is expected according to Theorem 1. If the  $d_m$  of the best agent also increases as  $m$  gets larger, the uniform team will actually behave like a diverse team and also converge to  $\tilde{p}_{best}$ .  $\tilde{p}_{best}^{uniform} \geq \tilde{p}_{best}^{diverse}$ , as the best agent has a higher probability of playing the optimal action. Hence, in the limit the uniform team will play better than the diverse team. However, as we saw in Corollary 1, the speed of con-

vergence is in the order of  $1/d_m$ . Therefore, the diverse team will converge faster, and can overcome the uniform team for moderately large  $m$ .

As Theorem 2 only holds when  $m \rightarrow \infty$ , we also explore the effect of increasing the number of agents for a large  $m$ . The  $\tilde{p}_{best}$  of a team of agents is shown as the dashed line in Figure 2. We are plotting for agents that have a probability of playing the best action of only 10%, but as we can see the probability quickly grows as the number of agents increases. We also calculate  $p_{best}$  for random teams from 2 to 6 agents (shown as the continuous line), when there are 300 available actions. As can be seen, the teams have a close performance to the expected.

### Computer Go

We show now results in a real system. We use 4 different Go software: Fuego 1.1 (Enzenberger et al. 2010), GnuGo 3.8, Pachi 9.01 (Baudiš and Gailly 2011), MoGo 4 (Gelly et al. 2006), and two (weaker) variants of Fuego (Fuego $\Delta$  and Fuego $\Theta$ ), in a total of 6 different, publicly available, agents. Fuego is considered the strongest agent among all of them. The description of Fuego $\Delta$  and Fuego $\Theta$  is available in the appendix. All our results are obtained by playing either 1000 games (to evaluate individual agents) or 2000 games (to evaluate teams), in a HP d1165 with dual dodeca core, 2.33GHz processors and 48GB of RAM. We compare results obtained by playing against a fixed opponent. Therefore, we evaluate systems playing as white, against the original Fuego playing as black. We removed all databases and specific board size knowledge of the agents, including the opponent. We call Diverse as the team composed of all 6 agents, and Uniform as the team composed of 6 copies of Fuego. Each agent is initialized with a different random seed, therefore they will not vote for the same action all the time in a given world state, due to the characteristics of the search algorithms. In the graphs we show in this section, the error bars show the confidence interval (99% of confidence).

We evaluate the performance of the teams over 7 different board sizes. We changed the time settings of individual agents as we increased the board size, in order to keep their strength as constant as possible. The average winning rates of the team members, and also the winning rates of the individual agents, is available in the appendix.<sup>1</sup>

We can see our results in Figure 3 (a). Diverse improves from 58.1% on 9x9 to 72.1% on 21x21, an increase in winning rate that is statistically significant with  $p < 2.2 \times 10^{-16}$ . This result is expected according to Theorem 1. The Uniform team changes from 61.0% to 65.8%, a statistically significant improvement with  $p = 0.0018$ . As we saw before, an increase in the performance of Uniform can also be expected, as the best agent might not be a perfect *NST* agent. A linear regression of the results of both teams gives

<sup>1</sup>In our first experiment, Diverse improved from 56.1% on 9x9 to 85.9% on 19x19. We noted, however, that some of the diverse agents were getting stronger in relation to the opponent as the board size increased. Hence, by changing the time setting to keep the strength constant, we are actually making our claims harder to show, not easier.

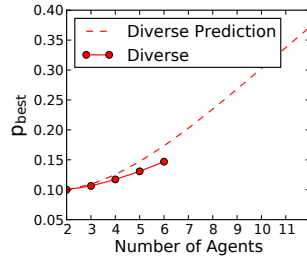


Figure 2:  $p_{best}$  of a diverse team as the number of agents increases.

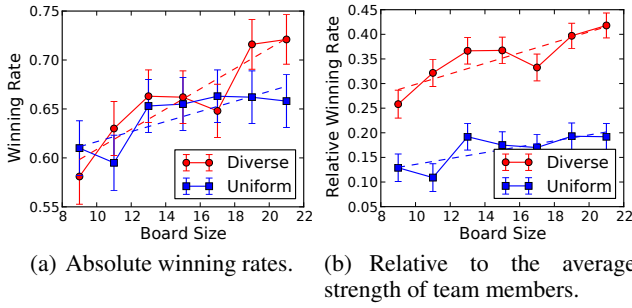


Figure 3: Winning rate in the real Computer Go system.

