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

Leandro Soriano Marcolino[1], **Haifeng Xu**[1], **Albert Xin Jiang**[1], **Milind Tambe**[1], **Emma Bowring**[2]
[1]University of Southern California, Los Angeles, CA, 90089, USA
{sorianom,haifengx,jiangx,tambe}@usc.edu
[2]University of the Pacific, Stockton, CA, 95211, USA
ebowring@pacific.edu

Appendix

Model for Analysis of Diversity in Teams

We present here the proof of Theorem 3. First we restate the theorem:

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

Proof. Let s be one of the agents. If we break ties in favor of s, the probability of voting for the optimal choice will be given by:

$$\tilde{p}_{best} = 1 - \prod_{i=1}^{n}(1-p_{i0}) - (1-p_{s0})\Big(\sum_{i=1,i\neq s}^{n} p_{i0} \prod_{j=1,j\neq i,j\neq s}^{n}(1-p_{j0})\Big) \quad (2)$$

It is clear that Equation 2 is maximized by choosing agent s with the highest p_{s0}. However, we still have to show that it is better to break ties in favor of the strongest agent than breaking ties randomly. That is, we have to show that Equation 2 is always higher than Equation 1.

Equation 2 differs from Equation 1 only on the last term. Therefore, we have to show that the last term of Equation 2 is smaller than the last term of Equation 1. Let's begin by rewriting the last term of Equation 1 as: $\frac{n-1}{n}\sum_{i=1}^{n} p_{i0}\prod_{j=1,j\neq i}^{n}(1-p_{j0}) = \frac{n-1}{n}(1-p_{s0})\sum_{i=1,i\neq s}^{n} p_{i0}\prod_{j=1,j\neq i,j\neq s}^{n}(1-p_{j0}) + \frac{n-1}{n}p_{s0}\prod_{j=1,j\neq s}^{n}(1-p_{j0})$

This implies that $\frac{n-1}{n}\sum_{i=1}^{n} p_{i0}\prod_{j=1,j\neq i}^{n}(1-p_{j0}) \geq \frac{n-1}{n}(1-p_{s0})\sum_{i=1,i\neq s}^{n} p_{i0}\prod_{j=1,j\neq i,j\neq s}^{n}(1-p_{j0})$. We know that $(1-p_{s0})\sum_{i=1,i\neq s}^{n} p_{i0}\prod_{j=1,j\neq i,j\neq s}^{n}(1-p_{j0}) = \frac{n-1}{n}(1-p_{s0})\sum_{i=1,i\neq s}^{n} p_{i0}\prod_{j=1,j\neq i,j\neq s}^{n}(1-p_{j0}) + \frac{1}{n}(1-p_{s0})\sum_{i=1,i\neq s}^{n} p_{i0}\prod_{j=1,j\neq i,j\neq s}^{n}(1-p_{j0})$

Therefore, for the last term of Equation 2 to be smaller than the last term of Equation 1 we have to show that: $\frac{n-1}{n}p_{s0}\prod_{j=1,j\neq s}^{n}(1-p_{j0}) \geq \frac{1}{n}(1-p_{s0})\sum_{i=1,i\neq s}^{n} p_{i0}\prod_{j=1,j\neq s,j\neq i}^{n}(1-p_{j0})$

It follows that this equation will be true if $p_{s0} \geq (1-p_{s0})\frac{\sum_{i=1,i\neq s}^{n} p_{i0}\prod_{j=1,j\neq i,j\neq s}^{n}(1-p_{j0})}{(n-1)\prod_{j=1,j\neq s}^{n}(1-p_{j0})} \rightsquigarrow p_{s0} \geq (1-p_{s0})\frac{1}{n-1}\sum_{i=1,i\neq s}^{n}\frac{p_{i0}}{(1-p_{i0})} \rightsquigarrow \frac{p_{s0}}{(1-p_{s0})} \geq \frac{\sum_{i=1,i\neq s}^{n}\frac{p_{i0}}{(1-p_{i0})}}{n-1}$

As s is the strongest agent the previous inequality is always be true. This is because $\frac{p_{s0}}{1-p_{s0}} = \frac{\sum_{i=1,i\neq s}^{n}\frac{p_{s0}}{(1-p_{s0})}}{n-1}$ and $\frac{p_{s0}}{1-p_{s0}} \geq \frac{p_{i0}}{(1-p_{i0})}\forall i \neq s$. Therefore, it is always better to break ties in favor of the strongest agent than breaking ties randomly. □

Now we present the proof of Theorem 4:

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

Proof. Let an event be the resulted choice set of actions of these n agents. We denote by $P(V)$ the probability of occurrence of any event in V (hence, $P(V) = \sum_{v\in V} p(v)$). We call it a winning event if in the event the action chosen by plurality is action 0 (including ties). We assume that for all agents i, if $U(a_j) \geq U(a_{j'})$, then $p_{i,j} \geq p_{i,j'}$.