a slope of 0.010 for the diverse team (adjusted  $R^2$ : 0.808,  $p = 0.0036$ ) and 0.005 for the uniform team (adjusted  $R^2$ : 0.5695,  $p = 0.0305$ ). Therefore, the diverse team improves its winning rate faster than the uniform team. To check if this is a significant difference, we evaluate the interaction term in a linear regression with multiple variables. We find that the influence of board size is higher on Diverse than on Uniform with  $p = 0.0797$  (estimated coefficient of “size of the board \* group type”:  $-10.321$ , adjusted  $R^2$ : 0.7437). Moreover, on the 9x9 board Diverse is worse than Uniform ( $p = 0.0663$ ), while on the 21x21 board Diverse is better with high statistical significance ( $p = 1.941 \times 10^{-5}$ ). We also analyze the performance of the teams subtracted by the average strength of their members (Figure 3 (b)), in order to calculate the increase in winning rate achieved by “team-work” and compensate fluctuations on the winning rate of the agents as we change the board size. Again, the diverse team improves faster than the uniform team. A linear regression results in a slope of 0.0104 for Diverse (adjusted  $R^2$ : 0.5549,  $p = 0.0546$ ) and 0.0043 for Uniform (adjusted  $R^2$ : 0.1283,  $p = 0.258$ ).

We also evaluate the performance of teams of 4 agents (Diverse 4 and Uniform 4) and 6 agents (Diverse 6 and Uniform 6). For Diverse 4, we removed Fuego $\Delta$  and Fuego $\Theta$  from the Diverse team. As can be seen in Figure 4, the impact of adding more agents is higher for the diverse team in a larger board size (21x21). In the 9x9 board, the difference between Diverse 4 and Diverse 6 is only 4.4%, while in 21x21, it is 14%. Moreover, we can see a higher impact of adding agents for the diverse team, than for the uniform team. These results would be expected according to Theorem 2.

As can be seen, the prediction of our theory holds: the diverse team improves significantly as we increase the action space. The improvement is enough to make it change from playing worse than the uniform team on 9x9 to playing better than the uniform team with statistical significance on the 21x21 board. Furthermore, we show a higher impact of adding more agents when the size of the board is larger.

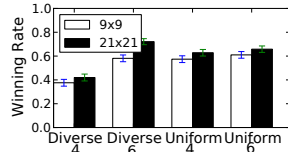


Figure 4: Winning rates for 4 and 6 agents teams.

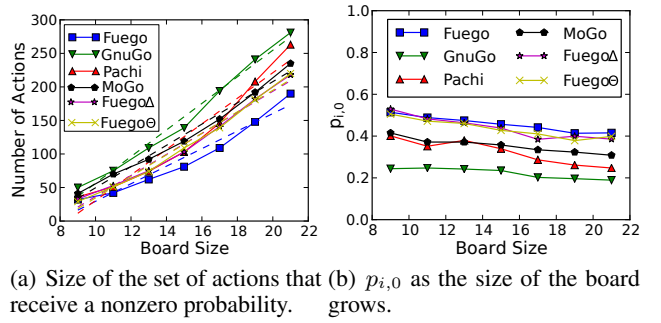


Figure 5: Verifying the assumptions in the real system.

## Analysis

To test the assumptions of our model, we estimate a pdf for each one of the agents. For each board size, and for each one of 1000 games from our experiments, we randomly choose a board state between the first and the last movement. We make Fuego evaluate the chosen board, but we give it a time limit 50x higher than the default one. Therefore, we use this much stronger version of Fuego to approximate the true ranking of all actions. For each board size, we run all agents in each board sample and check in which position of the approximated true ranking they play. This allow us to build a histogram for each agent and board size combination. Some examples can be seen in the appendix.

We study how the pdfs of the agents change as we increase the action space. Our hypothesis is that weaker agents will have a behavior closer to  $ST$  agents, while stronger agents to  $NST$  agents. In Figure 5(a) we show how many actions receive a probability higher than 0. As can be seen, Fuego does not behave exactly like an  $NST$  agent. However, it does have a slower growth rate than the other agents. A linear regression gives the following slopes: 13.08, 19.82, 19.05, 15.82, 15.69, 16.03 for Fuego, Gnugo, Pachi, Mogo, Fuego $\Delta$  and Fuego $\Theta$ , respectively ( $R^2$ : 0.95, 0.98, 0.94, 0.98, 0.98, 0.98, respectively). It is clear, therefore, that the probability mass of weak agents is distributed into bigger sets of actions as we increase the action space; and even though the strongest agent does not behave in the idealized way, it does have a slower growth rate.

We also verify how the probability of playing the best action ( $p_{i,0}$ ) changes for each one of the agents as the number of actions increase. Figure 5(b) shows that even though  $p_{i,0}$  decreases for all agents, it does not decrease much (on average, they decreased about 25% from 9x9 to 21x21).

## Conclusion

We present a new model to analyze diversity in teams. It allows us to show that the performance of diverse teams increase as the size of the action space gets larger, and also that diverse teams converge faster than uniform teams. Besides, in large action spaces the performance of a diverse team converges exponentially fast to the optimal one as the number of agents increases. Experimental results with real Computer Go agents match the predictions of our theory.

## Acknowledgments

This research was supported by MURI grant W911NF-11-1-0332. The authors would also like to thank Ariel Procaccia, Tuomas Sandholm, Scott Page, Fei Fang and Abram Demski for useful discussions.

## References

- Agmon, N., and Stone, P. 2012. Leading ad hoc agents in joint action settings with multiple teammates. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, AAMAS '12*, 341–348.
- Barrett, S.; Stone, P.; Kraus, S.; and Rosenfeld, A. 2013. Teamwork with limited knowledge of teammates. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence, AAAI'13*.
- Baudiš, P., and Gailly, J.-l. 2011. Pachi: State of the Art Open Source Go Program. In *Advances in Computer Games 13*.
- Caragiannis, I.; Procaccia, A. D.; and Shah, N. 2013. When do noisy votes reveal the truth? In *Proceedings of the fourteenth ACM conference on Electronic commerce, EC '13*, 143–160. New York, NY, USA: ACM.
- Conitzer, V., and Sandholm, T. 2005. Common voting rules as maximum likelihood estimators. In *Proceedings of the Twentieth Conference on Uncertainty in Artificial Intelligence, UAI'05*, 145–152. Morgan Kaufmann Publishers.
- Enzenberger, M.; Müller, M.; Arneson, B.; and Segal, R. 2010. Fuego - An open-source framework for board games and go engine based on Monte Carlo Tree Search. *IEEE Transactions on Computational Intelligence and AI in Games* 2(4):259–270.
- Fu, B.; Wang, Z.; Pan, R.; Xu, G.; and Dolog, P. 2012. An integrated pruning criterion for ensemble learning based on classification accuracy and diversity. In Uden, L.; Herrera, F.; Perez, J. B.; and Rodriguez, J. M. C., eds., *Proceedings of the 7th International Conference on Knowledge Management in Organizations: Service and Cloud Computing, KMO'12*, volume 172 of *Advances in Intelligent Systems and Computing*, 47–58. Springer.
- Gelly, S.; Wang, Y.; Munos, R.; and Teytaud, O. 2006. Modification of UCT with patterns in Monte-Carlo Go. Technical report, Institut National de Recherche en Informatique et en Automatique.
- Guttman, C. 2008. Making allocations collectively: Iterative group decision making under uncertainty. In Bergmann, R.; Lindemann, G.; Kirn, S.; and Pechoucek, M., eds., *Proceedings of the 6th German Conference on Multiagent System Technologies*, volume 5244 of *Lecture Notes in Computer Science*, 73–85. Kaiserslautern, Germany: Springer.
- Hong, L., and Page, S. E. 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America* 101(46):16385–16389.
- Lamberson, P. J., and Page, S. E. 2012. Optimal forecasting groups. *Management Science* 58(4):805–810.
- LiCalzi, M., and Surucu, O. 2012. The power of diversity over large solution spaces. *Management Science* 58(7):1408–1421.
- Liemhetcharat, S., and Veloso, M. 2012. Modeling and learning synergy for team formation with heterogeneous agents. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, AAMAS '12*, 365–374.
- List, C., and Goodin, R. E. 2001. Epistemic democracy: Generalizing the Condorcet Jury Theorem. *Journal of Political Philosophy* 9:277–306.
- Marcolino, L. S.; Jiang, A. X.; and Tambe, M. 2013. Multi-agent team formation: Diversity beats strength? In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence, IJCAI'13*, 279–285.
- Nair, R., and Tambe, M. 2005. Hybrid BDI-POMDP framework for multiagent teaming. *Journal of Artificial Intelligence Research* 23(1):367–420.
- Polikar, R. 2012. *Ensemble Machine Learning: Methods and Applications*. Springer. Chapter: Ensemble Learning.
- Soufiani, H. A.; Parkes, D. C.; and Xia, L. 2012. Random utility theory for social choice. In Bartlett, P. L.; Pereira, F. C. N.; Burges, C. J. C.; Bottou, L.; and Weinberger, K. Q., eds., *Proceedings of the Neural Information Processing Systems Conference, NIPS'12*, 126–134.