We show by mathematical induction that we can divide the probability of multiple suboptimal actions into a new action and $p_{best}(m+1) \geq p_{best}(m)$. Let λ be the number of actions whose probability is being divided. The base case holds trivially when $\lambda = 0$. That is, there is a new action, but all agents have a 0 probability of voting for that new action. In this case we have that p_{best} does not change, therefore $p_{best}(m+1) \geq p_{best}(m)$.

Now assume that we divided the probability of λ actions and it is true that $p_{best}(m+1) \geq p_{best}(m)$. We show that it is also true for $\lambda+1$. Hence, let's pick one more action to divide the probability. Without loss of generality, assume it is action a_{d_m}, for agent c, and its probability is being divided into action a_{d_m+1}. Therefore, $p'_{c,d_m} = p_{c,d_m} - \beta$ and $p'_{c,d_m+1} = p_{c,d_m+1} + \beta$, for $0 \leq \beta \leq p_{c,d_m}$. Let $p_{best}^{after}(m+1)$ be the probability of voting for the best action after this new division, and $p_{best}^{before}(m+1)$ the probability before this new division. We show that $p_{best}^{after}(m+1) \geq p_{best}^{before}(m+1)$.

Let Γ be the set of all events where all agents voted, except for agent c (the order does not matter, so we can consider agent c is the last one to post its vote). If $\gamma \in \Gamma$ will be a winning event no matter if agent c votes for a_{d_m} or a_{d_m+1}, then changing agent c's pdf will not affect the probability of these winning events. Hence, let $\Gamma' \subset \Gamma$ be the set of all events that will become a winning event depending if agent c does not vote for a_{d_m} or a_{d_m+1}. Given that $\gamma \in \Gamma'$ already happened, the probability of winning or losing is equal to the probability of agent c not voting for a_{d_m} or a_{d_m+1}.

Now let's divide Γ' in two exclusive subsets: $\Gamma_{d_m+1} \subset \Gamma'$, where for each $\gamma \in \Gamma_{d_m+1}$ action a_{d_m+1} is in tie with action a_0, so if agent c does not vote for a_{d_m+1}, γ will be a winning event; $\Gamma_{d_m} \subset \Gamma'$, where for each $\gamma \in \Gamma_{d_m}$ action a_{d_m} is in tie with action a_0, so if agent c does not votes for a_{d_m}, γ will be a winning event. We do not consider events where both a_{d_m+1} and a_{d_m} are in tie with a_0, as in that case the probability of a winning event does not change (it is given by $1 - p'_{c,d_m} - p'_{c,d_m+1} = 1 - p_{c,d_m} - p_{c,d_m+1}$).

Note that for each $\gamma \in \Gamma_{d_m+1}$, the probability of a winning event equals $1 - p'_{c,d_m+1}$. Therefore, after changing the pdf of agent c, for each $\gamma \in \Gamma_{d_m+1}$, the probability of a winning event decreases by β. Similarly, for each $\gamma \in \Gamma_{d_m}$, the probability of a winning event equals $1 - p'_{c,d_m}$. Therefore, after changing the pdf of agent c, for each $\gamma \in \Gamma_{d_m}$, the probability of a winning event increases by β.

Therefore, $p^{after}_{best}(m+1) \geq p^{before}_{best}(m+1)$ if and only if $P(\Gamma_{d_m}) \geq P(\Gamma_{d_m+1})$. Note that $\forall \gamma \in \Gamma_{d_m+1}$ there are more agents that voted for a_{d_m+1} than for a_{d_m}. Also, $\forall \gamma \in \Gamma_{d_m}$ there are more agents that voted for a_{d_m} than for a_{d_m+1}. If, for all agents i, $p_{i,d_m} \geq p_{i,d_m+1}$, we have that $P(\Gamma_{d_m}) \geq P(\Gamma_{d_m+1})$. Therefore, $p^{after}_{best}(m+1) \geq p^{before}_{best}(m+1)$, so we still have that $p_{best}(m+1) \geq p_{best}(m)$. Also note that for the next step of the induction be valid, so that we can still divide the probability of one more action, it is necessary that $p'_{c,d_m} \geq p'_{c,d_m+1}$. $\square$

Experimental Analysis

Synthetic Experiments We show here more details about the synthetic experiments. We use a uniform distribution to generate all random numbers. When creating a pdf, we rescale the values assigned randomly, so that the overall sum of the pdf is equal to 1. In the experiments presented in Figure 1, we used teams of 4 agents. For each agent of the diverse team, $p_{i,0}$ is chosen uniformly random between 0.6 and 0.7. The remaining is distributed randomly from 10% to 20% of the next best actions (the number of actions that will receive a positive probability is also decided randomly). For the uniform team, we make copies of the best agent (with highest $p_{i,0}$) of the diverse team, but distribute the remaining probability randomly from 1% to 3% of the next best actions. In the experiment shown in Figure 2 each agent had a probability of playing the best action of 10%, and the remaining probability was randomly distributed over the 10% next best actions.

Computer Go Experiments We used modified versions of Fuego, called FuegoΔ and FuegoΘ in the Computer Go experiments. Fuego is an implementation of the UCT Monte Carlo Go algorithm, therefore it uses heuristics to simulate games in order to evaluate board configurations. Fuego uses mainly 5 heuristics during these simulations, and they are executed in a hierarchical order. The original Fuego agent follows the order <Atari Capture, Atari Defend, Lowlib, Pattern> (The heuristic called Nakade is not enabled by default). Our variation called FuegoΔ follows the order <Atari Defend, Atari Capture, Pattern, Nakade, Lowlib>, while FuegoΘ follows the order <Atari Defend, Nakade, Pattern, Atari Capture, Lowlib>. Also, FuegoΔ and FuegoΘ have half of the memory available when compared with the original Fuego.

In Table 2 we show the winning rates of each one of the agents, as we increase the board size. In Table 3 we see the average winning rates of the team members.

Agent	9x9	11x11	13x13	15x15
Fuego	48.1%	48.6%	46.1%	48 %
GnuGo	1.1%	1.1%	1.9%	1.9%
Pachi	25.7%	22.9%	25.8%	26.9%
MoGo	27.6%	26.4%	22.7%	22 %
FuegoΔ	45.7%	45.8%	42.2%	40.4%
FuegoΘ	45.5%	40.2%	39.2%	37.6%

Agent	17x17	19x19	21x21
Fuego	49.3%	46.9%	46.6%
GnuGo	4.5%	6.8%	6.1%
Pachi	23.5%	20.8%	11 %
MoGo	27.1%	30.1%	27.1%
FuegoΔ	43 %	44.5%	47.4%
FuegoΘ	41.8%	42.3%	43.6%

Table 2: Winning rates of each one of the agents

Team	9x9	11x11	13x13	15x15
Diverse	32.2%	30.8%	29.6%	29.4%
Uniform	48.1%	48.6%	46.1%	48.0%

Team	17x17	19x19	21x21
Diverse	31.5%	31.9%	30.3%
Uniform	49.3%	46.9%	46.6%

Table 3: Average winning rates of the individual team members across different board sizes.

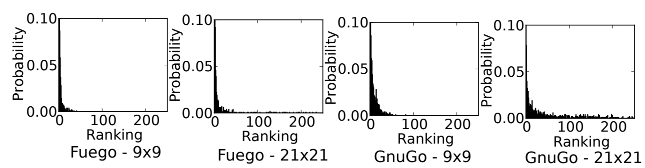

Figure 6: Histograms of agents for different board sizes.

Analysis We can see some of the generated histograms in Figure 6. We can see that a strong agent, like Fuego, has most of its probability mass in the higher ranked actions, while weaker agents, like GnuGo, has the mass of its pdf distributed over a larger set of actions, creating a larger tail. Moreover, the probability mass of GnuGo is spread over a larger number of actions when we increase the size of the board.